OPTIMAL STOPPING FOR MARKOV MODULATED ITO-DIFFUSIONS WITH APPLICATIONS TO FINANCE

by

THOMAS WILLIAM SEAQUIST

Presented to the Faculty of the Graduate School of
The University of Texas at Arlington in Partial Fulfillment
of the Requirements
for the Degree of

DOCTOR OF PHILOSOPHY

THE UNIVERSITY OF TEXAS AT ARLINGTON

May 2013

This dissertation is dedicated to my family and friends. Most especially to my partner Maifoua Thao who has never left my side and to my father Carl Seaquist who sparked my love for mathematics. A special dedication to my late uncle Thomas Canterbury whose wisdom will be forever missed.

ACKNOWLEDGEMENTS

First and foremost, I would like to thank my advisor Professor Andrzej Korzeniowski. His kind help, encouragement, and insight have made the writing of this dissertation a real pleasure. I am grateful for the tireless hours he spent discussing mathematics with me from which I have gained a new perspective on quality research and thinking. I would also like to thank Professors Chien-Pai Han and Shan Sun-Mitchell along with Drs D.L. Hawkins and Gaik Ambartsoumian for being members of my committee.

The friends I have made at UTA have been tremendously kind and helpful. It has been a particular privilege to know Paul Dawkins who has been a dear friend since our meeting; Ibrahim Diakite who has been endlessly encouraging and always fun to talk with; Carl Looney whose philosophical conversations I always enjoy; Padmini Veerapen for her encouragement and friendship; and Juan Licea-Salazar for keeping me healthy and running.

Lastly I would like to express my gratitude to my family. My partner Maifoua Thao has been by my side and encouraging me since I began graduate school. A special thanks to my father Carl Seaquist who sparked my interest in mathematics and whose support ever since has been invaluable. I would also like to thank my mother Margaret Brichta and my aunt Peggy Canterbury for their love and support, without which I would not be where I am today.

April 15, 2013

ABSTRACT

OPTIMAL STOPPING FOR MARKOV MODULATED ITO-DIFFUSIONS WITH APPLICATIONS TO FINANCE

Thomas William Seaquist, Ph.D.

The University of Texas at Arlington, 2013

Supervising Professor: Andrzej Korzeniowski

Despite the outstanding success of the Black-Scholes model, it relies on the assumption that drift and volatility of the underlying equity remain constant throughout time. This inaccuracy has motivated a number of interesting and innovative refinements, one of the most natural being Markov modulation. In this dissertation we analyze a variety of financially motivated optimal stopping problems under Markov modulated Ito-Diffusions. In Chapter 3, we generalize and refine a technique developed in [13] pricing an infinite time horizon American put option and we present a rigorous proof of optimality. In Chapter 4 we use this generalized technique to discover an optimal selling strategy for an infinite horizon American style forward contract. In so doing, we extend the work done in [12]. Finally in Chapter 5 we price the infinite horizon American put using a non-traditional model of a mean reverting Ornstein-Uhlenbeck process, further illustrating the broad scope of applicability of the technique developed herein.

TABLE OF CONTENTS

CHAPTER 1

INTRODUCTION

The theory of optimal stopping is fundamental to a variety of time dependent financial instruments. A key question to answer is this: when is the best time to buy or sell an underlying option contract in order to maximize potential return? Similarly, it is equally important in the pricing of contracts which can be bought or sold at any time. Contracts which can be exercised at a time of the owners choosing are referred to as "American" contracts and play a central part of our studies in this dissertation.

The Black-Scholes-Merton formula for pricing European options is a celebrated result in economics for which Myron Scholes and Robert Merton won the Nobel prize in economics in 1997. This formula is based on modeling stock as a geometric Brownian motion from which a hedging scheme for replicating an option price results in a partial differential equation that can be solved. Fischer Black and Myron Scholes first articulated the model in 1973, see [2], and it was extended up by Robert Merton, see [21]. Certain American type options have also been successfully priced under the same Black-Scholes geometric Brownian motion model using a variety of probabilistic and differential equation techniques, see [4, 7, 15, 20, 26]. Despite the outstanding success of the Black-Scholes model, it relies on the assumption that the drift and volatility of the stock remain constant throughout the lifetime of the contract. This is a reasonable assumption for short time periods but fails to hold true over longer periods of time as evidenced by the "volatility smile," (*e.g.*, [22], p. 221). There have been many attempts in refining the model to account for this discrepancy, see [5] for a few examples of established techniques. The techniques we seek to investigate involve allowing the drift and volatility of a stock to transition between states by a Markov chain.

There have been numerous studies of option pricing for Markov modulated, or regime switching models. A few of the more relevant studies to this dissertation will now be discussed. In [3], a European option is priced in an n state Markov modulated model and the finite horizon American put is approximated in a $n = 2$ state model. The technique used relies on some innovative ways of expressing Markov modulation from Robert Elliot's "Hidden Markov Models," [8]. In [11], a closed form solution for the infinite time horizon American put option is found for an $n = 2$ state model. This is extended upon in [13] where a closed form solution is found for any n state model. In Chapter 3, we seek to improve upon and generalize the technique presented in [13]. For further work in pricing a variety of options in a Markov Modulated frame work, we refer the reader to [10, 19, 32–34].

We begin with a general probabilistic framework in Chapter 2, where the key concepts are defined and relevant fundamental results are stated. In Chapter 3, we generalize and refine the technique developed in [13] and present a rigorous proof of optimality. In Chapter 4 we use this generalized technique to discover an optimal selling strategy for an infinite horizon American style forward contract in an n state model. In so doing, we extend the work done in [12] where an $n = 2$ state model is used. Finally in chapter 5 we price the infinite horizon American put using a non-traditional model of a mean reverting Ornstein-Uhlenbeck process, further illustrating the broad scope of applicability of the technique we have developed.

CHAPTER 2

GENERAL FRAMEWORK

2.1 The Ito-Integral

Let $(\Omega, \mathbb{F}, P)$ be a fixed probability space. All random objects are measurable functions from the probability space into the real line $\mathbb{R}$. A random process X_t is an abbreviation for $X_t(\omega)$, $\omega \in \Omega$, and $t \geq 0$, which for fixed ω is called a path, realization, or trajectory and is a function of time t. We also have an increasing family of σ - algebras $\{\mathcal{F}_t\}_{t\geq 0}, \mathcal{F}_t \subset \mathbb{F}$ referred to as a filtration. The filtration $\{\mathcal{F}_t\}_{t\geq 0}$ specifies how information is revealed through time. Each random process is $\{\mathcal{F}_t\}_{t\geq 0}$ adapted, *i.e.*, X_t is measurable with respect to $\mathcal{F}_t$ for all $t \geq 0$.

At times we will utilize the so called natural filtration of (or filtration generated by) a process X_t defined $\mathcal{H}_t = \sigma(X_u, u \leq t)$ the smallest σ - algebra that contains all sets of the form $\{X_u^{-1}[B] : B \text{ is a Borel set on } \mathbb{R}\}$ for $0 \leq u \leq t$.

We start by defining an essential concept of stochastic modeling. The Ito-integral is motivated by the desire to model a process which experiences some form of noise

$$\frac{dX_t}{dt} = b(t, X_t) + \sigma(t, X_t) \cdot noise.$$

For many purposes, we assume that the noise is a generalized stochastic process called the the *white noise process*, formally $\frac{dW_t}{dt}$ where W_t is the Wiener process or Brownian motion. Interpreting the above in integral form

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

Since W_t has unbounded variation almost surely on every time interval, path-wise integration cannot be carried out. Instead, integration with respect to Brownian motion is defined as the celebrated Ito-integral. An outline of how this is done is as follows.

First the integral will be defined on a restrictive class of simple functions defined

$$\phi(t,\omega) = \sum_{j\geq 0} e_j(\omega)\mathbb{1}_{[j2^{-n},(j+1)2^{-n})}(t)$$

where $e_j(\omega)$ is measurable with respect the σ - algebra $\mathcal{F}_{j2^{-n}}$. For a simple function, we define the integral

$$\int_S^T \phi(t,\omega)dW_t = \sum_{j\geq 0} e_j(\omega)[B_{t_{j+1}} - B_{t_j}](\omega)$$

where

$$t_k = \begin{cases} k2^{-n} & if \quad S \leq k2^{-n} \leq T \\ S & if \quad k2^{-n} \leq S \\ T & if \quad k2^{-n} \geq T \end{cases}.$$

From this definition it can be shown that there exists a sequence of simple functions converging to any function in the Ito integrable class defined:

Definition 2.1.1.

Let $\mathcal{V} = \mathcal{V}(S,T)$ be the class of functions

$$f(t,\omega) : [0,\infty) \times \Omega \to \mathbb{R}$$

such that

1. $f(t,\omega)$ is $\mathcal{B} \times \mathcal{F}$ - measurable, where $\mathcal{B}$ represents the Borel σ - algebra on $[0,\infty)$
2. $f(t,\omega)$ is $\mathcal{F}_t$ - adapted
3. $E[\int_S^T f(t,\omega)^2 dt] < \infty$.

The convergence of the above is L^2 convergence in the product space $[0,\infty)\times\Omega$, *i.e.*,

$$E[\int_S^T (f-\phi_n)^2 dt] \to 0 \quad as \quad n \to \infty.$$

The Ito-isometry plays an essential role in defining the Ito-integral. The Ito-isometry states

$$E[\int_S^T \phi(t,\omega)^2 dt] = E[(\int_S^T \phi(t,\omega)dW_t)^2].$$

This allows for the existence of the limit which defines the Ito-integral for $f \in \mathcal{V}(S,T)$

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

in the $L^2(P)$ sense.

For further details, we refer the reader to [23].

Remark 2.1.1.

Assumption 2 in Definition 2.1.1 puts an important (yet natural) restriction on the class of Ito integrands, that are often referred to as predictable or non-anticipating processes.

Next we provide the key theorem of Ito calculus.

Theorem 2.1.1 (Ito's Formula).

Let X_t be an Ito process given by

$$dX_t = u(t,\omega)dt + v(t,\omega)dW_t$$

where W_t is a Wiener process (Brownian motion). Let $g(t,x) \in C^2([0,\infty)\times\mathbb{R})$. Then

$$dg(t,X_t) = g_t(t,X_t)dt + g_x(t,X_t)dX_t + \frac{1}{2}g_{xx}(t,X_t)\cdot(dX_t)^2,$$

where $(dX_t)^2 = dX_t \cdot dX_t$ is computed according to the following rules

$$dt\cdot dt = dt\cdot dW_t = dW_t\cdot dt = 0, \quad dW_t\cdot dW_t = dt.$$

For proof of Ito's formula, see [23].

Finallay, we end this section with an important definition that will be essential in generalizing Ito's Formula.

Definition 2.1.2 (Quadratic Variation [18]).

The quadratic variation $[X, X]_t$ of a process X_t is defined

$$[X, X]_t = \lim_{\|P\| \to 0} \sum_{k=1}^{n} (X_{t_k} - X_{t_{k-1}})^2$$

where $P = \{0 = t_0, t_1, \ldots, t_n = t\}$ is a partition of $[0, t]$ and $\|P\|$ is the norm of the partition and the limit is in probability.

More generally we have

Definition 2.1.3 (Quadratic Covariation [18]).

The quadratic covariation $[X, Y]_t$ of two processes X_t and Y_t is defined

$$\begin{aligned}[X, Y]_t &= \lim_{\|P\| \to 0} \sum_{k=1}^{n} (X_{t_k} - X_{t_{k-1}})(Y_{t_k} - Y_{t_{k-1}}) \\ &= \frac{1}{2}([X + Y, X + Y]_t - [X, X]_t - [Y, Y]_t)\end{aligned}$$

where $P = \{0 = t_0, t_1, \ldots, t_n = t\}$ is a partition of $[0, t]$ and $\|P\|$ is the norm of the partition and the limit is in probability.

2.2 Martingales and Stopping Times

A martingale is defined as follows:

Definition 2.2.1.

Let $\{\mathcal{F}_t\}_{t \geq 0}$ be a filtration on $(\Omega, \mathcal{F})$. The process $\{M_t\}_{t \geq 0}$ is said to be a martingale if

1. M_t is $\mathcal{F}_t$ - measurable
2. $E[|M_t|] \leq \infty$ for all $t \geq 0$
3. $E[M_t | \mathcal{F}_s] = M_s$ for all $0 \leq s \leq t$.

A process is said to be a submartingale if condition 3 is replaced by

$$E[M_t|\mathcal{F}_s] \geq M_s$$

and a supermartingale when condition 3 is replaced by

$$E[M_t|\mathcal{F}_s] \leq M_s.$$

Definition 2.2.2.

Let $\{\mathcal{F}_t\}_{t\geq 0}$ be a filtration on (Ω, P). The random variable $\tau : \Omega \to [0, \infty)$ is called a stopping time w.r.t. $\{\mathcal{F}_t\}_{t\geq 0}$ if

$$\{\omega : \tau(\omega) \leq t\} \in \mathcal{F}_t.$$

A filtration can be thought of as the history of a process. With this in mind, a stopping time is a strategy for stopping a process which can only see the past and cannot look into the future.

Definition 2.2.3.

Let X_t be a stochastic process. A property $\mathcal{P}$ is said to hold locally if there exists a sequence of stopping times τ_n increasing to infinity a.s. such that $X_{\tau_n \wedge t}\mathbb{1}_{\{\tau_n > 0\}}$ has property $\mathcal{P}$ for each $n \geq 1$.

A process is called a local martingale if it satisfies the condition of a martingale locally.

Definition 2.2.4.

Let S_t be stochastic process. S_t is said to be a semimartingale if it can be decomposed

$$S_t = A_t + M_t$$

where M_t is a martingale and A_t is a process of finite variation.

An important property of the Ito-integral is that it is a martingale:

Theorem 2.2.1.

Let $f(t,\omega) \in \mathcal{V}(0,T)$ for all T. Then

$$M_t(\omega) = \int_0^t f(s,\omega)dW_s$$

is a martingale w.r.t. the natural filtration $\mathcal{F}_t$ generated by W_t. Furthermore,

$$E[M_t] = 0.$$

Next a few important theorems dealing with martingales and stopping times will be given.

Theorem 2.2.2 (Stopped Martingale Theorem [18]).

If M_t is a martingale and τ is a stopping time, the the stopped process $M_{\tau\wedge t}$ is a martingale and moreover

$$E[M_{\tau\wedge t}] = M_0.$$

Theorem 2.2.3 (Optional Sampling Theorem Version 1 [27]).

Let $\sigma \leq \tau$ be bounded stopping times and let M_t be a martingale (or supermartingale), then

$$E[M_\tau \mid \mathcal{F}_\sigma] = (\leq) M_\sigma$$

and in particular

$$E[M_\tau] = (\leq) M_0.$$

Theorem 2.2.4 (Optional Sampling Theorem Version 2 [18]).

Let $\sigma \leq \tau$ be any stopping times and M_t be a uniformly integrable martingale, then

$$E[M_\tau \mid \mathcal{F}_\sigma] = M_\sigma$$

and in particular

$$E[M_\tau] = M_0.$$

2.3 Continuous Time Markov Chains

Definition 2.3.1.

Consider the stochastic process $X_t(\omega) : [0, \infty) \times \Omega \to \mathbb{Z}^+$. We say X_t is a continuous time Markov chain if for all $s, t \geq 0$

$$P\left(X_{t+s} = j \mid X_t = i, \{X_u : 0 \leq u < t\}\right) = P(X_{t+s} = j \mid X_t = i).$$

Here we only consider time homogeneous Markov chains: $P(X_{t+s} = j \mid X_t = i)$ is independent of t. The Markov chain is a process that lacks "memory". Its future only depends upon it's current state. As a result we see that if τ_i denote the amount of time that the process spends in state i then

$$P(\tau_i > t + s \mid \tau_i > s) = P(\tau > t).$$

Hence τ_i must be exponentially distributed. To fully characterize a finite state Markov chain, we need only define

$$\begin{aligned} q_{ij} &= \nu_i P_{ij} \\ &= \lim_{t \searrow 0} \frac{P_{ij}(t)}{t} \end{aligned}$$

where ν_i is the rate (in the exponential distribution) of exiting state i, P_{ij} is the probability that given a transition out of i that the process goes to state j, and $P_{ij}(t)$ is the probability of being in state j at time t starting from state i and so q_{ij} is the transition rate into state j from i. Kolmogorov's Backward Equation describes how to relate these transition rates to the probability of being in a given state at a given time.

Theorem 2.3.1 (Kolmogorov's Backward Equation [30]).

$$P'(t) = QP(t)$$

where $P(t)$ is the matrix whose element in the i, j position is $P_{ij}(t)$. Q is the matrix whose element in the i, j position is q_{ij} where

$$q_{ii} = -\nu_i = -\sum_{j \neq i} q_{ij}.$$

Q is referred to as the transition rates matrix or generating matrix.

For further information on Markov chains we reference the reader to [14,30]

2.4 Strong Markov Property

First we define the shift operator along with some properties and then we will use this to define the strong Markov property. Without loss of generality we assume $\Omega = \mathbb{R}^{[0,\infty)}$ so that for each $\omega \in \Omega$, $\omega = (\omega_t)_{t \geq 0}$ and $X_t(\omega) = \omega_t$. The shift operator $\theta_t : \Omega \to \Omega$ is defined

$$\theta_t(\omega) = (\omega_{s+t})_{s \geq 0},$$

and for a random time σ

$$\theta_\sigma(\omega) = (\omega_{s+\sigma(\omega)})_{s \geq 0}.$$

Below several properties of the shift operator will be defined that will be useful in the sequel.

