UNIVERSITY OF CALIFORNIA, SAN DIEGO

The Russian Option in a Jump-Diffusion Model

A dissertation submitted in partial satisfaction of the
requirements for the degree
Doctor of Philosophy

in

Mathematics

by

Michael Scullard

Committee in charge:

Professor Patrick Fitzsimmons, Chair
Professor Graham Elliott
Professor Ian Galton
Professor Jason Schweinsberg
Professor Ruth Williams

2011

The dissertation of Michael Scullard is approved, and it is acceptable in quality and form for publication on microfilm and electronically:

Chair

University of California, San Diego

2011

DEDICATION

To my family.

TABLE OF CONTENTS

ACKNOWLEDGEMENTS

Foremost, I would like to thank my advisor Patrick Fitzsimmons, without whose help this thesis would not be possible. Your help and guidance over the past three years have been indispensable.

I'd like to thank the many great mathematics teachers I have had over the years, and in particular those who have trained me to be a probabilist: Professor Driver, Professor Saloff-Coste, Professor Schweinsberg and Professor Williams.

Finally, I would like to thank my mom, my dad, my sister Kim and my brother Mark for their love, encouragement, and patience during all the times I struggled during my studies.

VITA

2004	B. S. in Mathematics, Cornell University
2006	M. S. in Mathematics, University of California, San Diego
2011	Ph. D. in Mathematics, University of California, San Diego

ABSTRACT OF THE DISSERTATION

The Russian Option in a Jump-Diffusion Model

by

Michael Scullard

Doctor of Philosophy in Mathematics

University of California, San Diego, 2011

Professor Patrick Fitzsimmons, Chair

The Russian option is a lookback option which pays the maximum-to-date of the underlying, subject to some discounting factor. In this thesis we examine the properties of the value function for the Russian option problem in a jump-diffusion model, generalizing the results of Peskir [25]. In particular, we use the theory of viscosity solutions to show that the value function is smooth inside the continuation region. Furthermore, we show that optimal stopping boundary can be characterized as the unique solution to a free boundary problem under the same assumptions given in Pham [28] for the American put option.

Chapter 1

Introduction

The subfield of Mathematical Finance concerned with the pricing of derivatives has inspired a number of interesting problems in probability. Although the foundations of the field go back to the beginning of the twentieth century with Louis Bachelier's thesis *The Theory of Speculation* [3], the seminal work concerning options pricing was the 1973 paper *The Pricing of Options and Corporate Liabilities* [4] by Fischer Black and Myron Scholes. Robert C. Merton soon afterward expanded on their work, publishing the paper *Theory of Rational Option Pricing* [22].

The market model introduced by Black, Scholes, and Merton assumes the returns of the risky asset in the model (e.g. a stock or index) follow a Wiener process, so that the risky asset S_t is given by the geometric Brownian motion solving the stochastic differential equation

$$\frac{dS_t}{S_t} = \mu\, dt + \sigma\, dW_t, \tag{1.1}$$

where W_t is a Brownian motion and μ and σ are constants. Note that an important consequence of this model is that the risky asset's path is almost surely continuous. The Black-Scholes model given by (1.1) remains the dominant model even today, but it is widely accepted that actual market returns do not follow a log-normal distribution, as is implied by (1.1). Immediately obvious is the fact that prices are discrete, so that any movement is given by (possibly small) jumps and hence cannot be continuous. This is insignificant, however, when one considers that market returns have been shown empirically to have heavy tails (see Cont [9]), which cannot be modeled with a log-normal distribution. In particular, the geometric Brownian model performs well as a model most of the time, but it is unable to produce the large price shifts which occur during market crashes.

Another issue with using a diffusion process to simulate a risky asset is that such a model is unable to replicate the *volatility smile* which is observed in markets. To see what is meant by this, first observe that for any option price which is strictly increasing as a function of volatility (as is the case for vanilla options) there will be a bijection between the price of the option and the volatility if we hold all other variables and parameters fixed. Hence given the current market value of the option, we can invert the formula to obtain the *implied volatility*, which is the unique volatility that gives us the market price when substituted into the pricing function. The Black-Scholes model given by (1.1) assumes that the volatility σ is constant, which should imply a constant implied volatility. In particular, we should obtain the same implied volatility regardless of the strike price. However, it is well known that implied volatility depends heavily on the strike; for example, equity options typically display a downward sloping graph when implied volatility is plotted against strike price. This presents a problem since the volatility implied by commonly traded, vanilla options is frequently used to price exotic options which are not available on the open market.

A third issue with using a diffusion model is that such a model implies that the market is *complete*, which means any derivative product in that market can be perfectly hedged. One consequence of this is that all pricing is done under a unique risk-neutral measure, and so the price of the derivative is also unique. Ostensibly this is a desirable property, since otherwise the price depends on the choice of risk-neutral measure and so one needs to establish criteria for choosing the measure. However, market completeness is not particularly realistic. Options cannot be perfectly hedged in the real world, and diffusion models tend to understate the amount of risk inherent in pricing an option. See Cont and Tankov [8] for further discussion of the limitations of diffusion models.

As a result of these issues, there has recently been interest in using Lévy processes to price derivatives. Lévy processes offer a generalization of Brownian motion, and in particular allow jumps to occur. In general Lévy models, the jumps permit sudden market movements, giving rise to heavy tails for asset returns and resulting in volatility smiles. In addition, markets are no longer complete in a Lévy framework.

In this thesis, we consider option pricing with the class of Lévy processes known as *jump diffusions*, which are the sum of a Brownian motion with drift and a compound Poisson process. One can think of the Brownian motion as replicating everyday price movements of the asset, and compound Poisson process corresponding to rare events, where the price experiences a sudden large movement. In particular we will look at pricing a specific type of option, known as a Russian option, with a finite expiration date. A Russian option is a particular type of look-back option which pays out the higher of the to-date maximum of the underlying risky asset and some fixed amount m, subjected to some discounting factor. It is an *American style* option, which means it can be exercised at any time up to the expiry date.

The Russian option was first introduced by Shepp and Shiryaev [34], who priced the perpetual Russian option in a diffusion model. Soon afterward, the same authors published [33], where they considered a new method of pricing which involved reducing the value function from a two dimension optimal stopping time problem to a one-dimensional problem. The finite horizon case for the diffusion model was examined by Duistermatt et al. [14], Erkstrom [15], and Peskir [25], who published their works independently at about the same time. Like the finite time horizon American option, there is no known closed form for price of a Russian Option, even in a diffusion model.

For jump processes, perpetual Russian options were examined separately by Gapeev [17], Gerber et al. [18], and Mordecki and Moreira [23] for the case of compound Poisson processes where the jump sizes are given by mixed exponentials. Avram [2] considered the Russian option for the case of general spectrally negative Lévy processes in the perpetual case. Asmussen [1] considered the perpetual Russian option in the case of exponential phase-type Lévy models, where the jumps are no longer required to be one sided.

In this thesis, we apply the arguments used in Peskir [25] and those in Pham [28] for American options to price a finite horizon Russian option in a jump-diffusion model with general jump distributions. The outline of the thesis is as follows. In Chapter 2, we will introduce the background material, assumptions, and notation necessary for the rest of the paper. In addition, we will work through the change of measure necessary to translate from the real world measure to a given risk-neutral measure, as performed in Pham [28]. This will allow us to characterize the set of risk-neutral measures with which we will be working. In Chapter 3, we show how performing another change of measure converts the problem from a two-dimensional optimal stopping problem to a one-dimensional problem. We conclude the chapter by examining some of the properties of the new process which results from this measure change.

Chapter 4 looks at some properties of the value function $V(t, x)$ in the reformulated problem. In particular, we prove continuity and growth conditions for $V(t, x)$ and show the existence of an early exercise boundary $b(t)$. In Chapter 5, we discuss the variational inequality associated with the Russian option problem. This variational inequality must be interpreted in a weak sense since it not known that a smooth solution exists. To this end, we introduce the concept of viscosity solutions and show that $V(t, x)$ is a viscosity solution of the inequality. In Chapter 6, we use this fact and uniqueness results for viscosity solutions to show our main result that the value function is smooth inside the continuation region.

Finally, in Chapters 7 and 8 we further examine the behavior of $V(t, x)$ and the boundary function $b(t)$. In particular, in Chapter 7 we show the continuity of the boundary function under the same assumptions made in Pham [28]. In order to prove the main results of Chapter 8, we must first develop a generalization of Itô's formula which is applicable to our problem. Using this generalization, under the same assumption as in Chapter 7 we prove the uniqueness of the Russian option as a solution to a free boundary problem. Additionally, we derive an early exercise premium representation for the Russian option problem.

Chapter 2

Preliminaries

In this chapter we establish the framework needed to understand the Russian option, including the necessary notation and motivation. Throughout this and later chapters we assume that we are given a probability space $(\Omega, \mathcal{F}, P)$, where Ω is the state space, $\{\mathcal{F}_t : t \geq 0\}$ is a filtration satisfying the usual conditions (see Rogers and Williams [30] Vol 2, pg 172), and P is a probability measure.

2.1 Lévy Processes

We begin by recalling the definition of a Lévy process:

Definition 1. *A* Lévy Process *is a stochastic process which satisfies the following properties:*

1. *The increments of X are stationary. That is, for* $0 \leq s < t$, $X_t - X_s$ *is equal in distribution to* X_{t-s}.
2. *The increments of X are independent. That is, for* $0 \leq t_0 < t_1 < ... < t_{n-1} < t_n$, *the increments* $X_{t_1} - X_{t_0}, X_{t_2} - X_{t_1}, ..., X_{t_n} - X_{t_{n-1}}$ *are independent.*
3. $X_0 = 0$ *almost surely.*

Recall that a function is called *cadlag* if it is right continuous with left limits. It is well known that any Lévy process X_t admits a version such that $t \to X_t$ is cadlag;

hence from now on, we will assume without loss of generality that all Lévy processes we encounter are cadlag.

Two particularly important Lévy processes are Brownian motion and the compound Poisson process. Brownian motion is the only continuous stochastic process which is also a Lévy process, and for a given measure P we will let W^P denote a standard Brownian motion under P. We define Y^P to be a compound Poisson process under P, independent of W^P, so that $Y_t^P := \sum_{i=1}^{N_t} Z_n$. Here N is a Poisson process with intensity λ under P and $\{Z_i\}_{i=1}^{\infty}$ is a sequence of independent, identically distributed random variables, independent of N, with probability measure m^P under P. Recall that a compound Poisson process can be thought of as a continuous time random walk with random jump sizes, which for Y^P are given by the random variables Z_i. The jumps occur at random times, and we will denote the time of the nth jump by T_n. Hence the Poisson process N_t equals the number of jumps which occur up to time t.

Next, we recall the notion of a Poisson random measure:

Definition 2. *Let E be a Borel subset of $\mathbb{R}^d$, let $\mathcal{E}$ be the set of Borel subsets of E, and let μ be a sigma-finite measure on E. A function $M : \Omega \times \mathcal{E} \to \mathbb{Z}_+$ is called a Poisson random measure with intensity measure μ if*

1. *For every fixed $A \in \mathcal{E}$, $\omega \to M(\omega, A)$ is a Poisson random variable with mean $\mu(A)$.*
2. *For almost every fixed $\omega \in \Omega$, $A \to M(\omega, A)$ is a measure on $(E, \mathcal{E})$.*
3. *If $A_1, ..., A_n \subset \mathcal{E}$ are disjoint, the random variables $M(A_1), ..., M(A_n)$ are independent.*

We define $\nu(dt, dz)$ to be the Poisson random measure on $\mathbb{R}_+ \times \mathbb{R}$ associated to the Marked point process (T_n, Z_n) mentioned above, so that for $A \subset \mathbb{R}_+$ and $B \subset \mathbb{R}$, ν is given by

$$\nu(\omega, A \times B) = \sum_{n \geq 1} \delta_{(T_n(\omega), Z_n(\omega))}(A \times B),$$

where δ denotes the Dirac delta function. Hence for fixed ω, ν is a counting measure such that $\nu(A \times B)$ equals the number of jumps of Y_t which occur at a time in the set A and whose size is in the set B. Finally, we remark that the intensity measure of ν under

P is a measure of the form $q^P(dt, dz) = \lambda^P\, m^P(dz)\, dt$, where m^P is a probability measure. Thus the compensated Poisson random measure under P is given by

$$\tilde{\nu}^P(dt, dz) = \nu(dt, dz) - q^P(dt, dz),$$

which is a martingale measure. The constant λ^P, known as the jump intensity of Y_t, controls the frequency of jumps while m^P controls the size of the jumps.

2.2 The Market Model

We will assume that we have a financial market consisting of two continuously traded assets given by the pair (B, S). Here B represents a riskless asset, such as a bond, which satisfies the differential equation

$$\frac{dB_t}{B_t} = r\, dt,$$

with constant interest rate r. Hence the value of B at time t is given by $B_t = B_0 e^{rt}$. The asset S is a risky asset, such as a stock, which follows a geometric jump diffusion process given by the stochastic differential equation

$$\frac{dS_t}{S_{t-}} = \mu\, dt + \sigma\, dW_t^P + \int_{\mathbb{R}} \gamma(z)\, \tilde{\nu}^P(dt, dz). \tag{2.1}$$

The above is shorthand for

$$S_t = \int_0^t \mu S_{u-}\, du + \int_0^t \sigma S_{u-}\, dW_u^P + \int_0^t \int_{\mathbb{R}} S_{u-} \gamma(z)\, \tilde{\nu}^P(du, dz),$$

where the final two terms are integration against Brownian motion and the compensated Poisson random measure, respectively. For details about stochastic integration, see Chung and Williams [6]. In (2.1), μ denotes the drift and σ the volatility of the risky asset, both of which we take to be constant. Further, we will make the assumption that $\sigma > 0$. The function $\gamma(z)$ denotes the relative jump size of S, so that if a jump occurs at time t we have

$$S_t = S_{t-} + S_{t-}\gamma(Z_{N(t)}).$$

We assume $\gamma \in L^2(m^P)$ and that $1 + \gamma(z) > 0$ for all z. This last assumption is necessary for the riskless asset S to be positive at all times.

Before we proceed, let us recall Itô's Lemma for semi-martingales, which tells us that, for $f \in C^{1,2}$,

$$\begin{aligned} &f(t,S_t) - f(0,S_0) \\ &\quad = \int_0^t \frac{\partial f}{\partial s}(u,S_{u-})\,du + \int_0^t \frac{\partial f}{\partial x}(u,S_{u-})\,dS_u + \frac{1}{2}\int_0^t \frac{\partial^2 f}{\partial x^2}(u,S_{u-})\,d[S,S]^c_u \\ &\quad + \int_0^t \int_{\mathbb{R}} f(u,S_{u-}+\Delta S_u) - f(u,S_{u-}) - \Delta S_u \frac{\partial f}{\partial x}(u,S_{u-})\,\nu(du,dz). \end{aligned} \tag{2.2}$$

where $[S,S]^c$ is the continuous part of the quadratic variation of S. When the jumps of the semimartingale S satisfy $\sum_{0<s\le t} |\Delta S_s| < \infty$, as is the case for our process, then the formula can be written in the equivalent form

$$\begin{aligned} &f(t,S_t) - f(0,S_0) \\ &\quad = \int_0^t \frac{\partial f}{\partial s}(u,S_{u-})\,du + \int_0^t \frac{\partial f}{\partial x}(u,S_{u-})\,dS^c_u + \frac{1}{2}\int_0^t \frac{\partial^2 f}{\partial x^2}(u,S_{u-})\,d[S,S]^c_u \\ &\quad + \int_0^t \int_{\mathbb{R}} f(u,S_{u-}+\Delta S_u) - f(u,S_{u-})\,\nu(du,dz). \end{aligned} \tag{2.3}$$

We will use the notation $S_t(s_0)$ to denote the process S_t with $S_0 = s_0$. Thus, applying Itô's formula with $f(x) = \log(x)$, we see that the solution of (2.1) can be represented explicitly by the equation

$$S_t(s_0) = s_0 \exp\left(\left[\mu - \frac{1}{2}\sigma^2 - \lambda^P k^P\right] t + \sigma W^P_t + \int_0^t \int_{\mathbb{R}} \ln(1+\gamma(z))\,\nu(du,dz)\right),$$

where we define $k^P := \int_{\mathbb{R}} \gamma(z)\, m^P(dz)$.

2.3 Derivatives and Risk Neutral Measures

In finance, a derivative product is a financial instrument whose value is derived from an underlying asset. An option is a specific type of derivative; it gives the purchaser of the option the right, but not the obligation, to either purchase or sell the underlying asset at some future time for a particular price. While there are many characteristics which distinguish different options, two particularly important ones are the *option style* and *option type*. The option style typically refers to when the option can be exercised, with the two most common styles being European and American. European options can be executed only at some fixed expiration time, whereas American options can be exercised any time up to expiry.

Option type, on the other hand, typically refers to the payoff of the option. Call and put options are two simple examples of option types. A *European call*, for instance, gives its holder the right to buy the underlying asset (such as a stock) for some fixed price K at expiration time T. Conversely, a *European put* gives its holder the right to sell the underlying for a fixed price K at expiration time T. In both these examples, K is known as the *strike price*. Hence the European call expiring at time T has a payoff of

$$g(S_T) = (S_T - K)_+ = \max\{S_T - K, 0\}$$

since the holder will only execute the option if the asset price is above the strike. On the other hand, the European put expiring at time T has a payoff of

$$g(S_T) = (K - S_T)_+ = \max\{K - S_T, 0\}$$

since the holder will only execute the option if the asset price is below the strike. In general, derivative products exist for a wide variety of payoff functions.

2.4 Arbitrage Free Pricing

Of fundamental importance in mathematical finance is the concept of arbitrage. An arbitrage is a trading strategy which requires no initial investment, always has non-negative value, and has a positive probability of being greater than zero. Intuitively, we can think of an arbitrage as "money for nothing". One of the major assumptions in mathematical finance is that arbitrage does not exist, as any observed arbitrage opportunities will quickly be taken advantage of by market participants. The laws of supply and demand will soon change the asset values until the arbitrage possibilities are removed.

As mentioned above, we assume that the risky asset has price dynamics determined by the probability measure P. The measure P is often called the "real world measure", since it determines the behavior of the risky asset that is seen by market participants. When pricing an option however, it is convenient to use what is known as an equivalent martingale measure:

Definition 3. *A probability measure Q is called an equivalent martingale measure (EMM) with respect to the measure P if*

1. *The measures P and Q have the same sets of measure zero.*
2. *The discounted risky asset price, $e^{-rt}S_t$, is a martingale under Q.*

An EMM Q is often called a risk neutral measure, since under Q the expected value of investing some amount S_0 in the risky asset is equal to investing S_0 in the riskless asset. An investor who only cares about the expected return of an investment and not the risk involved will be indifferent to choosing between the risky asset and the riskless asset when the dynamics of the risky asset are determined by Q.

It is well known (see Shreve [32]) that the arbitrage free price for a European option with payoff function $g(x)$ expiring at time T is given by the discounted final payoff of the option under a given risk-neutral measure; that is, for an option on an underlying with initial value s purchased at time 0 and expiring at time T, possible prices are given by

$$C(s) = E^Q[e^{-rT}g(S_T(s))], \tag{2.4}$$

where Q is some risk-neutral measure. Let us discuss heuristically the idea behind (2.4) in the diffusion case (for an argument in the jump diffusion case, see Cont and Tankov[8], Chapter 9). By the fundamental theorem of arbitrage-free pricing, market completeness in the diffusion case implies that every derivative can be hedged with a self-financing trading strategy with value H_t at time t – that is, a portfolio consisting of some amount of (possibly shorted) risky and riskless asset that requires no further capital injection aside from the starting capital. The fact that Q is a risk-neutral measure implies that the discounted value of H is a $\mathcal{F}_t$-martingale and so letting E_s^Q denote the expectation under Q when $S_0 = s$ we have

$$H_0 = E_s^Q[e^{-rT}H_T] = E_s^Q[e^{-rT}g(S_T)].$$

For the final equality we used the fact that $H_T = g(S_T)$ since H is a hedging strategy for the derivative with payoff $g(x)$. Hence we define H_0, the amount of capital necessary to fund our hedging strategy, to be the price of the option, as any other choice would lead to arbitrage due to our ability to perfectly hedge the option.

In a diffusion model, there exists a unique risk neutral measure and hence the arbitrage free price of a European option must be unique as well. In more general models, such as Lévy models and stochastic volatility models, the risk neutral measure need not be unique, so there does not exist a unique arbitrage free price. Instead, when pricing an option we are required to choose a particular risk-neutral measure. A wide variety of methods have been devised for doing this; for an overview of these methods, see Chapter 10 of Cont and Tankov [8].

2.5 American Style Options

The European options discussed above give the holder the right to execute the option at a fixed time T. In contrast, a finite time American option is one which allows the holder the right to execute the option at any point up to time T. A perpetual American option is one which does not expire, so that $T = \infty$. Since the payoff of an American style option depends on the time that the holder chooses to execute the option, it can no longer be perfectly replicated. Instead, the seller of the option should be able to find a *super-replicating porfolio*. A super-replicating portfolio is a trading strategy in the risky and riskless asset which is always worth more than the option payoff, regardless of which time the seller chooses to execute the option. Thus the price of an American style option is given by a European option expiring at the time with highest possible value. More formally, an arbitrage free price of an American option with payoff function $g(x)$, initial risky asset price $S_0 = s$, and time remaining until expiry t is given by

$$W(t,s) = \sup_{0 \le \tau \le t} E_s^Q[e^{-r\tau} g(S_\tau)] = \sup_{0 \le \tau \le t} E^Q[e^{-r\tau} g(S_\tau(s))], \tag{2.5}$$

where τ is required to be a stopping time and Q is some equivalent martingale measure. As mentioned in the European option case above, the price will not be unique if more than one equivalent martingale measure exists.

2.6 Optimal Stopping Time Problems

The pricing of an American style option is a specific example of what is known as an *optimal stopping problem*, and the function $W(t, s)$ in (2.5) is known as the *value* *function* for that problem. In a typical one-dimensional, finite time optimal stopping problem, we are given a *gains* function $f : [0,T] \times \mathbb{R}_+ \to \mathbb{R}$ and a value function of the form

$$W(t,x) = \sup_{0 \le \tau \le t} E_x[f(\tau, X_\tau)],$$

where X is some stochastic process with state space $\mathbb{R}$. We can think of such a problem as though, given some process X, and we are allowed to stop X at any time τ, and our goal is to maximize the expected value of $f(\tau, X_\tau)$. The fact that we require τ to be a stopping time means that our decision to stop can depend on the behavior of the process, but only up to the current time; we are not allowed to "see the future" of the process. A solution to such a problem involves finding the function $W(t,x)$ and an optimal stopping time $\tau^*(t,x)$ with $0 \le \tau^*(t,x) \le t$ such that $W(t,x) = E_x^Q[f(\tau^*, X_{\tau^*})]$. For a given optimal stopping time problem to be well defined, it is enough for

$$E\left[\sup_{0 \le t \le T} |f(t, X_t)|\right] < \infty$$

to hold (see Peskir [26], Section I.2.1).

When the process X is a Markov process, our decision to continue or stop the process at time t depends only on the current value of X_t. In this case, $[0,T] \times \mathbb{R}_+$ can be broken up into the set of points where the process should be continued, known as the continuation region C, and the set of points where the process should be stopped, known as the stopping region D. Hence the optimal stopping time τ^* can be written as

$$\tau^*(t,x) = \inf\{t \ge s \ge 0 : (s, X_s(x)) \notin C\}.$$

Furthermore, it is well known (see Peskir [26], Section I.2.2) that in the case when X is a Markov process, the continuation and stopping regions can be written as

$$C = \{(t,x) : W(t,x) > f(t,x)\}$$
$$D = \{(t,x) : W(t,x) = f(t,x)\}.$$

Recall that a measurable function $F(t,x) : [0,T] \times \mathbb{R}$ is called *superharmonic* for the process X if $E_{(t,x)}[F(\tau, X_\tau)] \le F(t,x)$ for all stopping times τ and $(t,x) \in [0,T] \times \mathbb{R}$. It is a result from Optimal Stopping Theory (Peskir [26], Section I.2.2) that $W(t,x)$ is the smallest superharmonic function which dominates the gains function f. Furthermore, the process X, stopped upon exiting the continuation region C, is a martingale. This implies that $\mathcal{L}W = 0$ in C, where $\mathcal{L}$ is the infinitesimal operator of X. Assuming W is smooth, these properties result in the following two conditions:

$$\begin{aligned} &\mathcal{L}W \leq 0 \ (W \text{ superharmonic}) \\ &W \geq f \quad (W > f \text{ on } C \text{ and } W = f \text{ on } D). \end{aligned}$$

Note that these facts can be combined to give the variational inequality

$$\min\{-\mathcal{L}W, W - f\} = 0. \tag{2.6}$$

As an example, consider the diffusion case, where the process X solves the SDE given by $dX_t = \mu(t, X_t)\,dt + \sigma(t, X_t)\,dW_t$. Then the operator $\mathcal{L}$ is the second order parabolic partial differential operator

$$\mathcal{L}g(t,x) = \frac{\partial g}{\partial t}(t,x) + \mu(t,x)\frac{\partial g}{\partial x}(t,x) + \frac{1}{2}\sigma(t,x)^2\,\frac{\partial^2 g}{\partial x^2}(t,x)$$

for any $g \in C^{1,2}$, the set of all functions $g(t, x)$ which are once differentiable in t and twice differentiable in x. If we let $g(t, x) = e^{-rt}W(t, x)$, the discounted price of an American style option with value function $W(t, x)$, we obtain

$$\begin{aligned} \mathcal{L}g(t,x) = e^{-rt}\bigg[&-rW(T-t,x) - \frac{\partial W}{\partial t}(T-t,x) + \mu(T-t,x)\frac{\partial W}{\partial x}(T-t,x) \\ &+ \frac{1}{2}\sigma(T-t,x)^2\frac{\partial^2 W}{\partial x^2}(T-t,x)\bigg], \end{aligned}$$

As mentioned above, in the continuation region C we have $\mathcal{L}g(t, x) = 0$. Observe that if we let $V(t, x) := W(T - t, x)$ denote the price of the option at time t, then this fact implies that V solves the PDE

$$\frac{\partial V}{\partial t}(t,x) + \mu(t,x)\frac{\partial V}{\partial x}(t,x) + \frac{1}{2}\sigma(t,x)^2\frac{\partial^2 V}{\partial x^2}(t,x) = rV(t,x),$$

which is the classical Black-Scholes equation.

From the above facts, we see that we need to find a function $W(t, x)$ which solves

$$\begin{aligned} &\mathcal{L}W(t,x) = 0 \text{ for } (t,x) \in C \\ &W(t,x) > f(t,x) \text{ for } (t,x) \in C \\ &W(t,x) = f(t,x) \text{ for } (t,x) \in D. \end{aligned} \tag{2.7}$$

The system given by (2.7) is an example of what is known as a free boundary problem, since the boundary of C is initially unknown and typically evolves with time. Additional conditions are needed to be able to determine a unique solution, and they are obtained from the specific payoff function $f(t, x)$.

2.7 Definition of the Russian Option

First introduced by Shepp and Shiryaev [34], a Russian option is an American style option which, when executed, pays the higher of the maximum to-date stock price and some fixed constant, subject to a discounting factor. In particular, the payoff function is given by

$$g(t, m, s_0) = e^{-\lambda t} M_t(m, s_0),$$

where we define $S_t^*(s_0) := \sup_{0 \le u \le t} S_u(s_0)$ and $M_t(m, s_0) := S_t^*(s_0) \vee m$, and $\lambda > 0$ is the discounting factor. Here the notation $a \vee b$ denotes the maximum of a and b. Thus we are presented with an optimal stopping problem with value function given by

$$W(t, m, s_0) = \sup_{0 \le \tau \le t} E^Q_{m,s_0}[e^{(-r+\lambda)\tau} M_\tau(m, s_0)], \tag{2.8}$$

where $t \in [0, T]$ is the amount of time remaining before the option must be exercised, $s_0 \in (0, \infty)$ is the initial risky asset price at time 0 (so that $S(0) = s_0$) and $m \ge s_0$. Similarly to before, the notation E^Q_{m,s_0} denotes the expectation taken under an equivalent martingale measure Q with $S_0 = s_0$ and $M_0 = m$.

2.8 The Equivalent Martingale Measure

It is well known that in the Lévy process framework, there are infinitely many possible equivalent martingale measures (see Cont and Tankov [8], Chapter 9). A characterization of EMMs in terms of their Radon-Nykodym densities is given by Colwell and Elliott[7], pg. 298 or Jacod and Shiryaev [19] , Chapter III:

Proposition 1. *Suppose we are given a probability space $(\Omega, \mathcal{F}, P)$ such that W^P is a Brownian motion and ν is a Poisson random measure, independent of W^P, with intensity* *measure $\lambda^P m^P(dz)\,dt$. Let $\xi_t(x)$ be a process of the form*

$$\xi_t(x) = x + \int_0^t \mu(u, \xi_{u-}(x))\,du + \int_0^t \sigma(u, \xi_{u-}(x))\,dW_u + \int_0^t \int_{\mathbb{R}} \gamma(u, \xi_{u-}(x), z)\,\tilde{\nu}(du, dz)$$

Define the Girsanov density G_t by

$$\begin{aligned} G_t(x_0) =& 1 + \int_0^t G_{u-}(x_0) g(u, \xi_{u-}(x_0))\,dW_u^P \\ &+ \int_0^t \int_{\mathbb{R}} G_{u-}(x_0)[h(u, \xi_{u-}(x_0), z) - 1]\,\tilde{\nu}(du, dz) \end{aligned}$$

where g and h are functions such that G is a square integrable martingale, and both g and h have continuous first derivatives in their second argument. Define the measure Q as the Radon-Nykodym derivative

$$\left.\frac{dQ}{dP}\right|_{\mathcal{F}_t} = G_t(x_0).$$

Then under Q, $W_t^Q \equiv W_t^P - \int_0^t g(u, \xi_{u-}(x_0))\,du$ is a standard Wiener process and ν is a Poisson random measure with intensity measure $\lambda^P h(t, \xi_{t-}(x_0), z)\,m^P(dz)\,dt$.