- For any stopping times σ and τ,

$$X_\tau \circ \theta_\sigma = X_{\sigma + \tau \circ \theta_\sigma}.$$

- If $\sigma \leq \tau$ where τ is an entry time into a set, then

$$\tau = \sigma + \tau \circ \theta_\sigma.$$

Definition 2.4.1 (Strong Markov Property [26]).

X_t is said to have the strong Markov property, or to be a strong Markov process if for every bounded Borel function ϕ and any stopping times σ and τ

$$E^x[\phi(X_\tau) \circ \theta_\sigma \mid \mathcal{F}_\sigma] = E^{X_\sigma}[\phi(X_\tau)].$$

A finite state continuous time Markov chain is a strong Markov process. Ito-processes also possess this property:

Theorem 2.4.1 (Strong Markov Property for Ito Diffusions [23]).

Define the Ito-process by the SDE

$$dX_t = b(t, X_t)dt + \sigma(t, X_t)dB_t$$

then X_t has the strong Markov property.

2.5 Infinitesimal Generators

Definition 2.5.1 (Infinitesimal Generator).

The infinitesimal generator L_X of a process X_t is defined by

$$L_X[f](x) = \lim_{t \searrow 0} \frac{E^x[f(X_t)] - f(x)}{t}$$

the domain of which is the set of functions $f : \mathbb{R} \to \mathbb{R}$ such that the above limit exists.

A similar operator can be defined

Definition 2.5.2 (Characteristic Operator [23]).

The characteristic operator $\mathcal{A}_X$ of a process X_t is defined by

$$\mathcal{A}_X[f](x) = \lim_{U \searrow x} \frac{E^x[f(X_{\tau_U})] - f(x)}{E^x[\tau_U]}$$

where the $U's$ are open sets U_k decreasing to the point x, in the sense that $U_{k+1} \subset U_k$ and $\bigcap_k U_k = \{x\}$ and $\tau_U = \{t \geq 0 : X_t \notin U\}$ is the first exit time from U for X_t. The domain is the set of functions $f : \mathbb{R} \to \mathbb{R}$ such that the above limit exists.

It is a fact that the infinitesimal generator and the characteristic operator coincide on the more restrictive domain of the infinitesimal generator, [23].

Theorem 2.5.1 (Infinitesimal Generator of an Ito-Process [23]).

Let X_t be an Ito-process defined

$$dX_t = b(t, X_t)dt + \sigma(t, X_t)dB_t.$$

The generator L_X of this process is given by

$$L_X[f](x) = b(t,x)f'(x) + \frac{1}{2}\sigma^2(t,x)f''(x).$$

Next, the infinitesimal generator for a Markov chain will be given.

Theorem 2.5.2 (Infinitesimal Generator of a Markov Chain).

Let ξ_t be Markov chain with n states defined by its generating matrix Q whose element in the i, j position is q_{ij}. The generator L_ξ of this process is given by

$$L_\xi[f](i) = \sum_{j=1}^{n} q_{ij} f(j).$$

This is equivalent to the i-th row of $Q\mathbf{f}$ where $\mathbf{f}$ is the vector whose j-th element is $f(j)$.

For additional reference in stochastic processes, we refer the reader to [16–18, 23, 30].

CHAPTER 3

THE AMERICAN PUT OPTION IN A MARKOV MODULATED MARKET

3.1 Introduction

In this chapter we will present a nontraditional framework under which to price an American put option. The traditional approach is limited to a market with a drift and volatility which are constant throughout time. This in an unrealistic assumption which fails over long periods of time. In what follows, we will develop a model which allows the market to switch between n different states at exponentially distributed times.

After the stock model is developed, the value of an infinite time horizon American put option will be posed as optimal stopping problem. This optimal stopping problem will then be solved by relating the logarithmic stock process to a system of ordinary differential equations (ODEs) which are readily solved after imposing a smoothness condition. Finally, the solution to the ODEs is proven to be the optimal solution with a corresponding optimal stopping strategy.

3.2 A Markov Modulated Ito Diffusion and the American Put Option as an Optimal Stopping Problem

The standard model for the evolution of a stock is presented as the following stochastic differential equation (SDE)

$$dS_t = \mu S_t dt + \sigma S_t dW_t$$

where S_t is the stochastic process describing stock price and W_t is a Wiener process or Brownian motion, where μ and σ are the drift and volatility of the stock respectively. The solution to this SDE can be found using the standard Ito formula and results in

$$S_t = S_0 e^{(\mu - \frac{1}{2}\sigma^2)t + \sigma W_t}$$

referred to as a geometric Brownian motion. In our model we introduce a continuous time Markov chain ξ_t of n states corresponding to the n states we allow the market to switch between. This Markov chain is assumed to be time homogeneous. The n state market is modeled by the SDE

$$dS_t = \mu(\xi_t)S_t dt + \sigma(\xi_t)dW_t \tag{3.1}$$

resulting in the stock process,

$$S_t = S_0 e^{\int_o^t \left(\mu(\xi_s) - \frac{1}{2}\sigma^2(\xi_s)\right)ds + \int_0^t \sigma(\xi_s)S_t dW_s}.$$

The key issue in option pricing is to derive a price producing no arbitrage opportunities. An arbitrage opportunity is the situation under which an investor can make a guaranteed profit while incurring no risk. In order to avoid arbitrage, we must price the option under the so called "risk neutral" measure or the martingale measure. It is well established that the pricing of options under the "risk neutral" measure produces a price eliminating arbitrage opportunities. Under this "artificial" probability measure, we would like for

$$e^{-\int_0^t r(\xi_s)ds} S_t$$

to be a martingale. Here $r(\xi_t)$ is the risk free interest rate. We will call the original probability measure P under which W_t is a brownian motion. To get the discounted stock process to be a martingale under a new measure , we will define a equivalent measure Q with the Radon-Nikodym derivative:

$$\frac{dQ}{dP} = \exp\left(-\int_0^t \frac{\mu(\xi_s) - r(\xi_s)}{\sigma(\xi_s)}ds - \frac{1}{2}\int_0^t \frac{(\mu(\xi_s) - r(\xi_s))^2}{\sigma^2(\xi_s)}dWs\right).$$

By Girsanov's change of measure theorem, the process

$$\tilde{W}_t = \int_0^t \frac{\mu(\xi_s) - r(\xi_s)}{\sigma(\xi_s)}ds + Wt$$

is a Q-brownian motion. Expressing this in differential form we get,

$$d\tilde{W}_t = \frac{\mu(\xi_t) - r(\xi_t)}{\sigma(\xi_t)}dt + dWt.$$

Combining this with (3.1), we get

$$dS_t = r(\xi_t)S_t dt + \sigma(\xi_t)S_t d\tilde{W}_t.$$

From this it is obvious that the discounted stock process,

$$e^{-\int_0^t r(\xi_s)ds}S_t = S_0 e^{-\frac{1}{2}\int_o^t \left(\sigma^2(\xi_s)\right)ds + \int_0^t \sigma(\xi_s)d\tilde{W}_s}$$

is a Q-martingale as desired. From this point forward, we will only consider the Q probability space and will ignore the tilde on W_t.

The value of the infinite horizon American put option is defined probabilistically by the following optimal stopping problem

$$V(s,i) = \sup_\tau E^{(s,i)}\left[e^{-\int_0^\tau r(\xi_s)ds}(K - S_\tau)^+\right]$$

where τ is a stopping time and K is the strike price and the expectation is taken in Q. This is an elegant definition since it can be interpreted as the expected payoff of the option discounted for time under the optimal stopping strategy. Of course, this expectation is taken under the synthetic measure Q, but it is under this risk free measure only that an arbitrage free price is established.

At this point we will take a critical change of perspective. We define the logarithmic stock process

$$X_t = \ln(S_t).$$

With the standard Ito formula we arrive at

$$dX_t = \left(r(\xi_t) - \frac{1}{2}\sigma^2(\xi_t)\right)dt + \sigma(\xi_t)dW_t$$

and the value function as a function of X_t

$$V(x,i) = \sup_{\tau} E^{(s,i)}\left[e^{-\int_0^{\tau} r(\xi_s)ds}(K - e^{X_\tau})^+\right].$$

This critical change of perspective is what will later allow us to solve this optimal stopping problem, see Remark 3.5.1.

Notice that in the above, no assumption is made about the finiteness of τ. Suppose that $P(\tau = \infty) > 0$. In this case see that $e^{-\int_0^t r(\xi_s)ds} \to 0$ as $t \to \infty$ since $r(j) > 0$ for all j. Also notice that $(K - e^{X_\tau})^+$ is bounded. Because of this the value function is sometimes accurately written

$$V(x,i) = \sup_{\tau} E^{(s,i)}\left[e^{-\int_0^{\tau} r(\xi_s)ds}(K - e^{X_\tau})^+ \mathbb{1}_{\{\tau<\infty\}}\right]$$

making any definition for X_τ when $\tau = \infty$ redundant. In the remainder of this article, $\mathbb{1}_{\{\tau<\infty\}}$ will not be written inside the expectation for brevity, however it should be understood that when $\tau = \infty$ then $e^{-\int_0^{\tau} r(\xi_s)ds}(K - e^{X_\tau})^+ = 0$.

To confirm that the above observation is critical, it will be shown that under certain conditions, the optimal stopping time will be infinite with probability greater than zero. For ease of illustration, the following explanation will be done with a one state market. It is well demonstrated in ([26], p. 375) that the optimal stopping time for a one state market is an entry time of the form

$$\tau_b = \inf\{t \geq 0 : X_t \leq b\}$$

for some threshold b. Here the logarithmic stock process is

$$X_t = \left(r - \frac{1}{2}\sigma^2\right)dt + \sigma(\xi_t)dW_t$$

which is simply a Brownian motion with drift $\eta = \left(r - \frac{1}{2}\sigma^2\right)$. It is important to notice that we do not know the sign of η. If $\eta < 0$ then the drift of the Brownian motion will be pushing towards the boundary b and X_t will strike the boundary with probability 1. The same is true with $\eta = 0$. However, if $\eta > 0$ the drift of the Brownian motion will be pushing away from the boundary b and there is a positive probability that the process X_t will never strike b. Thus we have

$$P(\tau_b = \infty) > 0$$

mandating that we be very careful with the possibility that the optimal stopping time for the multistate process might also infinite.

3.3 Generalized Ito's Formula including a Pure Jump Markov Process

First we develop the mathematical tools necessary for the later needed analysis. In this section a generalized Ito-formula will be presented and used to find the infinitesimal generator of the two dimensional process (X_t, ξ_t). These two results will then be combined to present a very elegant version of Ito's formula with a Markov chain.

The general Ito's formula for an n-tuple of possibly discontinuous semimartingales can be found in a monograph of stochastic integration ([27], p. 74, Th. 33), and is given below.

Theorem 3.3.1 (Generalized Ito's Formula [27]).

Let $X = (X^1, \ldots, X^n)$ be collection of semimartingales and let $f : \mathbb{R}^n \to \mathbb{R}$ have continuous *second order partial derivatives. Then $f(X_t)$ is a semimartingale and the following formula holds:*

$$f(X_t) = f(X_0) + \sum_{i=1}^{n} \int_{0^+}^{t} \frac{\partial f}{\partial x_i}(X_{s^-}) dX_s^i + \frac{1}{2} \sum_{1 \le i,j}^{i,j \le n} \int_{0^+}^{t} \frac{\partial^2 f}{\partial x_i \partial x_j}(X_{s^-}) d[X^i, X^j]_s^c$$
$$+ \sum_{0 < s \le t} \left\{ \Delta f(X_s) - \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(X_{s^-}) \Delta X_s^i \right\},$$

where, $X_{t^-} = \lim_{t \to 0^-} X_t$, $\Delta X_t = X_t - X_{t^-}$ and $\Delta f(X_t) = f(X_t) - f(X_{t^-})$. $[X, X]_t$ and $[X, Y]_t$ denote the quadratic variation and the quadratic covariation respectively and $[X, Y]_t^c$ denotes the path by path continuous part of $[X, Y]_t$.

A theory of extending Ito-integrals to a broader class integration with respect to martingales and semimartingales is well developed and presented thoroughly in [28, 29]

Now Ito's Formula will be applied to the function $f(Z_t)$ where $Z_t = (X_t, \xi_t)$. Recall that X_t is the Markov modulated Ito-process defined

$$dX_t = \left(r(\xi_t) - \frac{1}{2}\sigma^2(\xi_t) \right) dt + \sigma(\xi_t) dW_t$$

and ξ_t is the n state Markov chain.

$$f(Z_t) = f(Z_0) + \int_{0^+}^{t} f_x(Z_{s^-}) dX_s + \int_{0^+}^{t} f_\xi(Z_{s^-}) d\xi_s + \frac{1}{2} \int_{0^+}^{t} f_{xx}(Z_{s^-}) d[X, X]_s^c$$
$$+ \int_{0^+}^{t} f_{x\xi}(Z_{s^-}) d[X, \xi]_s^c + \frac{1}{2} \int_{0^+}^{t} f_{\xi\xi}(Z_{s^-}) d[\xi, \xi]_s^c$$
$$+ \sum_{s \le t} [\Delta f(Z_s) - f_x(Z_{x^-}) \Delta X_s - f_\xi(Z_{s^-}) \Delta \xi_s].$$

Since ξ_s is of bounded variation and X_s is continuous, $[X, \xi]_s^c = [\xi, \xi]_s^c = \Delta X_s = 0$. In addition, since ξ_s is a pure jump process, $\int_{0^+}^{t} f_\xi(Z_{s^-}) d\xi_s = \sum_{s \le t} f_\xi(Z_{s^-}) \Delta \xi_s$. By $dX_s = (r(\xi_s) - \frac{1}{2}\sigma^2(\xi_s)) ds + \sigma(\xi_s) dW_s$ and $d[X, X]_s^c = \sigma^2(\xi_s) ds$ one obtains

$$f(Z_t) = f(Z_0) + \int_{0}^{t} \left[\left(r(\xi_s) - \frac{1}{2}\sigma^2(\xi_s) \right) f_x(Z_s) + \frac{1}{2}\sigma^2(\xi_s) f_{xx}(Z_s) \right] ds$$
$$+ \int_{0^+}^{t} \sigma(\xi_{s^-}) f_x(Z_{s^-}) dW_s + \sum_{s \le t} \Delta f(Z_s).$$

Remark 3.3.1.

The left limit was ignored in the ds integral since $X_s \neq X_{s^-}$ at only a finite number of times almost surely. Left limits cannot be ignored in the Ito integral since the integrand without it would not be a predictable process.

To simplify notation and for clarity, the observation is made that the infinitesimal operator L_X of the process X_t is

$$L_X[f](x,i) = \left(r(i) - \frac{1}{2}\sigma^2(i)\right) f_x(x,i) + \frac{1}{2}\sigma^2(i) f_{xx}(x,i),$$

and yields

$$f(Z_t) = f(Z_0) + \int_0^t L_X[f](Z_s)ds + \int_{0^+}^t \sigma(\xi_{s^-}) f_x(Z_{s^-})dW_s + \sum_{s\leq t} \Delta f(Z_s). \tag{3.2}$$

Proposition 3.3.1.

The infinitesimal generator of the process $Z_t = (X_t, \xi_t)$ *for a bounded function* $f(\cdot, \xi) \in C^2(\mathbb{R})$*is given by*

$$L_{(X,\xi)}[f](x,i) = L_X[f](x,i) + L_\xi[f](x,i)$$

where

$$L_X[f](x,i) = \left(r(i) - \frac{1}{2}\sigma^2(i)\right) f_x(x,i) + \frac{1}{2}\sigma^2(i) f_{xx}(x,i)$$
$$L_\xi[f](x,i) = \sum_{j=1}^n q_{ij} f(x,j)$$

and q_{ij} *is the infinitesimal transition rate from state* i *to* j.