In our notation, taking $\xi = S$, $h(t, \xi_{t-}(x_0), z) = p(z)$, and $g(t, \xi_{t-}(x_0)) = -\theta$ and letting $Z \equiv \frac{dQ}{dP}$ we get

$$\frac{dZ_t}{Z_{t-}} = -\theta\,dW_t^P + \int_{\mathbb{R}} [p(z) - 1]\,\tilde{\nu}^P(dt, dz)$$

as our Girsanov density. For now we will assume only that $p(z) > 0$ and that $p \in \mathcal{L}^2(m^P)$; additional assumptions on p will be given in Section 2.9. Note that Z_t can also be written as the stochastic exponential

$$Z_t = \mathcal{E}\left(-\theta\,W_t^P + \int_0^t \int_{\mathbb{R}} [p(z) - 1]\,\tilde{\nu}^P(du, dz)\right).$$

Since the process given by $M_t = -\theta\,W_t^P + \int_0^t \int_{\mathbb{R}} [p(z) - 1]\,\tilde{\nu}^P(du, dz)$ is a Lévy process and a martingale, the stochastic exponential of M_t is also a martingale (see Cont and Tankov [8], Proposition 8.23) and thus for our choice of g and h we see that conditions of Proposition 1 are satisfied.

Under Q, Proposition 1 tells us that $dW_t^Q = dW_t^P + \theta\,dt$ is a standard Brownian motion and that ν is a Poisson random measure with intensity $\lambda^P p(z)\,m^P(dz)\,dt$. Thus we can write the characteristics $\left(\lambda^Q, m^Q(dz)\right)$ of ν under Q as

$$\begin{aligned} \lambda^Q &:= \lambda^P \int_{\mathbb{R}} p(z)\, m^P(dz) \\ m^Q(dz) &:= \frac{p(z)\, m^P(dz)}{\int_{\mathbb{R}} p(z)\, m^P(dz)}. \end{aligned} \tag{2.9}$$

Substituting W^Q and the intensity measure of ν under Q into equation (2.1) we obtain

$$\begin{aligned} \frac{dS_t}{S_{t-}} &= \mu\, dt + \sigma\, dW_t^P + \int_{\mathbb{R}} \gamma(z)\, \tilde{\nu}^P(dt, dz) \\ &= \left\{\mu - \sigma\theta + \lambda^P \int_{\mathbb{R}} \gamma(z)[p(z) - 1]\, m^P(dz)\right\} dt + \sigma\, dW_t^Q + \int_{\mathbb{R}} \gamma(z)\, \tilde{\nu}^Q(dt, dz), \end{aligned}$$

where $\tilde{\nu}^Q$ denotes the compensated Poisson random measure which has characteristics $\left(\lambda^Q, m^Q(dz)\right)$.

In order for Q to be an EMM, we choose p and θ such that the discounted risky asset process $\frac{S_t}{B_t}$ is a martingale. This means that we need

$$\mu - \sigma\theta + \lambda^P \int_{\mathbb{R}} \gamma(z)[p(z) - 1]\, m^P(dz) - r = 0.$$

Thus we obtain infinitely many equivalent martingale measures, which we can index by the parameter $p(z)$ since θ is determined once we choose $p(z)$. Under Q, the process S satisfies the stochastic differential equation

$$\frac{dS_t}{S_{t-}} = rdt + \sigma\, dW_t^Q + \int_{\mathbb{R}} \gamma(z)\, \tilde{\nu}^Q(dt, dz).$$

or, more explicitly,

$$S_t(s_0) = s_0 \exp\left(\left[r - \frac{1}{2}\sigma^2 - \lambda^Q k^Q\right] t + \sigma W_t^Q + \int_0^t \int_{\mathbb{R}} \ln(1 + \gamma(z))\, \nu(du, dz)\right) \tag{2.10}$$

where $k^Q := \int_{\mathbb{R}} \gamma(z)\, m^Q(dz)$. Note that $\gamma \in \mathcal{L}^1(m^Q)$ by Hölder's inequality since $\gamma, p \in \mathcal{L}^2(m^P)$.

2.9 Assumptions on $p(z)$

In what follows, it will be necessary to make some additional assumptions on the function $p(z)$. In particular, we we require

- $\gamma^2[1+\gamma]p \in \mathcal{L}^1(m^P)$
- $\dfrac{\gamma^2 p}{1+\gamma} \in \mathcal{L}^1(m^P)$

Note that our first assumption is implied if $p, \gamma \in \mathcal{L}^4(m^P)$, by Hölder's inequality applied twice. For the second assumption, since

$$\frac{\gamma^2 p}{1+\gamma} = -p + \gamma p + \frac{p}{1+\gamma},$$

the assumption will hold if $\frac{1}{1+\gamma} \in \mathcal{L}^2(m^P)$, again by Hölder's inequality.

Chapter 3

The Reformulated Problem

3.1 Change of Measure

As stated, the Russian option problem given by (2.8) is a two-dimensional problem since the value function depends on both the stock price and the running maximum. In general, the greater the number of dimensions of an optimal stopping problem, the more difficult it is to solve. For the diffusion case, Shepp and Shiryaev [33] simplified the problem significantly by performing a change of measure, resulting in a one-dimensional problem. Imitating this measure change, we will be able to simplify the problem in the jump-diffusion case as well.

To this end, define the random variable $X_t := \frac{M_t}{S_t}$. We can rewrite the value function as

$$W(t, m, s_0) = \sup_{0 \le \tau \le t} E^Q_{m,s_0}[e^{-(r+\lambda)\tau} M_\tau(m, s_0)] = s_0 \sup_{0 \le \tau \le t} E^Q_{m,s_0}\left[e^{-\lambda\tau} X_\tau \frac{e^{-r\tau} S_\tau}{s_0}\right].$$

We now perform the change of measure mentioned above, which we take to have density

$$\frac{e^{-r\tau} S_\tau(s_0)}{s_0} = \exp\left(\sigma W^Q_\tau - \frac{1}{2}\sigma^2 \tau\right) \cdot \exp\left(-\lambda^Q k^Q \tau + \int_0^\tau \int_{\mathbb{R}} \ln(1 + \gamma(z))\, \nu(du, dz)\right).$$

We will need a generalized version of Girsanov's Theorem to perform this change of measure. A version is given by Theorem 2.5 of Runggaldier [31]:

Proposition 2. *On the interval* $[0, T]$, *let* $\nu(dt, dz)$ *be a Poisson random measure with* $(Q, \mathcal{F}_t)$*-characteristics* $(\lambda, m(dz))$. *Let* $\psi_t \ge 0$, $h_t(y) \ge 0$ *be* $\mathcal{F}_t$*-predictable processes such* *that*

$$\int_0^t \psi_s\, ds < \infty; \qquad \int_{\mathbb{R}} h_t(z)\, m(dz) = 1. \tag{3.1}$$

Define $L_t^{(1)}$ and $L_t^{(2)}$ by

$$\begin{aligned}
\frac{dL_t^{(1)}}{L_t^{(1)}} &= \sigma dW_t \\
\frac{dL_t^{(2)}}{L_{t-}^{(2)}} &= \int_{\mathbb{R}} [\psi_t h_t(y) - 1]\, \tilde{\nu}(dt, dz) \\
&= \int_{\mathbb{R}} [\psi_t h_t(z) - 1]\, \nu(dt, dz) - \lambda \int_{\mathbb{R}} [\psi_t h_t(z) - 1]\, m(dz)\, dt.
\end{aligned}$$

If $E^Q[L_t^{(1)}] = 1$ and $E^Q[L_t^{(2)}] = 1$, then there exists a probability measure R, given by

$$dR = L_T^{(1)} L_T^{(2)} dQ,$$

which is equivalent to Q and such that under R, $W^R = W^Q - \sigma t$ is a standard Brownian motion and ν is a Poisson random measure with $(R, \mathcal{F}_t)$-characteristics $(\lambda\psi_t, h_t(z)\, m(dz))$.

In order to apply the theorem, we write $\frac{e^{-rt} S_t(s_0)}{s_0} = C_t J_t$, where we let

$$L_t^{(1)} \equiv C_t := \exp\left(-\frac{1}{2}\sigma^2 t + \sigma W_t^Q\right),$$

and

$$L_t^{(2)} \equiv J_t := \exp\left(-\lambda^Q k^Q t + \int_0^t \int_{\mathbb{R}} \ln(1 + \gamma(z))\, \nu(du, dz)\right)$$

so that

$$\begin{aligned}
\frac{dC_t}{C_{t-}} &= \sigma\, dW_t^Q \\
\frac{dJ_t}{J_{t-}} &= -\lambda_Q k_Q\, dt + \int_{\mathbb{R}} \gamma(z)\, \nu(dt, dz) = \int_{\mathbb{R}} \gamma(z)\, \tilde{\nu}^Q(dt, dz).
\end{aligned}$$

Thus we want to take $\psi_t h_t(z) - 1 = \gamma(z)$ in Proposition 2, and we also need (3.1) to be satisfied. Hence we let

$$\begin{aligned}
\psi_t = \psi &= \int_{\mathbb{R}} (\gamma(z) + 1)\, m^Q(dz) \\
h_t(z) = h(z) &= \frac{[\gamma(z) + 1]}{\int_{\mathbb{R}} (\gamma(z) + 1)\, m^Q(dz)}.
\end{aligned}$$

Note that both C_t and J_t are Q-martingales since each is the stochastic exponential of a Lévy process which is a martingale. Thus $E[C_t] = E[J_t] = 1$, and the conditions for the theorem are satisfied.

Applying this version of Girsanov's theorem we obtain a new measure R such that under R, the process $W_t^R = W_t^Q - \sigma t$ is a Brownian motion and $\nu(dt, dz)$ is a Poisson random measure with characteristics

$$\begin{aligned} \lambda^R &:= \lambda^Q \int_{\mathbb{R}} [\gamma(z) + 1]\, m^Q(dz) \\ m^R(dz) &:= \frac{[\gamma(z) + 1]\, m^Q(dz)}{\int_{\mathbb{R}} (\gamma(z) + 1)\, m^Q(dz)}. \end{aligned} \tag{3.2}$$

Note that $\lambda^R > 0$ since $\gamma(z) > -1$. Using these facts we see that under R, the process S_t satisfies the stochastic differential equation

$$\begin{aligned} \frac{dS_t}{S_{t-}} &= [\sigma^2 + r - \lambda^Q k^Q]dt + \sigma dW_t^R + \int_{\mathbb{R}} \gamma(z)\, \nu(dt, dz) \\ &= \left[\sigma^2 + r + \lambda^Q \int_{\mathbb{R}} \gamma(z)^2\, m^Q(dz)\right] dt + \sigma\, dW_t^R + \int_{\mathbb{R}} \gamma(z)\, \tilde{\nu}^R(dt, dz) \end{aligned} \tag{3.3}$$

or, alternatively, the equation

$$\begin{aligned} S_t(s_0) &= s_0 \exp\left(\left[\frac{\sigma^2}{2} + r - \lambda^Q k^Q\right] t + \sigma W_t^R + \int_0^t \int_{\mathbb{R}} \ln(1 + \gamma(z))\, \nu(dt, dz)\right) \\ &= s_0 \exp\left(\left[\frac{\sigma^2}{2} + r - \lambda^Q k^Q + \lambda^R \int_0^t \int_{\mathbb{R}} \ln(1 + \gamma(z))\, m^R(dz)\right] t\right) \\ &\quad \cdot \exp\left(\sigma W_t^R + \int_0^t \int_{\mathbb{R}} \ln(1 + \gamma(z))\, \tilde{\nu}^R(dt, dz)\right). \end{aligned} \tag{3.4}$$

Having performed such a change of measure, we see that

$$W(t, m, s_0) = s_0 \sup_{0 \le \tau \le t} E_{m,s_0}^R[e^{-\lambda\tau} X_\tau]. \tag{3.5}$$

Denote by $X_t(m, s_0)$ the process X_t such that $X_0(m, s_0) = \frac{M_0(m,s_0)}{S_0(s_0)} = \frac{m}{s_0} =: x$. Letting $S_t^*(s_0) = \sup_{0 \le u \le t} S_u(s_0)$ and using the fact that $S_t(s_0) = s_0\, S_t(1)$ we can write

$$\begin{aligned} X_t(m, s_0) &= \frac{M_t(m, s_0)}{S_t(s_0)} = \frac{S_t^*(s_0) \vee m}{S_t(s_0)} = \frac{s_0 S_t^*(1) \vee m}{s_0 S_t(1)} \\ &= \frac{S_t^*(1) \vee x}{S_t(1)} = \frac{M_t(x, 1)}{S_t(1)} = X_t(x, 1). \end{aligned}$$

Thus when dealing with process X_t alone, we can write $X_t(x)$ for the process X which satisfies $X_0 = x$. Furthermore, we can and will assume without loss of generality that $s_0 = 1$ and $m = x$. From (3.5), we see that we now only need to work with the new value function $V(t, x)$ given by

$$V(t, x) := \sup_{0 \leq \tau \leq t} E_x^R[e^{-\lambda\tau} X_\tau].$$

Thus we have reduced the dimension of our optimal stopping problem at the cost of needing to work with the slightly more complicated process X.

3.2 Properties of X_t

Before we proceed, let us discuss some basic properties of the random process X_t which will be needed later. As we shall see, X_t behaves like a geometric Lévy process in the interval $(1, \infty)$. At the boundary, when the stock is at a maximum, a boundary term forces X_t to remain in the interval.

3.2.1 Markov Property

There are many results from optimal stopping theory which require that the process to be stopped is a Markov process. The next theorem shows that this is true for X.

Proposition 3. *The process X is a time-homogeneous strong Markov process under R.*

Proof. From (3.4), we see that we can write $S_t = \exp(L_t)$, where L is a Lévy process. Hence

$$X_{t+s} = \frac{\sup_{0 \leq u \leq t+s} \exp(L_u) \vee x}{\exp(L_{t+s})}$$

and so

$$\begin{aligned}
X_{\tau+s} &= \frac{\sup_{0\le u\le \tau+s}\exp(L_u)\vee x}{\exp(L_{\tau+s})}\\
&= \frac{[\sup_{0\le u\le\tau}\exp(L_u)\vee x]\vee \sup_{0<u\le s}\exp(L_{\tau+u})}{\exp(L_{\tau+s}-L_\tau+L_\tau)}\\
&= \frac{1}{\exp(L_{\tau+s}-L_\tau)}[X_\tau \vee \sup_{0\le u\le s}\exp(L_{\tau+u}-L_\tau)]\\
&= \frac{1}{\exp(\tilde{L}_s)}[X_\tau\vee \sup_{0\le u\le s}\exp(\tilde{L}_u)]
\end{aligned}$$

where we define $\tilde{L}_s = L_{\tau+s} - L_\tau$, which is independent of $\mathcal{F}_\tau$ since L is a Lévy process. Thus $X_{\tau+s}$ depends on $\mathcal{F}_\tau$ only through X_τ and so we see that it is a time-homogeneous strong Markov process. □

3.2.2 Differential Form

Next we derive the differential form of X.

Proposition 4. *Under R, the process X satisfies the SDE*

$$dX_t = \alpha(X_{t-})\,dt + \beta(X_{t-})\,dW_t^R + \int_{\mathbb{R}} \zeta(X_{t-},z)\,\tilde{\nu}^R(dt,dz) + \mathbf{1}_{[\Delta M_t=0]}\frac{dM_t}{S_{t-}} \tag{3.6}$$

where the coefficients are given by

$$\begin{aligned}
\alpha(x) &= -rx + \lambda^R\int_{\{1+\gamma(z)>x\}}\left[1-\frac{x}{1+\gamma(z)}\right] m^R(dz)\\
\beta(x) &= -\sigma x\\
\zeta(x,z) &= -x\left[\frac{\gamma(z)}{1+\gamma(z)}\right] + \mathbf{1}_{[1+\gamma(z)>x]}\left[1-\frac{x}{1+\gamma(z)}\right].
\end{aligned} \tag{3.7}$$

Proof. By Itô's product formula, we know that

$$dX_t = d\left(\frac{M_t}{S_t}\right) = M_{t-}\,d\left(\frac{1}{S_t}\right) + \frac{1}{S_{t-}}\,dM_t + \left[M,\frac{1}{S}\right]_t, \tag{3.8}$$

where $[M,\frac{1}{S}]_t$ denotes the quadratic variation of M and $\frac{1}{S_t}$. Applying (2.3) with $f(x)=\frac{1}{x}$

we obtain, using the assumptions of Section 2.9 and (3.3),

$$
\begin{aligned}
d\left(\frac{1}{S}\right)_t &= -\frac{1}{S_{t-}^2}\,dS_t^c + \frac{1}{2}\left(\frac{2}{S_{t-}^3}\right)d[S,S]_t^c + \int_{\mathbb{R}} \frac{1}{S_t} - \frac{1}{S_{t-}}\,\nu(dt,dz) \\
&= -\frac{1}{S_{t-}}[\sigma^2 + r - \lambda^Q k^Q]dt - \frac{\sigma}{S_{t-}}\,dW_t^R + \frac{\sigma^2}{S_{t-}}\,dt + \frac{1}{S_{t-}}\int_{\mathbb{R}} \frac{1}{1+\gamma(z)} - 1\,\nu(dt,dz) \\
&= -\frac{1}{S_{t-}}[r - \lambda^Q k^Q]dt - \frac{\sigma}{S_{t-}}\,dW_t^R + \int_{\mathbb{R}} \frac{1}{S_{t}-}\left[\frac{-\gamma(z)}{1+\gamma(z)}\right]\nu(dt,dz).
\end{aligned}
\tag{3.9}
$$

Computing the quadratic covariation we get

$$
\begin{aligned}
\left[M, \frac{1}{S}\right]_t &= \sum_{0\le u\le t,\,\Delta M_u>0} \Delta M_u\,\Delta\left(\frac{1}{S}\right)_u \\
&= \sum_{0\le u\le t,\,\Delta M_u>0} (M_u - M_{u-})\left(\frac{1}{S_u} - \frac{1}{S_{u-}}\right) \\
&= \sum_{0\le u\le t,\,\Delta M_u>0} (S_u - M_{u-})\left(\frac{1}{S_{u-}(1+\gamma(z))} - \frac{1}{S_{u-}}\right) \\
&= \int_0^t \int_{\mathbb{R}} (S_{u-}(1+\gamma(z)) - M_{u-})\left(\frac{1}{S_{u-}(1+\gamma(z))} - \frac{1}{S_{u-}}\right)\mathbf{1}_{[\Delta M_u>0]}\,\nu(du,dz) \\
&= \int_0^t \int_{\mathbb{R}} (1+\gamma(z) - X_{u-})\left(\frac{1}{1+\gamma(z)} - 1\right)\mathbf{1}_{[\Delta M_u>0]}\,\nu(du,dz) \\
&= \int_0^t \int_{\mathbb{R}} (1+\gamma(z) - X_{u-})\left(\frac{-\gamma(z)}{1+\gamma(z)}\right)\mathbf{1}_{[\Delta M_u>0]}\,\nu(du,dz).
\end{aligned}
\tag{3.10}
$$

Here the third inequality follows since $S_u = M_u$ when $\Delta M_u > 0$. Note that the event

$$
\begin{aligned}
\{\Delta M_t > 0\} = \{M_t > M_{t-}\} = \{S_t > M_{t-}\} &= \{S_{t-}(1+\gamma(Z_{N(t)})) > M_{t-}\} \\
&= \{1+\gamma(Z_{N(t)}) > X_{t-}\},
\end{aligned}
\tag{3.11}
$$

where Z_i and $N(t)$ are as defined in Section 2.2. We similarly calculate

$$
\begin{aligned}
\int_0^t \frac{dM_u}{S_{u-}} &= \int_0^t \mathbf{1}_{[\Delta M_u=0]}\frac{dM_u}{S_{u-}} + \sum_{0\le u\le t,\,\Delta M_u>0} \frac{\Delta M_u}{S_{u-}} \\
&= \int_0^t \mathbf{1}_{[\Delta M_u=0]}\frac{dM_u}{S_{u-}} + \int_0^t \int_{\mathbb{R}} \mathbf{1}_{[\Delta M_u>0]}\frac{S_{u-}(1+\gamma(z)) - M_{u-}}{S_{u-}}\,\nu(du,dz) \qquad (3.12) \\
&= \int_0^t \mathbf{1}_{[\Delta M_u=0]}\frac{dM_u}{S_{u-}} + \int_0^t \int_{\mathbb{R}} \mathbf{1}_{[\Delta M_u>0]}(1+\gamma(z) - X_{u-})\,\nu(du,dz)
\end{aligned}
$$

Combining (3.8), (3.9), (3.10), (3.11) and (3.12) we get

$$\begin{aligned} dX_t =& - X_{t-}[r - \lambda^Q k^Q]\,dt - \sigma X_{t-}\,dW_t^R + \mathbf{1}_{[\Delta M_t=0]}\,\frac{dM_t}{S_{t-}} \\ &+ \int_{\mathbb{R}} X_{t-}\left[\frac{-\gamma(z)}{1+\gamma(z)}\right] + \mathbf{1}_{[1+\gamma(z)>X_{t-}]}\left(1 - \frac{X_{t-}}{1+\gamma(z)}\right)\nu(dt,dz) \\ =& \left[-rX_{t-} + \lambda^R \int_{\{1+\gamma(z)>X_{t-}\}}\left[1 - \frac{X_{t-}}{1+\gamma(z)}\right] m^R(dt,dz)\right] dt - \sigma X_{t-}\,dW_t^R + \mathbf{1}_{[\Delta M_t=0]}\,\frac{dM_t}{S_{t-}} \\ &+ \int_{\mathbb{R}} -X_{t-}\left[\frac{\gamma(z)}{1+\gamma(z)}\right] + \mathbf{1}_{[1+\gamma(z)>X_{t-}]}\left(1 - \frac{X_{t-}}{1+\gamma(z)}\right)\tilde{\nu}^R(dt,dz). \end{aligned} \tag{3.13}$$

Here we have used the fact that $\lambda^Q k^Q = \lambda^R \int_{\mathbb{R}}\left[\frac{\gamma(z)}{1+\gamma(z)}\right] m^R(dz)$. □

Remark 1. *Note that $\frac{\gamma(z)}{1+\gamma(z)} \in L^2(m^R)$ since, for some constant C,*

$$\begin{aligned} \int_{\mathbb{R}}\left(\frac{\gamma(z)}{1+\gamma(z)}\right)^2 m^R(dz) &= C\int_{\mathbb{R}} \frac{\gamma(z)^2}{(1+\gamma(z))^2}(1+\gamma(z))\,m^Q(dz) \\ &= C\int_{\mathbb{R}} \frac{\gamma(z)^2 p(z)}{1+\gamma(z)}\,m^P(dz) < \infty \end{aligned}$$

by the assumptions in Section 2.9.

Remark 2. *The jump term $\zeta(x,z)$ above behaves as we would expect. When S jumps to a new maximum, so that $1+\gamma(Z_{N(t)}) \geq X_{t-}$, we have $\Delta X_t = \zeta(X_{t-}, Z_{N(t)}) = 1 - X_{t-}$ so that X jumps to 1. When S jumps to a level that is below the maximum, we have*

$$\Delta X_t = \frac{M_t}{S_t} - \frac{M_{t-}}{S_{t-}} = \frac{M_{t-}}{S_{t-}(1+\gamma(Z_{N(t)}))} - \frac{M_{t-}}{S_{t-}} = -X_{t-}\left(\frac{\gamma(Z_{N(t)})}{1+\gamma(Z_{N(t)})}\right) \tag{3.14}$$

which matches $\zeta(X_{t-}, Z_{N(t)})$ when $1+\gamma(Z_{N(t)}) < X_{t-}$.

3.2.3 Regularity of the Coefficients

When working with a stochastic differential equation, we typically require that the coefficients be continuous and satisfy a linear growth condition. Fortunately, these conditions are true for the coefficients of the SDE given in (3.6). This will be quite useful later on.

Lemma 1. *The coefficients $\alpha(x)$, $\beta(x)$, and $\zeta(x, y)$ given in Proposition 4 are continuous in x for fixed z. Furthermore, for some constant K and some function $\rho : \mathbb{R} \to \mathbb{R}_+$, with $\rho \in \mathcal{L}^2(m^R)$, the global Lipschitz conditions*

$$\begin{aligned} |\alpha(x) - \alpha(y)| &\leq K|x - y| \\ |\beta(x) - \beta(y)| &\leq K|x - y| \\ |\zeta(x, z) - \zeta(y, z)| &\leq \rho(z)|x - y| \end{aligned} \tag{3.15}$$

and the linear growth conditions

$$\begin{aligned} |\alpha(x)| &\leq K(1 + |x|) \\ |\beta(x)| &\leq K(1 + |x|) \\ |\zeta(x, z)| &\leq \rho(z)(1 + |x|) \end{aligned} \tag{3.16}$$

hold for all $x, y \in [1, \infty)$ and $z \in \mathbb{R}$.

Proof. The inequalities and continuity obviously hold for β. For ζ, note that

$$\zeta(x, z) = \begin{cases} -x\left[\frac{\gamma(z)}{1+\gamma(z)}\right] & \text{if } 1 + \gamma(z) \leq x \\ 1 - x & \text{if } 1 + \gamma(z) > x \end{cases}.$$

From this, we see that ζ is continuous for fixed z since

$$\lim_{x \to (1+\gamma(z))^+} \zeta(x, z) = \lim_{x \to (1+\gamma(z))^-} \zeta(x, z) = -\gamma(z). \tag{3.17}$$

Let $\rho(z) = \left|\frac{\gamma(z)}{1+\gamma(z)}\right| \vee 1$, and observe that $\rho \in \mathcal{L}^2(m^R)$. Then with this choice of ρ the linear growth condition is satisfied since when $1 + \gamma(z) > x$, we have $|1 - x| \leq |\rho(z)|(1 + |x|)$ and the result is obvious when $1 + \gamma(z) \leq x$. Next we will show the Lipschitz condition holds for ζ. Without loss of generality, let $x > y$. The result is obvious if $x > y \geq 1 + \gamma(z)$ or $1 + \gamma(z) > x > y$. Assume $x \geq 1 + \gamma(z) > y \geq 1$. Then

$$\begin{aligned} |\zeta(x, z) - \zeta(y, z)| &= \left| - x\left(\frac{\gamma(z)}{1 + \gamma(z)}\right) - (1 - y)\right| \\ &= \left|\frac{1}{1 + \gamma(z)}\right| \cdot \left|(x - y)\gamma(z) + (1 + \gamma(z) - y)\right| \\ &\leq \left|\frac{1}{1 + \gamma(z)}\right| \cdot \left|(x - y)\gamma(z) + x - y\right| \\ &\leq \rho(z)|x - y| \end{aligned} \tag{3.18}$$

where, for the second to last inequality, we used the fact that $x \geq 1 + \gamma(z) > y$.

Finally, we will show the results for $\alpha(x)$. Let $\phi(x, z) := \mathbf{1}_{[1+\gamma(z)>x]} \left[1 - \frac{x}{1+\gamma(z)}\right]$. Then if we can show that ϕ is continuous in x for fixed z and that both

$$|\phi(x, z) - \phi(y, z)| \leq \tilde{\rho}(z)(|x - y|) \tag{3.19}$$

and

$$|\phi(x, z)| \leq \tilde{\rho}(z)(1 + |x|) \tag{3.20}$$

hold for some $\tilde{\rho} \in \mathcal{L}(m^R)$, then the results for α will follow from the Dominated Convergence Theorem since $\phi(x, z)$ is bounded by 1. Let $\tilde{\rho}(z) = \frac{1}{1+\gamma(z)}$, and observe that $\tilde{\rho} \in \mathcal{L}(m^R)$ by the assumptions in Chapter 2. Then condition (3.20) and the continuity of ϕ are obvious, so it only remains to prove (3.19). Again assume without loss of generality that $x > y$, and note that the result is trivial when $x > y \geq 1 + \gamma(z)$. For $x \geq 1 + \gamma(z) > y$, we have

$$|\zeta(x, z) - \zeta(y, z)| = \left|1 - \frac{y}{1 + \gamma(z)}\right| = \frac{1 + \gamma(z) - y}{1 + \gamma(z)} \leq \frac{x - y}{1 + \gamma(z)} \leq \tilde{\rho}(z)|x - y|. \tag{3.21}$$

Finally, if $1 + \gamma(z) > x > y$, we have

$$|\zeta(x, z) - \zeta(y, z)| = \left|\left(1 - \frac{x}{1 + \gamma(z)}\right) - \left(1 - \frac{y}{1 + \gamma(z)}\right)\right| \leq \frac{|x - y|}{1 + \gamma(z)} \leq \frac{x - y}{1 + \gamma(z)} \leq \tilde{\rho(z)}|x - y|. \tag{3.22}$$

and the result is proved. □

3.2.4 Estimates on the Moments of X_t

In later sections we will need bounds on the moments of X. Lemma 3.1 of Pham [27] provides estimates for general processes given by a stochastic differential equation with coefficients satisfying the same conditions as those in Lemma 1. The difference between our process X and those satisfying the conditions of Pham [27] is that X has a boundary term which must be handled. Fortunately, a similar argument works, and we obtain the following lemma whose proof is provided in the Appendix.

Lemma 2. *For any $k \in [0, 2]$, there exists a constant C, depending on k and T, such that for any $h \in [0, T]$ and any stopping time τ satisfying $0 \leq \tau \leq h$, we have* *(i)* $E_x^R|X_\tau|^k \leq C(1+|x|^k)$

(ii) $E_x^R|X_\tau - x|^k \leq C(1+|x|^k)h^{\frac{k}{2}}$

(iii) $E_x^R\left[\sup_{0\leq\tau\leq h}|X_\tau - x|^k\right] \leq C(1+|x|^k)h^{\frac{k}{2}}$

Remark 3. *The proof of part (iii) of Lemma 2 uses the martingale property of the next section. The argument below references Lemma 2, but only part (i), which does not require the martingale property. Hence we observe that there is no circular reasoning in the proofs of the two results.*

3.3 Martingale Property

As a consequence of part (i) of Lemma 2, we now show that the stochastic integrals

$$M_t = \int_0^t \beta(X_u)\, dW_u^R \tag{3.23}$$

and

$$N_t = \int_0^t \int_{\mathbb{R}} \zeta(X_{u-}, z)\, \tilde{\nu}(du, dz) \tag{3.24}$$

are square-integrable $\mathcal{F}_t$-martingales, which we will need later. Recall the fact that any stochastic integral of the form $\int_0^t \phi(u, X_{u-})\, dW_u^R$ will be a martingale provided

$$E_x^R\left[\int_0^T |\phi(t, X_{u-})|^2\, du\right] < \infty$$

and any stochastic integral of the form $\int_0^t \int_{\mathbb{R}} \psi(t, X_{u-}, z)\, \tilde{\nu}^R(dz)\, du$ will be a martingale provided

$$E_x^R\left[\int_0^T \int_{\mathbb{R}} |\psi(t, X_{u-}, z)|^2\, m^R(dz)\, du\right] < \infty,$$

where $\phi(t,x)$ and $\psi(t,x,z)$ are $\mathcal{F}_t$-predictable.

This is clearly true for (3.23), since

$$E_x^R\left[\int_0^T |\beta(X_u)|^2\, du\right] = E_x^R\left[\int_0^T \sigma^2 X_u^2\, du\right] \leq C E_x^R\left[\int_0^T X_u^2\, du\right] < \infty \tag{3.25}$$

for some constant C by part (i) of Lemma 2 and Fubini's Theorem. For (3.24), we similarly have

$$
\begin{aligned}
&E_x^R\left[\int_0^T \int_{\mathbb{R}} |\zeta(X_{u-},z)|^2\, m^R(dz)\, du\right] \\
&\quad = \left(E\left[\int_0^T \int_{\mathbb{R}} |\zeta(X_{u-}z)|^2 \mathbf{1}_{[1+\gamma(z)\le x]}\, m^R(dz)\, du + \int_0^T \int_{\mathbb{R}} |\zeta(X_{u-}z)|^2 \mathbf{1}_{[1+\gamma(z)>x]}\, m^R(dz)\, du\right]\right) \\
&\quad \le \left(E\left[\int_0^T \int_{\mathbb{R}} X_{u-}^2 \frac{|\gamma(z)|^2}{|1+\gamma(z)|^2}\, m^R(dz)\, du\right] + E\left[\int_0^T \int_{\mathbb{R}} |1 - X_{u-}|^2\, m^R(dz)\, du\right]\right) \\
&\quad \le C\left(\int_{\mathbb{R}} 1 + \frac{|\gamma(z)|^2}{|1+\gamma(z)|^2}\, m^R(dz)\right) E_x^R\left[\int_0^T X_{u-}^2\, du\right] < \infty
\end{aligned}
$$

by (i) of Lemma 2 and the remark following Proposition 4. Finally, observe that a nearly identical argument shows that

$$\tilde{M}_t = \int_0^t e^{-\lambda u}\beta(X_u)\, dW_u^R$$

and

$$\tilde{N}_t = \int_0^t \int_{\mathbb{R}} e^{-\lambda u}\zeta(X_{u-},z)\, \tilde{\nu}(du,dz)$$

are also $\mathcal{F}_t$ martingales.

3.4 Statement of the Reformulated Problem

Before proceeding, for convenience let us restate our optimal stopping problem after reformulation. The value function is given by

$$V(t,x) = \sup_{0\le\tau\le t} E_x^R[e^{-\lambda\tau}X_\tau]. \tag{3.26}$$

where X satisfies the stochastic differential equation

$$dX_t = \alpha(X_{t-})\, dt + \beta(X_{t-})\, dW_t^R + \int_{\mathbb{R}} \zeta(X_{t-},z)\, \tilde{\nu}^R(dt,dz) + \mathbf{1}_{[\Delta M_t=0]}\frac{dM_t}{S_{t-}}$$

which has coefficients given by

$$
\begin{aligned}
\alpha(x) &= -rx + \lambda^R \int_{\{1+\gamma(z)>x\}} \left[1 - \frac{x}{1+\gamma(z)}\right] m^R(dz) \\
\beta(x) &= -\sigma x \\
\zeta(x,z) &= -x\left[\frac{\gamma(z)}{1+\gamma(z)}\right] + \mathbf{1}_{[1+\gamma(z)>x]}\left[1 - \frac{x}{1+\gamma(z)}\right]
\end{aligned}
$$

As mentioned in Chapter 2, in order for an optimal stopping problem with value function $W(x,t) = \sup_{0\leq\tau\leq t} E_x^R[e^{-\lambda\tau} g(\tau, X_\tau)]$ with $g(t,x)$ is some measurable function to be well defined a condition such as

$$E\left[\sup_{0\leq t\leq T} |g(t, X_t)|\right] < \infty.$$

is required. Note that this is true for $V(t,x)$ since, by the triangle inequality,

$$E_x^R\left[\sup_{0\leq t\leq T} |e^{-\lambda t} X_\tau|\right] \leq E_x^R\left[\sup_{0\leq t\leq T} |X_\tau - x|\right] + |x|,$$

which is finite by Lemma 2.

Chapter 4

Properties of the Value Function

In this section we give some fundamental properties of the value function $V(t, x)$, as well as some less important properties which will be needed later. The corresponding proofs given in Peskir [25] for the Russian option in the diffusion case work either without change or with some modifications for the jump diffusion case. We will also prove some additional properties of $V(t, x)$ not given in Peskir [25] which will be needed in Proposition 6 when we prove $V(t, x)$ is smooth.

Before we begin, let us observe that we can write

$$V(t, x) = \sup_{0 \le \tau \le t} E_x^R[e^{-\lambda\tau} X_\tau] = \sup_{0 \le \tau \le t} E^R \left[e^{-\lambda\tau} \frac{(x - S_\tau^*)^+ + S_\tau^*}{S_\tau} \right]. \tag{4.1}$$

From (4.1) we can immediately see that $V(t, x)$ is convex and increasing in x. This expression for $V(t, x)$ will also be useful in the proof of Proposition 5 below.

4.1 Continuity of the Value Function

In this subsection we prove the continuity of the value function. We begin by proving a stronger form of continuity for $x \to V(t, x)$ when t is fixed.

Proposition 5. *The value function $V(t, x)$ is Lipschitz in x (with Lipschitz constant 1), and thus uniformly continuous in x, for fixed t.*

Proof. Using the facts that $\sup(f) - \sup(g) \le \sup(f - g)$, that $(y - z)^+ - (x - z)^+ \le (y - x)^+$, and (4.1) we have, for $1 \leq y \leq x$,

$$0 \leq V(t,x) - V(t,y) \leq \sup_{0\leq\tau\leq t} E^R\left[e^{-\lambda\tau}\frac{(x-S^*_\tau)^+ - (y-S^*_\tau)^+}{S_\tau}\right]$$
$$\leq \sup_{0\leq\tau\leq t} E^R\left[e^{-\lambda\tau}\frac{(x-y)^+}{S_\tau}\right] = (x-y)^+ \sup_{0\leq\tau\leq t} E^R\left[\frac{e^{-\lambda\tau}}{S_\tau}\right].$$

Now by (3.9), letting $K_t := \frac{1}{S_t}$ we can write

$$\frac{dK_t}{K_t} = -[r - \lambda^Q k^Q]dt - \sigma dW^R_t - \int_{\mathbb{R}} \frac{\gamma(z)}{1+\gamma(z)}\nu(dz,dt)$$
$$= -rdt - \sigma dW^R_t - \int_{\mathbb{R}} \frac{\gamma(z)}{1+\gamma(z)}\tilde{\nu}^R(dz,dt)$$

since $\lambda^Q \int_{\mathbb{R}} \frac{\gamma(z)}{1+\gamma(z)} m^R(dz) = \lambda^Q \int_{\mathbb{R}} \gamma(z) m^Q(dz) = \lambda^Q k^Q$. Thus $e^{rt}K_t = \frac{\exp(rt)}{S_t}$ is a martingale, and so we have

$$V(t,x) - V(t,y) \leq (x-y)^+ \sup_{0\leq\tau\leq t} E\left[\frac{e^{-\lambda\tau}}{S_\tau}\right] = (x-y)^+ \sup_{0\leq\tau\leq t} E\left[\frac{e^{-(\lambda+r)\tau}e^{r\tau}}{S_\tau}\right]$$
$$\leq (x-y)^+ \sup_{0\leq\tau\leq t} E\left[\frac{e^{r\tau}}{S_\tau}\right] \leq (x-y)^+,$$

where the final inequality follows from the Optimal Sampling Theorem since the stopping time τ is bounded. □

Using the uniform continuity in x for fixed t, we can now show that $V(t,x)$ is continuous in (t,x):

Theorem 1. *The value function $V(t,x)$ is jointly continuous in x and t.*

Proof. Since $V(t,x)$ is uniformly Lipschitz continuous in x, it is enough to show continuity in t. Let $s,t \in [0,T]$ and assume $s < t$. Let $\tau_1 := \tau_1^\epsilon(t,x)$, where $0 < \tau_1 < t$, be a stopping time such that $E^R_x[e^{-\lambda\tau_1}X_{\tau_1}] \geq V(t,x) - \epsilon$. Furthermore, define $\tau_2 := \tau_2^\epsilon(t,s,x) = \tau_1 \wedge s$, and note that $E^R_x[e^{-\lambda\tau_2}X_{\tau_2}] \leq V(s,x)$. Then since $t \to V(t,x)$ is non-decreasing we have

$$\begin{aligned} 0 \leq V(t,x) - V(s,x) &\leq E^R_x[e^{-\lambda\tau_1}X_{\tau_1} - e^{-\lambda\tau_2}X_{\tau_2}] + \epsilon \\ &= E^R_x[e^{-\lambda\tau_1}X_{\tau_1} - e^{-\lambda\tau_1}X_{\tau_1}\mathbf{1}_{[\tau_1\leq s]} - e^{-\lambda s}X_s\mathbf{1}_{[\tau_1>s]}] + \epsilon \\ &= E^R_x[(e^{-\lambda\tau_1}X_{\tau_1} - e^{-\lambda s}X_s)\mathbf{1}_{[\tau_1>s]}] + \epsilon. \end{aligned} \tag{4.2}$$

Now since

$$E_x^R[e^{-\lambda\tau_1}X_{\tau_1}] \leq V(t,x) \leq V(T,x),$$

we can apply the Dominated Convergence Theorem to the final term of (4.2). If $t \searrow s$, we have

$$\lim_{t\searrow s} E_x^R\left[(e^{-\lambda\tau_1}X_{\tau_1} - e^{-\lambda s}X_s)\mathbf{1}_{[\tau_1>s]}\right] = 0$$

since $\lim_{t\searrow s}\tau_1 \leq s$, so either $\mathbf{1}_{[\tau_1>s]} = 0$ or $X_{\tau_1} \to X_s$ from the fact that $t \to X_t$ is cadlag and hence is continuous from the right. If $s \nearrow t$, the cadlag property of X implies that $X_s \to X_{t-}$, so the fact that this limit exists implies that

$$\lim_{s\nearrow t} E_x^R[(e^{-\lambda\tau_1}X_{\tau_1} - e^{-\lambda s}X_s)\mathbf{1}_{[\tau_1>s]}] = 0$$

since $\lim_{s\nearrow t}\mathbf{1}_{[\tau_1>s]} = 0$. Thus letting $\epsilon \to 0^+$ in (4.2) after taking these limits above we see that $V(t,x) - V(s,x) \to 0$ so the joint continuity is proved. □

4.2 The Continuation Region and Stopping Regions

As mentioned in Section 2.6, from standard results of Optimal Stopping Theory we know there exists a continuation set $C = \{(t,x) \in (0,T] \times [1,\infty) : V(t,x) > g(x)\}$ and a stopping set $D = \{(t,x) \in (0,T] \times [1,\infty) : V(t,x) = g(x)\}$, where $g(x) := x$ is defined to be the payoff function for our optimal stopping problem. Furthermore, we know that the first hitting time of D given by

$$\begin{aligned}\tau_D(t,x) &= \inf\{s : 0 \leq s \leq t, (t-s, X_s(x)) \in D\} \\ &= \inf\{s : 0 \leq s \leq t, V(t-s, X_s(x)) = g(X_s(x))\}\end{aligned} \tag{4.3}$$

is optimal. It is well known that for any such problem the stopped process $e^{-\lambda(u\wedge\tau_D)}V(t-(u\wedge\tau_D), X_{u\wedge\tau_D})$ is an R-martingale (see Shirayev and Peskir [26]).

4.2.1 Existence of an Early Exercise Boundary

Consider the Russian option problem before our change of measure, where the value function is given by

$$W(t,m,s_0) = \sup_{0\leq\tau\leq t} E_{m,s_0}^Q[e^{(-r+\lambda)\tau}M_\tau(m,s_0)] \tag{4.4}$$

and $M_t(m, s_0) := S_t^* \vee m$. Intuitively, we want to continue to run our process as long as we are "sufficiently close" to the current maximum of the stock price. Once the stock price is far enough from the maximum, we are unlikely to hit the maximum again and by waiting to execute the option, the $e^{(-r+\lambda)t}$ term in (4.4) causes the value of the option to decrease. Note that if S_t is close to M_t, then X_t is close to 1, so for our reformulated problem where the value function is given by $V(t, x)$, we should stop once X_t is exceeds some sufficiently large value. This value of course should depend on the amount of time remaining, and so we expect that there should exist some boundary function $b(t)$ for $V(t, x)$, where we continue to run the process if $X_t < b(t)$ and stop the process if $X_t \geq b(t)$. The existence of such a boundary is given by our next result:

Theorem 2. *There exists a function $b(t)$ such that the continuation region has the form*

$$C = \{(t, x) \in (0, T] \times [1, \infty) : x < b(t)\}.$$

Proof. Applying Itô's Product Formula to $e^{-\lambda s}X_s$ we obtain

$$\begin{aligned} e^{-\lambda s}X_s &= X_0 - \lambda \int_0^s e^{-\lambda u} X_{u-}\, du + \int_0^s e^{-\lambda u}\, dX_u \\ &= X_0 + \int_0^s e^{-\lambda u}(-\lambda X_{u-} + \alpha(X_{u-}))\, du + \int_0^s e^{-\lambda u} \mathbf{1}_{[\Delta M_u = 0]} \frac{dM_u}{S_{u-}} + N_s \end{aligned} \tag{4.5}$$

where

$$N_s = \int_0^s e^{-\lambda u}\beta(X_u)\, dW_u^R + \int_0^s \int_{\mathbb{R}} e^{-\lambda u}\zeta(X_{u-}, z)\, \tilde{\nu}^R(du, dz),$$

which is a martingale for $0 \leq s \leq T - t$ by the results of Section 3.3

Let $x > y \geq 1$ be fixed, let $\tau_* := \tau_D(t, x)$ be optimal for $V(t, x)$, and suppose $(t, x) \in C$. Using (4.5) and applying the Optimal Sampling Theorem we claim that

$$\begin{aligned} V(t, y) - y &\geq E_y^R[e^{-\lambda\tau_*}X_{\tau_*}] - y \\ &= X_0 + E_y^R\left[\int_0^{\tau_*} e^{-\lambda u}(-\lambda X_u + \alpha(X_{u-}))\, du + \int_0^{\tau_*} e^{-\lambda u}\mathbf{1}_{[\Delta M_u = 0]} \frac{dM_u}{S_{u-}}\right] - y \\ &\geq X_0 + E_x^R\left[\int_0^{\tau_*} e^{-\lambda u}(-\lambda X_u + \alpha(X_{u-}))\, du + \int_0^{\tau_*} e^{-\lambda u}\mathbf{1}_{[\Delta M_u = 0]} \frac{dM_u}{S_{u-}}\right] - x \\ &= E_x^R[e^{-\lambda\tau_*}X_{\tau_*}] - x = V(t, x) - x > 0, \end{aligned} \tag{4.6}$$

since $(t, x) \in C$. If this is true, the result will be proved since (4.6) then implies that $V(t, y) > g(y) = y$, so that $(t, y) \in C$. Let us justify our claim. The first inequality in (4.6) follows since $V(t,y) \geq E^R_y[e^{-\lambda\tau}X_\tau]$ for any stopping time τ. The second inequality above is due to the following two facts. First, when $x > y$ we have that $\alpha(X^y_{u-}) \geq \alpha(X^x_{u-})$. This is true since $X_{u-}(x) \geq X_{u-}(y)$ and

$$\mathbf{1}_{[1+\gamma(z)>X_{u-}(y)]} \geq \mathbf{1}_{[1+\gamma(z)>X_{u-}(x)]}.$$

The second fact we used was that for all s satisfying $0 \leq s \leq t$,

$$\int_0^s e^{-\lambda u}\mathbf{1}_{[\Delta M^y_u=0]}\frac{dM^y_u}{S_{u-}} \geq \int_0^s e^{-\lambda u}\mathbf{1}_{[\Delta M^x_u=0]}\frac{dM^x_u}{S_{u-}},$$

which follows since $d(S^*_u \vee y) \geq d(S^*_u \vee x)$ when $x > y$ and the fact that $\mathbf{1}_{[\Delta M^y_u=0]} = 0$ implies either $\mathbf{1}_{[\Delta M^x_u=0]} = 0$ or $\int_0^s e^{-\lambda u}\mathbf{1}_{[\Delta M^x_u=0]}\frac{dM^x_u}{S_{u-}} = 0$. □

Let us also observe that $b(t)$ is non-decreasing. To see this, suppose there exists x, s, and t with $s < t$ and such that $b(s) > x > b(t)$. Then from the definition of the continuation region and Theorem 2, $V(s,x) > x$ and $V(t,x) \leq x$, which contradicts the fact that $t \to V(t,x)$ is non-decreasing.

4.3 Growth Conditions

In this section we obtain bounds on the growth of $V(t,x)$. From (i) of Lemma 2 in Section 3.2.4 we immediately see that

$$V(t,x) \leq C(1+|x|) \tag{4.7}$$

for some constant C. Using this fact, we can obtain an even stronger statement on the growth of $V(t,x)$.

Lemma 3. *For fixed T there exists a constant C such that for any $s,t \in [0,T]$ and $x \in [1,\infty)$,*

$$|V(t,x) - V(s,x)| \leq C(1+|x|)|t-s|^{\frac{1}{2}}$$

Proof. Assume without loss of generality that $s \leq t$ and denote the optimal stopping time $\tau_D(t,x)$ defined in (4.3) by τ_D. Recall that $e^{-\lambda(u\wedge\tau_D)}V\left(t-(u\wedge\tau_D), X_{u\wedge\tau_D}\right)$ is a martingale, so taking $u = t-s$ and using the facts that $V(t,x)$ is non-decreasing in t and that

$V(t-\tau_D, X_{\tau_D}) = g(X_{\tau_D})$ we have

$$
\begin{aligned}
0 &\le V(t,x) - V(s,x) \\
&= E_x^R[e^{-\lambda((t-s)\wedge\tau_D)}V(t-((t-s)\wedge\tau_D), X_{(t-s)\wedge\tau_D}) - V(s,x)] \\
&= E_x^R\left[\mathbf{1}_{[t-s\le\tau_D]}[e^{-\lambda(t-s)}V(s,X_{t-s}) - V(s,x)] + \mathbf{1}_{[\tau_D<t-s]}[e^{-\lambda\tau_D}V(t-\tau_D, X_{\tau_D}) - V(s,x)]\right] \\
&= E_x^R\Big[\mathbf{1}_{[t-s\le\tau_D]}[e^{-\lambda(t-s)}(V(s,X_{t-s}) - V(s,x)) + (e^{-\lambda(t-s)}-1)V(s,x)] \\
&\qquad + \mathbf{1}_{[\tau_D<t-s]}\left(e^{-\lambda\tau_D}(g(X_{\tau_D}) - g(x)) + e^{-\lambda\tau_D}(g(x) - V(s,x)) + (e^{-\lambda\tau_D}-1)V(s,x)\right)\Big].
\end{aligned}
$$

Using the fact that $e^{-\lambda y} - 1 \le \lambda y$ for $y \ge 0$, that $V(t,x)$ is Lipschitz in x, the fact that $g(x) \le V(t,x)$, and linear growth condition (4.7) for V, we have

$$
\begin{aligned}
0 &\le V(t,x) - V(s,x) \\
&\le C \cdot E_x^R\Big[\mathbf{1}_{[t-s\le\tau_D]}\left(e^{-\lambda(t-s)}|X_{t-s} - x| + \lambda(t-s)V(s,x)\right) \\
&\qquad\qquad + \mathbf{1}_{[\tau_D\le t-s]}\left(e^{-\lambda\tau_D}|X_{\tau_D} - x| + \lambda\tau_D V(s,x)\right)\Big] \\
&\le C\left(E_x^R[|X_{t-s} - x|] + \lambda(t-s)(1+|x|)\right) \\
&\le C(1+|x|)|t-s|^{\frac{1}{2}}
\end{aligned}
$$

where we let C denote some generic constant which depends on T. Note that for the final inequality above we used (ii) of Lemma 1 in Section 3.2.4 and the fact that, for $s, t \in [0,T]$, $|t-s| \le |t-s|^{\frac{1}{2}}T$. □

4.4 Hölder Continuity

When we prove the value function is smooth in Chapter 6, we will need joint Hölder continuity of $V(t,x)$ in t and x in order to apply a result from classical PDE theory. We prove the following lemma for general functions $f(t,x)$ which are α-Hölder continuous in each variable separately.

Lemma 4. *Let $0 < \alpha \le 1$ and suppose $f(t,x)$ is α-Hölder continuous in t, uniformly in x and also α-Hölder continuous in x, uniformly in t. Then f is α-Hölder continuous in (t,x).*

Proof. The result will follow if we can show that for some generic constant C,

$$|f(t,x) - f(s,y)|^{\frac{2}{\alpha}} \leq C|(t-s)^2 + (x-y)^2|.$$

This is true since

$$\begin{aligned} |f(t,x) - f(s,y)|^{2/\alpha} &\leq \Big(|f(t,x) - f(s,x)| + |f(s,x) - f(s,y)|\Big)^{2/\alpha} \\ &\leq C[|t-s|^{\alpha} + |x-y|^{\alpha}]^{2/\alpha} \leq C|(t-s)^2 + (x-y)^2|, \end{aligned}$$

where for the second inequality we used the Hölder continuity of f in t and in x and for the final inequality we used the fact that the function $z \to z^{2/\alpha}$ is convex whenever $\alpha \leq 2$. □

Note that by Lemma 3, on any bounded set $V(t,x)$ is $\frac{1}{2}$-Hölder continuous in t. Furthermore, since $V(t,x)$ is Lipschitz continuous in x, it is $\frac{1}{2}$-Hölder continuous in x on any bounded set. Thus we have the following corollary:

Corollary 1. *On any bounded subset of* $[0,T] \times [1,\infty]$ *the value function* $V(t,x)$ *is* $\frac{1}{2}$*-Hölder continuous.*

Chapter 5

Viscosity Solutions

Before proceeding, we will need to introduce the notion of a *viscosity solution* of a partial differential equation, which we will use below to prove regularity results for $V(t,x)$. Viscosity solutions were first introduced by M.G. Crandall and P.L. Lions [12] for first order partial differential equations and later expanded to second-order partial differential equations by Lions [21].

Given a typical differential equation of order n, it is natural to require any solutions of such a differential equation be n-times continuously differentiable. Such a solution is known as a classical solution. Unfortunately, many naturally arising partial differential equations do not have classical solutions. Instead, we need to interpret the solution to such a PDE in a different sense which does not require differentiability. Such a solution is known as a “generalized" or “weak solution", and viscosity solutions are an example of a generalized solution.

5.1 Motivation and Definition of Viscosity Solutions

Before we define viscosity solutions for second-order parabolic partial integro-differential equations, let us look at the motivation for them in the case of general parabolic equations given by

$$D_t u(t,x) + F\Big(t, x, u(t,x), D_x u(t,x), D^2_{xx} u(t,x)\Big) = 0. \tag{5.1}$$

We will require $F(t, x, r, p, X) : [0, T] \times \mathbb{R}^4 \to \mathbb{R}$ to satisfy the monotonicity conditions given by

$$F(t, x, m, p, X) \leq F(t, x, n, p, X) \text{ whenever } m \leq n, \tag{5.2}$$

and

$$F(t, x, m, p, X) \leq F(t, x, m, p, Y) \text{ whenever } Y \leq X. \tag{5.3}$$

When F satisfies such conditions we say that F is *proper*. In fact, in our motivation below, we will really only need F to satisfy (5.3) but (5.2) is required in order for further theory, such as uniqueness, to hold.

For the moment, let us assume that (5.1) has a $C^{1,2}$ *subsolution* u; that is, u satisfies

$$D_t u(t, x) + F\Big(t, x, u(t, x), D_x u(t, x), D^2_{xx} u(t, x)\Big) \leq 0 \tag{5.4}$$

for all $(t, x) \in [0, T] \times \mathbb{R}$. Next let $\phi(t, x) : [0, T] \times \mathbb{R} \to \mathbb{R}$ be a $C^{1,2}$ function such that $u - \phi$ has a local maximum at the point (t_0, x_0). Then from classical calculus we know that $D_x u(t_0, x_0) = D_x\phi(t_0, x_0)$, $D_t u(t_0, x_0) = D_t\phi(t_0, x_0)$ and $D_{xx} u(t_0, x_0) \leq D_{xx}\phi(t_0, x_0)$. Using these facts, (5.3), and (5.4), we see that

$$\begin{aligned} 0 &\geq D_t u(t_0, x_0) + F\Big(t_0, x_0, u(t_0, x_0), D_x u(t_0, x_0), D^2_{xx} u(t_0, x_0)\Big) \\ &\geq D_t\phi(t_0, x_0) + F\Big(t_0, x_0, u(t_0, x_0), D_x\phi(t_0, x_0), D^2_{xx}\phi(t_0, x_0)\Big). \end{aligned} \tag{5.5}$$

Alternatively, let u be a *supersolution* of (5.1), so that

$$D_t u(t, x) + F\Big(t, x, u(t, x), D_x u(t, x), D^2_{xx} u(t, x)\Big) \geq 0.$$

If ϕ is such that $u - \phi$ has a local minimum at (t_0, x_0), then by similar reasoning to that above we have

$$\begin{aligned} 0 &\leq D_t u(t_0, x_0) + F\Big(t_0, x_0, u(t_0, x_0), D_x u(t_0, x_0), D^2_{xx} u(t_0, x_0)\Big) \\ &\leq D_t\phi(t_0, x_0) + F\Big(t_0, x_0, u(t_0, x_0), D_x\phi(t_0, x_0), D^2_{xx}\phi(t_0, x_0)\Big). \end{aligned} \tag{5.6}$$

The important point is that the right-most sides of inequalities (5.5) and (5.6) do not depend on the derivatives of u, which suggests that we might use this idea to define a generalized solution to our equation. With this idea in mind, let us give a formal definition of a viscosity solution for second-order parabolic partial differential equations:

Definition 4. *For a proper function F and any open set Ω, we say $u \in C^0([0,T] \times \Omega)$, the set of continuous functions defined on $[0,T] \times \Omega$, is a viscosity supersolution (subsolution) of the second-order parabolic partial differential equation given by* (5.1) *if*

$$D_t\phi(t_0, x_0) + F(t_0, x_0, u(t_0, x_0), D_x\phi(t_0, x_0), D^2_{xx}\phi(t_0, x_0)) \geq 0 \tag{5.7}$$

(≤ 0) whenever $\phi \in C^{1,2}((0,T) \times \Omega)$ and $u - \phi$ has a local minimum (maximum) at $(t_0, x_0) \in (0,T) \times \Omega$. We say u is a viscosity solution of (5.7) *if it is both a viscosity supersolution and a viscosity subsolution.*

Note that this definition is for partial differential equations, which are only able to capture the local behavior of our process. Since we are working in a jump model, we will need viscosity solutions for second-order parabolic integro-differential equations, which have an additional integral term that arises as a result of the non-local behavior caused by jumps. The discussion in this section is meant only to serve as a motivation for viscosity solutions; we will develop a definition of viscosity solutions specialized to our problem in the next section.