Proof. The generator is defined by

$$L_{(X,\xi)}[f](x,i) = \lim_{t\searrow 0} \frac{E^{(x,i)}[f(X_t,\xi_t)] - f(x,i)}{t}.$$

We apply Ito's Formula (3.2) and observe that by standard properties of Ito integrals $\int_{0^+}^t \sigma(\xi_{s^-}) f_x(Z_{s^-}) dW_s$ is a martingale and thus $E[\int_{0^+}^t \sigma(\xi_{s^-}) f_x(Z_{s^-}) dW_s] = 0$. Now

$$\begin{aligned}
L_{(X,\xi)}[f](x,i) &= \lim_{t\searrow 0} \frac{E^{(x,i)}[\int_0^t L_X[f](X_s,\xi_s)ds + \sum_{s\le t} \Delta f(X_s,\xi_s)]}{t} \\
&= L_X[f](x,i) + \lim_{t\searrow 0} \frac{E^{(x,i)}[\sum_{s\le t} \Delta f(X_s,\xi_s)]}{t} \\
&= L_X[f](x,i) + \lim_{t\searrow 0} \frac{E^{(x,i)}\left[(f(X_T,\xi_t) - f(X_T,i))\mathbb{1}_{\{N(t)\le 1\}}\right]}{t} \\
&\quad + \lim_{t\searrow 0} \frac{E^{(x,i)}\left[\sum_{s\le t} \Delta f(X_s,\xi_s)\mathbb{1}_{\{N(t)\ge 2\}}\right]}{t} \qquad (3.3)
\end{aligned}$$

where $N(t)$ counts the number of jumps of ξ_t and T is the time of the first jump. Next, it will be shown that $\lim_{t\searrow 0} \frac{1}{t} E^{(x,i)}[\sum_{s\le t} \Delta f(X_t,\xi_t)\mathbb{1}_{\{N(t)\ge 2\}}] = 0$. It was assumed that f is a bounded function, so let $\sup_{(x,i)}|f(x,i)| \le M$. Let $N^*(t)$ be a Poisson process with rate $\lambda = \max_i(-q_{ii})$, recalling that $-q_{ii}$ is the rate of leaving state i. It is clear to see that choosing this maximal rate yields the inequality $E[N(t)] \le E[N^*(t)]$. Now we have that

$$\begin{aligned}
E^{(x,i)}\left[\left|\sum_{s\le t} \Delta f(X_s,\xi_s)\mathbb{1}_{\{N(t)\ge 2\}}\right|\right] &\le E[2MN(t)\mathbb{1}_{\{N(t)\ge 2\}}] \\
&\le E[2MN^*(t)\mathbb{1}_{\{N(t)\ge 2\}}] \\
&= 2M\sum_{k=2}^{\infty} kP(N^*(t) = k) \\
&= 2M\Big(E[N^*(t)] - P(N^*(t) = 1)\Big) \\
&= 2M\lambda t(1 - e^{-\lambda t}).
\end{aligned}$$

Utilizing this result, we see that

$$\lim_{t\searrow 0} \frac{E^{(x,i)}[|\sum_{s\le t} \Delta f(X_s,\xi_s)\mathbb{1}_{\{N(t)\ge 2\}}|]}{t} \le \lim_{t\searrow 0} 2M\lambda(1 - e^{-\lambda t}) = 0$$

and thus

$$\lim_{t\searrow 0} \frac{E^{(x,i)}[\sum_{s\le t} \Delta f(X_s,\xi_s)\mathbb{1}_{\{N(t)\ge 2\}}]}{t} = 0.$$

Implementing this in (3.3) we have

$$
\begin{aligned}
L_{(X,\xi)}[f](x,i) &= L_X[f](x,i) + \lim_{t \searrow 0} \frac{E^{(x,i)}\left[\left(f(X_T,\xi_t) - f(X_T,i)\right)\mathbb{1}_{\{N(t)\le 1\}}\right]}{t} \\
&= L_X[f](x,i) + \lim_{t \searrow 0} \frac{E^{(x,i)}\left[E\left[\left(f(X_T,\xi_t) - f(X_T,i)\right)\mathbb{1}_{\{N(t)\le 1\}} \mid X_t\right]\right]}{t} \\
&= L_X[f](x,i) + \lim_{t \searrow 0} E^{(x,i)}\left[\sum_{\substack{j=1 \\ j\neq i}}^{n} \frac{P_{ij}(t)}{t}\left(f(X_T,j) - f(X_T,i)\right)\mathbb{1}_{\{N(t)\le 1\}}\right] \\
&= L_X[f](x,i) + E^{(x,i)}\left[\sum_{\substack{j=1 \\ j\neq i}}^{n} \lim_{t \searrow 0} \frac{P_{ij}(t)}{t}\left(f(X_T,j) - f(X_T,i)\right)\mathbb{1}_{\{N(t)\le 1\}}\right] \\
&= L_X[f](x,i) + \sum_{\substack{j=1 \\ j\neq i}}^{n} q_{ij}\left(f(x,j) - f(x,i)\right) \\
&= L_X[f](x,i) + \sum_{\substack{j=1 \\ j\neq i}}^{n} q_{ij} f(x,j) - \sum_{j=1}^{n} q_{ij} f(x,i) \\
&= L_X[f](x,i) + \sum_{\substack{j=1 \\ j\neq i}}^{n} q_{ij} f(x,j) + q_{ii} f(x,i) \\
&= L_X[f](x,i) + \sum_{j=1}^{n} q_{ij} f(x,j) \\
&= L_X[f](x,i) + L_\xi[f](x,i)
\end{aligned}
$$

where $P_{ij}(t) = P(\xi_t = j|\xi_0 = i)$, and q_{ij} is defined by $\lim_{t\to 0} \frac{P_{ij}(t)}{t}$. The interchange of limit and expectation is justified since f and $\frac{P_{ij}(t)}{t}$ are bounded and thus Lebesgue's dominated convergence theorem applies. □

From Proposition 1.7 in ([9], p. 162), we have

$$
f(X_t,\xi_t) - f(X_0,\xi_0) - \int_0^t L_{(X,\xi)}[f](X_s,\xi_s)ds
$$

is a martingale with respect to the natural filtration generated by (X_t, ξ_t). Combining Proposition 3.3.1 and Ito's formula (3.2) we get

$$\int_{0^+}^{t} \sigma(\xi_{s^-}) f_x(Z_{s^-}) dW_s + \sum_{s\le t} \Delta f(X_s, \xi_s) - \int_0^t L_\xi[f](X_s, \xi_s) ds$$

is a martingale. On the other hand, by standard properties of Ito-integrals, $\int_0^t \sigma(\xi_s) f_x(X_s, \xi_s) dW_s$ is a martingale, thus

$$M_t^f := \sum_{s\le t} \Delta f(X_s, \xi_s) - \int_0^t L_\xi[f](X_s, \xi_s) ds$$

is a martingale. This produces a nice semimartingale decomposition:

$$\sum_{s\le t} \Delta f(X_s, \xi_s) = \int_0^t L_\xi[f](X_s, \xi_s) ds + M_t^f.$$

As a result, a very elegant and useful version of Ito's formula for the process (X_t, ξ_t) is obtained as follows:

$$\begin{aligned} f(X_t, \xi_t) &= f(X_0, \xi_0) + \int_0^t L_{(X,\xi)}[f](X_s, \xi_s) ds + \int_{0^+}^t \sigma(\xi_{s^-}) f_x(Z_{s^-}) dW_s + M_t^f \qquad (3.4) \\ &= f(X_0, \xi_0) + \int_0^t L_{(X,\xi)}[f](X_s, \xi_s) ds + Martingale. \end{aligned}$$

Another useful result that will be needed later is to apply Ito's Formula to the function defined $F(e^{-\int_0^t r(\xi_s)ds}, Z_t) := e^{-\int_0^t r(\xi_s)ds} f(Z_t)$. Applying Theorem 3.3.1 in the previous manner, we obtain

$$\begin{aligned} e^{-\int_0^t r(\xi_s)ds} f(Z_t) &= f(Z_0) + \int_0^t f(Z_s) d\left(e^{-\int_0^s r(\xi_u)du}\right) + \int_0^t e^{-\int_0^s r(\xi_u)du} L_{(X,\xi)}[f](Z_s) ds \\ &\quad + \int_{0^+}^t e^{-\int_0^s r(\xi_u)du} \sigma(\xi_{s^-}) f_x(Z_{s^-}) dW_s + M_t^F \\ &= f(Z_0) + \int_0^t e^{-\int_0^s r(\xi_u)du} \left(L_{(X,\xi)}[f](Z_s) - r(\xi_s) f(Z_s)\right) ds \\ &\quad + \int_{0^+}^t e^{-\int_0^s r(\xi_u)du} \sigma(\xi_{s^-}) f_x(Z_{s^-}) dW_s + M_t^F \qquad (3.5) \\ &= f(Z_0) + \int_0^t e^{-\int_0^s r(\xi_u)du} \left(L_{(X,\xi)}[f](Z_s) - r(\xi_s) f(Z_s)\right) ds + Martingale. \end{aligned}$$

3.4 American Put and the Dirichlet Problem

We start this section with a conjecture about the nature of the optimal stopping time used in determining the value function of an Amercan put option,

$$V(x,i) = \sup_{\tau} E^{(x,i)}\left[e^{-\int_0^{\tau} r(\xi_s)ds}(K - e^{X_\tau})^+\right].$$

This in turn will allow us to draw a connection to the Dirichlet problem.

Conjecture 3.4.1.

The optimal stopping time of

$$V(x,i) = \sup_{\tau} E^{(x,i)}\left[e^{-\int_0^{\tau} r(\xi_s)ds}(K - e^{X_\tau})^+\right]$$

is of the form

$$\tau = \inf\{t \geq 0 : V(X_t, \xi_t) = (K - e^{X_t})^+\}.$$

Justification. It seems intuitive that the optimal stopping strategy should be to wait for the first time at which $V(x,i) = (K - e^x)^+$, in other words waiting to hit the price at which there is no expected gain for waiting longer. So the stopping time $\tau = \inf\{t \geq 0 : V(X_t, \xi_t) = (K - e^{X_t})^+\}$ would appear to be optimal.

Remark 3.4.1.

This conjecture will be proven to be correct in section 3.6.

Lemma 3.4.1.

Supposing a stopping time of the form given in conjecture 3.4.1 there exists a series of thresholds, $\{b(1), \ldots, b(n)\}$ forming a region $C = \{(x,i) : x > b(i)\}$ referred to as the continuation region and a region $D = \{(x,i) : x \leq b(i)\}$ referred to as the stopping region such that the entry time $\tau_D = \inf\{t \geq 0 : (X_t, \xi_t) \in D\} = \inf\{t \geq 0 : X_t \leq b(\xi_t)\}$ is the optimal stopping time:

$$V(x,i) = E^{(x,i)}\left[e^{-\int_0^{\tau_D} r(\xi_s)ds}(K - e^{X_{\tau_D}})^+\right].$$

Without loss of generality, we will assume $b(n) < \cdots < b(1)$.

Proof. The fact that $\{b(i)\}$ can be chosen in descending order is a simple consequence of relabeling (*i.e.*, renaming) the original Markov states $\{1, 2, \ldots, n\}$ since doing so does not change the dynamics of state to state transitions. To prove this lemma, we follow [13] and show that $V(x, i) - \phi(x)$, where $\phi(x) = (K - e^x)^+$, is an increasing function on $(-\infty, \ln(K))$. From this we conclude that there are unique thresholds depending on the Markov state below which $V(x, i) = \phi(x)$ and the lemma is proven. Let $x, x + \delta \in (-\infty, \ln(K))$ and let τ be the optimal stopping time when $X_0 = x$.

$$
\begin{aligned}
V(j, x + \delta) &\geq E^{(x+\delta,j)}\left[e^{-\int_0^\tau r(\xi_s)ds}\phi(X_\tau)\right] \\
&= E^{(x,j)}\left[e^{-\int_0^\tau r(\xi_s)ds}\left(K - e^{X_\tau+\delta}\right)^+\right] \\
&= E^{(x,j)}\left[e^{-\int_0^\tau r(\xi_s)ds}\left(K - e^{X_\tau} - (e^\delta - 1)e^{X_\tau}\right)^+\right] \\
&\geq E^{(x,j)}\left[e^{-\int_0^\tau r(\xi_s)ds}\left((K - e^{X_\tau})^+ - (e^\delta - 1)e^{X_\tau}\right)\right] \\
&\geq E^{(x,j)}\left[e^{-\int_0^\tau r(\xi_s)ds}(K - e^{X_\tau})^+\right] - (e^\delta - 1)E^{(x,j)}\left[e^{-\int_0^\tau r(\xi_s)ds}e^{X_\tau}\right] \\
&= V(x, j) - (e^\delta - 1)E^{(x,j)}\left[e^{-\int_0^\tau r(\xi_s)ds}e^{X_\tau}\right] \\
&\geq V(x, j) - (e^\delta - 1)\liminf_{t\to\infty} E^{(x,j)}\left[e^{-\int_0^{t\wedge\tau} r(\xi_s)ds}e^{X_{t\wedge\tau}}\right] && (3.6) \\
&= V(x, j) - (e^\delta - 1)e^x && (3.7) \\
&= V(x, j) + \phi(x + \delta) - \phi(x).
\end{aligned}
$$

In (3.6) Fatou's lemma is applied. Finally (3.7) is valid since the discounted stock process $e^{-\int_0^t r(\xi_s)ds}e^{X_t}$ is a martingale which implies that the stopped process $e^{-\int_0^{t\wedge\tau} r(\xi_s)ds}e^{X_{t\wedge\tau}}$ is also a martingale by theorem 2.2.2. Notice that no assumptions are made about the finiteness of τ. □

Next, the techniques from [26] will be used to show that solving the entry time problem,

$$V(x, i) = E^{(x,i)}\left[e^{-\int_0^{\tau_D} r(\xi_s)ds}\left(K - e^{X_{\tau_D}}\right)^+\right]$$

is equivalent to solving the Dirichlet problem

$$
\begin{aligned}
L_{(X,\xi)}[V](x,i) &= r(i)V(x,i) \quad in \quad C \qquad (3.8)\\
V(x,i) &= (K-e^x)^+ \quad in \quad D.
\end{aligned}
$$

$C = \{(x,i) : V(x,i) > (K-e^x)^+\}$ and is referred to as the continuation region and D is the its complement in $\mathbb{R} \times \{1, 2, \cdots, n\}$, and is referred to as the stopping region. The entrance time into D is defined by $\tau_D = \inf\{t \geq 0 : (X_t, \xi_t) \in D\}$. Here, no assumptions are made as to the finiteness of τ_D.

To verify (3.8), we first define a "killed" process $\tilde{X}_t$. Let T be a killing time, with rate of killing defined as follows

$$
r(\xi_t) = \lim_{\delta \searrow 0} \frac{P(t < T \leq t+\delta \mid \{\xi_u : 0 \leq u < t\}, T > t)}{\delta} = \frac{f(t \mid \{\xi_u : 0 \leq u < t\})}{1 - F(t \mid \{\xi_u : 0 \leq u < t\})}
$$

with $F(t \mid \{\xi_u : 0 \leq u < t\}) = P(T \leq t \mid \{\xi_u : 0 \leq u < t\})$ and $f(t \mid \{\xi_u : 0 \leq u < t\})$ its derivative. Now if $r(\xi_t)$ is integrated from 0 to t, we get

$$
\begin{aligned}
\int_0^t r(\xi_s) ds &= \int_0^t \frac{f(s \mid \{\xi_u : 0 \leq u < s\})}{1 - F(s \mid \{\xi_u : 0 \leq u < s\})} ds \\
&= -\ln\Big(1 - F(s \mid \{\xi_u : 0 \leq u < s\})\Big)
\end{aligned}
$$

giving,

$$
\begin{aligned}
P(T \leq t \mid \{\xi_u : 0 \leq u < t\}) &= 1 - e^{-\int_0^t r(\xi_s)ds} \\
P(t \leq T \mid \{\xi_u : 0 \leq u < t\}) &= e^{-\int_0^t r(\xi_s)ds}. \qquad (3.9)
\end{aligned}
$$

Define the killed process

$$
\tilde{X}_t := \begin{cases} X_t & t < T \\ \Delta & t \geq T \end{cases},
$$

where Δ is referred to as the cemetery (hence the terminology of killed process and killing time T). Define a function $\phi(\Delta) = 0$ and observe that from (3.9) we have

$$\begin{aligned} E[\phi(\tilde{X}_t)] &= E[\phi(X_t)\mathbb{1}_{t<T}] \\ &= E\left[E[\phi(X_t)\mathbb{1}_{t<T} \mid \mathcal{F}_t]\right] \\ &= E\left[\phi(X_t)E[\mathbb{1}_{t<T} \mid \mathcal{F}_t]\right] \\ &= E\left[\phi(X_t)e^{-\int_0^t r(\xi_s)ds}\right]. \end{aligned}$$

The above provides a convenient way to write the value function

$$\begin{aligned} V(x,i) &= E\left[\phi(\tilde{X}_{\tau_D})\right] \\ \phi(x) &= (K - e^x)^+. \end{aligned}$$

Next, the strong Markov property, theorem 2.4.1, will be used to show that $L_{(X,\xi)}[F](x,i) = r(i)F(x,i)$. In addition the properties of the shift operator presented in section 2.4 will be used in the following proof.

Choose $(x,i) \in C$ and a bounded open set $U \subset C$ and define $\sigma = \inf\{t : (X_t, \xi_t) \notin U\}$ when (X_t, ξ_t) starts at (x,i). Notice that $\sigma \le \tau_D$.

$$\begin{aligned} E^{(x,i)}[V(\tilde{X}_\sigma, \xi_\sigma)] &= E^{(x,i)}\left[E^{(\tilde{X}_\sigma,\xi_\sigma)}[\phi(\tilde{X}_{\tau_D})]\right] \\ &= E^{(x,i)}\left[E^{(x,i)}[\phi(\tilde{X}_{\tau_D}) \circ \theta_\sigma \mid \mathcal{F}_\sigma]\right] \\ &= E^{(x,i)}\left[E^{(x,i)}[\phi(\tilde{X}_{\sigma+\tau_D\circ\theta_\sigma}) \mid \mathcal{F}_\sigma]\right] \\ &= E^{(x,i)}\left[E^{(x,i)}[\phi(\tilde{X}_{\tau_D}) \mid \mathcal{F}_\sigma]\right] \\ &= E^{(x,i)}[\phi(\tilde{X}_{\tau_D})] \\ &= V(x,i). \end{aligned}$$

Thus the characteristic operator is identically zero:

$$\lim_{U \searrow (x,i)} \frac{E^{(x,i)}[V(\tilde{X}_\sigma, \xi_\sigma)] - V(x,i)}{E[\sigma]} = 0.$$

Since the characteristic and infinitesimal operator coincide on the domain of the infinitesimal operator, we have

$$L_{(\tilde{X},\xi)}[V](x,i) = 0 \quad for \quad (x,i) \in C.$$

The infinitesimal generator of the killed process is given by

$$L_{(\tilde{X},\xi)}[V](x,i) = L_{(X,\xi)}[V](x,i) - r(i)V(x,i).$$

This is seen as follows.

$$\begin{aligned}
L_{(\tilde{X},\xi)}[V](x,i) &= \lim_{t\searrow 0} \frac{E^{(x,i)}[V(\tilde{X}_t,\xi_t)] - V(x,i)}{t} \\
&= \lim_{t\searrow 0} \frac{E^{(x,i)}[V(X_t,\xi_t)] - V(x,i)}{t} + \frac{E^{(x,i)}[V(\tilde{X}_t,\xi_t)] - E^{(x,i)}[V(X_t,\xi_t)]}{t} \\
&= L_{(X,\xi)}[V](x,i) + \lim_{t\searrow 0} \frac{E^{(x,i)}[e^{-\int_0^t r(\xi_s)ds} V(X_t,\xi_t)] - E^{(x,i)}[V(X_t,\xi_t)]}{t} \\
&= L_{(X,\xi)}[V](x,i) + \lim_{t\searrow 0} E^{(x,i)} \left[\frac{(e^{-\int_0^t r(\xi_s)ds} - 1)}{t} V(X_t,\xi_t) \right] \\
&= L_{(X,\xi)} - r(i)V(x,i).
\end{aligned}$$

In the last step the interchange of limit and expectation is justified by Lebesgue's dominated convergence theorem since $V(x,i)$ is bounded. Lastly the definition of derivative is applied to $e^{-\int_0^t r(\xi_s)ds}$.