5.2 The Variational Inequality

Central to Optimal Stopping Theory is the Hamilton-Jacobi-Bellman (HJB) equation associated with a given value function which, as discussed briefly in Section 2.6, is given by a variational inequality. Let us discuss the variational inequality arising for our problem. Define the set

$$\mathcal{A}([0,T] \times [1,\infty)) = C^{1,2}((0,T) \times (1,\infty)) \cap C^0([0,T] \times [1,\infty)),$$

the set of all functions which are continuous on $[0,T] \times [1,\infty)$ and are twice differentiable in x and once once differentiable in t on $(0,T) \times (1,\infty)$. Furthermore, for $u \in \mathcal{A}([0,T] \times [1,\infty))$ define the operators

$$Au(t,x) \coloneqq \alpha(x)\frac{\partial u}{\partial x}(t,x) + \frac{1}{2}\beta^2(x)\frac{\partial^2 u}{\partial x^2}(t,x)$$
$$Bu(t,x) \coloneqq \lambda^R \int_{\mathbb{R}} u(t, x + \zeta(x,z)) - u(t,x) - \zeta(x,z)\frac{\partial u}{\partial x}(t,x)\, m^R(dz).$$

where $\alpha(x)$, $\beta(x)$, and $\zeta(x,z)$ are given by (3.7). Then for $V(t,x)$, the HJB equation is given by the variational inequality

$$\min\{-\mathcal{L}V(t,x); V(t,x)-g(x)\} = 0 \tag{5.8}$$

in $[0,T]\times[1,\infty)$, where, for $u \in C^{1,2}((0,T)\times(1,\infty))$, $\mathcal{L}$ is the second-order parabolic partial integro-differential operator defined by

$$\mathcal{L}u = -\lambda u(t,x) - \frac{\partial u}{\partial t} + Au(t,x) + Bu(t,x), \tag{5.9}$$

and as before $g(x) = x$ is our payoff function.

It is well known that the corresponding HJB equation for the American option in a jump diffusion model does not have a classical solution, and so we should not expect a classical solution to exist for our problem above. Instead, we will use the notion of viscosity solutions in order to give meaning to (5.8). This allows us to work with the HJB equation without knowing a priori if $V(t,x)$ is differentiable. However, in order for (5.8) to be defined even for smooth test functions ϕ we need to restrict our attention to ϕ for which the integral term in $B\phi(t,x)$ is convergent. Let us define the set

$$C_2([0,T]\times[1,\infty)) := \left\{\phi \in C^0([0,T]\times[1,\infty)) : \sup_{[0,T]\times[1,\infty)} \frac{|\phi(t,x)|}{1+|x|^2} < \infty\right\}$$

and the set $\mathcal{A}_2([0,T]\times[1,\infty)) = \mathcal{A}([0,T]\times[1,\infty)) \cap C_2([0,T]\times[1,\infty))$. Then if $\phi \in \mathcal{A}_2([0,T]\times[1,\infty))$, we see that there exists a constant $C_{p,x}$ depending only on $p := \frac{\partial u}{\partial x}(t,x)$ and x such that the integrand

$$\phi(t, x+\zeta(x,z)) - \phi(t,x) - \zeta(x,z)\frac{\partial u}{\partial x}(t,x)$$

is bounded by $C_{p,x}(1+|\zeta(x,z)|^2)$. Thus $B\phi(t,x)$ is convergent by Lemma 1, and so (5.8) is well defined for $\phi \in \mathcal{A}_2([0,T]\times[1,\infty))$. With this in mind, let us give the revised definition of a viscosity solution which we will be working with:

Definition 5. *We say $u \in C^0([0,T]\times[1,\infty))$ is a viscosity supersolution (subsolution) of* (5.8) *if*

$$\min\left\{\lambda u(t,x) + \frac{\partial \phi}{\partial t}(t,x) - A\phi(t,x) - B\phi(t,x); u(t,x) - g(x)\right\} \geq 0 \tag{5.10}$$

(≤ 0) whenever $\phi \in \mathcal{A}_2([0,T]\times[1,\infty))$ and $u-\phi$ has a global minimum (maximum) at $(t,x) \in (0,T)\times(1,\infty)$. We say u is a viscosity solution of (5.8) *if it is both a viscosity supersolution and a viscosity subsolution.*

Notice that in (5.10) we obtained a partial integro-differential equation where all derivatives are taken only with respect to the test function. The integral term in $B\phi(t,x)$ depends on both ϕ and $\frac{\partial\phi}{\partial x}$, and one might ask if we can replace the undifferentiated terms by our original function u. Our next result shows that this is indeed the case if u is sufficiently regular. Before we proceed, let us define the operator

$$\tilde{B}(t,x,D_x\phi(t,x),u) := \lambda^R \int_{\mathbb{R}} u(t,x+\zeta(x,z)) - u(t,x) - \zeta(x,z)D_x\phi(t,x)\, m^R(dz).$$

Using this definition we obtain the following equivalent formulation of a viscosity solution for $u \in \mathcal{A}_2([0,T]\times[1,\infty))$:

Proposition 6. *A function $u \in \mathcal{A}_2([0,T]\times[1,\infty))$ is a viscosity supersolution (subsolution) of* (5.8) *if and only if*

$$\min\left\{\lambda u(t,x) + \frac{\partial\phi}{\partial t}(t,x) - A\phi(t,x) - \tilde{B}(t,x,D_x\phi(t,x),u);\ u(t,x) - g(x)\right\} \geq 0 \qquad (5.11)$$

(≤ 0) whenever $\phi \in \mathcal{A}_2([0,T]\times[1,\infty))$ and $u-\phi$ has a global minimum (maximum) at $(t,x) \in (0,T)\times(1,\infty)$.

Proof. We show the result only for supersolutions as the proof for subsolutions is similar. We begin by assuming (5.11) holds whenever $\phi \in \mathcal{A}_2([0,T]\times[1,\infty))$ and $u-\phi$ has a global minimum at (t,x). Let (t_0,x_0), u and $\phi \in \mathcal{A}_2([0,T]\times[1,\infty))$ be such that

$$u(t_0,x_0) - \phi(t_0,x_0) = \min_{(t,x)\in(0,T)\times(1,\infty)} u(t,x) - \phi(t,x),$$

which implies that

$$u(t,x) - u(t_0,x_0) \geq \phi(t,x) - \phi(t_0,x_0) \qquad (5.12)$$

for all $(t,x) \in (0,T)\times(1,\infty)$. Since (5.11) holds, we have

$$\min\{\lambda u(t_0,x_0) + \frac{\partial\phi}{\partial t}(t_0,x_0) - A\phi(t_0,x_0) - \tilde{B}(t,x,D_x\phi(t,x),u(t,x)); u(t_0,x_0) - g(x_0)\} \geq 0.$$

Combining this fact with the fact that

$$\begin{aligned}\tilde{B}(t_0,x_0,D_x\phi(t,x),u) &= \lambda^R \int_{\mathbb{R}} u(t_0,x_0+\zeta(x_0,z)) - u(t_0,x_0) - \zeta(x_0,z)D_x\phi(t_0,x_0)\, m^R(dz)\\ &\geq \lambda^R \int_{\mathbb{R}} \phi(t_0,x_0+\zeta(x_0,z)) - \phi(t_0,x_0) - \zeta(x_0,z)D_x\phi(t_0,x_0)\, m^R(dz)\\ &= B\phi(t_0,x_0)\end{aligned}$$

by (5.12) we see that u is a supersolution.

Conversely, suppose u is a supersolution. Let $\phi \in \mathcal{A}_2([0,T]\times[1,\infty))$ and (t_0, x_0) be such that

$$(u-\phi)(t_0,x_0) = \min_{(t,x)\in(0,T)\times(1,\infty)} (u-\phi)(t,x).$$

We may suppose without loss of generality that $(u-\phi)(t_0,x_0) = 0$ since otherwise we can use $\tilde{\phi}(t,x) = \phi(t,x) + u(t_0,x_0) - \phi(t_0,x_0)$ as $\tilde{\phi}$ has the same derivatives as ϕ. Finally, let $\epsilon > 0$ and let $\chi^\epsilon(t,x)$ be a smooth function on $[0,T]\times[1,\infty)$ satisfying

$$\begin{aligned}
&0 \le \chi^\epsilon(t,x) \le 1 \text{ for all } (t,x) \in [0,T]\times[1,\infty)\\
&\chi^\epsilon(t,x) = 1 \text{ if } x \in (x_0-\epsilon, x_0+\epsilon) \cap [1,\infty)\\
&\chi^\epsilon(t,x) = 0 \text{ if } x \in \Big((-\infty, x_0-2\epsilon] \cup [x_0+2\epsilon,\infty)\Big) \cap [1,\infty).
\end{aligned}$$

For $\delta > 0$, let $\phi_\delta \in \mathcal{A}_2([0,T]\times[1,\infty))$ be such that

$$\|u - \phi_\delta\|_\infty < \delta \tag{5.13}$$

and $\phi_\delta \le u$ which we can find since $u \in \mathcal{A}_2([0,T]\times[1,\infty))$. Define the function $\Phi^\epsilon_\delta(t,x) = \chi^\epsilon(t,x)\phi(t,x) + (1-\chi^\epsilon(t,x))\phi_\delta(t,x)$, and note $\Phi^\epsilon_\delta(t_0,x_0) = \phi(t_0,x_0) = u(t_0,x_0)$ and that

$$u(t,x) - \Phi^\epsilon_\delta(t,x) = \chi^\epsilon(t,x)(u(t,x)-\phi(t,x)) + (1-\chi^\epsilon(t,x))(u(t,x)-\phi_\delta(t,x)) \ge 0$$

since $u \ge \phi_\delta$ and $u \ge \phi$. Thus the supersolution conditions are satisfied for $\Phi^\epsilon_\delta(t,x)$ so

$$\begin{aligned}
0 &\le \lambda u(t_0,x_0) + \frac{\partial \Phi^\epsilon_\delta}{\partial t}(t_0,x_0) - A\Phi^\epsilon_\delta(t_0,x_0) - B\Phi^\epsilon_\delta(t_0,x_0)\\
&= \lambda u(t_0,x_0) + \frac{\partial \phi}{\partial t}(t_0,x_0) - A\phi(t_0,x_0) - \tilde{B}(t_0,x_0,D_x\phi(t_0,x_0),\Phi^\epsilon_\delta(t_0,x_0))
\end{aligned}$$

since the derivatives of $\Phi^\epsilon_\delta(t,x)$ equal the derivatives of $\phi(t,x)$ at (t_0,x_0). Now

$$\begin{aligned}
&|\tilde{B}(t_0,x_0,D_x\phi(t_0,x_0),u(t_0,x_0)) - \tilde{B}(t_0,x_0,D_x\phi(t_0,x_0),\Phi^\epsilon_\delta(t_0,x_0))| \le\\
&\lambda^R \int_{\mathbb{R}} |u(t_0,x_0+\zeta(x_0,z)) - \Phi^\epsilon_\delta(t_0,x_0+\zeta(x_0,z))| + |u(t_0,x_0) - \Phi^\epsilon_\delta(t_0,x_0)|\, m^R(dz)
\end{aligned} \tag{5.14}$$

and we can apply the Dominated Convergence Theorem to (5.14) since, for ϵ and δ both small,

$$|u(t,x) - \Phi^\epsilon_\delta(t,x)| \le |\chi^\epsilon(t,x)(u(t,x)-\phi(t,x))| + |(1-\chi^\epsilon(t,x))(u(t,x)-\phi_\delta(t,x))| \le C$$

where C is some constant which does not depend on ϵ or δ. Here we used (5.13) and the fact that $\chi^\epsilon(t,x)(u(t,x)-\phi(t,x))$ is nonzero only on a compact interval. Since $\Phi^\epsilon_\delta \to u$ if we let δ and then ϵ approach 0,

$$\tilde{B}(t_0, x_0, D_x\phi(t_0, x_0), \Phi^\epsilon_\delta(t_0, x_0)) \to \tilde{B}(t_0, x_0, D_x\phi(t_0, x_0), u(t_0, x_0))$$

and we obtain the result.

□

5.3 Proof that $V(t,x)$ is a Viscosity Solution to the HJB Equation

In this section we will show that $V(t,x)$ is a viscosity solution to (5.8). Before proceeding, we will need the following lemma which provides a generalized version of the Fundamental Theorem of Calculus when the variable we are differentiating with respect to is a stopping time.

Lemma 5. *For $\Omega \subseteq \mathbb{R}$, let Y be a process taking values in $\mathbb{R}$ and let $f(t,x) : [0,T]\times\Omega \to \mathbb{R}$ be such that $E_y\left[\int_0^a |f(u, Y_{u-})|\, du\right] < \infty$ for some $a > 0$ and satisfy*

$$f(u, Y_{u-}) \to f(0, Y_0) \text{ as } u \to 0^+. \tag{5.15}$$

Furthermore, let τ be a stopping time with $\tau > 0$ almost surely. Then

$$\lim_{b\to 0^+} \frac{E_y\left[\int_0^{\tau\wedge b} f(u, Y_{u-})\, du\right]}{E_y[\tau \wedge b]} = f(0, y).$$

Proof. We first show the result for bounded f. In this case, by Fubini's theorem we have

$$\frac{E_y\left[\int_0^{\tau\wedge b} f(u, Y_{u-})\, du\right]}{E_y[\tau \wedge b]} = \frac{\frac{1}{b}\int_0^b E_y\left[f(u, Y_{u-})\mathbf{1}_{[u<\tau]}\right] du}{E_y\left[\frac{\tau}{b} \wedge 1\right]}. \tag{5.16}$$

By the Dominated Convergence Theorem and (5.15), we see that

$$E_y[f(u, Y_{u-})\mathbf{1}_{[u<\tau]}] \to f(0, y)$$

as $u \to 0^+$, so letting $b \to 0^+$ above, using the Fundamental Theorem of Calculus, and using the fact that $E_y\left[\frac{\tau}{b} \wedge 1\right] \to 1$ by dominated convergence, we get that (5.16) approaches $f(0, y)$.

Next assume $f(t, x) \geq 0$, but is possibly unbounded. Let $f_k(t, x) = f(t, x) \wedge k$. Then by the Monotone Convergence Theorem and the result above for bounded f we have for $b \leq a$

$$\lim_{b \to 0^+} \frac{E_y\left[\int_0^{\tau\wedge b} f(u, Y_{u-})\, du\right]}{E_y[\tau \wedge b]} = \lim_{b\to 0^+} \lim_{k\to\infty} \frac{E_y\left[\int_0^{\tau\wedge b} f_k(u, Y_{u-})\, du\right]}{E_y[\tau \wedge b]}$$
$$= \lim_{k\to\infty} f(0, y) \wedge k = f(0, y).$$

Here we were able to interchange the limits since

$$\left| \frac{E_y\left[\int_0^{\tau\wedge b} f_k(u, Y_{u-})\, du\right]}{E_y[\tau \wedge b]} - f_k(0, y) \right| \leq \frac{E_y\left[\int_0^{\tau\wedge b} |f_k(u, Y_{u-}) - f_k(0, y)|\, du\right]}{E_y\,[\tau \wedge b]} < \epsilon$$

for any $\epsilon > 0$ if we choose b small enough, by (5.15). Thus the convergence of

$$\frac{E_y\left[\int_0^{\tau\wedge b} f_k(u, Y_{u-})\, du\right]}{E_y[\tau \wedge b]} \to f_k(0, y) \text{ as } b \to 0^+$$

is uniform in k, and the limit interchange above is permitted.

Finally, for general f satisfying our hypothesis, the result follows by applying the above to the positive and negative parts of f. □

For the remaining results of this section, we will need the following notation. For the value function $V(t, x)$ and $\phi \in \mathcal{A}_2([0, T] \times [1, \infty))$ such that $V \geq \phi\,(V \leq \phi)$, define $\mathcal{B}_n^+(\phi)\,(\mathcal{B}_n^-(\phi))$ to be the set of functions $\phi_n \in \mathcal{A}_2([0, T] \times [1, \infty))$ for which $\frac{\partial \phi_n}{\partial x}$ is bounded and which satisfy the following conditions:

i) $\phi_n(t, x) = \phi(t, x)$ for all $(t, x) \in [0, T] \times [1 + \frac{1}{n}, n)$.

ii) $|\phi_n - \phi|_\infty < \frac{1}{n}$ for all $(t, x) \in [0, T] \times [1, n)$.

iii) $\phi_n(t, x) \geq V(t, x)\,(\leq)$ for all $(t, x) \in [1, \infty)$.

iv) The right hand derivative $\frac{\partial \phi_n}{\partial x}(t, 1+)$ exists and equals 0 for all $t \in [0, T]$.

Observe in particular that we are able to find functions with bounded first derivative in x and satisfying the third condition above since $V(t,x) \leq C(1+|x|)$, so that $\mathcal{B}_n^+(\phi)$ and $\mathcal{B}_n^-(\phi)$ are nonempty.

Lemma 6. *Let $\phi_n \in \mathcal{B}_n^+(\phi) \cup \mathcal{B}_n^-(\phi)$. Then $A\phi_n(t_0,x_0) \to A\phi(t_0,x_0)$ and $B\phi_n(t_0,x_0) \to B\phi(t_0,x_0)$ for $(t_0,x_0) \in (0,T)\times(1,\infty)$ as $n \to \infty$.*

Proof. For n large, $\phi_n = \phi$, $D_t\phi_n = D_t\phi$, $D_x\phi_n = D_x\phi$, and $D_{xx}\phi_n = D_{xx}\phi$ at (t_0,x_0) since $(t_0,x_0) \in (0,T)\times(1,\infty)$ and when n is large enough, $\phi_n(t,x) = \phi(t,x)$ in a neighborhood of (t_0,x_0). Thus we have

$$\begin{aligned}|A\phi_n(t_0,x_0) - A\phi(t_0,x_0)| \leq & |\alpha(x_0)| \cdot |D_x\phi_n(t_0,x_0) - D_x\phi(t_0,x_0)| \\ & + \frac{1}{2}|\beta(x_0)|^2 \cdot |D_{xx}^2\phi_n(t_0,x_0) - D_{xx}^2\phi(t_0,x_0)|\end{aligned}$$

and the terms on the right equal 0 for n large enough. For B we have, when n is large,

$$\begin{aligned}&|B\phi_n(t_0,x_0) - B\phi(t_0,x_0)| \\ &\quad\leq \lambda^R \int_{\mathbb{R}} |\phi_n(t_0,x_0+\zeta(x_0,z)) - \phi(t_0,x_0+\zeta(x_0,z))| \\ &\qquad + |\phi_n(t_0,x_0) - \phi(t_0,x_0)| + |\zeta(x_0,z)||D_x\phi_n(t_0,x_0) - D_x\phi(t_0,x_0)|\, m^R(dz) \\ &\quad= \lambda^R \int_{\mathbb{R}} |\phi_n(t_0,x_0+\zeta(x_0,z)) - \phi(t_0,x_0+\zeta(x_0,z))| \cdot \mathbf{1}_{[1+\frac{1}{n}>x_0+\zeta(x_0,z)]}\, m^R(dz) \\ &\qquad + \lambda^R \int_{\mathbb{R}} |\phi_n(t_0,x_0+\zeta(x_0,z)) - \phi(t_0,x_0+\zeta(x_0,z))| \cdot \mathbf{1}_{[n\leq x_0+\zeta(x_0,z)]}\, m^R(dz).\end{aligned}$$

Using the fact that $|\phi_n - \phi| < \frac{1}{n}$ on $[1,n)$ we obtain

$$\begin{aligned}&|B\phi_n(t_0,x_0) - B\phi(t_0,x_0)| \\ &\quad\leq \frac{C}{n} + \lambda^R \int_{\mathbb{R}} |\phi_n(t_0,x_0+\zeta(x_0,z)) - \phi(t_0,x_0+\zeta(x_0,z))| \cdot \mathbf{1}_{[n\leq x_0+\zeta(x_0,z)]}\, m^R(dz)\end{aligned} \tag{5.17}$$

for some constant C. Next note that since ϕ and ϕ_n are in $\mathcal{A}_2([0,T]\times[1,\infty))$ we have

$$\begin{aligned}&|\phi_n(t_0,x_0+\zeta(x_0,z)) - \phi(t_0,x_0+\zeta(x_0,z))| \cdot \mathbf{1}_{[n\leq x_0+\zeta(x_0,z)]} \\ &\quad\leq C(1+|x_0+\zeta(x_0,z)|^2) \leq C_{x_0}(1+|\zeta(x_0,z)|^2) \in \mathcal{L}^2(m^R(dz))\end{aligned}$$

where C_{x_0} denotes a constant depending on x, but independent of z. Thus we can apply the Dominated Convergence Theorem in (5.17) and letting $n \to \infty$ the result follows. □

We now show the main result of this section. While the proof below uses standard techniques from Optimal Stopping Theory, note that special care has to be taken to deal with the boundary term at 1.

Theorem 3. *The value function $V(t, x)$ is a viscosity solution (in the sense of Definition 5) to* (5.8)*.*

Proof. First we show that $V(t, x)$ is a supersolution. Let $\phi \in \mathcal{A}_2([0, T] \times [1, \infty))$ be such that $V - \phi$ has a global minimum at $(t_0, x_0) \in (0, T) \times (1, \infty) \cap C$. Without loss of generality, we may assume that $V(t_0, x_0) = \phi(t_0, x_0)$, since otherwise we can work with $\tilde{\phi}(t, x) := \phi(t, x) + V(t_0, x_0) - \phi(t_0, x_0)$. Next, let $\{\phi_j\}_{j=1}^{\infty}$ be a family of test functions such that $\phi_n \in \mathcal{B}_n^-(\phi)$.

Let $\tau_D \equiv \tau_D(t_0, x_0)$. Since $e^{-\lambda(u\wedge\tau_D)}V(t-u\wedge\tau_D, X_{u\wedge\tau_D})$ is a martingale, after taking $u = \tau$ we have by the Optimal Stopping Theorem that $E_x^R[e^{-\lambda\tau}V(t-\tau, X_\tau)] = V(t, x)$ for any stopping time $\tau \leq \tau_D$. Thus

$$V(t_0, x_0) = E_{x_0}^R[e^{-\lambda\tau}V(t_0 - \tau, X_\tau)] \geq E_{x_0}^R[e^{-\lambda\tau}\phi_n(t_0 - \tau, X_\tau)]. \tag{5.18}$$

Applying Itô's Lemma (see (2.3)) to $e^{-\lambda t}\phi_n(t_0 - t, x)$ and letting $t = \tau$ we obtain

$$\begin{aligned}
&e^{-\lambda\tau}V(t_0 - \tau, X_\tau(x_0) \geq e^{-\lambda\tau}\phi_n(t_0 - \tau, X_\tau(x_0))\\
&=\phi_n(t_0, x_0) + \int_0^\tau e^{-\lambda u}\Big[- \lambda\phi_n(t_0 - u, X_{u-}) - \frac{\partial\phi_n}{\partial t}(t_0 - u, X_{u-})\\
&\qquad + \frac{\partial\phi_n}{\partial x}(t_0 - u, X_{u-})\alpha(X_{u-}) + \frac{1}{2}\frac{\partial^2\phi_n}{\partial x^2}(t_0 - u, X_{u-})\beta(X_{u-})^2\\
&\qquad - \zeta(X_{u-}, z)\frac{\partial\phi_n}{\partial x}(t_0 - u, X_{u-})\Big]\, du\\
&\quad + \int_0^\tau e^{-\lambda u}\frac{\partial\phi_n}{\partial x}(t_0 - u, X_{u-})\beta(X_{u-})\, dW_u + \int_0^\tau e^{-\lambda u}\frac{\partial\phi_n}{\partial x}(t_0 - u, X_{u-})\mathbf{1}_{[\Delta M_u=0]}\,\frac{dM_u}{S_{u-}}\\
&\quad + \int_0^\tau\int_{\mathbb{R}} e^{-\lambda u}\Big[\phi_n(t_0 - u, X_{u-} + \zeta(X_{u-}, z)) - \phi_n(t_0 - u, X_{u-})\Big]\nu(du, dz)\\
&=\phi_n(t_0, x_0) + \int_0^\tau e^{-\lambda u}\Big[- \lambda\phi_n(t_0 - u, X_{u-}) - \frac{\partial\phi_n}{\partial t}(t_0 - u, X_{u-})\\
&\qquad + A\phi_n(t_0 - u, X_{u-}) + B\phi_n(t_0 - u, X_{u-})\Big]\, du\\
&\quad + N_\tau + \int_0^\tau e^{-\lambda u}\frac{\partial\phi_n}{\partial x}(t_0 - u, X_{u-})\mathbf{1}_{[\Delta M_u=0]}\,\frac{dM_u}{S_{u-}}
\end{aligned} \tag{5.19}$$

where, by the same arguments as in Section 3.3,

$$N_t := \int_0^t e^{-\lambda u}\frac{\partial \phi_n}{\partial x}(t_0-u, X_{u-})\beta(X_{u-})\,dW_u \\ + \int_0^t\int_{\mathbb{R}} e^{-\lambda u}\left[\phi_n(t_0-u, X_{u-}+\zeta(X_{u-},z)) - \phi_n(t_0-u, X_{u-})\right]\,\tilde{\nu}^R(du,dz)$$

is a martingale since $D_x\phi_n(t,x)$ is bounded for fixed n.

Now since $\frac{\partial\phi_n}{\partial x}(t_0-u, X_{u-}) = 0$ when $X_{u-} = 1$ and $\mathbf{1}_{[\Delta M_u=0]}\,dM_u = 0$ when $X_{u-} \neq 1$ we have

$$\int_0^\tau e^{-\lambda u}\frac{\partial\phi_n}{\partial x}(t_0-u, X_{u-})\mathbf{1}_{[\Delta M_u=0]}\,\frac{dM_u}{S_{u-}} = 0.$$

Using this, the fact that $\phi_n(t_0,x_0) = V(t_0,x_0)$, (5.18), and combining the fact that N_t is a martingale with the Optional Sampling Theorem we obtain, after taking expectations in (5.19),

$$0 \geq E^R_{x_0}\Big[\int_0^\tau e^{-\lambda u}\Big(-\lambda\phi_n(t_0-u, X_{u-}) - \frac{\partial\phi_n}{\partial t}(t_0-u, X_{u-}) \\ + A\phi_n(t_0-u, X_{u-}) + B\phi_n(t_0-u, X_{u-})\Big)\,du\Big].$$

Let $\tau = \tau_j := \tau_D \wedge \frac{1}{j}$ above, and note that $\tau_j > 0$ almost surely since X starts in C. Dividing by $E[\tau_j]$ and taking $j \to \infty$ we see from Lemma 5 that

$$0 \geq -\lambda\phi_n(t_0,x_0) - \frac{\partial\phi_n}{\partial t}(t_0,x_0) + A\phi_n(t_0,x_0) + B\phi_n(t_0,x_0)$$

for $(t_0,x_0) \in (0,T)\times(1+\frac{1}{n},\infty)\cap C$. Letting $n\to\infty$ and using Lemma 6 we have that

$$0 \leq \lambda\phi(t_0,x_0) + \frac{\partial\phi}{\partial t}(t_0,x_0) - A\phi(t_0,x_0) - B\phi(t_0,x_0) \\ = \lambda V(t_0,x_0) + \frac{\partial\phi}{\partial t}(t_0,x_0) - A\phi(t_0,x_0) - B\phi(t_0,x_0)$$

for $(t_0,x_0) \in (0,T)\times(1,\infty)\cap C$. Thus the variational inequality (5.10) is satisfied for $(t_0,x_0) \in (0,T)\times(1+\frac{1}{n},\infty)\cap C$, and since $V(t_0,x_0) = g(x_0)$ for $(t_0,x_0)\in D$ we see the supersolution part of (5.10) holds.

The subsolution part of the inequality is similar. Let ϕ be such that $V-\phi$ has a global maximum at $(t_0,x_0)\in(0,T)\times(1,\infty)\cap C$ and assume without loss of generality that $V(t_0,x_0) = \phi(t_0,x_0)$. Let $\{\phi_j\}_{j=1}^\infty$ be a sequence of functions such that $\phi_n \in \mathcal{B}^+_n(\phi)$.

Applying a similar argument to the one above, we see that

$$V(t_0, x_0) \leq \phi_n(t_0, x_0) + E^R_{x_0}\left[\int_0^{\tau} e^{-\lambda u}\Big(- \lambda\phi_n(t_0 - u, X_u) - \frac{\partial \phi_n}{\partial t}(t_0 - u, X_u) + A\phi_n(t_0 - u, X_u) + B\phi_n(t_0 - u, X_u)\Big)\, du\right].$$

Using the fact that $V(t_0, x_0) = \phi(t_0, x_0) = \phi_n(t_0, x_0)$ and applying Lemma 5 with a similar argument to the supersolution case we have

$$0 \leq -\lambda\phi_n(t_0, x_0) - \frac{\partial \phi_n}{\partial t}(t_0, x_0) + A\phi_n(t_0, x_0) + B\phi_n(t_0, x_0)$$

for $(t_0, x_0) \in (0, T) \times (1 + \frac{1}{n}, \infty) \cap C$. Applying Lemma 6 and using the fact that $\phi(t_0, x_0) = V(t_0, x_0)$ we see that

$$0 \geq \lambda\phi_n(t_0, x_0) + \frac{\partial \phi_n}{\partial t}(t_0, x_0) - A\phi_n(t_0, x_0) - B\phi_n(t_0, x_0)$$

for $(t_0, x_0) \in (0, T) \times (1 + \frac{1}{n}, \infty) \cap C$. For $(t_0, x_0) \in D$, we have $V(t_0, x_0) = g(t_0, x_0)$ so the subsolution part of the variational inequality holds and the result is proved.

□

Chapter 6

Smoothness of the Value Function

In this chapter we prove the smoothness of $V(t, x)$ inside the continuation region C using a modification of a technique from Pham [27]. As shown in Chapter 5, $V(t, x)$ satisfies an integro-differential equation in a viscosity sense. Working with this equation is difficult because of the integral term, but since V is Lipschitz we can rearrange the PIDE and fix the integral part to depend explicitly on $V(t, x)$, allowing us to obtain a boundary value problem for a second-order parabolic partial differential equation. Then, using standard results from PDE theory we will be able to show their exists a classical solution to this equation, which will also be a viscosity solution. The uniqueness of viscosity solutions will then allow us to conclude that $V(t, x)$ must equal this smooth solution in the continuation region C, and hence it must also be smooth.