Thus we have

$$L_{(X,\xi)}[V](x,i) = r(i)V(x,i) \quad for \quad (x,i) \in C$$

as claimed.

3.5 Logarithmic form of the Black-Scholes Equations

The Dirichlet problem (3.8) produces the ODE's below

$$\left(r(i) - \frac{1}{2}\sigma^2(i)\right) f_x(x,i) + \frac{1}{2}\sigma^2(i) f_{xx}(x,i) + \sum_{j=1}^{n} q_{ij} f(x,j) = r(i)f(x,i) \quad x > b_i \tag{3.10}$$

$$f(x,i) = \phi(x) \quad x \le b_i.$$

Remark 3.5.1.

It is interesting to note that we have obtained in 3.10 a kind of Black-Scholes equation that is easier for analysis than the traditional Black-Scholes equation. Namely, by taking the logarithmic stock process $X_t = ln(S_t)$ our generator lacks multiplication by x in the drift coefficient and multiplication by x^2 in the diffusion coefficient which are present in the standard Black-Scholes equations. Thus we have a system of ODEs with constant coefficients which is easier to solve and analyze. The non-logarithmic form is used in [11], and it is remarked there that a closed-form solution is possible only in the case of a two state Markov chain. Using the logarithmic form, the ODEs are greatly simplified and, more importantly, a solution for n states can be found which will be shown in the sequel.

In solving the above system of ODEs we will start by looking at the system when $x > b_1$ to have a system with n unknown functions which can be written in matrix form

$$\mathcal{S}\boldsymbol{f}_x + \frac{1}{2}\Sigma\boldsymbol{f}_{xx} + (Q - R)\boldsymbol{f} = 0, \tag{3.11}$$

where Σ and R are the diagonal matrices whose i-th diagonal elements are $\sigma^2(i)$ and $r(i)$ respectively and Q is the infinitesimal generating matrix of the Markov chain. $\mathcal{S} = R - \frac{1}{2}\Sigma$ and $\boldsymbol{f}$ is the vector whose i-th element is $f(x,i)$. In the standard way, we seek a solution of the form $f(x,i) = g(i)e^{-\lambda x}$ leading to

$$(Q - R)\boldsymbol{g} - \lambda\mathcal{S}\boldsymbol{g} + \frac{1}{2}\lambda^2\Sigma\boldsymbol{g} = 0$$

with g being the vector whose i-th element is $g(i)$. This method was presented in ([13], p. 2065) for solving the "quadratic eigenvalue" problem. Multiplying the above equation on the left by $2\Sigma^{-1}$ and reformulating as a system of equations yields,

$$\begin{cases} \lambda g & = h \\ \lambda h & = 2\Sigma^{-1}\mathcal{S}h - 2\Sigma^{-1}(Q-R)g, \end{cases}$$

which can be written as a the standard linear eigenvalue problem

$$\begin{pmatrix} 0 & I \\ -2\Sigma^{-1}(Q-R) & 2\Sigma^{-1}R - I \end{pmatrix} \begin{pmatrix} g \\ h \end{pmatrix} = \lambda \begin{pmatrix} g \\ h \end{pmatrix}.$$

As stated in [13], there are exactly n eigenvalues with a positive real part and n with a negative real part. For $x > b_1$ we will only consider the n values of λ with positive real part since we do not want the solution to (3.11), sought in the form $f(x,i) = g(i)e^{-\lambda x}$, to grow unbounded. If g_i and λ_i for $i = 1, \ldots, n$ all solve the above eigenvalue problem, we arrive at part of the solution to (3.10).

$$f(x,j) = \sum_{i=1}^{n} \omega_i g_i(j) e^{-\lambda_i x} \quad for \quad x > b_1 \quad j = 1, \ldots, n.$$

In particular, we have an entire solution for $f(x,1)$:

$$f(x,1) = \begin{cases} \sum_{i=1}^{n} \omega_i g_i(1) e^{-\lambda_i x} & ; x > b_1 \\ K - e^x & ; x \leq b_1. \end{cases}$$

Now the region $b_2 < x \leq b_1$ will be considered. In this region, $f(x,1) = K - e^x$, thus the size of the ODE system reduces by one. Let Q_1 denote the matrix Q with the first row and column removed and let $\tilde{Q}_1$ be matrix (vector in this case) composed of the first column of Q with the first row removed. Let $\mathcal{S}_1$, R_1, and Σ_1 be defined as the corresponding matrix with the first row and column removed and let g_1 be the vector $\boldsymbol{g}$ with the first element removed. We arrive at the following matrix form of the $n-1$ dimensional system of ODE's

$$(Q_1 - R_1)\boldsymbol{g}_1 - \lambda \mathcal{S}_1 \boldsymbol{g}_1 + \frac{1}{2}\lambda^2 \Sigma \boldsymbol{g}_1 + \tilde{Q}_1(K - e^x) = 0 \quad for \quad b_2 < x \le b_1. \tag{3.12}$$

A particular solution to the above system of ODEs is of the form $B_1 + C_1 e^x$ where B_1 and C_1 are $n-1$ dimensional vectors. The values of B_1 and C_1 are completely determined by the ODEs. Let $B_1(j)$ and $C_1(j)$ represent the j-th component of the corresponding vector. The solution to the homogeneous equation,

$$(Q_1 - R_1)\boldsymbol{g}_1 - \lambda \mathcal{S}_1 \boldsymbol{g}_1 + \frac{1}{2}\lambda^2 \Sigma \boldsymbol{g}_1 = 0 \quad for \quad b_2 < x \le b_1,$$

is solved precisely like (3.11). Since the region $b_2 < x \le b_1$ is bounded, we will consider all $2(n-1)$ eigenvalues and eigenvectors of the homogeneous ODE. We arrive at another part of the solution to (3.10) and a complete solution for $f(x, 2)$.

$$f(x, j) = \sum_{i=1}^{n} \omega_i^1 g_i^1(j) e^{-\lambda_i^1 x} + B_1(j-1) + C_1(j-1)e^x \quad for \quad b_2 < x \le b_1 \quad j = 2, \dots, n$$

$$f(x, 2) = \begin{cases} \sum_{i=1}^{n} \omega_i g_i(2) e^{-\lambda_i x} & ; b_1 < x \\ \sum_{i=1}^{2(n-1)} \omega_i^1 g_i^1(2) e^{-\lambda_i^1 x} + B_1(1) + C_1(1)e^x & ; b_2 < x \le b_1 \\ K - e^x & ; x \le b_2 \end{cases} .$$

Continuing on in the same manner we arrive at the region $b_{k+1} < x \le b_k$. In this region, $f(x, j) = K - e^x$ for $j = 1, \dots, k$, thus the size of the ODE system is reduced by k. Let Q_k denote the matrix Q with the first k rows and columns removed and let $\tilde{Q}_k$ be matrix composed of the first k columns of Q with the first k rows removed. Let $\boldsymbol{\Phi}_k$ be vector of size k where each component is $K - e^x$. Let $\mathcal{S}_k$, R_k, and Σ_k be defined as the corresponding matrix with the first k rows and columns removed and let $\boldsymbol{g}_k$ be the vector $\boldsymbol{g}$ with the first k elements removed. We arrive at the following matrix form of the $n-k$ dimensional system of ODE's,

$$(Q_k - R_k)\boldsymbol{g}_k - \lambda \mathcal{S}_k \boldsymbol{g}_k + \frac{1}{2}\lambda^2 \Sigma \boldsymbol{g}_k + \tilde{Q}_k \boldsymbol{\Phi}_k = 0 \quad for \quad b_{k+1} < x \le b_k$$

solved the same way as (3.12). We arrive at another part of the solution to (3.10) and a complete solution for $f(x, k+1)$.

$$f(x, j) = \sum_{i=1}^{n} \omega_i^k g_i^k(j) e^{-\lambda_i^k x} + B_k(j-k) + C_k(j-k)e^x \quad for \quad b_{k+1} < x \leq b_k \quad j = k+1, \ldots, n$$

$$f(x, k+1) = \begin{cases} \sum_{i=1}^{n} \omega_i g_i(k+1) e^{-\lambda_i x} & ; b_1 < x \\ \sum_{i=1}^{2(n-1)} \omega_i^1 g_i^1(k+1) e^{-\lambda_i^1 x} + B_1(k) + C_1(k)e^x & ; b_2 < x \leq b_1 \\ \vdots & ; \vdots \\ \sum_{i=1}^{2(n-k)} \omega_i^k g_i^k(k+1) e^{-\lambda_i^k x} + B_k(1) + C_k(1)e^x & ; b_{k+1} < x \leq b_k \\ K - e^x & ; x \leq b_{k+1}. \end{cases} \qquad (3.13)$$

This procedure is continued up to the point when $k = n - 1$. Here, the system of ODEs becomes one ODE and the last function $f(x, n)$ is found in its entirety.

For the first function $f(x, 1)$ there are n unknown weights ω_i. The next function $f(x, 2)$ has $2(n-1)$ unknown weights, until the last function $f(x, n)$ contributes only 2 unknown weights. There are also n unknown boundary values b_i. Overall there are $2(1 + \cdots + n) = n(n+1)$ unknown parameters that need to be determined. Here, we assume that the function is C^1 everywhere. At this point, this assumption may seem restrictive, however, it will later be proven that the solution derived from this smoothness assumption is indeed the optimal solution. Assuming that $f(x, 1)$ is continuous and differentiable at b_1 will result in 2 conditions for $f(x, 1)$, 4 for $f(x, 2)$ with its two boundaries, and finally $2n$ for $f(x, n)$ and its n boundaries. In total, there are $n(n+1)$ conditions to be satisfied. We see that the $n(n+1)$ unknown parameters, including all weights and the n unknown boundaries, are completely determined by imposing the smoothness assumption.

3.6 Optimality of the Solution

Theorem 3.6.1 (Optimality).

Suppose that thresholds $b_n < \cdots < b_1 < \ln(K)$ *have been found such that the unique solution to*

$$\left(r(i) - \frac{1}{2}\sigma^2(i)\right) f_x(x,i) + \frac{1}{2}\sigma^2(i) f_{xx}(x,i) + \sum_{j=1}^{n} q_{ij} f(x,j) = r(i) f(x,i) \quad x > b_i \tag{3.14}$$

$$f(x,i) = \phi(x) \quad x \le b_i$$

is C^1 *on its domain and bounded on C. The solution* $f(x,i)$ *and the stopping time* $\tau_D = \{t : X_t \le b(\xi_t)\}$ *correspond to the value function*

$$V(x,i) = \sup_{\tau} E^{(x,i)} \left[e^{-\int_0^{\tau} r(\xi_s) ds} (K - e^{X_\tau})^+ \right]$$

and its optimal stopping time, i.e.,

$$V(x,i) = f(x,i) = E^{(x,i)} \left[e^{-\int_0^{\tau_D} r(\xi_s) ds} (K - e^{X_{\tau_D}})^+ \right].$$

The proof will be established through several steps.

We will start by looking at the process $e^{-\int_0^t r(\xi_s)ds} f(X_t, \xi_t)$. Since by definition $f(x,i)$ is twice differentiable everywhere except when $x = b_i$ for $i = 1, 2, \cdots, n$ where X_t spends zero time, we can apply the generalized Ito-formula (3.5) to get

$$e^{-\int_0^t r(\xi_s)ds} f(X_t, \xi_t) = f(X_0, \xi_0) + \int_0^t e^{-\int_0^s r(\xi_u)du} \left(L_{(X,\xi)}[f](X_s, \xi_s) - r(\xi_s) f(X_s, \xi_s) \right) ds \tag{3.15}$$
$$+ \mathit{Martingale}.$$

To show optimality we need to prove an auxiliary.

Proposition 3.6.1.

The inequality for the function defined below is true for all (x, i)

$$\begin{aligned}\Phi(x,i) &:= L_{(X,\xi)}[f](x,i) - r(i)f(x,i) \\ &= \left(r(i) - \frac{1}{2}\sigma^2(i)\right) f_x(x,i) + \frac{1}{2}\sigma^2(i) f_{xx}(x,i) - r(i)f(x,i) + \sum_{j=1}^{n} q_{ij} f(x,j) \\ &\leq 0. \end{aligned} \tag{3.16}$$

Proof. We will start by defining the two regions mentioned in section 3.4,

$$C = \{(x,i) : x > b(i)\}$$

$$D = \{(x,i) : x \leq b(i)\}$$

and notice that $\tau_D = \inf\{t \geq 0 : (X_t, \xi_t) \in D\} = \inf\{t \geq 0 : X_t \leq b(\xi_t)\}$.

In region D we will look at properties of the *left continuous* process $e^{-\int_0^{\tau_C\wedge t-} r(\xi_s)ds} f(X_{\tau_C\wedge t-}, \xi_{\tau_C\wedge t-})$ where $\tau_C = \inf\{t \geq 0 : (X_t, \xi_t) \in C\}$ and $Z_{\tau\wedge t-} := \lim_{s\nearrow\tau\wedge t} Z_s$. Now it will be shown that $e^{-\int_0^{\tau_C\wedge t-} r(\xi_s)ds} f(X_{\tau_C\wedge t-}, \xi_{\tau_C\wedge t-})$ is a supermartingale. Let $s \leq t$ and $(x,i) \in D$.

$$\begin{aligned} E^{(x,i)}[e^{-\int_0^{\tau_C\wedge t-} r(\xi_s)ds} f(X_{\tau_C\wedge t-}, \xi_{\tau_C\wedge t-}) \mid \mathcal{F}_{\tau_C\wedge s}] &= E^{(x,i)}[e^{-\int_0^{\tau_C\wedge t-} r(\xi_s)ds} \phi(X_{\tau_C\wedge t-}) \mid \mathcal{F}_{\tau_C\wedge s}] \\ &= E^{(x,i)}[e^{-\int_0^{\tau_C\wedge t-} r(\xi_s)ds} (K - e^{X_{\tau_C\wedge t}}) \mid \mathcal{F}_{\tau_C\wedge s}] \\ &= K e^{-\int_0^{\tau_C\wedge s-} r(\xi_u)du} E^{(x,i)}[e^{-\int_{\tau_C\wedge s-}^{\tau_C\wedge t-} r(\xi_s)ds} \mid \mathcal{F}_{\tau_C\wedge s}] \\ &\qquad - E^{(x,i)}[e^{-\int_0^{\tau_C\wedge t-} r(\xi_s)ds} e^{X_{\tau_C\wedge t}}) \mid \mathcal{F}_{\tau_C\wedge s}] \\ &\leq K e^{-\int_0^{\tau_C\wedge s-} r(\xi_u)du} - e^{-\int_0^{\tau_C\wedge s-} r(\xi_u)du} e^{X_{\tau_C\wedge s}} \\ &= e^{-\int_0^{\tau_C\wedge s-} r(\xi_u)du} \phi(X_{\tau_C\wedge s-}) \\ &= e^{-\int_0^{\tau_C\wedge s-} r(\xi_u)du} f(X_{\tau_C\wedge s-}, \xi_{\tau_C\wedge s-}). \end{aligned}$$

In the above, the following facts are used: $e^{-\int_0^{t-} r(\xi_s)ds} e^{X_t}$ is a martingale which by (2.2.2) implies that the stopped process $e^{-\int_0^{\tau_C\wedge t-} r(\xi_s)ds} e^{X_{\tau_C\wedge t}}$ is a martingale, X_t is continuous, and $X_{\tau_D} \leq b_1 < \ln(K)$ implying that $\phi(X_{\tau_D}) = K - e^{X_{\tau_D}}$.