6.1 The Boundary Value Problem

We begin by establishing a Dirichlet problem on a bounded region which has $V(t, x)$ as its boundary value. Let $(t, x) \in \mathcal{H} := C\backslash([0, T] \times \{1\} \cup \{0, T\} \times [1, \infty))$ and let $\mathcal{R} = (t_1, t_2) \times (x_1, x_2) \subset C$ be an open rectangle such that $(t, x) \in \mathcal{R}$. Note that we can find such an $\mathcal{R}$ since $\mathcal{H}$ is open. Consider the Dirichlet problem

$$\begin{cases} \tilde{\mathcal{L}}u(t, x) = f_V(t, x) \text{ for } (t, x) \in \mathcal{R} \\ u(t, x) = V(t, x) \text{ for } (t, x) \in \partial\mathcal{R} \end{cases} \tag{6.1}$$

where $\tilde{\mathcal{L}}$ is the second-order parabolic differential operator given by

$$\tilde{\mathcal{L}}u = \lambda u(t,x) + \frac{\partial u}{\partial t}(t,x) - \frac{1}{2}\beta(x)^2\frac{\partial^2 u}{\partial x^2}(t,x) - \left[\alpha(x) - \lambda^R \int_{\mathbb{R}} \zeta(x,z)\, m^R(dz)\right]\frac{\partial u}{\partial x}(t,x)$$

and $f_V(t,x) = \lambda^R \int_{\mathbb{R}} V(t, x+\zeta(x,z)) - V(t,x)\, m^R(dz)$. Note that $f_V(t,x)$ is well defined since $V(t,x)$ is Lipschitz continuous. Alternatively, we can write (6.1) in the equivalent form

$$\begin{cases} D_t u(t,x) + F(t,x,u(t,x),D_x u(t,x),D_{xx}u(t,x)) = 0 \text{ for } (t,x) \in \mathcal{R} \\ u(t,x) = V(t,x) \text{ for } (t,x) \in \partial\mathcal{R} \end{cases} \tag{6.2}$$

where we define the function

$$F(t,x,m,p,X) := \lambda m - \frac{1}{2}\beta^2(x)X - \left[\alpha(x) - \lambda^R \int_{\mathbb{R}} \zeta(x,z)\, m^R(dz)\right]p - f_V(t,x). \tag{6.3}$$

The key thing to notice about the boundary value problem given by (6.2) is that we have reduced our partial integro-differential equation to a partial differential equation by fixing V in the integral term. As we shall see, this fact, coupled with the fact that we fixed V on the boundary, will force V to be the only solution of (6.2).

With this in mind, let us show that $V(t,x)$ is a viscosity solution of (6.2), but in the sense of Definition 4 which defined viscosity solutions for partial differential equations instead of integro-differential equations. Having shown this, we will then be able to use well known uniquness results for viscosity solutions of PDEs.

The result follows almost immediately from Theorem 3, but one subtlety is that Theorem 3 says that $V(t,x)$ is a viscosity solution in the sense of Definition 5, which requires that all test functions ϕ be such that $V - \phi$ has a global maximum at (t_0, x_0), whereas Definition 4 merely requires $V - \phi$ to have a local maximum at (t_0, x_0). We prove the result:

Proposition 7. *The value function $V(t,x)$ is a viscosity solution (in the sense of Definition 4) to* (6.2) *on $\mathcal{R}$.*

Proof. We will show that V is a supersolution; the proof that V is a subsolution is similar. Let $\phi \in C^{1,2}(\mathcal{R})$ be such that $u - \phi$ has a local minimum at (t_0, x_0). Next, let $\tilde{\phi} \in \mathcal{A}_2([0,T]\times[1,\infty))$ be a function such that $u - \tilde{\phi}$ has a global minimum at (t_0, x_0) and such that $\tilde{\phi} = \phi$ in a neighborhood of (t_0, x_0). Thus $\tilde{\phi}$ and ϕ, along with all their derivatives, agree at (t_0, x_0). Using this fact, Propostion 6, and Theorem 3 we have

$$\begin{aligned}
0 &\leq D_t\tilde{\phi}(t_0, x_0) + \lambda V(t_0, x_0) - \frac{1}{2}\beta(x_0)^2 D^2_{xx}\tilde{\phi}(t_0, x_0) \\
&\quad - \left[\alpha(x_0) - \lambda^R \int_{\mathbb{R}} \zeta(x_0, z)\, m^R(dz)\right] D_x\tilde{\phi}(t_0, x_0) \\
&\quad - \lambda^R \int_{\mathbb{R}} V(t_0, x_0 + \zeta(x_0, z)) - V(t_0, x_0)\, m^R(dz) \\
&= D_t\phi(t_0, x_0) + \lambda V(t_0, x_0) - \frac{1}{2}\beta(x_0)^2 D^2_{xx}\phi(t_0, x_0) \\
&\quad - \left[\alpha(x_0) - \lambda^R \int_{\mathbb{R}} \zeta(x_0, z)\, m^R(dz)\right] D_x\phi(t_0, x_0) - f_V(t_0, x_0) \\
&= D_t\phi(t_0, x_0) + F(t_0, x_0, V(t_0, x_0), D_x\phi(t_0, x_0), D_{xx}(t_0, x_0)),
\end{aligned}$$

which gives the result. □

Next, we will conclude (6.2) has a smooth solution. By Chapter 3 of Friedman [16], since $V(t, x)$ is continuous and $\tilde{\mathcal{L}}$ has uniformly Lipschitz continuous coefficients by Lemma 1, a smooth solution will exist provided we can show f_V is α-Hölder continuous on $\mathcal{R}$ for some $\alpha \leq 2$. Note that by Corollary 1 in Chapter 4, $V(t, x)$ is Hölder continuous with exponent $\frac{1}{2}$ in $\mathcal{R}$. The following lemma shows that $f_V(t, x)$ is also Hölder Continuous with exponent $\frac{1}{2}$.

Lemma 7. *The function $f_V(t, x)$ defined above is $\frac{1}{2}$-Hölder continuous in (t, x) on $\mathcal{R}$, and Lipschitz continuous in x uniformly in t.*

Proof. Let the points (t, x) and $(s, y) \in \mathcal{R}$. Then since $V(t, x)$ is α-Hölder continuous with $\alpha = \frac{1}{2}$ in $\mathcal{R}$ we have, for some generic constant C,

$$\begin{aligned}
&|f_V(t, x) - f_V(s, y)| \\
&\leq \quad \lambda^R \int_{\mathbb{R}} |V(t, x + \zeta(x, z)) - V(s, y + \zeta(y, z))| + |V(t, x) - V(s, y)|\, m^R(dz) \\
&\leq C \int_{\mathbb{R}} \left(|x - y|^2 + |\zeta(x, z) - \zeta(y, z)|^2 + |s - t|^2\right)^{\alpha/2}\, m^R(dz) + C\left(|x - y|^2 + |s - t|^2\right)^{\alpha/2} \\
&\leq C(|x - y|^2 + |s - t|^2)^{\alpha/2} \int_{\mathbb{R}} (\rho(z)^2 + 1)^{\alpha/2}\, m^R(dz) + C\left(|x - y|^2 + |s - t|^2\right)^{\alpha/2} \\
&\leq C\left(|x - y|^2 + |s - t|^2\right)^{\alpha/2}
\end{aligned}$$

where, for the third inequality we used Lemma 1. □

As a result of the previous discussion and Lemma 7 we see that a smooth solution to (6.2) exists. We will denote this solution by $u^V(t, x)$.

6.2 A Uniqueness Theorem for Viscosity Solutions

Our goal is to show that $u^V(t, x) = V(t, x)$ in $\mathcal{R}$, which will allow us to conclude that $V(t, x)$ is smooth in $\mathcal{R}$ and hence in C. This will require a uniqueness result for viscosity solutions. In particular, we will have use for the following uniqueness result from Crandall and Lions [13]:

Theorem 4. *Let $\Omega \subset \mathbb{R}$ be open and bounded, $T > 0$, and $\phi \in C(\bar{\Omega})$. Consider the boundary value problem given by*

$$\begin{cases} u_t(t, x) + H(t, x, u, D_x u(t, x), D^2_{xx} u(t, x)) = 0 & \text{for } (t, x) \in (0, T) \times \Omega \\ u(t, x) = 0 & \text{for } 0 \leq t < T \text{ and } x \in \partial\Omega \\ u(0, x) = \phi(x) & \text{for } x \in \bar{\Omega} \end{cases} \tag{6.4}$$

where $H : [0, T] \times \bar{\Omega} \times \mathbb{R}^3$ is continuous. Furthermore, suppose H is proper, i.e. that

$$H(t, x, m, p, X) \leq H(t, x, n, p, Y) \text{ if } Y \leq X \text{ and } m \leq n \tag{6.5}$$

holds. Finally, suppose H also satisfies, for any $\rho > 0$,

$$H(t, y, m, \rho(x - y), Y) - H(t, x, m, \rho(x - y), X) \leq \omega(\rho|x - y|^2 + |x - y|) \tag{6.6}$$

whenever

$$-3\rho \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \leq \begin{pmatrix} X & 0 \\ 0 & -Y \end{pmatrix} \leq 3\rho \begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix} \tag{6.7}$$

for some $\omega : [0, \infty) \to [0, \infty)$ such that $\omega(0+) = 0$. Then if u is a subsolution and v a supersolution of (6.4) (in the sense of Definition 4), then $u \leq v$ on $[0, T] \times \Omega$.

Our goal is to apply Theorem 4 with $H \equiv G$, where

$$\begin{aligned} G(x, m, p, X) &:= \lambda m - \frac{1}{2}\beta^2(x) X - \left[\alpha(x) - \lambda^R \int_{\mathbb{R}} \zeta(x, z)\, m^R(dz)\right] p \\ &= F(t, x, m, p, X) + f_V(t, x), \end{aligned} \tag{6.8}$$

but we must first check that G satisfies the conditions of the theorem. It is clear that G satisfies (6.5). As for (6.6), observe that if (6.7) holds then for any vector (ζ_1, ζ_2) we have

$$\begin{pmatrix} \zeta_1 \\ \zeta_2 \end{pmatrix}^T \begin{pmatrix} 3\rho - X & -3\rho \\ -3\rho & 3\rho + Y \end{pmatrix} \begin{pmatrix} \zeta_1 \\ \zeta_2 \end{pmatrix} \geq 0$$

which, after expanding, implies that

$$3\rho(\zeta_1 - \zeta_2)^2 \geq X\zeta_1^2 - Y\zeta_2^2.$$

Taking $\zeta_1 = \beta(x)$ and $\zeta_2 = \beta(y)$ we obtain

$$3\rho(\beta(x) - \beta(y))^2 \geq X\beta(x)^2 - Y\beta(y)^2. \tag{6.9}$$

Returning to (6.6), we can use (6.9) and Lemma 1 to see that for some constant C,

$$\begin{aligned} &G(t, y, m, \rho(x-y), Y) - G(t, x, m, \rho(x-y), X) \\ &\quad = \frac{1}{2}\beta^2(x)X - \frac{1}{2}\beta^2(y)Y + \left[\alpha(x) - \alpha(y) + \lambda^R \int_{\mathbb{R}} \zeta(y,z) - \zeta(x,z)\, m^R(dz)\right]\rho(x-y) \\ &\quad \leq \frac{3}{2}\rho(\beta(x) - \beta(y))^2 + C\rho|x-y|^2 + C|x-y| \leq C[\rho|x-y|^2 + |x-y|]. \end{aligned}$$

Thus taking $\omega(x) = Cx$ we see that (6.6) is satisfied for G, so that the conditions of Theorem 4 are met.

6.3 Proof the Value Function is Smooth Inside the Continuation Region

Finally, we prove the following result which will allow us to obtain the main theorem of this chapter:

Proposition 8. *Let $\Omega \subset \mathbb{R}$ be open, $T > 0$, and let $H : [0,T] \times \Omega \times \mathbb{R}^3$ be linear. Furthermore, let u and v be viscosity solutions (in the sense of Definition 4) to*

$$w_t(t,x) + H(t, x, w, D_x w(t,x), D^2_{xx} w(t,x)) = 0 \tag{6.10}$$

and suppose $u \in C^{1,2}((0,T) \times \Omega)$. Then $v - u$ is also a viscosity solution to (6.10).

Proof. We will show $w = v - u$ is a subsolution of (6.10). Let $\phi \in C^{1,2}((0,T)\times\Omega)$ and let $(t_0, x_0) \in \mathcal{R}$ be a local maximum of $w - \phi$. Then $\phi_1 := -u$ and $\phi_2 := \phi + u$ are $C^{1,2}((0,T)\times\Omega)$ functions such that $(-u) - \phi_1$ and $v - \phi_2$ have local maximums at (t_0, x_0). Since u and v are subsolutions, this implies that

$$D_t\phi_1(t_0,x_0) + H(t_0,x_0,-u(t_0,x_0),D_x\phi_1(t_0,x_0),D^2_{xx}\phi_1(t_0,x_0)) \leq 0 \tag{6.11}$$

$$D_t\phi_2(t_0,x_0) + H(t_0,x_0,v(t_0,x_0),D_x\phi_2(t_0,x_0),D^2_{xx}\phi_2(t_0,x_0)) \leq 0. \tag{6.12}$$

Then, by the linearity of H and the fact that $\phi = \phi_1 + \phi_2$ we have

$$\begin{aligned} &D_t\phi(t_0,x_0) + H(t_0,x_0,w(t_0,x_0),D_x\phi(t_0,x_0),D^2_{xx}\phi(t_0,x_0)) \\ &\quad = [D_t\phi_1(t_0,x_0) + H(t_0,x_0,v(t_0,x_0),D_x\phi_1(t_0,x_0),D^2_{xx}\phi_1(t_0,x_0))] \\ &\quad\quad + [D_t\phi_2(t_0,x_0) + H(t_0,x_0,-u(t_0,x_0),D_x\phi_2(t_0,x_0),D^2_{xx}\phi_2(t_0,x_0))] \leq 0 \end{aligned}$$

which implies that w is a subsolution. The result for supersolutions follows in the same manner. □

Remark 4. *In general, if u is a viscosity solution of* $H = 0$*, it is not necessarily true that* $-u$ *is a solution to* $H = 0$ *(see Crandall, et al [11]). We used the fact that u is a classical solution to conclude that* $-u$ *is also a solution to* (6.11).

Finally, let us use Proposition 8 to show $u^V(t,x) = V(t,x)$ in $\mathcal{R}$. Define

$$w(t,x) := V(t+t_1,x) - u^V(t+t_1,x).$$

Then since G is linear, by Proposition 7 and the fact that u^V is a smooth solution to (6.2) we have

$$\begin{aligned} &D_t w(t,x) + G(x,w(t,x),D_x w(t,x),D^2_{xx}w(t,x)) \\ &\quad = V_t(t+t_1,x) + G(x,V(t+t_1,x),D_xV(t+t_1,x),D^2_{xx}V(t+t_1,x)) \\ &\quad\quad - \left[u^V_t(t+t_1,x) + G(x,u^V(t+t_1,x),D_xu^V(t+t_1,x),D^2_{xx}u^V(t+t_1,x))\right] \\ &\quad = V_t(t+t_1,x) + F(t+t_1,x,V(t+t_1,x),D_xV(t+t_1,x),D^2_{xx}V(t+t_1,x)) \\ &\quad\quad - \left[u^V_t(t+t_1,x) + F(t+t_1,x,u^V(t+t_1,x),D_xu^V(t+t_1,x),D^2_{xx}u^V(t+t_1,x))\right] = 0 \end{aligned}$$

for $(t,x) \in (0, t_2 - t_1) \times (x_1, x_2)$. Thus w is a solution of

$$\begin{cases} v_t(t,x) + G(x,v(t,x),D_xv(t,x),D^2_{xx}v(t,x)) = 0 & \text{for } (t,x) \in (0,t_2-t_1)\times(x_1,x_2) \\ v(t,x) = 0 & \text{for } 0 \leq t \leq t_2 - t_1 \text{ and } x \in \{x_1,x_2\} \\ v(0,x) = 0 & \text{for } x \in [x_1,x_2]. \end{cases} \tag{6.13}$$

But since 0 is also a solution to (6.13), by Theorem 4 we can conclude $w = 0$ and thus $u^V(t,x) = V(t,x)$ in $\mathcal{R}$. By the continuity of V, we conclude $u^V = V$ on $\partial\mathcal{R}$ as well. Thus $V(t,x)$ is $C^{1,2}$ in $\bar{\mathcal{R}}$ and therefore in C.

As a consequence of this smoothness, we observe that in the continuation region $\mathcal{L}V = 0$ in the classical sense. This property will be very useful for later results.

Chapter 7

Properties of the Boundary Function

In this chapter we examine the behavior of the boundary function. We first show that the boundary function $b(t)$ must be finite, which means that for any fixed time t there must be some value the stock can take for which it is optimal to execute the option. Next we show that $b(t) = 1$ only when $t = 0$, which implies that it is never optimal to execute the option when it is at the maximum unless the option has expired.

Our third result concerns the behavior of $V_x(t, x)$ across the boundary. We show the so-called "Principle of Smooth Fit", which means that the function $x \to V_x(t, x)$ is smooth across the boundary. Finally, we show the continuity of the boundary function under an assumption similar to that of Pham [28]. This will be useful in Chapter 8 when we derive an early exercise premium representation for $V(t, x)$.

7.1 Finiteness of the Boundary

Arguing with a proof similar to that of Peskir [25], we can show that the boundary must be finite:

Proposition 9. *The boundary function $b(t)$ cannot take the value ∞.*

Proof. Using the fact that $b(t)$ is non-decreasing, first assume there exists $t_0 \in [0, T)$ such that $b(t) = \infty$ for all $t_0 \le t \le T$. This implies $(T, x) \in C$ for all $x \ge 1$, and thus $V(T, x) > x$. Let $\tau^* := \tau^*(T, x)$ denote the optimal stopping time for $V(T, x)$. Then by

(4.5) we have

$$x < V(T,x) = E_x^R\left[e^{-\lambda\tau^*}X_{\tau^*}\right]$$
$$= x + E_x^R\left[\int_0^{\tau^*} e^{-\lambda u}(-\lambda X_{u-} + \alpha(X_{u-}))\,du\right] + E_x^R\left[\int_0^{\tau^*} e^{-\lambda u}\mathbf{1}_{[\Delta M_u=0]}\,\frac{dM_u}{S_{u-}}\right]$$

which implies that

$$E_x^R\left[\int_0^{\tau^*} e^{-\lambda u}\mathbf{1}_{[\Delta M_u=0]}\,\frac{dM_u}{S_{u-}}\right] > E_x^R\left[\int_0^{\tau^*} e^{-\lambda u}(\lambda X_u - \alpha(X_{u-}))\,du\right]. \tag{7.1}$$

Now recalling that $M_t = S_t^* \vee x$, we must have

$$\begin{aligned} E_x^R\left[\int_0^{\tau^*} e^{-\lambda u}\mathbf{1}_{[\Delta M_u=0]}\,\frac{dM_u}{S_{u-}}\right] &\leq E_x^R\left[\left(\sup_{0\leq u\leq T}\frac{1}{S_u}\right)(M_{\tau^*} - M_0)\right] \\ = E^R\left[\left(\sup_{0\leq u\leq T}\frac{1}{S_u}\right)(S_{\tau^*}^* \vee x - x)\right] &\leq E^R\left[\left(\sup_{0\leq u\leq T}\frac{1}{S_u}\right)S_T^*\mathbf{1}_{[S_T>x]}\right] \end{aligned} \tag{7.2}$$

for all $x \geq 1$. Observe that (3.3) implies that S is a submartingale under R, and also observe that $e^{rt}S_t^{-1}$ is an R-martingale from the proof of Proposition 5. Hence using Cauchy's Inequality, Doob's Inequality in Mean, and these two facts we have

$$E^R\left[\left(\sup_{0\leq u\leq T}\frac{1}{S_u}\right)S_T^*\mathbf{1}_{[S_T>x]}\right] \leq E^R\left[\left(\sup_{0\leq u\leq T}\frac{1}{S_u}\right)S_T^*\right] < \infty.$$

Thus we can apply the Dominated Convergence Theorem and so the final term of (7.2) approaches 0 as $x \to \infty$. Next, setting $K := \lambda^R m^R(\mathbb{R})T$ and using the fact that $\tau^* \geq T - t_0$ we have

$$\begin{aligned} &E_x^R\left[\int_0^{\tau^*} e^{-\lambda u}(\lambda X_{u-} - \alpha(X_{u-}))\,du\right] \\ &= E_x^R\left[\int_0^{\tau^*} e^{-\lambda u}\left(\lambda X_{u-} + rX_{u-} - \lambda^R\int_{\{1+\gamma(z)>X_{u-}\}} 1 - \frac{X_{u-}}{1+\gamma(z)}\,m^R(dz)\,du\right)\right] \\ &\geq (\lambda + r)\,E_x^R\left[\int_0^{\tau^*} e^{-\lambda u}X_{u-}\,du\right] - E_x^R\left[\lambda^R\int_0^T\int_{\mathbb{R}} m^R(dz)\,du\right] \\ &\geq e^{-\lambda(T-t_0)}(\lambda + r)\,E_x^R\left[\int_0^{T-t_0} X_{u-}\,du\right] - K \\ &= e^{-\lambda(T-t_0)}(\lambda + r)\,E^R\left[\int_0^{T-t_0}\frac{S_{u-}^* \vee x}{S_{u-}}\,du\right] - K \\ &\geq e^{-\lambda(T-t_0)}(\lambda + r)x\,E^R\left[\int_0^{T-t_0}\frac{du}{S_{u-}}\right] - K \end{aligned}$$

which approaches ∞ as $x \to \infty$. Thus the inequality (7.1) cannot hold and we have a contradiction.

To see that the case $t_0 = T$ is impossible, note that the value problem with time horizon given by $T + \epsilon$ for $\epsilon > 0$ has the same boundary up to time T since the payoff function $g(x) = x$ is time homogeneous. By the argument above applied to this value function, $b(T) < \infty$. $\square$

7.2 Lower bound on the Boundary Function

Next, we will show that it is never optimal to stop when the stock process is at a maximum if $t > 0$. Our proof will be an extension of the one given for the diffusion case in Peskir [25], and we will require results about the scale function of spectrally negative Levy processes. From general results of Fluctuation Theory (see Kyprianou [20]), we know that for any spectrally negative Levy process Z_t there exists a function $w : [0, \infty) \to \mathbb{R}$ known as the *scale function* of Z_t. Letting $\tau_a := \inf\{t \geq 0 : Z_t \leq a\}$ and $\tau_b := \inf\{t \geq 0 : Z_t \geq b\}$ denote the first exit times of Z_t from a closed interval $[a, b]$, the scale function has the properties that

$$P_x(\tau_b \leq \tau_a) = \frac{w(x-a)}{w(b-a)}$$

and

$$P_x(\tau_a < \tau_b) = 1 - \frac{w(x-a)}{w(b-a)}.$$

While the specific form of w is not known for general spectrally negative Levy processes, it is well known that w is smooth and that if the diffusion component σ of Z_t is positive, then $w(0+) = 0$ and $w'(0+) = \frac{2}{\sigma^2}$. Using these properties we obtain the following lemma:

Lemma 8. *Let $\tau_n^- := \inf\{t \geq 0 : Z_t \leq -\frac{1}{n}\}$ and $\tau_n^+ := \inf\{t \geq 0 : Z_t \geq \frac{1}{n}\}$ and let $\tau_n = \tau_n^- \wedge \tau_n^+$, so that τ_n is the first exit time of Z_t from the interval $\left(-\frac{1}{n}, \frac{1}{n}\right)$. Then there exists positive constants K_1 and K_2, both independent of n, such that*

$$|P_x(\tau_n^+ \leq \tau_n^-) - P_x(\tau_n^- < \tau_n^+)| \leq \frac{K_1}{n}$$

and $E[\sup_{0 \leq t \leq \tau_n} Z_t] \geq \frac{K_2}{n}$ for n large.

Proof. Using the facts about the scale function above with $x = 0$ and taking the Taylor expansion of w we obtain

$$\begin{aligned}
|P_0(\tau_n^+ \le \tau_n^-) - P_0(\tau_n^- < \tau_n^+)| &= \left|\frac{2w(1/n) - w(2/n)}{w(2/n)}\right| \\
&= \left|\frac{2(w(0+) + w'(0+)(1/n) + O(1/n^2)) - (w(0+) + w'(0+)(2/n) + O(1/n^2))}{w(0+) + w'(0+)(2/n) + O(1/n^2)}\right| \\
&= \left|\frac{O(1/n^2)}{\frac{4}{\sigma^2}\frac{1}{n} + O(1/n^2)}\right| \le \frac{K_1}{n}
\end{aligned}$$

for some constant K_1. For the second part we have, after manipulating the Taylor expansion of $w(x)$,

$$E\left[\sup_{0\le t\le \tau_n} Z_t\right] \ge \frac{1}{n}P(\tau_{n+} < \tau_{n-}) = \frac{1}{n}\cdot\frac{w(1/n)}{w(2/n)} = \frac{1}{n}\cdot\frac{1+o(1)}{2+o(1)} \ge \frac{K_2}{n}$$

for n large, where for the first inequality we used the fact that Z_t is spectrally negative.

□

In addition to Lemma 8, we will need the following lemma:

Lemma 9. *Let Z_t be a spectrally negative jump diffusion given by*

$$Z_t := \mu t + \sigma B_t + \int_0^t \int_{\mathbb{R}} \rho(z)\,\tilde{\nu}(dz),$$

where B_t is a standard Brownian motion and $\rho(z) \le 0$. Furthermore, let $\rho(z) \in L^2(m^R)$ and let τ_n be as in Lemma 8. Then

$$E[\tau_n] \le \frac{K_2}{n^2}$$

for some positive constant K_2 which is independent of n.

Proof. First we show that for fixed n, we may assume that the jumps of Z_t are bounded below by $\frac{2}{n}$. Observe that a jump of size larger than $\frac{2}{n}$ forces Z_t to exit the interval $(-\frac{1}{n}, \frac{1}{n})$ immediately. Let $J_t := \int_0^t \int_{\mathbb{R}} \rho(z)\,\tilde{\nu}(dz)$ and for fixed n decompose J as $J = J^1 + J^2$ where

$$\begin{aligned}
J_t^1 &:= \int_0^t \int_{\mathbb{R}} \rho(z)\mathbf{1}_{[\rho(z)<-\frac{2}{n}]}\tilde{\nu}(dz) \\
J_t^2 &:= \int_0^t \int_{\mathbb{R}} \rho(z)\mathbf{1}_{[\rho(z)\ge-\frac{2}{n}]}\tilde{\nu}(dz),
\end{aligned}$$

so that J_t^1 represents the downward jumps of Z_t with magnitude larger than $\frac{2}{n}$ and J_t^2 represents the other jumps. Let $Z_t' := Z_t - J_t^1$ and let $\tau_n' := \inf\{t \geq 0 : |Z_t'| \geq \frac{1}{n}\}$ denote the first time Z_t' exits $\left(-\frac{1}{n}, \frac{1}{n}\right)$. Then $\tau_n \leq \tau_n'$ almost surely and so we can assume that for fixed n, the jumps of Z_t are bounded below by $\frac{2}{n}$.

We will divide the proof into three cases:

Case (i): $\mu > 0$. Since $L^2(m^R)$, J_t and B_t are $\mathcal{F}_t$-martingales. Thus $E[Z_{\tau_n \wedge T}] = \mu E[\tau_n \wedge T]$ for $T > 0$ by the Optional Sampling Theorem. Letting $T \to \infty$, we obtain from the Dominated and Monotone Convergence Theorems that

$$\begin{aligned} \mu E[\tau_n] = E[Z_{\tau_n}] &= E[Z_{\tau_n^+} \mathbf{1}_{[\tau_n^+ < \tau_n^-]}] + E[Z_{\tau_n^-} \mathbf{1}_{[\tau_n^- < \tau_n^+]}] \\ &\leq \frac{1}{n} P(\tau_n^+ \leq \tau_n^-) - \frac{1}{n} P(\tau_n^- < \tau_n^+) \leq \frac{K_1}{n^2}, \end{aligned} \tag{7.3}$$

by Lemma 8 and the fact that $Z_{\tau_n^+} = \frac{1}{n}$ since Z only has downward jumps. Dividing by μ and using the fact that $\mu > 0$ we obtain $E[\tau_n] \leq \frac{C}{n^2 \mu}$.

Case (ii): $\mu < 0$. This case is similar to case (i). Arguing as above we obtain

$$\mu E[\tau_n] = E[Z_{\tau_n}] = \frac{1}{n} P(\tau_n^+ \leq \tau_n^-) + E[Z_{\tau_n^-} \mathbf{1}_{\{\tau_n^- \leq \tau_n^+\} \cap A}] + E[Z_{\tau_n^-} \mathbf{1}_{\{\tau_n^- \leq \tau_n^+\} \cap A^c}]$$

where

$$A := \{Z_{t-} = Z_t \text{ for all } t \in [0, \tau_n]\}, \tag{7.4}$$

which is the event that Z_t doesn't jump before exiting $\left(-\frac{1}{n}, \frac{1}{n}\right)$. Thus

$$\begin{aligned} \mu E[\tau_n] &\geq \frac{1}{n} P(\tau_n^+ \leq \tau_n^-) - \frac{1}{n} P(\{\tau_n^- \leq \tau_n^+\} \cap A) - \frac{3}{n} P(\{\tau_n^- \leq \tau_n^+\} \cap A^c) \\ &\geq \frac{1}{n} P(\tau_n^+ \leq \tau_n^-) - \frac{1}{n} P(\tau_n^- < \tau_n^+) - \frac{3}{n} P(A^c) \end{aligned} \tag{7.5}$$

where for the first inequality we used the fact that the magnitude of the jumps of Z_t are bounded by $\frac{2}{n}$. Next let us assume the jumps of Z_t are controlled by a Poisson process with intensity λ. Then

$$P(A^c) = E[P(A^c | \tau_n)] = E[1 - e^{-\lambda \tau_n}] \leq \lambda E[\tau_n] \tag{7.6}$$

since $1 - e^{-x} \leq x$. Combining (7.5) with (7.6) and Lemma 8 we obtain

$$\mu E[\tau_n] \geq -\frac{K_1}{n^2} - \frac{3}{n} \lambda E[\tau_n]$$

which implies that $(\mu + \frac{3}{n}\lambda)E[\tau_n] \geq -\frac{K_1}{n^2}$. Let $n_0 := \inf\{n \in \mathbb{N} : \mu + \frac{3}{n}\lambda < 0\}$, which exists since $\mu < 0$. Then

$$E[\tau_n] \leq -\frac{1}{(\mu + \frac{3}{n}\lambda)}\frac{K_1}{n^2} \leq -\frac{1}{(\mu + \frac{3}{n_0}\lambda)}\frac{K_1}{n^2}$$

when $n \geq n_0$, and the result follows.

Case (iii): $\mu = 0$. In this case $Z_t = \sigma B_t + J_t$, and so $M_t := Z_t^2 - (\sigma^2 + 1)t$ is a $\mathcal{F}_t$-martingale since $B_t^2 - t$, $J_t^2 - t$, and $J_t B_t$ are $\mathcal{F}_t$-martingales. Thus by the Optimal Stopping Theorem $(\sigma^2 + 1)E[\tau_n \wedge T] = E[Z^2_{\tau_n \wedge T}]$. Similarly to case (ii), we let $T \to \infty$ to obtain

$$(\sigma^2 + 1)E[\tau_n] = E[Z^2_{\tau_n}] = E\left[Z^2_{\tau_n^+}\mathbf{1}_{[\tau_n^+ \leq \tau_n^-]}\right] + E\left[Z^2_{\tau_n^-}\mathbf{1}_{\{\tau_n^- < \tau_n^+\}\cap A}\right] + E\left[Z^2_{\tau_n^-}\mathbf{1}_{\{\tau_n^- < \tau_n^+\}\cap A^c}\right]$$

where A is defined by (7.4). Using the fact that $Z_{\tau_n^+} = \frac{1}{n}$ and $Z^2_{\tau_n^-} \leq \frac{9}{n^2}$ we obtain

$$(\sigma^2 + 1)E[\tau_n] \leq \frac{1}{n^2}P(\tau_n^+ \leq \tau_n^-) + \frac{1}{n^2}P(\tau_n^- < \tau_n^+) + \frac{9}{n^2}P(A^c) \leq \frac{1}{n^2} + \frac{9\lambda}{n^2}E[\tau_n]$$

by (7.6). Rearranging we obtain $\left(\sigma^2 + 1 - \frac{9\lambda}{n^2}\right)E[\tau_n] \leq \frac{1}{n^2}$. For n large, $\sigma^2 + 1 - \frac{9\lambda}{n^2} > 0$ and thus

$$E[\tau_n] \leq \frac{1}{(\sigma^2 + 1 - \frac{9\lambda}{n^2})}\frac{1}{n^2} \leq \frac{1}{\sigma^2}\frac{1}{n^2}$$

when $n \geq 3\sqrt{\lambda}$. □

Finally, let us prove that it is not optimal to stop the process early when we are at the maximum:

Proposition 10. *The boundary function $b(t)$ is strictly greater than 1 for $t \in (0, T]$.*

Proof. With the results above, the proof is the same as the proof in the diffusion case in Peskir [25] and is reproduced here for completeness. Reverting to the untransformed stopping problem under the measure Q given by (2.8), the result will follow if we can show that, for some $0 \leq \tau \leq T$,

$$E^Q_{s,s}[e^{-(r+\lambda)\tau}M_\tau(s, s)] > s$$

or, equivalently,

$$E^Q_{1,1}[e^{-(r+\lambda)\tau}M_\tau(1, 1)] > 1.$$

These inequalities imply that it is never optimal to stop on the diagonal, which is when $X_t = 1$.

We first show the result holds if we do not require τ to be bounded. By Jensen's Inequality, we have from (2.10)

$$E^{Q}_{1,1}[e^{-(\lambda+r)\tau} M_\tau] \geq \exp\left(E\left[\sup_{0\leq t\leq \tau} Z_t - (\lambda + r)\tau\right]\right), \tag{7.7}$$

where we define $Z_t := (r - \lambda^Q k^Q - \frac{1}{2}\sigma^2)t + \sigma W^Q_t + \int_0^t \int_{\{\gamma(z)<0\}} \ln(1 + \gamma(z))\, \nu(du, dz)$. Let τ_n be as in Lemma 8. Then for n large,

$$E\left[\sup_{0\leq t\leq \tau_n} Z_t\right] \geq \frac{K_1}{n} \text{ and } E[\tau_n] \leq \frac{K_2}{n^2}$$

for some constants K_1 and K_2 by Lemma 8 and Lemma 9 respectively. Thus for n large, by (7.7) we have

$$E_{1,1}[e^{-(\lambda+r)\tau_n} M_{\tau_n}] \geq \exp\left(\frac{K_1}{n} - \frac{K_2}{n^2}\right) > 1. \tag{7.8}$$

Hence the result holds when τ is not required to be bounded. Next, let $\sigma_n = \tau_n \wedge T$, so that

$$\begin{aligned} &E\left[\sup_{0\leq t\leq \tau_n} Z_t - \sup_{0\leq t\leq \sigma_n} Z_t\right] \leq \\ &\qquad E\left[\mathbf{1}_{[\tau_n > T]} \sup_{T\leq t\leq \tau_n} Z_t\right] \leq \frac{1}{n} P(\tau_n > T) \leq \frac{E[\tau_n]}{nT} \leq \frac{K_2}{n^3 T} \end{aligned} \tag{7.9}$$

where the third inequality follows from Markov's inequality and the final inequality follows from Lemma 9. Thus

$$\begin{aligned} E_{1,1}[e^{-(\lambda+r)\sigma_n} M_{\sigma_n}] &\geq \exp\left(E\left[\sup_{0\leq t\leq \sigma_n} Z_t - (\lambda + r)\sigma_n\right]\right) \\ &= \exp\left(E\left[\sup_{0\leq t\leq \tau_n} Z_t - (\lambda + r)\tau_n - \frac{K_2}{n^3 T}\right]\right) > 1 \end{aligned}$$

for n large, where we used (7.7), (7.9), the fact that $\tau_n > \sigma_n$, and the result above for unbounded τ. Hence $b(t) > 1$ for $t \in (0, T]$. □

7.3 Normal Reflection

With Proposition 10 proved, we now show that the value function has a right-hand derivative at $x = 1$ and that *normal reflection* holds. This will be useful later when proving uniqueness results for our free boundary problem.

Proposition 11. *(Normal Reflection): For $t \in (0, T]$, the right-hand derivative $V_x(t, 1+)$ exists and equals* 0*.*

Proof. It is clear $V_x(t, 1+)$ exists and $V_x(t, 1+) \geq 0$ since $x \to V(t, x)$ is increasing and convex on $[1, \infty)$. Suppose, by way of contradiction, that $V_x(t, 1+) > 0$ for some $t \in (0, T)$. By the continuity of $t \to V_x(t, 1+)$ there exists $\delta > 0$ such that $V_x(s, 1+) \geq \epsilon > 0$ for all $s \in [t, t+\delta]$ where $t + \delta < T$. Setting $\tau_\delta = \tau_D \wedge (t + \delta)$, it follows by Itô's Formula and the fact that $\mathcal{L}V = 0$ in C that

$$E_1^R[e^{-\lambda\tau_\delta} V(t - \tau_\delta, X_{\tau_\delta})] = V(t, 1) + E^R\left[\int_0^{\tau_\delta} e^{-\lambda u} V_x(t - u, X_u)\mathbf{1}_{[\Delta M_u = 0]} \frac{dM_u}{S_{u-}}\right].$$

Since $e^{-\lambda(s\wedge\tau_D)}V(t-(s\wedge\tau_D), X_{s\wedge\tau_D})$ is a martingale, the Optimal Sampling Theorem gives us that the left hand side above equals $V(t, 1)$ so

$$E^R\left[\int_0^{\tau_\delta} e^{-\lambda u} V_x(t - u, X_u)\mathbf{1}_{[\Delta M_u = 0]} \frac{dM_u}{S_{u-}}\right] = 0,$$

which implies

$$\int_0^{\tau_\delta} e^{-\lambda u} V_x(t - u, X_u)\mathbf{1}_{[\Delta M_u = 0]} \frac{dM_u}{S_{u-}} = 0$$

R-almost surely. The fact that $V_x(t - u, X_u) > 0$ means that M can only increase by jumping in $[0, \tau_\delta]$. This is a contradiction however, as there is always some probability in a finite interval that S doesn't jump and the Brownian part plus drift increases to a fixed level. Thus we obtain a contradiction and $V_x(t, 1+) = 0$ for all $t \in (0, T)$. The result for $t = T$ follows by taking limits and using the fact that V is $C^{1,2}$ in C. □

7.4 Smooth Fit at the Boundary

Next we show that the *principal of smooth fit holds*. This means that the value function is smooth across the boundary in x:

Proposition 12. *(Smooth Fit): For $t \in (0, T]$, $V_x(t, x) = 1$ for $x = b(t)$.*

Proof. By Proposition 10, we know that if $t > 0$ and $x = b(t)$ then $x > 1$. Thus there exists $\epsilon > 0$ such that $x - \epsilon > 1$. Since $V(t, x) = g(x)$ and $V(t, x - \epsilon) > g(x)$,

$$\frac{V(t, x) - V(t, x - \epsilon)}{\epsilon} < \frac{g(x) - g(x - \epsilon)}{\epsilon} = 1. \tag{7.10}$$

The convexity of $x \to V(t,x)$ implies that the left-hand derivative $V_x^-(t,x)$ exists, so letting $\epsilon \searrow 0$ we see from (7.10) that $V_x^-(t,x) \leq 1$.

For the reverse inequality, let τ_ϵ denote that optimal stopping time for $V(t, x-\epsilon)$. Then by (4.1),

$$\begin{aligned}
\frac{V(t,x) - V(t,x-\epsilon)}{\epsilon} &\geq \frac{1}{\epsilon} E^R \left[e^{-\lambda\tau_\epsilon} \left(\frac{(x - S^*_{\tau_\epsilon})^+ + S^*_{\tau_\epsilon}}{S_{\tau_\epsilon}} - \frac{(x - \epsilon - S^*_{\tau_\epsilon})^+ + S^*_{\tau_\epsilon}}{S_{\tau_\epsilon}} \right) \right] \\
&\geq \frac{1}{\epsilon} E^R \left[\frac{e^{-\lambda\tau_\epsilon}}{S_{\tau_\epsilon}} \left((x - S^*_{\tau_\epsilon})^+ - (x - \epsilon - S^*_{\tau_\epsilon})^+ \right) \mathbf{1}_{[S^*_{\tau_\epsilon} \leq x - \epsilon]} \right] \\
&= E^R \left[\frac{e^{-\lambda\tau_\epsilon}}{S_{\tau_\epsilon}} \mathbf{1}_{[S^*_{\tau_\epsilon} \leq x - \epsilon]} \right].
\end{aligned}$$

The final term above converges to 1 as $\epsilon \searrow 0$ by the Dominated Convergence Theorem, since as $\epsilon \searrow 0$, $\tau_\epsilon \to 0$ so that $S^*_{\tau_\epsilon} \to 1 < x - \epsilon$ and $S_{\tau_\epsilon} \to 1$. Thus $V_x^-(t,x) \geq 1$ and so $V_x^-(t,x) = 1$. For the right hand side, since $V(t,y) = g(y) = y$ for $y > x$ it is obvious that $V_x^+(t,x) = 1$.

□

7.5 Continuity of the Boundary Function

Next we consider the continuity of the boundary function $b(t)$. For the American option in a jump diffusion model, Pham [28] showed the continuity of the boundary under the assumption that

$$r - \lambda^P \int_{\{\gamma(z) \geq 0\}} \gamma(z) p(z)\, m^P(dz) > 0,$$

which is implied if the riskless interest rate minus the jump risk,

$$r - \lambda^P E^P[\gamma(Z_1)] \tag{7.11}$$

is positive, where Z_1 is a jump of our process as defined in Section 2.1. In the proof below, we will make the corresponding assumption for our problem:

$$(A) \qquad K := \lambda + r - \lambda^P \int_{\{\gamma(z) \geq 0\}} \gamma(z) p(z)\, m^P(dz) > 0. \tag{7.12}$$

Due to the presence of the λ term, this assumption is slightly weaker than the one given by Pham. Before we proceed with the proof of continuity, we will need the following lemmas:

Lemma 10. *Let* $t_0 \in (0, T)$*. For any* $\epsilon_1 > 0$*, there exists* $\epsilon_2 > 0$ *and* $y \in \mathbb{R}$ *satisfying* $1 < y < b(t_0)$ *such that* $V_x(t, x) > 1 - \epsilon_1$ *whenever* $t_0 < t < t_0 + \epsilon_2$ *and* $y < x < b(t)$*.*

Proof. Since $\lim_{x \to b(t_0)} V_x(t_0, x) = 1$ by Theorem 12, there exists $\epsilon_3 > 0$ and y satisfying $b(t_0) - \epsilon_3 < y < b(t_0)$ such that $V_x(t_0, y) > 1 - \epsilon_1/2$. Using the facts that $(t_0, y) \in C$ and that $V_x(t, x)$ is continuous in C, we see that there exists a neighborhood N of (t_0, y) such that $V_x(t, x) > 1 - \epsilon_1$ for $(t, x) \in N$. Hence we can find $\epsilon_2 > 0$ such that $(t_0 + \epsilon_2, y) \in N$, and so since $x \to V_x(t, x)$ is non-decreasing by the convexity of V, we have $V_x(t, x) > 1 - \epsilon_1$ for $t_0 < t < t_0 + \epsilon_2$ and $y < x < b(t)$. □

Lemma 11. *Let* $t_0 \in (0, T)$*. Under assumption (A), there exists* $\epsilon_1, \epsilon_2 > 0$ *and* $y \in \mathbb{R}$ *satisfying* $1 < y < b(t_0)$ *such that*

$$\frac{1}{2}\sigma^2 x^2 \frac{\partial^2 V}{\partial x^2}(t, x) \geq \epsilon_1$$

for $(t, x) \in A(t_0, \epsilon_2) := \{(t, x) : t_0 < t < t_0 + \epsilon_2$ *and* $y < x < b(t_0)\}$*.*

Proof. For (t, x) in the continuation region C, we know $\mathcal{L}V(t, x) = 0$ so

$$\begin{aligned}
&\frac{1}{2}\sigma^2 x^2 \frac{\partial^2 V}{\partial x^2}(t, x) \\
&\quad = \lambda V(t, x) + \frac{\partial V}{\partial t}(t, x) + rx \frac{\partial V}{\partial x}(t, x) \\
&\qquad - \lambda^R \int_{\mathbb{R}} V(t, x + \zeta(x, z)) - V(t, x) + x \frac{\gamma(z)}{1 + \gamma(z)} \frac{\partial V}{\partial x}(t, x) \, m^R(dz) \\
&\quad \geq \lambda x + x \left(r - \lambda^R \int_{\mathbb{R}} \frac{\gamma(z)}{1 + \gamma(z)} \, m^R(dz) \right) \frac{\partial V}{\partial x}(t, x) \\
&\qquad + \lambda^R \int_{\mathbb{R}} V(t, x) - V(t, x + \zeta(x, z)) \, m^R(dz)
\end{aligned}$$

since $V(t, x) > x$ in C and $t \to V(t, x)$ is non-decreasing so that $V_t \geq 0$. Using the fact that $x \to V(t,x)$ is increasing x we further obtain

$$
\begin{aligned}
&\frac{1}{2}\sigma^2 x^2 \frac{\partial^2 V}{\partial x^2}(t,x) \\
&\quad \geq \lambda x + x\left(r - \lambda^R \int_{\mathbb{R}} \frac{\gamma(z)}{1+\gamma(z)}\, m^R(dz)\right)\frac{\partial V}{\partial x}(t,x) \\
&\qquad + \lambda^R \int_{\{\zeta(x,z)\geq 0\}} V(t,x) - V(t, x+\zeta(x,z))\, m^R(dz) \\
&\quad \geq \lambda x + x\left(r - \lambda^R \int_{\mathbb{R}} \frac{\gamma(z)}{1+\gamma(z)}\, m^R(dz)\right)\frac{\partial V}{\partial x}(t,x) \\
&\qquad - \lambda^R \int_{\{\zeta(x,z)\geq 0\}} \zeta(x,z)\, m^R(dz) \\
&\quad \geq \lambda + \left(r - \lambda^R \int_{\mathbb{R}} \frac{\gamma(z)}{1+\gamma(z)}\, m^R(dz)\right)\frac{\partial V}{\partial x}(t,x) \\
&\qquad + \lambda^R \int_{\{\gamma(z)<0\}} \frac{\gamma(z)}{1+\gamma(z)}\, m^R(dz).
\end{aligned}
\tag{7.13}
$$

Here we used Proposition 5 for the second inequality and for the final inequality we used the fact that $x \geq 1$ and Proposition 4 along with the fact that $\{\zeta(x,z) \geq 0\} = \{\gamma(z) < 0\}$. Fix ϵ_1 with $0 < \epsilon_1 < K$, where K is given in (7.12). By assumption (A), we can find δ sufficiently small such that

$$
\begin{aligned}
&\lambda + \left(r - \lambda^R \int_{\mathbb{R}} \frac{\gamma(z)}{1+\gamma(z)}\, m^R(dz)\right)(1-\delta) + \lambda^R \int_{\{\gamma(z)<0\}} \frac{\gamma(z)}{1+\gamma(z)}\, m^R(dz) \\
&= \lambda + r - \lambda^R \int_{\{\gamma(z)\geq 0\}} \frac{\gamma(z)}{1+\gamma(z)}\, m^R(dz) - \delta\left(r - \lambda^R \int_{\mathbb{R}} \frac{\gamma(z)}{1+\gamma(z)}\, m^R(dz)\right) \\
&= \lambda + r - \lambda^P \int_{\{\gamma(z)\geq 0\}} \gamma(z)p(z)\, m^P(dz) - \delta\left(r - \lambda^P \int_{\mathbb{R}} \gamma(z)p(z)\, m^P(dz)\right) \geq \epsilon_1,
\end{aligned}
\tag{7.14}
$$

since

$$
\lambda^R \int_{\{\gamma(z)\geq 0\}} \frac{\gamma(z)}{1+\gamma(z)}\, m^R(dz) = \lambda^P \int_{\{\gamma(z)\geq 0\}} \gamma(z)p(z)\, m^P(dz) \tag{7.15}
$$

from (2.9) and (3.2). By Lemma 10, there exists ϵ_2 and y satisfying $1 < y < b(t_0)$ such that $V_x(t,x) > 1-\delta$ whenever $(t,x) \in A(t_0,\epsilon_2)$. Thus from (7.13) and (7.14), when $(t,x) \in A(t_0,\epsilon_2)$, we see that

$$
\frac{1}{2}\sigma^2 x^2 \frac{\partial^2 V}{\partial x^2}(t,x) \geq \epsilon_1.
$$

□

Finally, let us prove the main result of this section:

Theorem 5. *Under assumption (A), the boundary function $b(t)$ is continuous on $(0, T]$.*

Proof. We first show the left continuity of $b(t)$. Let $t_n \nearrow t_0 \in (0, T]$, so that

$$(t_n, b(t_n)) \to (t, b(t_0-)) \text{ as } t_n \nearrow t_0.$$

Since $(t_n, b(t_n)) \in D$ and D is closed, we must have $(t, b(t_0-)) \in D$ so $b(t_0-) \geq b(t_0)$. But since $b(t)$ is non-decreasing, $b(t_0-) \leq b(t_0)$ and hence $b(t_0-) = b(t_0)$.

For the right continuity, assume $b(t_0+) > b(t_0)$, and let $0 < \eta < b(t_0+) - b(t_0)$. Suppose $t_n \searrow t_0$ and using the notation in Lemma 11, without loss of generality assume $t_0 < t_n < t_0 + \epsilon_2$ for all n. Then

$$V(t_n, b(t_n) - \eta) - g(t_n, b(t_n) - \eta) = \int_{b(t_n)-\eta}^{b(t_n)} \int_y^{b(t_n)} \frac{\partial^2 F}{\partial x^2}(t_n, u)\, du\, dy \geq \frac{\epsilon_1}{\sigma^2 b(T)^2}\eta^2$$

which, after letting $n \to \infty$, implies

$$V(t, b(t_0+) - \eta) > g(t, b(t_0+) - \eta).$$

Thus $(t, b(t_0+) - \eta) \in C$ for all η such that $0 < \eta < b(t_0+) - b(t_0)$, so $b(t_0+) - \eta < b(t_0)$ which implies that $b(t_0+) \leq b(t_0)$. But since $b(t)$ is non-decreasing, we have $b(t_0) \leq b(t_0+)$ and so $b(t_0+) = b(t_0)$. Thus we conclude $b(t)$ is continuous. □

Chapter 8

Further Properties of the Value Function

8.1 A Generalized Itô Formula

In order to derive the early exercise premium representation and a uniqueness result for the Russion option later in the chapter, we would like to apply Itô's formula to our value function $V(t, x)$. Unfortunately, the usual Itô's formula requires that the function we apply it to is once differentiable in t and twice differentiable in x, and we do not have that amount of regularity for $V(t, x)$. Instead we know that there exist two sets, A and B, and a boundary curve $b(t)$ such that

$$A = \{(t, x) \in (0, T] \times [1, \infty) : V(t, x) < b(t)\}$$
$$B = \{(t, x) \in (0, T] \times [1, \infty) : V(t, x) > b(t)\},$$

and that $V(t, x)$ is $C^{1,2}$ on A and B individually. The difficulty arises in that we do not know the behavior of V_{xx} or V_t at the boundary. Peskir [24] considered this problem for a function satisfying these conditions in the diffusion case, where he developed a version of Itô's formula that can be applied when F_{xx} and F_t in some sense "cancel each other out" at the boundary; specifically, for any diffusion with drift μ and diffusion coefficient σ, Peskir's formula holds if $F_t + \mu(t, x)F_x + \frac{\sigma^2}{2}(t, x)F_{xx}$ is locally bounded on $A \cup B$. The same proof works with some modifications for our case, where the process we are considering is a jump-diffusion with an additional finite variation term.

With this in mind, let us assume that $b(t) : [0, t_1] \to \mathbb{R}$ is a continuous curve and let X be a semimartingale on a probability space $(\Omega, \mathcal{F}, P)$ given by the stochastic differential equation

$$dX_t = \mu(t, X_{t-})\,dt + \phi(t, X_{t-})\,dW_t + \psi(t, X_{s-}, z)\,\tilde{\nu}(dt, dz) + dA_t \tag{8.1}$$

where A_t is a continuous process of bounded variation, $\nu(dt, dz)$ is a Poisson random measure with intensity measure $\lambda\, m(dz)\, dt$, and we assume the coefficients μ, ϕ, and ψ are locally bounded. Furthermore, suppose $\phi > 0$ and X_t takes values in the interval $[c, \infty)$, where c is a real number (results similar to our main theorem below can also be obtained for intervals of the form $(-\infty, d)$, $[c, d]$, and $(-\infty, \infty)$). We consider two processes, one which equals X_t whenever X_t is sufficiently below the boundary and is constant otherwise, and one which equals X_t whenever X_t is sufficiently above the boundary and is constant otherwise. We define these processes as $Z_t^{1,\epsilon} := X_t \wedge b_{1,\epsilon}(t)$ and $Z_t^{2,\epsilon} := X_t \vee b_{2,\epsilon}(t)$, where $b_{1,\epsilon}(t) := (b(t) - \epsilon) \vee c$ and $b_{2,\epsilon}(t) := b(t) + \epsilon$. Here $b_{1,\epsilon}$ is truncated to ensure it never goes below the lower bound of our interval $[c, \infty)$. Next, note that

$$\begin{aligned} Z_t^{1,\epsilon} &= \frac{1}{2}(X_t + b_{1,\epsilon}(t) - |X_t - b_{1,\epsilon}(t)|) \\ Z_t^{2,\epsilon} &= \frac{1}{2}(X_t + b_{2,\epsilon}(t) + |X_t - b_{2,\epsilon}(t)|). \end{aligned}$$

Applying Tanaka's formula for semimartingales we see that

$$\begin{aligned} &|X_t - b_{i,\epsilon}(t)| \\ &= |X_0 - b_{i,\epsilon}(0)| + \int_0^t \operatorname{sign}(X_{s-} - b_{i,\epsilon}(s))\, d(X_s^c - b_{i,\epsilon}(s)) + L_t^{b_{i,\epsilon}}(X) \\ &\quad + \sum_{0<s\le t} (|X_s - b_{i,\epsilon}(s)| + |X_{s-} - b_{i,\epsilon}(s)|) \end{aligned} \tag{8.2}$$

for $i = 1, 2$, where sign is the signum function with $\operatorname{sign}(0) = 0$. The process $L^{b_{i,\epsilon}}(X)$ denotes the local time of X at the curve $b_{i,\epsilon}$ (or, equivalently, the local time of $X - b_{i,\epsilon}$ at 0) on $[0, t_1]$ and is given by

$$L_t^{b_{i,\epsilon}}(X) = \lim_{\delta\downarrow 0} \frac{1}{2\delta} \int_0^t \mathbf{1}_{[-\delta < X_{u-} - b_{i,\epsilon}(u) < \delta]}\, d[X - b_{i,\epsilon}]_u^c. \tag{8.3}$$

Here $[X - b_{i,\epsilon}]^c$ is the continuous part of the quadratic variation of $X - b_{i,\epsilon}$ and the limit in (8.3) is taken in probability. Thus we can write $Z_t^{1,\epsilon,c}$ and $Z_t^{2,\epsilon,c}$, the continuous parts of the semimartingales $Z_t^{1,\epsilon}$ and $Z_t^{2,\epsilon}$, in the differential form

$$\begin{aligned} dZ_t^{1,\epsilon,c} =& \frac{1}{2}\left(dX_t^c + db_{1,\epsilon}(t) - \operatorname{sign}(X_{t-} - b_{1,\epsilon}(t))\, d(X_t^c - b_{1,\epsilon}(t)) - dL_t^{b_{1,\epsilon}}(X)\right) \\ =& \frac{1}{2}\left((1 - \operatorname{sign}(X_{t-} - b_{1,\epsilon}(t)))\, dX_t^c + (1 + \operatorname{sign}(X_{t-} - b_{1,\epsilon}(t)))\, db_{1,\epsilon}(t) - dL_t^{b_{1,\epsilon}}(X)\right) \\ =& \left(\mathbf{1}_{[X_{t-} - b_{1,\epsilon}(t) < 0]} + \frac{1}{2}\mathbf{1}_{[X_{t-} - b_{1,\epsilon}(t) = 0]}\right) dX_t^c \\ &+ \left(\mathbf{1}_{[X_{t-} - b_{1,\epsilon}(t) > 0]} + \frac{1}{2}\mathbf{1}_{[X_{t-} - b_{1,\epsilon}(t) = 0]}\right) db_{1,\epsilon}(t) - \frac{1}{2}\, dL_t^{b_{1,\epsilon}}(X) \end{aligned} \tag{8.4}$$

and

$$\begin{aligned} dZ_t^{2,\epsilon,c} =& \frac{1}{2}\left(dX_t^c + db_{2,\epsilon}(t) + \operatorname{sign}(X_{t-} - b_{2,\epsilon}(t))\, d(X_t^c - b_{2,\epsilon}(t)) + dL_t^{b_{2,\epsilon}}(X)\right) \\ =& \frac{1}{2}\left((1 + \operatorname{sign}(X_{t-} - b_{2,\epsilon}(t)))\, dX_t^c + (1 - \operatorname{sign}(X_{t-} - b_{2,\epsilon}(t)))\, db_{2,\epsilon}(t) + dL_t^{b_{2,\epsilon}}(X)\right) \\ =& \left(\mathbf{1}_{[X_{t-} - b_{2,\epsilon}(t) > 0]} + \frac{1}{2}\mathbf{1}_{[X_{t-} - b_{2,\epsilon}(t) = 0]}\right) dX_t^c \\ &+ \left(\mathbf{1}_{[X_{t-} - b_{2,\epsilon}(t) < 0]} + \frac{1}{2}\mathbf{1}_{[X_{t-} - b_{2,\epsilon}(t) = 0]}\right) db_{2,\epsilon}(t) + \frac{1}{2}\, dL_t^{b_{2,\epsilon}}(X). \end{aligned} \tag{8.5}$$