Then, as desired, $e^{-\int_0^{\tau_C \wedge t} r(\xi_s)ds} f(X_{\tau_C \wedge t-}, \xi_{\tau_C \wedge t-})$ is a supermartingale. From (3.15), we see that

$$\int_0^{\tau_C \wedge t-} e^{-\int_0^s r(\xi_u)du} \Big(L_{(X,\xi)}[f](X_s, \xi_s) - r(\xi_s) f(X_s, \xi_s) \Big) ds \tag{3.17}$$

is also a supermartingale. Next note that by the bounds of this integral, (X_s, ξ_s) is not allowed to pass into C and thus $f(X_s, \xi_s) = \phi(X_s)$. Now we look at the function $\Phi(x, i)$ in the region D,

$$\begin{aligned}
\Phi(x, i) &= L_{(X,\xi)}[f](x, i) - r(i) f(x, i) \\
&= \sum_{j=1}^{n} q_{ij} f(x, j) - r(i)K \\
&= \sum_{\substack{j=1 \\ j \neq i}}^{n} q_{ij} \Big(f(x, j) - f(x, i) \Big) - r(i)K \\
&= \sum_{j=i+1}^{n} q_{ij} \Big(f(x, j) - \phi(x) \Big) - r(i)K.
\end{aligned} \tag{3.18}$$

Define the bounded process

$$A_s := e^{-\int_0^s r(\xi_u)du} \left(\sum_{j=i+1}^{n} q_{ij} \Big(f(X_s, j) - \phi(X_s) \Big) - r(\xi_s)K \right)$$

and observe that by (3.17) and (3.18) $\int_0^{\tau_C \wedge t-} A_s ds$ is a supermartingale. Next we will show that $A_s \mathbb{1}_{s \leq \tau_C} \leq 0$. Note that in the upper integral limit the left limit is ignored, since it will not affect the value of the integral.

$$\begin{aligned}
\int_0^{\tau_C \wedge t} A_s ds &\geq E^{(x,i)} \left[\int_0^{\tau_C \wedge (t+u)} A_s ds \mid \mathcal{F}_{\tau_C \wedge t} \right] \\
&= E^{(x,i)} \left[\int_0^{\tau_C \wedge t} A_s ds + \int_{\tau_C \wedge t}^{\tau_C \wedge (t+u)} A_s ds \mid \mathcal{F}_{\tau_C \wedge t} \right] \\
&= \int_0^{\tau_C \wedge t} A_s ds + E^{(x,i)} \left[\int_{\tau_C \wedge t}^{\tau_C \wedge (t+u)} A_s ds \mid \mathcal{F}_{\tau_C \wedge t} \right]
\end{aligned}$$

which upon cancellation and dividing by u gives

$$\begin{aligned}
0 &\geq \frac{1}{u}E^{(x,i)}\left[\int_{\tau_C\wedge t}^{\tau_C\wedge(t+u)} A_s ds \mid \mathcal{F}_{\tau_C\wedge t}\right] \\
&= \frac{1}{u}E^{(x,i)}\left[\int_t^{t+u} A_s ds \mathbb{1}_{\{t+u\leq\tau_C\}} + \int_t^{\tau_C\wedge(t+u)} A_s ds \mathbb{1}_{\{t<\tau_C\leq t+u\}} \mid \mathcal{F}_{\tau_C\wedge t}\right] \\
&= \frac{1}{u}E^{(x,i)}\left[\int_t^{t+u} A_s ds \mathbb{1}_{\{t+u\leq\tau_C\}} + \int_t^{\tau_C\wedge(t+u)} A_s ds \mathbb{1}_{\{t<\tau_C\leq t+u\}} \mid \mathcal{F}_{\tau_C\wedge t}\right] \\
0 &= \lim_{u\searrow 0} E^{(x,i)}\left[\frac{\int_t^{t+u} A_s ds}{u}\mathbb{1}_{\{t+u\leq\tau_C\}} + \frac{\int_t^{\tau_C\wedge(t+u)} A_s ds}{u}\mathbb{1}_{\{t<\tau_C\leq t+u\}} \mid \mathcal{F}_{\tau_C\wedge t}\right] \qquad (3.19) \\
&= E^{(x,i)}\left[A_t \mathbb{1}_{\{t\leq\tau_C\}} \mid \mathcal{F}_{\tau_C\wedge t}\right] \\
&= A_t \mathbb{1}_{\{t\leq\tau_C\}}.
\end{aligned}$$

In (3.19), the limit was interchanged with the expectation since A_s is bounded and thus Lebesgue's dominated convergence theorem applies. Then by elementary calculus the first term becomes $A_t\mathbb{1}_{\{t\leq\tau_C\}}$ by observing that A_s is continuous at t *a.s.* since ξ_s transitions precisely at t with probability zero. It remains to show that

$$\lim_{u\searrow 0} \frac{\int_t^{\tau_C\wedge(t+u)} A_s ds}{u}\mathbb{1}_{\{t<\tau_C\leq t+u\}} = 0$$

for every ω. To see this, notice that $t < \tau_C(\omega)$ and thus for all u sufficiently close to 0 we have $\mathbb{1}_{\{t<\tau_C\leq t+u\}} = 0$ and the proof is finished.

Finally,

$$A_s\mathbb{1}_{\{s\leq\tau_C\}} = e^{-\int_0^s r(\xi_u)du}\left(\sum_{j=i+1}^{n} q_{ij}\Big(f(X_s, j) - \phi(X_s)\Big) - r(\xi_s)K\right)\mathbb{1}_{\{s\leq\tau_C\}} \leq 0$$

which implies that

$$\Phi(x,i) = \sum_{j=i+1}^{n} q_{ij}\Big(f(X_s, j) - \phi(X_s)\Big) - r(\xi_s)K \leq 0 \quad for \quad X_s \in D$$

or $\Phi(x,i) \leq 0$ for $(x,i) \in D$. Additionally for $x \in C$, we have $\Phi(x,i) = 0$ by construction, see (3.10) and (3.16). Thus $\Phi(x,i) \leq 0$ for any (x,i) as desired and the proof is finished. □

We will need to prove one more proposition before optimality can be proven.

Proposition 3.6.2.

$f(x,i) \geq \phi(x)$ for all (x,i).

Proof. From (3.14) we see that

$$f_{xx}(x,i) = \frac{2}{\sigma^2(i)}\left(r(i)\Big(f(x,i) - f_x(x,i)\Big) - \sum_{j=1}^{n} q_{ij} f(x,j)\right) + f_x(x,i).$$

Now, using the C^1 property of f at b_i and the fact that $f(b_i,i) = \phi(b_i)$, we get

$$\begin{aligned} f_{xx}(b_i+,i) &= \frac{2}{\sigma^2(i)}\left(r(i)K - \sum_{j=i+1}^{n} q_{ij}\Big(f(b_i,j) - \phi(b_i)\Big)\right) - e^{b_i} \\ &= \frac{2}{\sigma^2(i)}\left(-\Phi(b_i,i)\right) - e^{b_i}. \end{aligned}$$

Since it was shown that $\Phi(x,i) \leq 0$ for (x,i), we have that $f_{xx}(b_i+,i) \geq -e^{b_i}$. Observe that

$$\begin{aligned} f(b_i,i) &= \phi(b_i) = K - e^{b_i} \\ f_x(b_i,i) &= \phi'(b_i) = -e^{b_i} \\ f_{xx}(b_i+,i) &\geq \phi''(b_i) = -e^{b_i}. \end{aligned}$$

Thus it is apparent that there exists a $\delta_i > 0$ such that

$$x \leq b_i + \delta_i \implies f(x,i) \geq \phi(x). \tag{3.20}$$

Next it will be shown that for any two points $x, x+\delta \in (-\infty, \ln(K))$ such that $0 < \delta \leq \min\{\delta_i\}_{i=1}^{n}$ then $f(x+\delta,i) - \phi(x+\delta) \geq f(x,i) - \phi(x)$ thus showing that $f(x,i) - \phi(x)$ is an increasing function on $(-\infty, \ln(K))$. This will of course imply that $f(x,i) \geq \phi(x)$ on $(-\infty, \ln(K))$.

In the following argument an outline of what is done in [13] will be followed. Two points $x, x+\delta \in (-\infty, \ln(K))$ are chosen and τ_D is taken to be the stopping time when (X_t, ξ_t) starts at (x, i). Before proceeding, notice that if X_t is started at $x + \delta$, then $X_{\tau_D} \le b(\xi_{\tau_D}) + \delta$ which by (3.20) implies that $f(X_{\tau_D}, \xi_{\tau_D}) \ge \phi(X_{\tau_D})$. Using this we get,

$$
\begin{aligned}
f(j, x+\delta) &= E^{(x+\delta,j)}[e^{-\int_0^{t\wedge\tau_D} r(\xi_s)ds} f(X_{t\wedge\tau_D}, \xi_{t\wedge\tau_D})] \\
&= E^{(x+\delta,j)}\left[e^{-\int_0^{\tau_D} r(\xi_s)ds} f(X_{\tau_D}, \xi_{\tau_D})\right] &(3.21)\\
&\ge E^{(x+\delta,j)}\left[e^{-\int_0^{\tau_D} r(\xi_s)ds} \phi(X_{\tau_D})\right] &(3.22)\\
&= E^{(x,j)}\left[e^{-\int_0^{\tau_D} r(\xi_s)ds} \left(K - e^{X_{\tau_D}+\delta}\right)^+\right] \\
&= E^{(x,j)}\left[e^{-\int_0^{\tau_D} r(\xi_s)ds} \left(K - e^{X_{\tau_D}} - (e^\delta - 1)e^{X_{\tau_D}}\right)^+\right] \\
&\ge E^{(x,j)}\left[e^{-\int_0^{\tau_D} r(\xi_s)ds} \left((K - e^{X_{\tau_D}})^+ - (e^\delta - 1)e^{X_{\tau_D}}\right)\right] \\
&\ge E^{(x,j)}\left[e^{-\int_0^{\tau_D} r(\xi_s)ds} (K - e^{X_{\tau_D}})^+\right] - (e^\delta - 1)E^{(x,j)}\left[e^{-\int_0^{\tau_D} r(\xi_s)ds} e^{X_{\tau_D}}\right] \\
&\ge E^{(x,j)}\left[e^{-\int_0^{\tau_D} r(\xi_s)ds} f(X_{\tau_D})\right] - (e^\delta - 1)E^{(x,j)}\left[e^{-\int_0^{\tau_D} r(\xi_s)ds} e^{X_{\tau_D}}\right] &(3.23)\\
&= f(x, j) - (e^\delta - 1)E^{(x,j)}\left[e^{-\int_0^{\tau_D} r(\xi_s)ds} e^{X_{\tau_D}}\right] &(3.24)\\
&\ge f(x, j) - (e^\delta - 1)\liminf_{t\to\infty} E^{(x,j)}\left[e^{-\int_0^{t\wedge\tau_D} r(\xi_s)ds} e^{X_{t\wedge\tau_D}}\right] &(3.25)\\
&= f(x, j) - (e^\delta - 1)e^x &(3.26)\\
&= f(x, j) + \phi(x+\delta) - \phi(x).
\end{aligned}
$$

In (3.21) and (3.24) the optional sampling Theorem 2.2.4 is applied after observing that the quantity in the expected value is bounded and thus uniformly integrable. (3.22) is explained in the previous paragraph. (3.23) is valid because $(X_{\tau_D}, \xi_{\tau_D}) \in C$ and $f(x, j) = \phi(x)$ when $(x, j) \in D$. In (3.25) Fatou's lemma is applied. Finally (3.26) is valid since the discounted stock process $e^{-\int_0^t r(\xi_s)ds} e^{X_t}$ is a martingale which implies that the stopped process $e^{-\int_0^{t\wedge\tau_D} r(\xi_s)ds} e^{X_{t\wedge\tau_D}}$ is also a martingale by Theorem 2.2.2. Notice that no assumptions are made about the finiteness of τ_D.

From the above we have that $f(x,i) \geq \phi(x)$ on $(-\infty, \ln(K))$. To address the region $(\ln(K), \infty)$ we will observe that

$$e^{-\int_0^{t\wedge\tau_D} r(\xi_s)ds} f(X_{t\wedge\tau_D}, \xi_{t\wedge\tau_D})$$

is a martingale. This is evident from combining (3.16) with the fact that $(X_{t\wedge\tau_D}, \xi_{t\wedge\tau_D})$ is never allowed to pass into the region D, and thus $\Phi(X_{t\wedge\tau_D}, \xi_{t\wedge\tau_D}) = 0$, see (3.10). Also note that $e^{-\int_0^{t\wedge\tau_D} r(\xi_s)ds} f(X_{t\wedge\tau_D}, \xi_{t\wedge\tau_D})$ is bounded and is thus a uniformly integrable martingale. This allows the use of the optional sampling theorem 2.2.4 in the following,

$$\begin{aligned} f(x,i) &= E^{(x,i)}\left[e^{-\int_0^{t\wedge\tau_D} r(\xi_s)ds} f(X_{t\wedge\tau_D}, \xi_{t\wedge\tau_D})\right] \\ &= E^{(x,i)}\left[e^{-\int_0^{\tau_D} r(\xi_s)ds} f(X_{\tau_D}, \xi_{\tau_D})\right] \\ &= E^{(x,i)}\left[e^{-\int_0^{\tau_D} r(\xi_s)ds} \phi(X_{\tau_D})\right] \\ &\geq 0 \end{aligned} \tag{3.27}$$

so we see that $f(x,i) \geq 0 = \phi(x)$ for $x \geq \ln(K)$. Thus we have the desired result that $f(x,i) \geq \phi(x)$ for all (x,i). □

To finalize the proof of optimality, propositions 3.6.1 and 3.6.2 will be combined with (3.15) to see that

$$e^{-\int_0^t r(\xi_s)ds} \phi(X_t) \leq e^{-\int_0^t r(\xi_s)ds} f(X_t, \xi_t) \leq f(x,i) + Martingale.$$

Thus for any stopping time (not necessarily finite) τ we have

$$E^{(x,i)}\left[e^{-\int_0^{t\wedge\tau} r(\xi_s)ds}\left(K - e^{X_{t\wedge\tau}}\right)^+\right] \leq f(x,i).$$

Now let $t \to \infty$ and apply Fatou's Lemma to get

$$E^{(x,i)}\left[e^{-\int_0^{\tau} r(\xi_s)ds}\left(K - e^{X_{\tau}}\right)^+\right] \leq f(x,i).$$

However from (3.27) we have

$$E^{(x,i)}\left[e^{-\int_0^{\tau_D} r(\xi_s)ds} f(X_{\tau_D}, \xi_{\tau_D})\right] = f(x,i).$$

Thus the upper bound is achieved and we have

$$\sup_{\tau} E^{(x,i)}\left[e^{\int_0^{\tau} r(\xi_s)ds}(K - e^{X_\tau})\right] = f(x,i).$$

Optimality of $f(x,i)$ is proven and the optimal stopping time is τ_D.

CHAPTER 4

FORWARD CONTRACTS

4.1 Introduction

In this chapter we will discuss the pricing of one the most typical claims on a stock, the forward contract, referred to from here on as a forward. A forward is a contract in which one party agrees to give another a share of stock for a fixed price at an agreed upon date in the future. The stock is being sold forward. Once this derivative is priced under a martingale measure, a new infinite horizon American version of this forward will be discussed. This extends upon the work of [12] where the same is done with only a two state model. It will be shown that this new contract cannot be priced under a risk neutral measure. Despite this, we will discuss the optimal exercise strategy and expected payoff of this contract under the actual real world probability measure whose expected return and volatility are usually estimated from historical data.

4.2 Pricing Forwards in a Markov Modulated Market

The pricing of a forward goes as follows: what price K should one party be obligated to pay for one share of stock $\{S_t\}_{t\geq 0}$ at a predetermined time T in the future in order to make the game fair for both parties? Mathematically, we say

$$E[e^{-rT}(S_T - K)] = 0$$

where S_0 is the present value of a share of stock. Again, we price K under the risk neutral measure to avoid arbitrage. The necessity of the risk neutral measure will be illustrated in the sequel, but first we find the proper value of K. As discussed in section 3.2, the stock process described by $dS_t = \mu S_t dt + \sigma S_t dW_t$ becomes $dS_t = rS_t dt + \sigma S_t dW_t$ under the risk neutral (martingale) measure. Thus we have

$$E[e^{-rT}(S_T - K)] = S_0 - Ke^{-rT}$$

and it becomes clear that the proper value is $K = S_0 e^{rT}$.

This value for K is perhaps not intuitive since it does not depend on drift μ of the stock. Lets illustrate why we must price under the risk neutral measure.

Suppose that $K > S_0 e^{rT}$. In this case, an arbitrageur would engage in the forward contract to sell one share of stock at price K at time T. To buy this share of stock, he would borrow S_0 dollars at a rate r with the obligation to pay back $S_0 e^{rT}$ at time T. His total earnings at time T would be

$$K - S_0 e^{rT} > 0.$$

He could of course buy n shares of the stock to multiply guaranteed profit restricted only by his credit limit.

Similarly in the case that $K < S_0 e^{rT}$ the arbitrageur would engage in a forward contract being obligated to buy a share of stock for K at time T. To get the K dollars required at time T he will short sell one share of stock, *i.e.*, sell a share of stock without owning it, for S_0. He would then invest this money at interest rate r guaranteeing $S_0 e^{rT}$ at time T. At time T he buys the stock for price K and returns it as the short sold stock. His profit is given by

$$S_0 e^{rT} - K > 0.$$

From this discussion it is clear that K must be found under a risk neutral measure.

Now we will focus on how the issue is complicated by a Markov modulated market. We start with the stock process

$$dS_t = \mu(\xi_t)S_t dt + \sigma(\xi_t)S_t dW_t$$

where ξ_t is a regular Markov chain representing the n states of the market. After switching to the risk neutral measure, see 3.2, we get the stock process

$$dS_t = r(\xi_t)S_t dt + \sigma(\xi_t)S_t dW_t.$$

As before, we find K by mandating that the discounted expected payoff be zero.

$$\begin{aligned} 0 &= E[e^{-\int_0^T r(\xi_s)ds}(S_t - K)] \\ &= S_0 - E[e^{-\int_0^T r(\xi_s)ds}]K. \end{aligned}$$

Thus finding

$$K = \frac{S_0}{E[e^{-\int_0^T r(\xi_s)ds}]}.$$

All that remains is to find the value of $E[e^{-\int_0^T r(\xi_s)ds}]$.

Proposition 4.2.1.

Define

$$M(t,i) = E^i[e^{-\int_0^t r(\xi_s)ds}].$$

$M(t,i)$ satisfies the following ODE system,

$$\mathbf{M}' = (Q - R)\mathbf{M} \quad M(0,i) = 1$$

where $\mathbf{M}$ is the vector whose i-th element is $M(t,i)$, Q is the generating matrix of ξ_t, and R is the diagonal matrix whose i-th diagonal element is $r(i)$.

Proof. We will derive the ODE for $M(t,i)$ by conditioning on the first transition time and a time Δt. Below T_1 is the time of the first transition and P_{ij} is the probability that ξ_t transitions from i to j given a transition at t.

$$\begin{aligned}
M(t,i) &= E^i[e^{-\int_0^t r(\xi_s)ds}] \\
&= E^i[e^{-r(i)\Delta t}e^{-\int_{\Delta t}^t r(\xi_s)ds}\mathbb{1}_{\{T_1>\Delta t\}}] + E^i[e^{-r(i)T_1}e^{-\int_{T_1}^t r(\xi_s)ds}\mathbb{1}_{\{T_1\le\Delta t\}}] \\
&= E^i\left[E[e^{-r(i)\Delta t}e^{-\int_{\Delta t}^t r(\xi_s)ds}\mathbb{1}_{\{T_1>\Delta t\}} \mid \mathcal{F}_{\Delta t}]\right] + E^i\left[E[e^{-r(i)T_1}e^{-\int_{T_1}^t r(\xi_s)ds}\mathbb{1}_{\{T_1\le\Delta t\}} \mid \mathcal{F}_{T_1}]\right] \\
&= \mathbb{1}_{\{T_1>\Delta t\}}e^{-r(i)\Delta t}E^i\left[E[e^{-\int_{\Delta t}^t r(\xi_s)ds} \mid \mathcal{F}_{\Delta t}]\right] + E^i\left[\mathbb{1}_{\{T_1\le\Delta t\}}e^{-r(i)T_1}E[e^{-\int_{T_1}^t r(\xi_s)ds} \mid \mathcal{F}_{T_1}]\right] \\
&= \mathbb{1}_{\{T_1>\Delta t\}}e^{-r(i)\Delta t}E^i\left[E[e^{-\int_{\Delta t}^t r(\xi_s)ds}]\right] + E^i\left[\mathbb{1}_{\{T_1\le\Delta t\}}e^{-r(i)T_1}E[e^{-\int_{T_1}^t r(\xi_s)ds}]\right] \\
&= P(T_1>\Delta t)e^{-r(i)\Delta t}M(t-\Delta t,i) + E^i\Big[\mathbb{1}_{\{T_1\le\Delta t\}}e^{-r(i)T_1}\sum_{\substack{j=1\\ j\neq i}}^n P_{ij}M(t-T_1,j)\Big] \\
&= e^{q_{ii}\Delta t}e^{-r(i)\Delta t}M(t-\Delta t,i) + P(T_1\le\Delta t)e^{-r(i)\eta\Delta t}\sum_{\substack{j=1\\ j\neq i}}^n P_{ij}M(t-\eta\Delta t,j) \qquad (4.1) \\
&= e^{(-r(i)+q_{ii})\Delta t}M(t-\Delta t,i) + (1-e^{q_{ii}\Delta t})e^{-r(i)\eta\Delta t}\sum_{\substack{j=1\\ j\neq i}}^n P_{ij}M(t-\eta\Delta t,j)
\end{aligned}$$

where $\eta \in (0,1)$. Now (4.1) will be justified. Define

$$f(s) := e^{-r(i)s}\sum_{\substack{j=1\\ j\neq i}}^n P_{ij}M(t-s,j)$$

and notice that $f(s)$ is continuous since $M(s,j)$ is continuous. Next see that

$$E^i[\mathbb{1}_{\{T_1\le\Delta t\}}f(T_1)] = \int_0^{\Delta t} f(s)dF_{T_1}(s)$$

where F_{T_1} is the distribution of T_1 and apply the mean value theorem for integrals to get

$$\begin{aligned}
E^i[\mathbb{1}_{\{T_1\le\Delta t\}}f(T_1)] &= f(\eta\Delta t)\Big(F_{T_1}(\Delta t) - F_{T_1}(0)\Big) \\
&= f(\eta\Delta t)P(T_1\le\Delta T)
\end{aligned}$$

where $\eta \in (0, 1)$ thus providing justification for the above.

Next we use a Taylor expansion around $t - \Delta t$ for $M(t, i)$ to arrive at

$$
\begin{aligned}
&M(t-\Delta t, i) + M'(t-\Delta t, i)\Delta t + o(\Delta t) \\
&\qquad = e^{(-r(i)+q_{ii})\Delta t} M(t-\Delta t, i) + (1 - e^{q_{ii}\Delta t}) e^{-r(i)\eta\Delta t} \sum_{\substack{j=1 \\ j\neq i}}^{n} P_{ij} M(t - \eta\Delta t, j).
\end{aligned}
$$

Subtracting $M(t, i)$ from both sides and dividing by $-\Delta t$ and reorganizing yields

$$
\begin{aligned}
&\frac{M(t-\Delta t, i) - M(t, i)}{-\Delta t} - M'(t-\Delta t, i) - \frac{o(\Delta t)}{\Delta t} \\
&\quad = \frac{e^{(-r(i)+q_{ii})\Delta t} M(t-\Delta t, i) - M(t, i)}{-\Delta t} + \frac{(e^{q_{ii}\Delta t} - 1)}{\Delta t} e^{-r(i)\eta\Delta t} \sum_{\substack{j=1 \\ j\neq i}}^{n} P_{ij} M(t - \eta\Delta t, j) \\
&\quad = \frac{M(t-\Delta t, i) - M(t, i)}{-\Delta t} - M(t-\Delta t, i) \frac{e^{(-r(i)+q_{ii})\Delta t} - 1}{\Delta t} \\
&\qquad\qquad + \frac{(e^{q_{ii}\Delta t} - 1)}{\Delta t} e^{-r(i)\eta\Delta t} \sum_{\substack{j=1 \\ j\neq i}}^{n} P_{ij} M(t - \eta\Delta t, j).
\end{aligned}
$$

Taking the limit $\Delta t \to 0$ we get the ODE

$$
0 = M'(t, i) - (-r(i) + q_{ii}) M(t, i) + \sum_{\substack{j=1 \\ j\neq i}}^{n} q_{ii} P_{ij} M(t, j).
$$

Recognizing that $q_{ii} P_{ij} = -q_{ij}$,

$$
\begin{aligned}
M'(t, i) &= -r(i) M(t, i) + \left(\sum_{\substack{j=1 \\ j\neq i}}^{n} q_{ij} M(t, j) \right) + q_{ii} M(t, i) \\
&= -r(i) M(t, i) + \sum_{j=1}^{n} q_{ij} M(t, j).
\end{aligned}
$$

This is written nicely in matrix form,

$$
\mathbf{M}' = (Q - R)\mathbf{M}
$$

where Q is the infinitesimal generator of ξ_t and R is the diagonal matrix whose i-th diagonal element is $r(i)$. Since

$$M(0,i) = E^i[e^{-\int_0^0 r(\xi_s)ds}] = 1,$$

the proof is complete. □

The solution of the ODE in proposition 4.2.1 is determined by the eigenvalues and eigenvectors of $Q - R$:

$$\mathbf{M}(t) = \sum_{i=1}^{n} C_i \mathbf{V_i} e^{\lambda_i t}$$

where V_i and λ_i are the i-th corresponding eigenvector and eigenvalue and the C_i's make up n arbitrary constants completely determined by the initial condition. For complex eigenvalues, we interpret the solution as the sum of sines and cosines in the standard way.

Now we can find the value K, dependent upon the initial Markov state and stock value:

$$\begin{aligned} K &= \frac{S_0}{E[e^{-\int_0^T r(\xi_s)ds}]} \\ &= \frac{S_0}{M(T,\xi_0)}. \end{aligned}$$

We have now priced a forward in a Markov modulated market.

4.3 Optimal Exercise Strategy for American Style Forwards

In this section we will develop a contract which will be referred to as a "perpetual American future". The contract is formed by one party agreeing to buy a share of stock for a price K at any time of his choosing in the future. Perpetual refers to the fact that the contract has an infinite time horizon and thus no fixed termination time. American meaning that the contract can be exercised at any time. The payoff of the contract at a time t is given by $(S_t - K)$.

Next we attempt to find the fair value of K. We must keep in mind one major difference: the contract can be exercised at any time. Ideally the contract will be exercised at an optimal time, *i.e.*, the time that maximizes expected future discounted profit. We conclude that the expected profit under the real world measure P from the contract is

$$\sup_{\tau} E_P[e^{-r\tau}(S_\tau - K)\mathbb{1}_{\tau<\infty}]$$

the indicator being necessary since if the contract is never exercised, we gain no profit. Without the indicator, it is unclear how to interpret $\lim_{t\to\infty} e^{-rt}(S_t - K)$. If we are to price this derivative on the open market, we must use the the risk neutral measure Q. We set

$$\sup_{\tau} E_Q[e^{-r\tau}(S_\tau - K)\mathbb{1}_{\tau<\infty}] = 0$$

to find K. It must be determined whether such a value exist.

We now have two questions to answer. The first, what is the optimal exercise strategy and expected profit under the real measure for the owner of a contract. The second, is there a fair price for K under the risk neutral measure. Both of these questions will be answered when an optimal stopping time is identified. Notice that the value of the contract is not dependent on time since there is no termination date. Thus we conclude that the optimal stopping time must be only dependent upon the stock price. We conclude that if an optimal stopping time exists, it is of the form

$$\tau_s = \inf\{t \geq 0 : S_t \geq c\}$$

for some threshold c. Lets first answer the question of expected profit under this optimal exercise strategy:

$$E_P[e^{-r\tau_c}(S_{\tau_c} - K)\mathbb{1}_{\tau_c<\infty}] = E_P[e^{-r\tau_c}\mathbb{1}_{\tau_c<\infty}](c - K).$$

We need only find the value of $E_P[e^{-r\tau_c}\mathbb{1}_{\tau_c<\infty}]$ to determine the expected profit.

First, we will take the perspective of the logarithmic stock process, which by Ito's formula 2.1.1 we find to be

$$\begin{aligned} X_t &= \ln(S_t) \\ &= X_0 + (\mu - \frac{1}{2}\sigma^2)t + \sigma W_t. \end{aligned}$$

The hitting time for S_t hitting c is the same as X_t hitting $\ln(c)$. Define $b = \ln(c)$ and $\alpha = \mu - \frac{1}{2}\sigma^2$. Define the stopping time

$$T_b = \inf\{t \geq 0 : X_t \geq b\}.$$

See that

$$\begin{aligned} E_P[e^{-r\tau_c}\mathbb{1}_{\tau_c<\infty}] &= E_P[e^{-rT_b}\mathbb{1}_{T_b<\infty}] \\ &= \int_0^\infty e^{-rt} f(t)dt \end{aligned}$$

where $f(t)$ is the density function of T_b. All that needs to be done is to find the density function. Notice that the last line above is the Laplace transform of the density function. This will allow us to find the expected profit of the contract. The following two results will provide the tools necessary to find the density function of T_b.

Lemma 4.3.1.

Let $M_t = \max\{W_s : 0 \leq s \leq t\}$.

$$P(M_t \geq b) = 2P(W_t \geq b) = 2\left(1 - \Phi\left(\frac{b}{\sqrt{t}}\right)\right)$$

where Φ *is standard normal distribution.*

For a proof we refer the reader to ([18], Theorem 3.15, p. 71).

Theorem 4.3.1.

Let $X_t = x + \alpha t + \sigma W_t$. The distribution of the hitting time $T_b = \inf\{t \geq 0 : X_t \geq b\}$ is

$$F(t) = 1 - \Phi\left(\frac{b - x - \alpha t}{\sigma\sqrt{t}}\right) - e^{\frac{2\alpha(b-x)}{\sigma^2}}\Phi\left(\frac{b - x - \alpha t}{\sigma\sqrt{t}}\right), \quad x < b, \quad t > 0$$

and the density function is given by

$$f(t) = \frac{(b-x)}{\sqrt{2\pi t^3\sigma^2}}\exp\left(\frac{-(b-x-\alpha t)^2}{2t\sigma^2}\right), \quad x < b, \quad t > 0.$$

Before we proceed to the proof of this theorem, observe that if $\alpha > 0$, then we have a proper probability distribution, but if $\alpha < 0$ we have a defective probability distribution, i.e.,

$$\lim_{t\to\infty} F(t) = e^{\frac{2\alpha(b-x)}{\sigma^2}} < 1.$$

Thus there would be a positive probability the arithmetic Brownian motion never hits level b when $\alpha < 0$. It is now clear that is very important that we include $\mathbb{1}_{\tau_b<\infty}$ in taking expected values.

Proof. In the case that $\alpha = 0$ we use lemma 4.3.1 to show that

$$\begin{aligned} P(T_b \leq t) &= P(\max\{X_s : 0 \leq s \leq t\} \geq b) \\ &= P\left(M_t \geq \frac{b-x}{\sigma}\right) \\ &= 2\left(1 - \Phi\left(\frac{b-x}{\sigma\sqrt{t}}\right)\right). \end{aligned} \tag{4.2}$$

Next Girsanov's theorem will be used to remove the drift from X_t in the case that α is not zero. Define the measure Q by a Radon-Nikodym derivative:

$$\frac{dQ}{dP} = \exp(-\frac{1}{2}\frac{\alpha^2}{\sigma^2}t - \frac{\alpha}{\sigma}W_t).$$

Under the new measure Q, $\tilde{W}_t = W_t + \frac{\alpha}{\sigma}t$ is a Wiener process. Now the distribution of T_b will be found. Define $G_t = \frac{dP}{dQ} = \exp(-\frac{1}{2}\frac{\alpha^2}{\sigma^2}t + \frac{\alpha}{\sigma}\tilde{W}_t)$. We use the fact that

$E_P[H] = E_Q[HG_T]$ for any bounded Borel function H and where G_T is the terminal value of the martingale G_s.

$$
\begin{aligned}
P(T_b \le t) &= E_P[\mathbb{1}_{\{T_b \le t\}}] \\
&= E_Q[\mathbb{1}_{\{T_b \le t\}} G_T] \\
&= E_Q[E[\mathbb{1}_{\{T_b \le t\}} G_T \mid \mathcal{F}_{T_b \wedge t}]] \\
&= E_Q[\mathbb{1}_{\{T_b \le t\}} G_{T_b \wedge t}] \\
&= E_Q[\mathbb{1}_{\{T_b \le t\}} G_{T_b}] \\
&= E_Q[\mathbb{1}_{\{T_b \le t\}} \exp\left(-\frac{1}{2}\frac{\alpha^2}{\sigma^2} T_b + \frac{\alpha}{\sigma}(\frac{b-x}{\sigma})\right)] \\
&= \int_0^t \exp(-\frac{1}{2}\frac{\alpha^2}{\sigma^2} s + \frac{\alpha}{\sigma}(\frac{b-x}{\sigma}) f(s) ds \qquad (4.3)
\end{aligned}
$$

where $f(s)$ is the density function of T_b under Q, *i.e.*, with no drift. From (4.2), we get

$$
f(s) = \frac{d}{ds} 2\Phi\left(\frac{b-x}{\sigma\sqrt{s}}\right) = \frac{(b-x)}{\sqrt{2\pi t^3 \sigma^2}} \exp\left(\frac{-(b-x)^2}{2t\sigma^2}\right).
$$

Differentiating (4.3) and simplifying we get the density function

$$
f(t) = \frac{(b-x)}{\sqrt{2\pi t^3 \sigma^2}} \exp\left(\frac{-(b-x-\alpha t)^2}{2t\sigma^2}\right).
$$

It is apparent upon differentiation that $F'(t) = f(t)$ and the proof is finished. □

All that remains is to find the Laplace transform of $f(s)$ and we have $E[e^{-rT_b}\mathbb{1}_{\tau_b<\infty}]$.

$$
\begin{aligned}
E[e^{-rT_b}\mathbb{1}_{\tau_b<\infty}] &= \int_0^\infty e^{-rs} f(s) ds \\
&= \exp(-\frac{b-x}{\sigma^2}(\sqrt{\alpha^2 + 2r\sigma^2} - \alpha)).
\end{aligned}
$$

Now we will answer the question of whether there is a fair value for K to put this derivative on the market. In the risk neutral measure Q, $\alpha = r - \frac{1}{2}\sigma^2$ and we get

$$
E_Q[e^{-rT_b}\mathbb{1}_{\tau_b<\infty}] = e^{-(b-x)}.
$$

Lets find the value of b which maximizes profit,

$$E_Q[e^{-rT_b}(S_{T_b} - K)\mathbb{1}_{\tau_b<\infty}] = e^{-(b-x)}(e^b - K) = e^x(1 - Ke^{-b}).$$

To maximize profit $b \to \infty$, thus it is impossible to price this derivative on the market. However, it is still useful to address the issue of optimal exercise of the contract supposing that one has such a contract. Even if the contract can't be sold freely on the market, it is still feasible that the contract might be given to an individual. For example, an employer might give such a contract to an employee as a form of payment or bonus.