Observe from these equations that we can write the quadratic variation terms as

$$\begin{aligned} d[Z^{1,\epsilon}, Z^{1,\epsilon}]_t^c &= \left(\mathbf{1}_{[X_{t-} - b_{1,\epsilon}(t) < 0]} + \frac{1}{4}\mathbf{1}_{[X_{t-} - b_{1,\epsilon}(t) = 0]}\right) d[X^c, X^c]_t \\ &= \mathbf{1}_{[X_{t-} - b_{1,\epsilon}(t) < 0]}\, d[X^c, X^c]_t \end{aligned} \tag{8.6}$$

and

$$\begin{aligned} d[Z^{2,\epsilon}, Z^{2,\epsilon}]_t^c &= \left(\mathbf{1}_{[X_{t-} - b_{2,\epsilon}(t) > 0]} + \frac{1}{4}\mathbf{1}_{[X_{t-} - b_{2,\epsilon}(t) = 0]}\right) d[X^c, X^c]_t \\ &= \mathbf{1}_{[X_{t-} - b_{2,\epsilon}(t) > 0]}\, d[X^c, X^c]_t, \end{aligned} \tag{8.7}$$

where the final equalities in (8.6) and (8.7) above follow from the occupation times formula, which implies that

$$\int_0^t \mathbf{1}_{[X_{s-} - b_{1,\epsilon}(s) = 0]}\, d[X^c, X^c]_s = \int_0^t \mathbf{1}_{[X_{s-} - b_{2,\epsilon}(s) = 0]}\, d[X^c, X^c]_s = 0.$$

By applying the standard version Itô's formula to these two processes individually, and letting $\epsilon \searrow 0$, we are able to obtain our modified version of Itô's formula:

Theorem 6. *Let X be a semimartingale given by* (8.1) *taking values in the interval* $[c, \infty)$. *Suppose $b : [0, t_1] \to \mathbb{R}$ is a continuous function of bounded variation and $F : [0, t_1] \times [c, \infty) \to \mathbb{R}$ is a continuous function which is $C^{1,2}$ on the sets*

$$A = \{(t, x) \in [0, t_1] \times [c, \infty) : x < b(t)\}$$

$$B = \{(t, x) \in [0, t_1] \times [c, \infty) : x > b(t)\}.$$

Furthermore, suppose for $i = 1, 2$ that F satisfies

$$\mathcal{L}' F(t, x) \text{ is locally bounded on } A \cup B \tag{8.8}$$

$$s \to F_x(s, b_{i,\epsilon}(s)) \text{ converges to } F_x(s, b(s)\pm) \text{ uniformaly on } [0, t_1] \text{ as } \epsilon \searrow 0 \tag{8.9}$$

$$\text{For all } s \in [0, t_1], \sup_{0<\epsilon<\delta} \mathrm{Var}(F(\cdot, b_{i,\epsilon}(\cdot)))(s) < \infty \text{ for some } \delta > 0, \tag{8.10}$$

where $\mathrm{Var}(F)$ *denotes the total variation of the function F and $\mathcal{L}' F(t, x)$ is the operator*

$$\begin{aligned}\mathcal{L}' F(t, x) :=& F_t(t, x) + \mu(t, x) F_x(t, x) + \frac{\phi^2(t, x)}{2} F_{xx}(t, x) \\ &+ \lambda \int_{\mathbb{R}} F(t, x + \psi(t, x, z)) - F(t, x) - \psi(t, x, z) F_x(t, x)\, m(dz).\end{aligned}$$

Then the following version of Itô's formula holds P-almost surely:

$$\begin{aligned}F(t, X_t) =& F(0, X_0) + \int_0^t \mathcal{L}' F(u, X_{u-}) \mathbf{1}_{[X_{u-} \neq b(u)]}\, du \\ &+ \int_0^t (\phi F_x)(u, X_{u-}) \mathbf{1}_{[X_{u-} \neq b(u)]}\, dW_u + \int_0^t F_x(u, X_{u-}) \mathbf{1}_{[X_{u-} \neq b(u)]}\, dA_u \\ &+ \frac{1}{2} \int_0^t F_x(u, b(u)+)\, dL_s^b(X) - \frac{1}{2} \int_0^t F_x(u, b(u)-)\, dL_u^b(X) \\ &+ \int_0^t \int_{\mathbb{R}} F(u, X_{u-} + \psi(u, X_{u-}, z)) - F(u, X_{u-})\, \tilde{\nu}(du, dz).\end{aligned}$$

Proof. First observe that

$$F(t, X_t) = F(t, X_t \wedge b(t)) + F(t, X_t \vee b(t)) - F(t, b(t)), \tag{8.11}$$

which will be useful later in our proof. Applying Itô's formula to $F(t, Z_t^{1,\epsilon})$ and $F(t, Z_t^{2,\epsilon})$ and using the fact that $P(X_{s-} = b_{i,\epsilon}(s)) = 0$ for $0 \leq s \leq t_1$ since $\phi > 0$ we obtain, after using (8.4), (8.5), (8.6) and (8.7),

$$F(t, X_t \wedge b_{1,\epsilon}(t)) + F(t, X_t \vee b_{2,\epsilon}(t)) \tag{8.12}$$

$$= F(0, X_0 \wedge b_{1,\epsilon}(0)) + F(0, X_0 \vee b_{2,\epsilon}(0)) \tag{8.13}$$

$$+ \int_0^t \mathcal{L}' F(u, X_{u-}) \mathbf{1}_{[X_{u-} \notin [b_{1,\epsilon}(u), b_{2,\epsilon}(u)]]} \, du \tag{8.14}$$

$$+ \int_0^t \phi F_x(u, X_{u-}) \mathbf{1}_{[X_{u-} \notin [b_{1,\epsilon}(u), b_{2,\epsilon}(u)]]} \, dB_u \tag{8.15}$$

$$+ \int_0^t F_x(u, X_{u-}) \mathbf{1}_{[X_{u-} \notin [b_{1,\epsilon}(u), b_{2,\epsilon}(u)]]} \, dA_u \tag{8.16}$$

$$+ I_1^\epsilon + I_2^\epsilon + I_3^\epsilon + I_4^\epsilon$$

for $t \in [0, t_1]$, where

$$I_1^\epsilon := \frac{1}{2} \int_0^t F_x(u, b_{2,\epsilon}(u)) \, dL_u^{b_{2,\epsilon}}(X) - \frac{1}{2} \int_0^t F_x(u, b_{1,\epsilon}(u)) \, dL_u^{b_{1,\epsilon}}(X)$$

$$I_2^\epsilon := \int_0^t F_t(u, b_{1,\epsilon}(u)) \mathbf{1}_{[X_{u-} > b_{1,\epsilon}(u)]} \, du + \int_0^t F_x(u, b_{1,\epsilon}(u)) \mathbf{1}_{[X_{u-} > b_{1,\epsilon}(u)]} \, db_{1,\epsilon}(u)$$
$$+ \int_0^t F_t(u, b_{2,\epsilon}(u)) \mathbf{1}_{[X_{u-} < b_{2,\epsilon}(u)]} \, du + \int_0^t F_x(u, b_{2,\epsilon}(u)) \mathbf{1}_{[X_{u-} < b_{2,\epsilon}(u)]} \, db_{2,\epsilon}(u)$$

$$I_3^\epsilon := \int_0^t \int_{\mathbb{R}} F(u, (X_{u-} + \psi(u, X_{u-}, z)) \wedge b_{1,\epsilon}(u)) - F(u, X_{u-} \wedge b_{1,\epsilon}(u)) \, \nu(du, dz)$$
$$+ \int_0^t \int_{\mathbb{R}} F(u, (X_{u-} + \psi(u, X_{u-}, z)) \vee b_{2,\epsilon}(u)) - F(u, X_{u-} \vee b_{2,\epsilon}(u)) \, \nu(du, dz)$$

$$I_4^\epsilon := - \int_0^t \int_{\mathbb{R}} [F(u, X_{u-} + \psi(u, X_{u-}, z)) - F(u, X_{u-})] \mathbf{1}_{[X_{u-} \notin [b_{1,\epsilon}(u), b_{2,\epsilon}(u)]]} \, m^R(dz) \, du$$

We will examine the convergence of these terms as $\epsilon \searrow 0$. First, note that by the continuity of F, (8.12) converges to $F(t, X_t \wedge b(t)) + F(t, X_t \vee b(t))$ and (8.13) converges to $F(0, X_0 \wedge b(0)) + F(0, X_0 \vee b(0))$ as $\epsilon \searrow 0$. Next observe that

$$\left| \mathcal{L}' F(u, X_{u-}) \mathbf{1}_{[X_{u-} \notin [b_{1,\epsilon}(u), b_{2,\epsilon}(u)]]} \right| \leq \left| \mathcal{L}' F(u, X_{u-}) \mathbf{1}_{[X_{u-} \neq b(u)]} \right|$$

which is integrable by (8.8). Hence we may apply the Dominated Convergence Theorem to see that (8.14) converges to

$$\int_0^t \mathcal{L}' F(u, X_{u-}) \mathbf{1}_{[X_{u-} \neq b(u)]} \, du$$

in probability. In a similar manner, we observe that (8.9) implies that F_x is locally bounded on $A \cup B$. Hence since ϕ is locally bounded, we can apply the stochastic (see

Protter [29], Chapter IV, Theorem 32) and regular Dominated Convergence Theorems to see that (8.15) and (8.16) converge in probability to

$$\int_0^t \phi F_x(u, X_{u-})\mathbf{1}_{[X_{u-}\neq b(u)]}\, dB_u$$

and

$$\int_0^t F_x(u, X_{u-})\mathbf{1}_{[X_{u-}\neq b(u)]}\, dA_u$$

respectively.

For I_1^ϵ, note that by applying integration by parts we have

$$\int_0^t F_x(u, b_{i,\epsilon}(u))\, dL_u^{b_{i,\epsilon}}(X) = \left[F_x(u, b_{i,\epsilon}(u))L_u^{b_{i,\epsilon}}(X)\right]_0^t - \int_0^t L_u^{b_{i,\epsilon}}(X)\, d_u F_x(s, b_{i,\epsilon}(u)).$$

If we can show

$$\int_0^t L_u^{b_{i,\epsilon}}(X)\, d_u F_x(u, b_{i,\epsilon}(u)) \to \int_0^t L_u^b(X)\, d_u F_x(u, b(u))$$

as $\epsilon \searrow 0$ P-almost surely over a subsequence, then we can let $\epsilon \searrow 0$ and apply integration by parts again to see that I_1^ϵ approaches

$$\frac{1}{2}\int_0^t F_x(u, b(u)+)\, dL_u^b(X) - \frac{1}{2}\int_0^t F_x(u, b(u)-)\, dL_u^b(X). \tag{8.17}$$

Observe that

$$\begin{aligned}
&\left|\int_0^t L_u^{b_{i,\epsilon}}(X)\, d_u F_x(u, b_{i,\epsilon}(u)) - \int_0^t L_u^b(X)\, d_u F_x(u, b(u))\right| \\
&\quad \leq \int_0^t |L_u^{b_{i,\epsilon}}(X) - L_u^b(X)|\, d_u|F_x|(u, b_{i,\epsilon}(u)) \qquad (8.18)\\
&\qquad + \left|\int_0^t L_u^b(X)\, d_u F_x(u, b_{i,\epsilon}(u)) - \int_0^t L_u^b(X)\, d_u F_x(u, b(u))\right|.
\end{aligned}$$

Now $\sup_{0\leq s\leq t} |L_s^{b_{i,\epsilon}}(X) - L_s^b(X)| \to 0$ as $\epsilon \searrow 0$ P-a.s. over a subsequence (see Lemma 13 in Appendix A for a proof) so the first term in (8.18) can be made arbitrarily small. For the second term, since $F(s, b_{i,\epsilon}(s)) \to F(s, b(s))$ as $\epsilon \searrow 0$ for every $s \in [0, t_1]$ and (8.10) is satisfied, Helly's Selection Theorem allows us to conclude $d_s F(s, b_{i,\epsilon}(s)) \to d_s F(s, b(s))$ weakly as $\epsilon \searrow 0$. Since $s \to L_s^b(X)$ is continuous and bounded on $[0, t_1]$, the second term in (8.18) converges to 0 and thus I_1^ϵ converges to (8.17).

The fact that I_2^ϵ converges to $F(t,b(t)) - F(0,b(0))$ follows from minor modifications of the same proof in Peskir [24]. For the convenience of the reader, the proof is provided in the appendix.

For I_3^ϵ we can apply the Dominated Convergence Theorem to see that as $\epsilon \searrow 0$

$$\begin{aligned}
&\int_0^t \int_{\mathbb{R}} F(u,(X_{u-} + \psi(u,X_{u-},z)) \wedge b_{1,\epsilon}(u)) + F(u,(X_{u-} + \psi(u,X_{u-},z)) \vee b_{2,\epsilon}(u)) \\
&\qquad - \left[F(u,X_{u-} \wedge b_{1,\epsilon}(u)) + F(u,X_{u-} \vee b_{2,\epsilon}(u))\right] \nu(du,dz) \\
&\quad \to \int_0^t \int_{\mathbb{R}} F(u,(X_{u-} + \psi(u,X_{u-},z)) \wedge b(u)) + F(u,(X_{u-} + \psi(u,X_{u-},z)) \vee b(u)) \\
&\qquad - [F(u,X_{u-} \wedge b(u)) + F(u,X_{u-} \vee b(u))] \nu(du,dz) \\
&\quad = \int_0^t \int_{\mathbb{R}} \Big[F(u,(X_{u-} + \psi(u,X_{u-},z)) \wedge b(u)) \\
&\qquad + F(u,(X_{u-} + \psi(u,X_{u-},z)) \vee b(u)) - F(u,b(u))\Big] \\
&\qquad - [F(u,X_{u-} \wedge b(u)) + F(u,X_{u-} \vee b(u)) - F(u,b(u))] \nu(du,dz) \\
&\quad = \int_0^t \int_{\mathbb{R}} F(u,X_{u-} + \psi(u,X_{u-},z)) - F(u,X_{u-}) \nu(du,dz)
\end{aligned}$$

by (8.11).

Finally, I_4^ϵ converges to

$$\begin{aligned}
&-\int_0^t \int_{\mathbb{R}} [F(u,X_{u-} + \psi(u,X_{u-},z)) - F(u,X_{u-})] \mathbf{1}_{[X_{u-} \neq b(u)]} m^R(dz)\, du \\
&\quad = -\int_0^t \int_{\mathbb{R}} [F(u,X_{u-} + \psi(u,X_{u-},z)) - F(u,X_{u-})] m^R(dz)\, du,
\end{aligned}$$

P−a.s., where the equality comes from the fact that outside a P-null set $X_{u-} \neq b(u)$ for λ-almost all $u \in [0,t_1]$, as shown in Lemma 13. Combining the above convergence facts and using (8.11) once again we obtain the result. □

Before proceeding, let us apply the Generalized Itô Formula given in Theorem 6. Fix $t \in [0,T)$ and let $F(u,x) := e^{-\lambda u} V(t-u,x)$, which is $C^{1,2}$ on

$$\{(u,x) \in [0,t] \times [1,\infty) : x < b(t-u)\}$$

and

$$\{(u,x) \in [0,t] \times [1,\infty) : x > b(t-u)\}$$

Note that from Theorem 5 we know that $b(t-u)$ is continuous on $[0,t)$ under Assumption (A). Thus if $s < t$, then with the exception of (8.9) and (8.10), it is immediate that all of the necessary conditions to apply Theorem 6 on an interval $[0,s]$ are satisfied for F. Observe that (8.9) holds by Dini's Theorem, since $\epsilon \to F_x(u, b_{i,\epsilon}(t-u))$ is non-increasing and $u \to F_x(u, b(t-u))$ is continuous on $[0,u]$. Furthermore, since $b(t-u)$ is decreasing and F is decreasing, $F(u, b_{i,\epsilon}(t-u))$ is decreasing we see that (8.10) also holds. Thus we can apply Theorem 6 to F and doing so we obtain

$$
\begin{aligned}
e^{-\lambda s}V(t-s, X_s) =& V(t, X_0) + \int_0^s \mathcal{L}'(e^{-\lambda u}V(t-u, X_{u-}))\mathbf{1}_{[X_{u-}\neq b(t-u)]}\,du \\
&+ \int_0^s e^{-\lambda u}(\beta V_x)(t-u, X_{u-})\mathbf{1}_{[X_{u-}\neq b(t-u)]}\,dW_u \\
&+ \int_0^s e^{-\lambda u}V_x(t-u, X_{u-})\mathbf{1}_{[X_{u-}\neq b(t-u)]}\,\mathbf{1}_{[\Delta M_t = 0]}\frac{dM_u}{S_{u-}} \\
&+ \frac{1}{2}\int_0^s [e^{-\lambda u}V_x(t-u, b(t-u)+) - V_x(t-u, b(t-u)-)]\,dL_u^{b(t-u)}(X) \\
&+ \int_0^s \int_{\mathbb{R}} e^{-\lambda u}[V(t-u, X_{u-} + \zeta(X_{u-}, z)) - V(t-u, X_{u-})]\,\tilde{\nu}(du, dz)
\end{aligned}
$$

R-almost surely for $0 \leq s < t$. Direct calculation shows that $\mathcal{L}'(e^{-\lambda u}V(t-u,x)) = e^{-\lambda u}\mathcal{L}V(t-u,x)$, where $\mathcal{L}$ is the operator given by (5.9), so we can use Proposition 11, Proposition 12, the continuity of V in x, and the fact that $\mathcal{L}V(u,x) = 0$ for $x < b(t)$ to obtain

$$
\begin{aligned}
e^{-\lambda s}V(t-s, X_s) =& V(t, X_0) + \int_0^s e^{-\lambda u}\mathcal{L}V(t-u, X_{u-})\mathbf{1}_{[X_{u-}\geq b(t-u)]}\,du \\
&+ \int_0^s e^{-\lambda u}(\beta V_x)(t-u, X_{u-})\mathbf{1}_{[X_{u-}\neq b(t-u)]}\,dW_u \\
&+ \int_0^s \int_{\mathbb{R}} e^{-\lambda u}[V(t-u, X_{u-} + \zeta(X_{u-}, z)) - V(t-u, X_{u-})]\,\tilde{\nu}(du, dz)
\end{aligned}
\tag{8.19}
$$

outside an R-null set, for $0 \leq s < t$.

8.2 Uniqueness

In this section, we prove that the function $V(t,x)$ and the boundary $b(t)$ are the unique solution pair to a free boundary problem under the assumption

$$(A') \quad \lambda + r - \lambda^P \int_{\{\gamma(z)\geq 0\}} \gamma(z)p(z)\, m^P(dz) \geq 0,$$

which is slightly stronger than assumption (A) given by (7.12). This assumption is necessary since we do not know a priori that $\mathcal{L}v \leq 0$ for functions v satisfying the free boundary problem presented in Theorem 7. This condition allows us to conclude that the function $e^{-\lambda s}v(t-s, X_s)$ is a supermartingale, which will be needed in the proof of Theorem 7.

Let $\mathcal{V}$ denote the set consisting of all pairs (v,c), where $c : (0,T] \to [1,\infty)$ is a continuous, non-decreasing curve and the function $v : [0,T]\times[1,\infty) \to \mathbb{R}$ is continuous, $C^{1,2}$ for $x < c(t)$ and $x > c(t)$, and such that $x \to v(t,x)$ is non-decreasing and convex. With this notation, we have the following theorem:

Theorem 7. *Assume (A') holds. Then (V,b) is the unique solution pair $(v,c) \in \mathcal{V}$ of the free boundary problem*

$$\mathcal{L}v(t,x) = 0 \;\; for \;\; x < c(t) \tag{8.20}$$

$$\lim_{x\to c(t)} v(t,x) = c(t) \;\; for \;\; t \in (0,T] \tag{8.21}$$

$$\lim_{x\to c(t)} v_x(t,x) = 1 \;\; for \;\; t \in (0,T] \tag{8.22}$$

$$\lim_{x\to 1+} v_x(t,x) = 0 \;\; for \;\; t \in (0,T] \tag{8.23}$$

$$v(0,x) = g(x) \;\; for \;\; x \in [1,\infty) \tag{8.24}$$

$$v(t,x) > g(x) \;\; if \;\; x < c(t) \tag{8.25}$$

$$v(t,x) = g(x) \;\; if \;\; x \geq c(t), \tag{8.26}$$

where $g(x) = x$ is the payoff function for $V(t,x)$.

Proof. That (V,b) is a solution pair follows from our previous results. To see that (V,b) is the unique solution pair when (A') holds, let (v,c) be another solution pair in $\mathcal{V}$. When

$x \geq c(t)$ we have $v(t,x) = x$ by (8.26) and so

$$\begin{aligned}\mathcal{L}v(t,x) = &-(\lambda + r)x + \lambda^R \int_{\{1+\gamma(z)>x\}} 1 - \frac{x}{1+\gamma(z)}\, m^R(dz) \\ &+ \lambda^R \int_{\{x+\zeta(x,z)\leq c(t)\}} v(t, x+\zeta(x,z)) - x - \zeta(x,z)\, m^R(dz) \\ = &-(\lambda + r)x + \lambda^R \int_{\{1+\gamma(z)>x\}} 1 - \frac{x}{1+\gamma(z)}\, m^R(dz) \\ &+ \lambda^R \int_{\{x+\zeta(x,z)\leq c(t)\}} v(t, x+\zeta(x,z)) - x + x\frac{\gamma(z)}{1+\gamma(z)}\, m^R(dz) \\ &- \lambda^R \int_{\{x+\zeta(x,z)\leq c(t)\}} \mathbf{1}_{[1+\gamma(z)>x]} \left(1 - \frac{x}{1+\gamma(z)}\right) m^R(dz).\end{aligned} \tag{8.27}$$

Now if $1+\gamma(z) > x$, then $\zeta(x,z) = 1 - x$ so $x + \zeta(x,z) \leq c(t)$. Thus

$$\{1+\gamma(z) > x\} \subset \{x + \zeta(x,z) \leq c(t)\}$$

and so

$$\begin{aligned}\mathcal{L}v(t,x) &= -(\lambda + r)x + \lambda^R \int_{\{x+\zeta(x,z)\leq c(t)\}} v(t, x+\zeta(x,z)) - x + x\frac{\gamma(z)}{1+\gamma(z)}\, m^R(dz) \\ &\leq -(\lambda + r)x + \lambda^R \int_{\{x+\zeta(x,z)\leq c(t)\}} v(t, c(t)) - x + x\frac{\gamma(z)}{1+\gamma(z)}\, m^R(dz) \\ &= -(\lambda + r)x + \lambda^R \int_{\{x+\zeta(x,z)\leq c(t)\}} c(t) - x + x\frac{\gamma(z)}{1+\gamma(z)}\, m^R(dz) \\ &\leq \left[-(\lambda + r) + \lambda^R \int_{\{\gamma(z)\geq 0\}} \frac{\gamma(z)}{1+\gamma(z)}\, m^R(dz)\right] x \leq 0.\end{aligned} \tag{8.28}$$

Here we used the fact that $v(t,x)$ is non-decreasing and that we are integrating over the set $\{x + \zeta(x,z) \leq c(t)\}$ for the first inequality, (8.21) for the second equality, the fact that $c(t) \leq x$ for the second inequality, and for the final inequality we used (7.15) and (A'). Combining this with (8.20) we have $\mathcal{L}v(t,x) \leq 0$ for all $(t,x) \in [0,T] \times [1,\infty)$.

The argument following Theorem 8.1 also applies to $e^{-\lambda u} v(t-u, X_u)$, so applying the Generalized Ito's Formula we obtain

$$\begin{aligned} e^{-\lambda s}v(t-s,X_s) =& v(t,X_0) + \int_0^s e^{-\lambda u}\mathcal{L}v(t-u,X_{u-})\mathbf{1}_{[X_{u-}\neq c(t-u)]}\,du \\ &+ \int_0^s e^{-\lambda u}(\beta v_x)(t-u,X_{u-})\mathbf{1}_{[X_{u-}\neq c(t-u)]}\,dW_u \\ &+ \int_0^s e^{-\lambda u}v_x(t-u,X_{u-})\mathbf{1}_{[X_{u-}\neq c(t-u)]}\,\mathbf{1}_{[\Delta M_u=0]}\frac{dM_u}{S_{u-}} \\ &+ \frac{1}{2}\int_0^s \left[e^{-\lambda u}v_x(t-u,c(t-u)+) - v_x(t-u,c(t-u)-)\right]\,dL_u^b(X) \\ &+ \int_0^s\int_{\mathbb{R}} e^{-\lambda u}[v(t-u,X_{u-}+\zeta(X_{u-},z)) - v(t-u,X_{u-})]\,\tilde{\nu}(du,dz). \end{aligned} \tag{8.29}$$

R-almost surely for $0 \le s < t$. Letting $s \nearrow t$ and using (8.22) and (8.23) we see this equals

$$\begin{aligned} e^{-\lambda t}v(0,X_t) =& v(t,X_0) + \int_0^t e^{-\lambda u}\mathcal{L}v(t-u,X_{u-})\mathbf{1}_{[X_{u-}\neq c(t-u)]}\,du \\ &+ \int_0^t e^{-\lambda u}(\beta v_x)(t-u,X_{u-})\mathbf{1}_{[X_{u-}\neq c(t-u)]}\,dW_u \\ &+ \int_0^t\int_{\mathbb{R}} e^{-\lambda u}[v(t-u,X_{u-}+\zeta(X_{u-},z)) - v(t-u,X_{u-})]\,\tilde{\nu}(du,dz) \end{aligned} \tag{8.30}$$

R-almost surely.

Observe that the fact that v is non-decreasing implies that $v_x \ge 0$, and the fact that v is convex implies v_x is non-decreasing. Since $v_x(t,x) = 1$ for $x > c(T)$ and $t \in [0,T]$, we see then that v_x is bounded by 1. Furthermore, from (8.25), (8.26) and the fact that $c(u)$ is non-decreasing we have

$$\left(v(t-u,X_{u-}+\zeta(X_{u-},z)) - v(t-u,X_{u-})\right)^2 \le \max\{\zeta^2(X_{u-},z), X_{u-}^2, c(T)^2\}.$$

These two facts, combined with Lemma 2 and the fact that $\zeta^2(x,z) \in \mathcal{L}^2(m^R)$ from (3.16) imply that the final two terms in (8.30) are martingales. Therefore using that $\mathcal{L}v(t,x) \le 0$ we see that $\{e^{-\lambda u}v(t-u,X_u), 0 \le u \le t\}$ is a supermartingale. Hence for any stopping time τ such that $0 \le \tau \le T$,

$$v(t,x) \ge E_x^R[e^{-\lambda\tau}v(t-\tau,X_\tau)] \ge E_x^R[e^{-\lambda\tau}g(X_\tau)] = E_x^R[e^{-\lambda\tau}X_\tau]$$

by Optimal Sampling Theorem and the fact that $v \ge g$ from (8.25) and (8.26). Taking the supremum over all stopping times $0 \le \tau \le T$, we have $v(t,x) \ge V(t,x)$.

Finally, let us show that $v(t,x) \leq V(t,x)$. If $x \geq c(t)$, then by (8.25) we see that $v(t,x) = g(x) \leq V(t,x)$. If $x < c(t)$ we define the stopping time

$$\tau^* := \inf\{0 \leq s \leq t : v(t-s, X_s(x)) = g(X_s(x))\}.$$

Note $v(0,x) = g(x)$ so τ^* exists and $0 \leq \tau^* \leq t$. It follows from (8.30) that

$$\{e^{-\lambda u} v(t-u, X_u(x)), 0 \leq u \leq \tau^*\}$$

is an R-martingale since (8.20), (8.25), and (8.26) imply $\mathcal{L}v(t-u, X_u) = 0$ for $u < c(t-u)$. Hence

$$v(t,x) = E_x^R[e^{-\lambda\tau^*} v(t-\tau^*, X_{\tau^*})] = E_x^R[e^{-\lambda\tau^*} g(X_{\tau^*})] \leq V(t,x).$$

by the Optimal Sampling Theorem and the definition of τ^*. □

8.3 The Early Early Exercise Premium Representation

Finally, we derive a decomposition for $V(t,x)$ known as the early exercise premium representation (EEPR). Consider an option identical to the Russian option except for the fact that the holder is not allowed to exercise it before time T. The EEPR expresses the Russian Option's value function in terms of the value function of this restricted option, and in particular gives us the additional premium we must pay in order to obtain the right to early exercise the restricted option.

Proposition 13. *Under assumption* (A) *given by* (7.12)*, the value function* $V(t,x)$ *admits the "early exercise premium" representation*

$$V(t,x) = \tilde{V}(t,x) + e_1(t,x) - e_2(t,x),$$

where $\tilde{V}(t,x) = e^{-\lambda t} E_x^R[X_t]$,

$$e_1(t,x) = (\lambda + r)\int_0^t e^{-\lambda u} E_x^R[X_{u-}\mathbf{1}_{[X_{u-} \geq b(t-u)]}]du$$
$$e_2(t,x) = \lambda^R E_x^R\left[\int_0^t \int_{A_{x,t,u}} e^{-\lambda u}\left(V(t-u, X_{u-} + \zeta(X_{u-}, z)) - X_{u-}\left[\frac{1}{1+\gamma(z)}\right]\right) m^R(dz)\, du\right],$$

and $A_{x,t,u} = \{z : x + \zeta(x,z) \leq b(t-u)\} \cap \{x \geq b(t-u)\}$.