Lets now address the issue of optimal exercise and expected profit. First we address the case when $\mu > r$

$$E^x[e^{-rt}(S_t - K)] = e^{(\mu-r)t}e^x - e^{-rt}K.$$

we see that in this case we want to hang on to this contract, because the discounted expected profit will gain value with time. Because of this, there is no optimal stopping time.

Next we address the case when $\mu < r$ when it is not in our best interest to keep the contract indefinitely. Lets find the optimal level for b

$$E^x[e^{-rT_b}(S_{T_b} - K)\mathbb{1}_{\tau_b<\infty}] = \exp\left(-\frac{b-x}{\sigma^2}(\sqrt{\alpha^2 + 2r\sigma^2} - \alpha)\right)(e^b - K) \quad for \quad x < b.$$

Maximizing this relative to b yields

$$e^b = \frac{K}{1 - \frac{\sigma^2}{z}}$$

where $z = \sqrt{\alpha^2 + 2r\sigma^2} - \alpha$ and $\alpha = \mu - \frac{1}{2}\sigma^2$. Notice that when $\mu < r$ then $\frac{\sigma^2}{z} < 1$ and there is a valid value for b. On the other hand if $\mu \geq r$ then $\frac{\sigma^2}{z} \geq 1$ and there does not exist a value for b as we would expect from the previous discussion.

Finally the expected discounted profit when $\mu < r$ is given by

$$\begin{aligned} E^x[e^{-rT_b}(S_{T_b} - K)\mathbb{1}_{\tau_b<\infty}] &= E^x[e^{-rT_b}\mathbb{1}_{\tau_b<\infty}]\left(\frac{K}{1-\frac{\sigma^2}{z}} - K\right) \\ &= K\exp\left(-\frac{z(b-x)}{\sigma^2}\right)\left(\frac{\sigma^2}{z-\sigma^2}\right). \end{aligned}$$

We have identified an optimal stopping strategy along with an accompanying discounted expected profit, which concludes this section.

4.4 American Style Forwards in a Markov Modulated Market

In this section we will allow the stock process to change between n states by a Markov chain as before. However, here we fix the risk-free interest rate r. Our goal in this section is to identify an optimal stopping strategy for the perpetual American style forward discussed in the previous section and to find the discounted expected payoff under the optimal strategy.

The Markov modulated stock process is defined

$$dS_t = \mu(\xi_t)S_t dt + \sigma(\xi_t)S_t dW_t$$

as before. Here we do not apply a change of measure since we are interested in the real optimal strategy and payoff. We define the optimal payoff as

$$V(x,i) = \sup_{\tau} E^{(s,i)}[e^{-r\tau}(S_\tau - K)\mathbb{1}_{\{\tau<\infty\}}]$$

again applying the indicator to make it clear that there is zero payoff if the contract is held indefinitely. See that

$$S_t = S_0 \exp\left(\int_0^t \left(\mu(\xi_s) - \frac{1}{2}\sigma^2(\xi_s)\right)ds + \int_0^t \sigma(\xi_s)dW_s\right).$$

Note that

$$S_0 \exp\left(\int_0^t -\frac{1}{2}\sigma^2(\xi_s)ds + \int_0^t \sigma(\xi_s)dW_s\right)$$

is a martingale.

Proposition 4.4.1.

Let $X_t = \ln(S_t)$ *to get*

$$dX_t = \left(\mu(\xi_t) - \frac{1}{2}\sigma^2(\xi_t)\right) + \sigma(\xi_t)dW_t$$

then

$$\begin{aligned} E^s[S_t] &= E^x[e^{X_t}] \\ &= E\left[\exp\left(\int_0^t \mu(\xi_s)ds\right)\right]e^x. \end{aligned}$$

Proof. We will start by applying the infinitesimal generator from proposition 3.3.1 to the function $v(t,x,i) = E^{(x,i)}[e^{X_t}]$ like so

$$\begin{aligned} L_{X,\xi}[v](t,x,i) &= \lim_{s\searrow 0}\frac{1}{s}E^{(x,i)}[v(t,X_s,\xi_s) - v(t,x,i)] \\ &= \lim_{s\searrow 0}\frac{1}{s}E^{(x,i)}\left[E^{(X_s,\xi_s)}[e^{X_t}] - E^{(x,i)}[e^{X_t}]\right] \\ &= \lim_{s\searrow 0}\frac{1}{s}E^{(x,i)}\left[E^{(x,i)}[e^{X_{t+s}} \mid \mathcal{F}_s] - E^{(x,i)}[e^{X_t} \mid \mathcal{F}_s]\right] \\ &= \lim_{s\searrow 0}\frac{1}{s}E^{(x,i)}\left[e^{X_{t+s}} - e^{X_t}\right] \\ &= \lim_{s\searrow 0}\frac{v(t+s,x,i) - v(t,x,i)}{s} \\ &= \frac{\partial v}{\partial t}(t,x,i). \end{aligned}$$

Now we make the observation that $\frac{\partial v}{\partial x} = v$ as shown below.

$$\begin{aligned} \frac{\partial v}{\partial x} &= \lim_{h\to 0}\frac{E^{(x+h,i)}[e^{X_t}] - E^{(x,i)}[e^{X_t}]}{h} \\ &= \lim_{h\to 0}\frac{E^{(x,i)}[e^{X_t+h} - e^{X_t}]}{h} \\ &= \lim_{h\to 0}\frac{e^h - 1}{h}E^{(x,i)}[e^{X_t}] \\ &= E^{(x,i)}[e^{X_t}] = v. \end{aligned}$$

Combining the above two results, we get

$$\begin{aligned}\frac{\partial v}{\partial t}(t,x,i) &= L_{(X,\xi)}[v](x,i)\\ &= \left(\mu(i) - \frac{1}{2}\sigma^2(i)\right)v_x(t,x,i) + \frac{1}{2}\sigma^2(i)v_{xx}(t,x,i) + \sum_{j=1}^{n} q_{ij}v(t,x,j)\\ &= \mu(i)v(t,x,i) + \sum_{j=1}^{n} q_{ij}v(t,x,j).\end{aligned}$$

Put in matrix form, $\mathbf{V}' = (Q+U)\mathbf{V} \quad v(0,x,i) = e^x$ where $\mathbf{V}$ is the vector whose i-th element is $v(t,x,i)$ and U is the diagonal matrix whose i-th diagonal element is $\mu(i)$. Comparing this to proposition 4.2.1, it becomes clear that

$$v(t,x,i) = E^{(x,i)}\left[\exp\left(\int_0^t \mu(\xi_s)ds\right)\right]e^x$$

and the proof is complete. □

Now, from proposition 4.4.1 we have that

$$E^{(s,i)}[e^{-rt}(S_t - K)] = e^{-rt}E^i\left[\exp\left(\int_0^t \mu(\xi_s)ds\right)\right]s - e^{-rt}K.$$

From proposition 4.2.1 we have that

$$E^i\left[\exp\left(\int_0^t \mu(\xi_s)ds\right)\right] = M(t,i).$$

We define $M(t,i)$ as the i-th element of the vector

$$\mathbf{M}(t) = \sum_{i=1}^{n} C_i\mathbf{V_i}e^{\lambda_i t}$$

where $\mathbf{V}_i$ and λ_i are the i-th corresponding eigenvector and eigenvalue of $Q+U$ with U being the diagonal matrix whos i-th diagonal element is $\mu(i)$. The C_i's make up n arbitrary constants completely determined by the initial condition $\mathbf{M}(0) = 1$.

Combing the above, we get

$$E^{(s,i)}[e^{-rt}(S_t - K)] = \sum_{i=1}^{n}\left(sC_i\mathbf{V_i}e^{(\lambda_i - r)t}\right) - e^{-rt}K$$

and it becomes clear that if the real part of any of the eigenvalues of $Q + U$ are larger than r, then we expect the value of the contract to grow without bound and there is no optimal stopping strategy.

Conjecture 4.4.1.

If r is greater than the real part of all eigenvalues of $Q + U$ and $r \geq \max\{\mu(1), \ldots, \mu(n)\}$ then there exists an optimal stopping time for

$$V(x, i) = \sup_{\tau} E^{(x,i)}[e^{-r\tau}(e^{X_\tau} - K)\mathbb{1}_{\{\tau<\infty\}}]$$

which is of the form

$$\tau = \inf\{t \geq 0 : V(X_t, \xi_t) = (e^{X_t} - K)\}.$$

As in lemma 3.4.1, we will show that this stopping time is equivalent to having a series of thresholds, above which it is optimal to exercise the contract.

Lemma 4.4.1.

Supposing a stopping time of the form given in conjecture 4.4.1 there exists a series of thresholds, $\{b(1), \ldots, b(n)\}$ forming a region $C = \{(x, i) : x < b(i)\}$ referred to as the continuation region and a region $D = \{(x, i) : x \geq b(i)\}$ referred to as the stopping region such that the entry time $\tau_D = \inf\{t \geq 0 : (X_t, \xi_t) \in D\} = \inf\{t \geq 0 : X_t \leq b(\xi_t)\}$ is the optimal stopping time:

$$V(x, i) = E^{(x,i)}\left[e^{-r\tau_D}(e^{X_{\tau_D}} - K)\mathbb{1}_{\{\tau_D<\infty\}}\right].$$

Without loss of generality, we will assume $b(1) < \cdots < b(n)$.

Proof. The proof is in effect a reprise of the proof of lemma 3.4.1. We will show that $V(x, i) - \phi(x)$, where $\phi(x) = (e^x - K)$, is a decreasing function on $\mathbb{R}$. From this we conclude that so long as there is a value above which $V(x, i) = \phi(x)$ for all i, then there are unique thresholds depending on the Markov state above which $V(x, i) = \phi(x)$ and the lemma is proven. Let $x, x + \delta \in \mathbb{R}$ and let τ be the optimal stopping time when $X_0 = x$. We need the fact that $r \geq \max\{\mu(1), \ldots, \mu(n)\}$ implies that $e^{-rt}e^{X_t}$ is a supermartingale, which follows directly from proposition 4.4.1.

$$
\begin{aligned}
V(j, x+\delta) &= E^{(x+\delta,j)}\left[e^{-r\tau}\phi(X_\tau)\right] \\
&= E^{(x,j)}\left[e^{-r\tau}\left(e^{X_\tau+\delta} - K\right)\right] \\
&= E^{(x,j)}\left[e^{-r\tau}\left(e^{X_\tau} - K + (e^\delta - 1)e^{X_\tau}\right)\right] \\
&= E^{(x,j)}\left[e^{-r\tau}(e^{X_\tau} - K)\right] + (e^\delta - 1)E^{(x,j)}\left[e^{-r\tau}e^{X_\tau}\right] \\
&= V(x,j) + (e^\delta - 1)E^{(x,j)}\left[e^{-r\tau}e^{X_\tau}\right] \\
&\leq V(x,j) + (e^\delta - 1)\liminf_{t\to\infty} E^{(x,j)}\left[e^{-r(t\wedge\tau)}e^{X_{t\wedge\tau}}\right] && (4.4) \\
&\leq V(x,j) + (e^\delta - 1)e^x && (4.5) \\
&= V(x,j) + \phi(x+\delta) - \phi(x).
\end{aligned}
$$

In (4.4) Fatou's lemma is applied. Finally (4.5) is valid by theorem 2.2.3 since the discounted stock process $e^{-rt}e^{X_t}$ is a supermartingale and $t \wedge \tau$ is a bounded stopping time. Notice that no assumptions are made about the finiteness of τ.

The ordering of the n thresholds would involve a simple renaming of the Markov states to achieve descending order and the proof is complete. □

Next the techniques from section 3.4 will be followed and it can be shown that the payoff function $V(x,i)$ solves the Dirichlet problem

$$
\begin{aligned}
L_{(X,\xi)}[V](x,i) &= rV(x,i) \quad in \quad C && (4.6) \\
V(x,i) &= e^x - K \quad in \quad D.
\end{aligned}
$$

This leads to the system of ODEs below

$$
\left(\mu(i) - \frac{1}{2}\sigma^2(i)\right)f_x(x,i) + \frac{1}{2}\sigma^2(i)f_{xx}(x,i) + \sum_{j=1}^{n} q_{ij}f(x,j) = rf(x,i) \quad x < b_i
$$

$$
f(x,i) = e^x - K \quad x \geq b_i.
$$

This system is solved in exactly the same manner as in section 3.5. We get the following solution

$$f(x,k+1) = \begin{cases} \sum_{i=1}^{n} \omega_i g_i(k+1) e^{-\lambda_i x} & ; x < b_1 \\ \sum_{i=1}^{2(n-1)} \omega_i^1 g_i^1(k+1) e^{-\lambda_i^1 x} + B_1(k) + C_1(k) e^x & ; b_1 \le x < b_2 \\ \vdots & ; \vdots \\ \sum_{i=1}^{2(n-k)} \omega_i^k g_i^k(k+1) e^{-\lambda_i^k x} + B_k(1) + C_k(1) e^x & ; b_k \le x < b_{k+1} \\ e^x - K & ; b_{k+1} \le x \end{cases} . \quad (4.7)$$

Below b_1 we only use negative eigenvalues since the solution should decay as x goes to negative infinity. Again, we will require that the solution be C^1 on its domain, and this will completely determine all weights and boundaries. As before, it remains to show that this solution matches the profit function of the contract.

Theorem 4.4.1 (Optimality).

Suppose that the thresholds $\ln(K) < b_1 < \cdots < b_n$ *have been found such that the unique solution to*

$$\left(\mu(i) - \frac{1}{2}\sigma^2(i)\right) f_x(x,i) + \frac{1}{2}\sigma^2(i) f_{xx}(x,i) + \sum_{j=1}^{n} q_{ij} f(x,j) = r f(x,i) \quad x < b_i$$

$$f(x,i) = \phi(x) \quad x \ge b_i$$

is C^1 *on its domain and bounded on C. Further, suppose the following assumptions hold*

1. $f(x,i) \ge e^x - K$ *for all* (x,i)
2. r *is greater than the real part of all eigenvalues of* $Q + U$
3. $r > \max\{\mu(1), \ldots, \mu(n)\}$
4. $r > \max_{1 \le i \le n} \left\{ \frac{M(i) + \mu(i) e^{b(i)}}{e^{b(i)} - K} \right\}$ *where* $M(i) = \sum_{j=1}^{i-1} q_{ij} K$.

Then the solution $f(x,i)$ *and the stopping time* $\tau_D = \{t : X_t \ge b(\xi_t)\}$ *correspond to the value function*

$$V(x,i) = \sup_{\tau} E^{(x,i)}[e^{-r\tau}(e^{X_\tau} - K) \mathbb{1}_{\{\tau < \infty\}}]$$

and its optimal stopping time, i.e.,

$$f(x,i) = V(x,i) = E^{(x,i)}[e^{-r\tau_D}(e^{X_{\tau_D}} - K)\mathbb{1}_{\{\tau_D<\infty\}}].$$

Proof. We will start by looking at the process $e^{-rt}f(X_t,\xi_t)$. Since by definition $f(x,i)$ is twice differentiable everywhere except when $x = b_i$ for $i = 1,2,\cdots,n$ where X_t spends zero time, we can apply the generalized Ito-formula (3.5) to it to get

$$e^{-rt}f(X_t,\xi_t) = f(X_0,\xi_0) + \int_0^t e^{-rs}\left(L_{(X,\xi)}[f](X_s,\xi_s) - rf(X_s,\xi_s)\right)ds + Martingale. \quad (4.8)$$

To show optimality we need the following

Proposition 4.4.2.

The inequality for the function defined below is true for all (x,i)

$$\begin{aligned}\Phi(x,i) &:= L_{(X,\xi)}[f](x,i) - rf(x,i)\\ &= \left(\mu(i) - \frac{1}{2}\sigma^2(i)\right)f_x(x,i) + \frac{1}{2}\sigma^2(i)f_{xx}(x,i) - rf(x,i) + \sum_{j=1}^{n} q_{ij}f(x,j)\\ &\leq 0.\end{aligned}$$

Proof. Using the hypothesis $f(x,i) \geq \phi(x)$ we first show that $f(x,i) - \phi(x)$ is decreasing in x:

$$\begin{aligned}
V(j, x-\delta) &= E^{(x-\delta,j)}\left[e^{-r(t\wedge\tau_D)} f(X_{t\wedge\tau_D}, \xi_{t\wedge\tau_D})\right] && (4.9)\\
&= E^{(x-\delta,j)}\left[e^{-r\tau_D} f(X_{\tau_D}, \xi_{\tau_D})\right] && (4.10)\\
&\geq E^{(x-\delta,j)}\left[e^{-r\tau_D} \phi(X_{\tau_D})\right]\\
&= E^{(x,j)}\left[e^{-r\tau_D}\left(e^{X_{\tau_D}-\delta} - K\right)\right]\\
&= E^{(x,j)}\left[e^{-r\tau_D}\left(e^{X_{\tau_D}} - K + (e^{-\delta}-1)e^{X_{\tau_D}}\right)\right]\\
&= E^{(x,j)}\left[e^{-r\tau_D}(e^{X_{\tau_D}} - K)\right] + (e^{-\delta}-1)E^{(x,j)}\left[e^{-r\tau_D}e^{X_{\tau_D}}\right]\\
&= V(x,j) + (e^{-\delta}-1)E^{(x,j)}\left[e^{-r\tau_D}e^{X_{\tau_D}}\right]\\
&\geq V(x,j) + (e^{-\delta}-1)\liminf_{t\to\infty} E^{(x,j)}\left[e^{-r(t\wedge\tau_D)}e^{X_{t\wedge\tau_D}}\right] && (4.11)\\
&\geq V(x,j) + (e^{-\delta}-1)e^{x} && (4.12)\\
&= V(x,j) + \phi(x-\delta) - \phi(x).
\end{aligned}$$

(4.9) is valid since $e^{-r(t\wedge\tau_D)} f(X_{t\wedge\tau_D}, \xi_{t\wedge\tau_D})$ is a martingale due to $\Phi(x,i) = 0$ for $(x,i) \in C$. We note that when $X_0 = x-\delta$ at any time t less than τ_D, X_t is still in region C despite the fact that τ_D is found as if $X_0 = x$. In (4.11) and (4.10) Fatou's lemma is applied. Finally (4.12) is valid from theorem 2.2.3 since the discounted stock process $e^{-rt}e^{X_t}$ is a supermartingale and $t \wedge \tau_D$ is a bounded stopping time.