Proof. Note that under assumption (A), (8.19) holds. Thus letting $s \nearrow t$ and taking expectations in (8.19) we obtain, after using the fact that the last two terms are martingales,

$$E_x^R[e^{-\lambda t}V(0,X_t)] = V(t,x) + E_x^R\left[\int_0^t e^{-\lambda u}\mathcal{L}V(t-u,X_{u-})\mathbf{1}_{[X_{u-}\geq b(t-u)]}\,du\right]. \tag{8.31}$$

Now since $V(0,x)=x$, after rearranging the above equation we have

$$V(t,x) = e^{-\lambda t}E_x^R[X_t] - E_x^R\left[\int_0^t e^{-\lambda u}\mathcal{L}V(t-u,X_{u-})\mathbf{1}_{[X_{u-}\geq b(t-u)]}\,du\right]$$

Using the first equality in (8.28) with $V(t-u,x)$ instead of $v(t,x)$ and b instead of c we see that

$$\begin{aligned}
&E_x^R\left[\int_0^t e^{-\lambda u}\mathcal{L}V(t-u,X_{u-})\mathbf{1}_{[X_{u-}\geq b(t-u)]}\,du\right]\\
&= -(\lambda+r)\int_0^t e^{-\lambda u}E_x^R[X_{u-}\mathbf{1}_{[X_{u-}\geq b(t-u)]}]\,du\\
&\quad+\lambda^R E_x^R\left[\int_0^t\int_{A_{x,t,u}} e^{-\lambda u}\left(V(t-u,X_{u-}+\zeta(X_{u-},z)) - X_{u-}\left(\frac{1}{1+\gamma(z)}\right)\right)m^R(dz)\,du\right]
\end{aligned}$$

which proves that $v=V$. That $c=b$ follows immediately from (8.25), (8.26), and the definition of b. □

Observe that the term $\tilde{V}(t,x)$ is the value function for the Russian option without early exercise rights, and $e_1 - e_2$ is the premium for these rights. The part of (8.31), given by the expression $\tilde{V}(t,x) + e_1(t,x)$, is the same as the early exercise premium representation given by Peskir [25] in the diffusion case, while the e_2 term is a result of the addition of jumps to the model.

Appendix A

Additional Lemmas

In this appendix we provide proofs of some ancillary results used in the main portion of the thesis.

A.1 Proof of Lemma 12

Lemma 3.1 of Pham [27] provides estimates for general processes given by a stochastic differential equation with coefficients satisfying the same conditions as those in Lemma 1. The difference between our process X and those satisfying the conditions of Pham [27] is that X has a boundary term which must be dealt with. We begin by proving a simple bound for this term.

Lemma 12. *For any stopping time τ such that $0 \leq \tau \leq h \leq T$ there exists some constant C, dependent on T but independent of h, such that*

$$E\left[\left(\int_0^\tau \mathbf{1}_{[\Delta M_u = 0]} \frac{dM_u}{S_{u-}}\right)^2\right] \leq Ch.$$

Proof. Since the process M only increases when S either jumps to a new maximum or equals M, we have

$$0 \leq \int_0^\tau \mathbf{1}_{[\Delta M_u = 0]} \frac{dM_u}{S_{u-}} = \int_0^\tau \mathbf{1}_{[\Delta M_u = 0]} \frac{dM_u}{M_{u-}} \leq \int_0^\tau \mathbf{1}_{[\Delta M_u = 0]} \, dS^*_u \leq S^*_\tau - S^*_0 \leq \sup_{0 \leq u \leq h} S_u - 1$$

where for the second inequality we used the fact that $M_{u-} \geq x \geq 1$ and the fact that the change in M_u equals the change in S^*_u if M has no jumps. Thus, since $M_u(x, 1)$ only increases when $S^*_u(1) > x$,

$$E_x\left[\left(\int_0^\tau \Delta M_{[u=0]}\frac{dM_u}{S_{u-}}\right)^2\right] \le E_1\left[\left(\int_0^\tau \Delta M_{[u=0]}\frac{dM_u}{S_{u-}}\right)^2\right] \le E_1\left[\sup_{0\le u\le h}|S_u - 1|^2\right] \le Ch, \tag{A.1}$$

where the final inequality follows from Lemma 3.1 of Pham [27]. □

Next we provide a proof of Lemma 2. The proof below is a modification of that of Lemma 3.1 in Pham [27].

Proof. For notational convenience, in the proof that follows we will use C to denote any generic constant dependent only on k and T. We begin by remarking that it suffices to prove the results for $k = 2$. To see this, assume the results holds for $k = 2$. Applying Hölder's inequality with $p = \frac{2}{k}$, $q = \frac{2}{2-k}$ we obtain

$$\begin{aligned} E^R_x|X_\tau|^k &\le (E^R_x|X^2_\tau|)^{\frac{k}{2}} \le C(1+|x|^2)^{\frac{k}{2}} \le C(1+|x|^k) \\ E^R|X_\tau - x|^k &\le (E^R|X_\tau - x|^2)^{\frac{k}{2}} \le C(1+|x|^2)^{\frac{k}{2}}h^{\frac{k}{2}} \le C(1+|x|^k)h^{\frac{k}{2}} \\ E^R\left[\sup_{0\le\tau\le h}|X_\tau - x|^k\right] &\le \left(E^R\left[\sup_{0\le\tau\le h}|X_\tau - x|^2\right]\right)^{\frac{k}{2}} \le C(1+|x|^2)^{\frac{k}{2}}h^{\frac{k}{2}} \le C(1+|x|^k)h^{\frac{k}{2}}. \end{aligned} \tag{A.2}$$

Let us now prove (i) for $k = 2$. Applying Lemma 1, Lemma 12 and the Itô-Lévy isometry we obtain

$$\begin{aligned} E^R_x|X_\tau|^2 \le &C\Bigg(|x|^2 + E^R_x\left[\int_0^\tau |\alpha(X_{u-})|^2\,du + \int_0^\tau |\beta(X_{u-})|^2\,du\right] \\ &+ E^R_x\left[\int_0^\tau\int_{\mathbb{R}} |\zeta(X_{u-},z)|^2\,m^R(dz)\,du + \int_0^\tau \mathbf{1}_{[\Delta M_u>0]}\frac{dM_u}{S_{u-}}\right]^2\Bigg) \\ \le &C\left(1 + |x|^2 + E^R_x\int_0^\tau |X_{u-}|^2\,du\right). \end{aligned} \tag{A.3}$$

From this, Fubini's Theorem and Gronwall's inequality implies that for any deterministic time $\tau = s$,

$$E^R_x|X_s|^2 \le C(1+|x|^2) \tag{A.4}$$

Substituting (A.4) into (A.3) we obtain

$$E|X_\tau|^2 \le C\left(1 + |x|^2 + (1+|x|^2)E[\tau]\right) \le C(1+|x|^2)$$

since $\tau \leq T$.

The proof for (ii) is similar. We have again by Lemma 12, and the fact that $\tau \leq h$, that

$$\begin{aligned} E_x^R|X_\tau - x|^2 \\ &\leq C\left(h + E_x^R\left[\int_0^h |\alpha(X_{u-})|^2\,du + \int_0^h |\beta(X_{u-})|^2\,du + \int_0^h \int_{\mathbb{R}} |\zeta(X_{u-}, z)|^2\, m^R(dz)\,du\right]\right) \\ &\leq C\left(h + \int_0^h E_x^R|X_u|^2\,du\right) \leq C(1 + |x|^2)h, \end{aligned} \tag{A.5}$$

where we used (A.4) for the final inequality.

Finally, let us show (iii). Similarly to the above proofs we have

$$\begin{aligned} E_x^R\left[\sup_{0\leq\tau\leq h} |X_\tau - x|^2\right] \leq C\Bigg(E_x^R\Bigg[&\left(\sup_{0\leq\tau\leq h}\int_0^\tau |\alpha(X_{u-})|\,du\right)^2 + \left(\sup_{0\leq\tau\leq h}\int_0^\tau |\beta(X_{u-})|\,du\right)^2 \\ &+ \left(\sup_{0\leq\tau\leq h}\int_0^\tau \int_{\mathbb{R}} |\zeta(X_{u-}, z)|\,\tilde{\nu}(dt, dz)\right)^2 + \left(\sup_{0\leq\tau\leq h}\int_0^\tau \mathbf{1}_{[\Delta M_u = 0]}\,\frac{dM_u}{S_{u-}}\right)^2\Bigg]\Bigg). \end{aligned} \tag{A.6}$$

Using the fact that the stochastic integrals are square-integrable martingales (see Section 3.3) we can apply Doob's Inequality in Mean, and combining this with the fact that the final integral above is increasing in τ we obtain

$$\begin{aligned} E_x^R&\left[\sup_{0\leq\tau\leq h} |X_\tau - x|^2\right] \\ &\leq C\Bigg(E_x^R\Bigg[\sup_{0\leq\tau\leq h}\int_0^h |\alpha(X_{u-})|^2\,du + \sup_{0\leq\tau\leq h}\int_0^h |\beta(X_{u-})|^2 du \\ &+ \sup_{0\leq\tau\leq h}\int_0^h \int_{\mathbb{R}} |\zeta(X_{u-}, z)|^2\, m^R(dz)\,du + \left(\int_0^h \mathbf{1}_{[\Delta M_u = 0]}\,\frac{dM_u}{S_{u-}}\right)^2\Bigg]\Bigg) \\ &\leq C\left(h^2 + E_x^R\left[\int_0^h X_{u-}^2 du\right]\right) \leq C(1 + T + |x|^2)h \leq C(1 + |x|^2)h \end{aligned} \tag{A.7}$$

by Lemma 1, Lemma 12, (A.4) and the fact that $h \leq T$. □

A.2 Results Used in the Proof of Theorem 6

Next we provide proofs of two facts used in proving the Generalized Itô Formula in Theorem 6, both similar to those used in Peskir [24]. Minor modifications have been made to deal with the fact that we are on a bounded interval and the addition of jumps and a bounded variation term in (8.1).

Lemma 13. *Suppose X_t is given by* (8.1) *and L_t^b is given by* (8.3) *with $b_{i,\epsilon}$ replaced by b. Then P-almost surely, for $i = 1, 2$ and $t \in [0, t_1]$ we have* $\sup_{0 \le s \le t} |L_s^{b_{i,\epsilon}}(X) - L_s^b(X)| \to 0$ *as $\epsilon \searrow 0$ over a subsequence.*

Proof. We prove the result for $b_{1,\epsilon}(s)$, as the proof for $b_{2,\epsilon}(s)$ is nearly identical. Applying Tanaka's formula as in (8.2) and rearranging, we see that

$$\sup_{0 \le s \le t} |L_s^{b_{1,\epsilon}}(X) - L_s^b(X)| \le I_1^\epsilon + I_2^\epsilon + I_3^\epsilon + I_4^\epsilon$$

where

$$\begin{aligned} I_1^\epsilon &= \sup_{0 \le s \le t} \Big| |X_s - b_{1,\epsilon}(s)| - |X_s - b(s)| \Big| + \Big| |X_0 - b_{1,\epsilon}(0)| - |X_0 - b(0)| \Big| \\ I_2^\epsilon &= \sup_{0 \le s \le t} \left| \int_0^s \operatorname{sign}(X_{u-} - b_{1,\epsilon}(u)) - \operatorname{sign}(X_{u-} - b(u))\, dX_u^c \right| \\ I_3^\epsilon &= \sup_{0 \le s \le t} \left| \int_0^s \operatorname{sign}(X_{u-} - b_{1,\epsilon}(u))\, db_{1,\epsilon}(u) - \int_0^s \operatorname{sign}(X_{u-} - b(u))\, db(u) \right| \\ I_4^\epsilon &= \sum_{0 \le u \le t} \Big| |X_u - b_{1,\epsilon}(u)| - |X_u - b(u)| \Big| + \Big| |X_{u-} - b_{1,\epsilon}(u)| - |X_{u-} - b(u)| \Big|. \end{aligned}$$

We will show that each of these terms converges to 0 as $\epsilon \searrow 0$.

For I_1^ϵ, observe that for all $s \in [0, t_1]$

$$\Big| |X_s - b_{1,\epsilon}(s)| - |X_s - b(s)| \Big| \le |(b(s) - \epsilon) \vee c - b(s)| \le \epsilon, \tag{A.8}$$

which implies $I_1^\epsilon \to 0$ as $\epsilon \searrow 0$.

Next, let $H_u^\epsilon := \operatorname{sign}(X_{u-} - b_{1,\epsilon}(u)) - \operatorname{sign}(X_{u-} - b(u))$ and observe that

$$\begin{aligned} I_2^\epsilon \le& \int_0^t |H_u^\epsilon| |\mu(u, X_{u-})|\, du + \lambda \int_0^t \int_{\mathbb{R}} |H_u^\epsilon| |\psi(u, X_{u-}, z)|\, m(dz)\, du \\ &+ \int_0^t |H_u^\epsilon|\, d|A|_u + \sup_{0 \le s \le t} \left| \int_0^s H_u^\epsilon \phi(u, X_{u-})\, dW_u \right| \end{aligned} \tag{A.9}$$

where $|A|$ denotes the variation of A. Since $P(X_{u-} = b(u)) = 0$ for all $u \in [0, t_1]$, we see that

$$E^P \left[\int_0^t \mathbf{1}_{[X_{u-} = b(u)]}\, du \right] = \int_0^t P(X_{u-} = b(u))\, du = 0.$$

Hence P-almost surely $\mathbf{1}_{[X_{u-}=b(u)]} = 0$ for λ-almost all $u \in [0,t]$, where λ denotes Lebesgue measure. Similarly, $\mathbf{1}_{[X_{u-}=b(u)]} = 0$ for ν_1-almost all $u \in [0,t]$, where ν_1 denotes the measure associated with $|A|$. Since $H_u^\epsilon = 0$ if $X_{u-} \notin [(b(u)-\epsilon) \vee c, b(u)]$, we can use the fact that $|H_u^\epsilon| \leq 2$ and the fact that μ and ψ are locally bounded to conclude that outside a P-null set the first three terms of (A.9) approach 0 as $\epsilon \searrow 0$.

For the forth term, the Burkholder-Davis-Gundy Inequality implies

$$E^P\left[\sup_{0\leq s\leq t}\left|\int_0^s H_u^\epsilon \phi(u, X_{u-})\, dW_u\right|\right] \leq E^P\left[\int_0^t (H_u^\epsilon)^2\phi^2(u, X_{u-})\, du\right] \tag{A.10}$$

which, from the local boundedness of ϕ, outside a P-null set approaches 0 as $\epsilon \searrow 0$. This implies that

$$\sup_{0\leq s\leq t}\left|\int_0^s H_u^\epsilon \phi(u, X_{u-})\, dW_u\right|$$

converges to 0 in probability and so $I_2^\epsilon \to 0$ P-a.s. as $\epsilon \searrow 0$ over a subsequence.

For I_3^ϵ, observe that since $db_{1,\epsilon}(u) = \frac{1}{2}(db(u) + \text{sign}(b(u) - \epsilon - c)\, db(u))$, P-almost surely

$$\begin{aligned} I_3^\epsilon &= \sup_{0\leq s\leq t}\left|\int_0^s \text{sign}(X_{u-} - b_{1,\epsilon}(u))\mathbf{1}_{[b(u)-\epsilon>c]}\, db(u)\right. \\ &\quad \left. + \frac{1}{2}\int_0^s \text{sign}(X_{u-} - b_{1,\epsilon}(u))\mathbf{1}_{[b(u)-\epsilon=c]}\, db(u) - \int_0^s \text{sign}(X_{u-} - b(u))\, db(u)\right| \\ &= \sup_{0\leq s\leq t}\left|\int_0^s \text{sign}(X_{u-} - b_{1,\epsilon}(u))\mathbf{1}_{[b(u)-\epsilon>c]}\, db(u) - \int_0^s \text{sign}(X_{u-} - b(u))\, db(u)\right|. \end{aligned}$$

To see second inequality above, let ν_2 denote the measure associated with $b(u)$ and note that by the same reasoning as for λ and ν_1 above, outside a P-null set $\mathbf{1}_{[b(u)-\epsilon=c]} = 0$ for ν_2 almost all $s \in [0,t]$. Continuing, we have

$$\begin{aligned} I_3^\epsilon &= \sup_{0\leq s\leq t}\left|\int_0^s H_u^\epsilon \mathbf{1}_{[b(u)-\epsilon>0]}\, db(u) - \int_0^s \text{sign}(X_{u-} - b(u))\mathbf{1}_{[b(u)-\epsilon\leq c]}\, db(u)\right| \\ &\leq \int_0^t |H_u^\epsilon|\mathbf{1}_{[b(u)-\epsilon>0]}\, d|b|_u + \int_0^t \mathbf{1}_{[X_{u-}\neq b(u)]}\mathbf{1}_{[b(u)-\epsilon\leq c]}\, d|b|_u \end{aligned}$$

which approaches 0 as $\epsilon \searrow 0$ since $H_u^\epsilon \to 0$ and $\mathbf{1}_{[b(u)-\epsilon\leq c]} \to 0$.

Finally, observe that by (A.8) we have $I_4^\epsilon \leq 2\epsilon N$, where N is the number of jumps of X which occur in the interval $[0,t]$, which is finite. Hence P-almost surely $I_4^\epsilon \to 0$ as $\epsilon \searrow 0$. □

Lemma 14. *Let I_2^ϵ be given by*

$$I_2^\epsilon := \int_0^t F_t(u, b_{1,\epsilon}(u))\mathbf{1}_{[X_{u-}>b_{1,\epsilon}(u)]}\,du + \int_0^t F_x(u, b_{1,\epsilon}(u))\mathbf{1}_{[X_{u-}>b_{1,\epsilon}(u)]}\,db_{1,\epsilon}(u)$$
$$+ \int_0^t F_t(u, b_{2,\epsilon}(u))\mathbf{1}_{[X_{u-}<b_{2,\epsilon}(u)]}\,du + \int_0^t F_x(u, b_{2,\epsilon}(u))\mathbf{1}_{[X_{u-}<b_{2,\epsilon}(u)]}\,db_{2,\epsilon}(u)$$

and suppose (8.10) *holds. Then for $t \in (0, t_1]$, P-almost surely $I_2^\epsilon \to F(t, b(t)) - F(0, b(0))$ as $\epsilon \searrow 0$ over a subsequence.*

Proof. First recall the change of variables formula

$$F(t, b_{i,\epsilon}(t)) = F(0, b_{i,\epsilon}(0)) + \int_0^t F_t(u, b_{i,\epsilon}(u))\,du + \int_0^t F_x(u, b_{i,\epsilon}(u))\,db_{i,\epsilon}(u) \qquad \text{(A.11)}$$

and note that to prove the result it is sufficient to show

$$J := \int_0^t F_t(u, b_{1,\epsilon}(u))\mathbf{1}_{[X_{u-}\le b_{1,\epsilon}(u)]}\,du + \int_0^t F_x(u, b_{1,\epsilon}(u))\mathbf{1}_{[X_{u-}\le b_{1,\epsilon}(u)]}\,db_{1,\epsilon}(u)$$
$$+ \int_0^t F_t(u, b_{2,\epsilon}(u))\mathbf{1}_{[X_{u-}\ge b_{2,\epsilon}(u)]}\,du + \int_0^t F_x(u, b_{2,\epsilon}(u))\mathbf{1}_{[X_{u-}\ge b_{2,\epsilon}(u)]}\,db_{2,\epsilon}(u)$$
$$\to F(t, b(t)) - F(0, b(0))$$

P-almost surely over a subsequence, since then

$$I_2^\epsilon = \int_0^t F_t(u, b_{1,\epsilon}(u))\,du + \int_0^t F_x(u, b_{1,\epsilon}(u))\,db_{1,\epsilon}(u)$$
$$+ \int_0^t F_t(u, b_{2,\epsilon}(u))\,du + \int_0^t F_x(u, b_{2,\epsilon}(u))\,db_{2,\epsilon}(u) - J$$
$$= F(t, b_{1,\epsilon}(t)) - F(0, b_{1,\epsilon}(t)) + F(t, b_{2,\epsilon}(t)) - F(0, b_{2,\epsilon}(0)) - J$$
$$\to F(t, b(t)) - F(0, b(0)).$$

Let μ_ϵ and μ denote the Lebesgue-Stiltjes signed measures associated with the functions $u \to F(u, b_{1,\epsilon}(u))$ and $u \to F(u, b(u))$ on $[0, t_1]$, respectively. Furthermore, let μ_ϵ^+ and μ_ϵ^- be the positive and negative parts of μ_ϵ, respectively. Helly's Selection Theorem applies since (8.10) holds, so over some subsequence ϵ_n such that $\epsilon_n \searrow 0$ we have $\mu_{\epsilon_n}^+ \to \mu_1$ and $\mu_{\epsilon_n}^- \to \mu_2$ weakly as $n \to \infty$ for some positive finite measures μ_1 and μ_2 on $[0, t_1]$. Since $F(u, b_{1,\epsilon_n}(u)) \to F(u, b(u))$ as $n \to \infty$ for $u \in [0, t_1]$ and $F(u, b(u))$ is continuous on $[0, t_1]$, we have $\mu = \mu_1 - \mu_2$.

Let $\mu_n^\pm := \mu_{\epsilon_n}^\pm$ and let $\mu_n := \mu_{\epsilon_n}$, and define the sets $A_n := \{0 \le u \le t | X_{u-} \le b_{1,\epsilon}(u)\}$ and $A := \{0 \le u \le t | X_{u-} < b(u)\}$. We will next show that $\mu_n(A_n) \to \mu(A)$ outside a P-null set, which will follow if we can show that $\mu_n^{\pm}(A_n) \to \mu_{1,2}(A)$ as $n \to \infty$ outside a P-null set. For this, set $a_{nm}^{\pm} := \mu_n^{\pm}(A_m)$. Then $\lim_{m\to\infty} a_{nm}^{\pm} = \mu_n^{\pm}(A) =: a_{n\infty}^{\pm}$ exists since $A_m \nearrow A$, and $\lim_{n\to\infty} a_{nm}^{\pm} = \mu_{1,2}(A_m) =: a_{\infty m}^{1,2}$ exists outside a P-null set since $\mu_n^{\pm} \to \mu_{1,2}$ weakly. We similarly see $\lim_{m\to\infty} a_{\infty m}^{1,2} = \lim_{n\to\infty} a_{n\infty}^{\pm} = \mu_{1,2}(A) =: a_{\infty\infty}^{1,2}$ outside a P-null set. We wish to show $\lim_{n\to\infty} a_{nn}^{\pm} = a_{\infty\infty}^{1,2}$. To this end, observe that since $A_n \nearrow A$, $a_{nn}^{\pm} \le a_{n\infty}^{\pm}$ so $\limsup_{n\to\infty} a_{nn}^{\pm} \le a_{\infty\infty}^{1,2}$. Conversely, $a_{nm}^{\pm} \le a_{nn}^{\pm}$ for $m \le n$ so $a_{\infty,m}^{1,2} \le \liminf_{n\to\infty} a_{nn}^{\pm}$. Thus we have

$$a_{\infty\infty}^{1,2} = \lim_{m\to\infty} a_{\infty m}^{1,2} \le \liminf_{n\to\infty} a_{nn}^{\pm} \le \limsup_{n\to\infty} a_{nn}^{\pm} \le a_{\infty\infty}^{1,2}$$

P-almost surely, which implies $\mu_n(A_n) \to \mu(A)$ as $n \to \infty$ outside a P-null set.

From (A.11) we have

$$\int_0^t \mathbf{1}_{A_n}(u) F_t(u, b_{1,\epsilon_n}(u))\, du + \int_0^t \mathbf{1}_{A_n}(u) F_t(u, b_{1,\epsilon_n}(u))\, db_1(u)$$
$$= \int_0^t \mathbf{1}_{A_n}(u)\, dF(u, b_{1,\epsilon_n}(u)) = \mu_n(A_n),$$

which converges to $\mu(A)$ from the above discussion. In the same way we can show that $\nu_{\epsilon_n}(B_n) \to \mu(B)$ where $B_n := \{0 \le u \le t | X_{u-} \ge b_{2,\epsilon}(u)\}$, $B := \{0 \le u \le t | X_{u-} > b(u)\}$, and ν_ϵ is the Lebesgue-Stiltjes measure associated with $u \to F(u, b_{2,\epsilon}(u))$ on $[0, t_1]$.

Hence $J \to \mu(A) + \mu(B) = \mu(A \cup B)$ over a subsequence. The result will follow if we can show $\mu(A \cup B) = \mu([0, t])$ outside a P-null set, since

$$\mu([0, t]) = F(t, b(t)) - F(0, b(0)).$$

For this, it suffices to show that $\mu_{1,2}(A^c \cap B^c) = 0$ outside a P-null set, which follows from the fact that

$$E\left[\int_0^t \mathbf{1}_{A^c \cap B^c}(u)\, \mu_{1,2}(du)\right] = E\left[\int_0^t \mathbf{1}_{[X_{u-} = b(u)]}\, \mu_{1,2}(du)\right]$$
$$= \int_0^t P(X_{u-} = b(u))\, \mu_{1,2}(du) = 0$$

since $P(X_{u-} = b(u)) = 0$.

□

Bibliography

[1] S. Asmussen, F. Avram, and M. Pistorius, Russian and American Put Options Under Exponential Phase-Type Lévy Models, *Stochastic Process and their Applications*, **109**, (2004) 79–111.

[2] F. Avram, A. Kyprianou, and M. Pistorius, Exit Problems for Spectrally Negative Lévy processes and applications to (Canadized) Russian Options, *The Annals of Applied Probability* , **14**, (2004) 215–238.

[3] L. Bachelier, Théorie de la spéculation, *Annales Scientifiques de l'École Normale Supérieure* **17** (1900), 21–86.

[4] F. Black and M. Scholes, The Pricing of Options and Corporate Liabilities, *Journal of Political Economy* **81** no. 3 (1973), 637–654.

[5] P. Brémaud, Point Processes and Queues: Martingale Dynamics, Spring-Verlag, New York (1981).

[6] K.L. Chung and R.J. Williams, An Introduction to Stochastic Integration, Birkhäuser Boston, Boston (1990).

[7] D.B Colwell and R.J. Elliott, Discontinuous Asset Prices and Non-attainable Contigent Claims, *Math. Finan.* **3** (1993), 295–308.

[8] R. Cont and P. Tankov, Financial Modelling with Jump Processes, *Chapman and Hall/CRC Press*, Boca Raton (2004).

[9] R. Cont, Empirical Properties of Asset Returns: Stylized Facts and Statistical Issues, *Quantitative Finance* **1** (2001), 223–236

[10] M.G. Crandall and H. Ishii. The Maximum Principle for Semicontinuous Functions, *Diff Int. Eqn* **3** (1990), 1001–1014.

[11] M.G. Crandall, Ishii H. and P.L. Lions. User's Guide to Viscosity Solutions of Second Order Partial Differential Equations, *Bull. Amer. Math. Soc.*, **27** (1992), 1–67.

[12] M.G. Crandall and P.L. Lions, Viscosity Solutions of Hamilton-Jacobi Equations, *Trans. A.M.S.*, **277** (1983), 1–42.

[13] M.G. Crandall and P.L. Lions, On Existence and Uniqueness of Solutions of Hamilton-Jacobi Equations, *Nonlin. Anal. Th. Meth. Appl.*, **10** (1986) 353–370.

[14] J.J. Duistermaat, A.E. Kyprianou, and K. Van Schaik, Finite Expiry Russian Options, *Stochastic Processes and their Applications* **115**, (2005) 609–638.

[15] E. Erkstrom, Russian Options with a Finite Time Horizon, *Journal of Applied Probability*, **3**, (2004) 313–326.

[16] A. Friedman, Partial Differential Equations of Parabolic Type, Prentice Hall (1983).

[17] P. Gapeev, Discounted optimal stopping for maxima of some jump-diffusion processes, *Journal of Applied Probability*, **44** 713–731.

[18] H. Gerber, F. Michaud and E. Shiu, "Pricing Russian Options with the Compound Poisson Process", *Transactions ICA Brussels*, (1995) **3** 243–264.

[19] J. Jacod and A.N. Shiryaev, Limit Theorems for Stochastic Processes, *Berlin, Heidelberg and New York: Spring-Verlag* (1987).

[20] A. Kyprianou, Introductory Lectures on Fluctuations of Lévy Processes With Applications, *Springer* (2006).

[21] P.L. Lions, Optimal Control of Diffusion Processes and Hamilton-Jacobi-Bellman Equations. Part 1: The Dynamic Programming Principal and Applications and Part 2: Viscosity Solutions and Uniqueness, *Comm. P.D.E*, **8** (1983) 1101 – 1174 and 1229 – 1276.

[22] R.C. Merton, The Theory of Rational Option Pricing, *Bell Journal of Economics and Management Science* **4** no. 1 (1973), 141–183.

[23] E. Morecki and W. Moreira, Russian Options for a Diffusion with Negative Jumps, *Publicaciones Matematicas del Uruguay* **9** (2001), 37–51.

[24] G. Peskir, A Change-of-Variable Formula with Local Time on Curves, *Journal of Theoretical Probability* **18** (2005), 499–535.

[25] G. Peskir, The Russian Option: Finite Horizon, *Finance and Stochastics* (2004), 251–267.

[26] G. Peskir and A. Shiryaev, Optimal Stopping and Free-Boundary Problems, Birkhäuser, Basel (2006).

[27] H. Pham, Optimal Stopping of Controlled Jump Diffusion Processes: A Viscosity Solution Approach, *Journal of Mathematical Systems, Estimation, and Control*, **8** (1998), 1–27

[28] H. Pham, Optimal Stopping, Free Boundary, and American Option in a Jump-Diffusion Model, *Applied Math and Optimization*, **35** (1997), 145–164.

[29] P. Protter, Stochastic Integration and Differential Equations, Springer (2003).

[30] L.C.G. Rogers and D. Williams, Markov Processes, Diffusion Processes and Martingales, Volume 2: Ito Calculus, Cambridge University Press, Cambridge (2000).

[31] W. Runggaldier, Jump Diffusion Models, In : Handbook of Heavy Tailed Distributions in Finance, Handbooks in Finance, Book 1, Elesevier/North-Holland (2003), pp.169-209.

[32] S. Shreve, Stochastic Calculus for Finance II: Continuous-Time Models, Springer Finance, New York (2004).

[33] L. Shepp, and A.N. Shiryaev, A New Look at Pricing of the "Russian Option", *Theory of Probability and its Applications*, **39**, Issue 1 (1994), 103–119

[34] L. Shepp and A.N. Shiryaev, The Russian Option: Reduced Regret, *Annals of Applied Probability*, **3**, Number 3 (1993), 631–640.