The above yeilds

$$\begin{aligned}
f(x,i) - \phi(x) &\leq \lim_{x\to-\infty} f(x,i) - \phi(x)\\
&= \lim_{x\to-\infty} \sum_{j=1}^{n} \omega_j g_j(i) e^{-\lambda_j x} - e^{x} - K\\
&= K
\end{aligned}$$

due to the fact that λ_i are chosen to be negative.

For $(x,i) \in C$, we have $\Phi(x,i) = 0$ by construction. For $(x,i) \in D$ we have $f(x,i) = \phi(x)$ and

$$\begin{aligned}
\Phi(x,i) &:= L_{(X,\xi)}[f](x,i) - rf(x,i) \\
&= \left(\mu(i) - \frac{1}{2}\sigma^2(i)\right) f_x(x,i) + \frac{1}{2}\sigma^2(i) f_{xx}(x,i) - rf(x,i) + \sum_{j=1}^{n} q_{ij} f(x,j) \\
&= e^x(\mu(i) - r) + rK + \sum_{\substack{j=1 \\ j\neq i}}^{n} q_{ij}(f(x,j) - \phi(x)) \\
&= e^x(\mu(i) - r) + rK + \sum_{j=1}^{i-1} q_{ij}(f(x,j) - \phi(x)) \\
&\leq e^x(\mu(i) - r) + rK + \sum_{j=1}^{i-1} q_{ij} K \\
&= -r(e^x - K) + \mu(i)e^x + M(i) \\
&\leq -\frac{M(i) + \mu(i)e^x}{e^x - K}(e^x - K) + \mu(i)e^x + M(i) \\
&= 0.
\end{aligned}$$

The last line follows from the fact that $\frac{M+\mu(i)e^x}{e^x-K}$ is decreasing function and thus

$$r > \frac{M + \mu(n)e^{b(i)}}{e^{b(i)} - K} \geq \frac{M + \mu(n)e^x}{e^x - K}.$$

Now we have that $\Phi(x,i) \leq 0$ as desired. □

Using proposition 4.4.2, we have that

$$e^{-rt}\phi(X_t, \xi_t) \leq e^{-rt} f(X_t, \xi_t) \leq f(x,i) + Martingale.$$

Therefore for any stopping time τ

$$E^{(x,i)}[e^{-r(t\wedge\tau)}\phi(X_{t\wedge\tau}, \xi_{t\wedge\tau})] \leq f(x,i)$$

and by Fatou's lemma

$$E^{(x,i)}[e^{-r\tau}\phi(X_\tau, \xi_\tau)] \leq f(x,i).$$

Finally we observe that $f(X_{\tau_D}, \xi_{\tau_D}) = \phi(X_{\tau_D}, \xi_{\tau_D})$ and that $e^{-r(t\wedge\tau)} f(X_{t\wedge\tau_D}, \xi_{t\wedge\tau_D})$ is a martingale due to $\Phi(x, i) = 0$ for $(x, i) \in C$. Furthermore, it is a uniformly integrable martingale since $x \leq b(n)$ when the process is stopped at τ_D and $f(x, i)$ is bounded below $b(n)$. Using theorem 2.2.4 we get

$$f(x, i) = E^{(x,i)}[e^{-r\tau_D} f(X_{\tau_D}, \xi_{\tau_D})] = E^{(x,i)}[e^{-r\tau_D} \phi(X_{\tau_D}, \xi_{\tau_D})]$$

and optimality is proven. □

CHAPTER 5

ORNSTEIN-UHLENBECK MODEL

5.1 Introduction

The Ornstein-Uhlenbeck process, thanks to its equilibrium or stationary measure, has been frequently used in finance. The commonly used Vasicek model [1] for example utilizes the mean reverting Ornstein-Uhlenbeck process. The Ornstein-Uhlenbeck is also sometimes used in modeling stochastic volatility as studied in [5]. In what follows we will devise an American option based on a commodity modeled by the Ornstein-Uhlenbeck process.

5.2 The American Put in an Ornstein-Uhlenbeck Model

Here we pose the question: how would we optimally exercise an American put option based on a commodity whose price evolution follows the Ornstein-Uhlenbeck process below?

$$dX_t = -\alpha X_t dt + \sigma dB_t$$

The solution to this SDE is obtained by applying Ito's formula 2.1.1 to $e^{\alpha t}X_t$ to get

$$d(e^{\alpha t}X_t) = e^{\alpha t}\sigma dB_t$$

so

$$X_t = X_0 e^{-\alpha t} + \sigma \int_0^t e^{-\alpha(t-s)} dB_s.$$

Next we observe that $\int_0^t e^{-\alpha(t-s)} dB_s$ is a martingale. From this we see that

$$E^x[X_t] = e^{-\alpha t}x \leq x$$

and also that X_t is a supermartingale. Here we take note that this commodity is mean reverting to zero. This would indeed be a strange property of real commodities. This American put option here has purely mathematical interest and the next section will deal with an Ornstein-Uhlenbeck process which does not revert to zero, and thus lends itself to practical applications. We now seek to optimally exercise an infinite time horizon American put option. This corresponds to the optimal stopping problem

$$V(x) = \sup_{\tau} E^x[e^{-r\tau}(K - X_\tau)^+].$$

As before, we conjecture that the form of the stopping time is as follows.

Conjecture 5.2.1.

There exists an optimal stopping time for the optimal stopping problem

$$V(x) = \sup_{\tau} E^x[e^{-r\tau}(K - X_\tau)^+]$$

which is of the form

$$\tau = \inf\{t \geq 0 : V(X_t) = (K - X_t)^+\}.$$

Lemma 5.2.1.

Supposing a stopping time of the form given in conjecture 5.2.1 there exists a threshold $b \leq K$ *forming a region* $C = \{x : x > b\}$ *referred to as the continuation region and a region* $D = \{x : x \leq b\}$ *referred to as the stopping region such that the entry time* $\tau_D = \inf\{t \geq 0 : X_t \in D\} = \inf\{t \geq 0 : X_t \leq b\}$ *is the optimal stopping time:*

$$V(x) = E^x[e^{-r\tau_D}(K - X_{\tau_D})^+].$$

Proof. To prove this lemma, we will show that $V(x,i) - \phi(x)$, where $\phi(x) = (K - e^x)^+$, is an increasing function on $(-\infty, K)$. From this we conclude that there is a unique threshold below which $V(x,i) = \phi(x)$ and the lemma is proven. Let $x, x+\delta \in (-\infty, K)$ and let τ be the optimal stopping time when $X_0 = x$. Then

$$\begin{aligned}
V(x+\delta) &\geq E^{x+\delta}[e^{-r\tau}(K - X_\tau)^+] \\
&= E^x[e^{-r\tau}(K - X_\tau - \delta)^+] \\
&\geq E^x[e^{-r\tau}(K - X_\tau)^+] - \delta E[e^{-r\tau}] \\
&\geq V(x) - \delta.
\end{aligned}$$

From this we get

$$\begin{aligned}
V(x+\delta) - \phi(x+\delta) &\geq v(x) - \delta - (K - x - \delta) \\
&= v(x) - \phi(x)
\end{aligned}$$

and the lemma is proven. □

Next, the techniques from section 3.4 show that the payoff function $V(x)$ solves the Dirichlet problem

$$\begin{aligned}
L_X[V](x) &= rV(x) \quad in \quad C \\
V(x) &= (K - x)^+ \quad in \quad D.
\end{aligned} \tag{5.1}$$

This leads to the ODE below

$$\begin{aligned}
-\alpha x f_x(x) + \frac{1}{2}\sigma^2 f_{xx}(x) &= rf(x) \quad x > b \\
f(x) &= K - x \quad x \leq b.
\end{aligned} \tag{5.2}$$

To solve this ODE, we assume that $f(x)$ is analytic and has a convergent Taylor series expansion. Let

$$f(x) = \sum_{k=0}^{\infty} a_k x^k.$$

Applying the differential equation yields

$$\begin{aligned}
0 &= -\alpha x \sum_{k=0}^{\infty}(k+1)a_{k+1}x^k + \frac{\sigma^2}{2}\sum_{k=0}^{\infty}(k+2)(k+1)a_{k+2}x^k - r\sum_{k=0}^{\infty}a_k x^k \\
&= \sum_{k=0}^{\infty}\left(-\alpha k a_k x^k + \frac{\sigma^2}{2}(k+2)(k+1)a_{k+2}x^k - r a_k x^k\right) \\
&= \sum_{k=0}^{\infty}\left(-\alpha k a_k + \frac{\sigma^2}{2}(k+2)(k+1)a_{k+2} - r a_k\right)x^k.
\end{aligned}$$

We arrive at the recursion relation

$$a_{k+2} = \frac{2}{\sigma^2}\frac{a_k(r+\alpha k)}{(k+2)(k+1)}.$$

Now, defining the even terms

$$\begin{aligned}
a_0 &= a_0 \\
a_2 &= a_0\frac{2}{\sigma^2}\frac{r}{2\cdot 1} \\
a_4 &= a_0\left(\frac{2}{\sigma^2}\right)^2\frac{r(r+2\alpha)}{4!} \\
\vdots &= \vdots \\
a_{2k} &= a_0\left(\frac{2}{\sigma^2}\right)^k\frac{r(r+2\alpha)(r+4\alpha)\cdots(r+2(k-1)\alpha)}{(2k)!}.
\end{aligned}$$

Next, the odd terms

$$\begin{aligned}
a_1 &= a_1 \\
a_3 &= a_1\frac{2}{\sigma^2}\frac{r+\alpha}{3\cdot 2} \\
a_5 &= a_1\left(\frac{2}{\sigma^2}\right)^2\frac{(r+\alpha)(r+3\alpha)}{5!} \\
\vdots &= \vdots \\
a_{2k+1} &= a_1\left(\frac{2}{\sigma^2}\right)^k\frac{(r+\alpha)(r+3\alpha)\cdots(r+(2k-1)\alpha)}{(2k+1)!}.
\end{aligned}$$

Next, for ease of notation we introduce the Pochhammer symbol

$$(x)_k = x(x+1)\cdots(x+k-1)$$

and define $(x)_0 = 1$. It is straightforward to show the following facts

$$r(r+2\alpha)(r+4\alpha)\cdots(r+2(k-1)\alpha) = (2\alpha)^{k-1}\left(\frac{r}{2\alpha}\right)_k$$

$$(r+\alpha)(r+3\alpha)\cdots(r+(2k-1)\alpha) = (2\alpha)^{k-1}\left(\frac{r+\alpha}{2\alpha}\right)_k$$

$$(2k)! = 4^k\left(\frac{1}{2}\right)_k k!$$

$$(2k+1)! = 4^k\left(\frac{3}{2}\right)_k k!\,.$$

Using the above we get

$$a_{2k} = \frac{a_0}{2\alpha}\frac{\left(\frac{\alpha}{\sigma^2}\right)^k\left(\frac{r}{2\alpha}\right)_k}{\left(\frac{1}{2}\right)_k k!}$$

$$a_{2k+1} = \frac{a_1}{2\alpha}\frac{\left(\frac{\alpha}{\sigma^2}\right)^k\left(\frac{r+\alpha}{2\alpha}\right)_k}{\left(\frac{3}{2}\right)_k k!}.$$

Now we write the two linearly independent solution functions

$$\frac{a_0}{2\alpha}\sum_{k=0}^{\infty}\frac{\left(\frac{\alpha}{\sigma^2}\right)^k\left(\frac{r}{2\alpha}\right)_k}{\left(\frac{1}{2}\right)_k}\frac{x^{2k}}{k!}$$

$$\frac{a_1}{2\alpha}\sum_{k=0}^{\infty}\frac{\left(\frac{\alpha}{\sigma^2}\right)^k\left(\frac{r+\alpha}{2\alpha}\right)_k}{\left(\frac{3}{2}\right)_k}\frac{x^{2k+1}}{k!}.$$

Lets look at a special function called Kummer's confluent hypergeometric function [24], which is defined by

$$M(a,b,x) = \sum_{k=0}^{\infty}\frac{(a)_k}{(b)_K}\frac{x^k}{k!}.$$

$M(a,b,x)$ is an entire function in x and thus the Taylor series is convergent everywhere. We will now rewrite the two solutions, ignoring the constants, in terms of Kummer functions

$$F_0(x) = M\left(\frac{r}{2\alpha},\frac{1}{2},\frac{\alpha x^2}{\sigma^2}\right)$$

$$F_1(x) = xM\left(\frac{r+\alpha}{2\alpha},\frac{3}{2},\frac{\alpha x^2}{\sigma^2}\right).$$

Since all parameters in the Kummer functions are positive, it is clear that the functions diverge when $x \to \infty$ which is not valid for the value function. We will try to determine if a linear combination of F_0 and F_1 that will go to zero when $x \to \infty$. To do so we look at an asymptotic function. The limiting form of $M(a, b, x)$ when $x \to \infty$ is

$$M(a, b, x) \sim \frac{\Gamma(b)}{\Gamma(a)} e^x x^{a-b}.$$

See [24] for details. Next we find the asymptotics of the functions of interest:

$$F_0(x) \sim \frac{\Gamma\left(\frac{1}{2}\right) \exp\left(\frac{\alpha x^2}{\sigma^2}\right) \left(\frac{\alpha x^2}{\sigma^2}\right)^{\frac{r}{2\alpha} - \frac{1}{2}}}{\Gamma(\frac{r}{2\alpha})} = \frac{\sqrt{\pi}}{\Gamma(\frac{r}{2\alpha})} \exp\left(\frac{\alpha x^2}{\sigma^2}\right) \left(\frac{\sqrt{\alpha} x}{\sigma}\right)^{\frac{r}{\alpha} - 1}$$

$$F_1(x) \sim x \frac{\Gamma\left(\frac{3}{2}\right) \exp\left(\frac{\alpha x^2}{\sigma^2}\right) \left(\frac{\alpha x^2}{\sigma^2}\right)^{\frac{r+\alpha}{2\alpha} - \frac{3}{2}}}{\Gamma(\frac{r+\alpha}{2\alpha})} = \frac{\sqrt{\pi}\sigma}{2\sqrt{\alpha}\Gamma(\frac{r+\alpha}{2\alpha})} \exp\left(\frac{\alpha x^2}{\sigma^2}\right) \left(\frac{\sqrt{\alpha} x}{\sigma}\right)^{\frac{r}{\alpha} - 1}.$$

For sake of compactness of notation we define the asymptotics

$$G_0(x) = \frac{\sqrt{\pi}}{\Gamma(\frac{r}{2\alpha})} \exp\left(\frac{\alpha x^2}{\sigma^2}\right) \left(\frac{\sqrt{\alpha} x}{\sigma}\right)^{\frac{r}{\alpha} - 1}$$

$$G_1(x) = \frac{\sqrt{\pi}\sigma}{2\sqrt{\alpha}\Gamma(\frac{r+\alpha}{2\alpha})} \exp\left(\frac{\alpha x^2}{\sigma^2}\right) \left(\frac{\sqrt{\alpha} x}{\sigma}\right)^{\frac{r}{\alpha} - 1}$$

and notice that

$$G_1(x) = BG_0(x) \quad \text{where} \quad B = \frac{\Gamma\left(\frac{r}{2\alpha}\right)\sigma}{2\sqrt{\alpha}\Gamma(\frac{r+\alpha}{2\alpha})}.$$

Suppose that

$$\lim_{x\to\infty} \Big(C_0 F_0(x) - C_1 F_1(x)\Big) = 0.$$

Then we have

$$\begin{aligned} \frac{C_0}{B} - C_1 &= \lim_{x\to\infty} \left(\frac{C_0 F_0(x)}{BG_0(x)} - \frac{C_1 F_1(x)}{G_1(x)}\right) \\ &= \lim_{x\to\infty} \left(\frac{C_0 F_0(x) - C_1 F_1(x)}{G_1(x)}\right) \\ &= 0. \end{aligned}$$

Thus we see that for the linear combination to go to zero it is necessary, but not sufficient, for

$$C_0 = BC_1 = \frac{\Gamma\left(\frac{r}{2\alpha}\right)\sigma}{2\sqrt{\alpha}\Gamma(\frac{r+\alpha}{2\alpha})}C_1.$$

We look at another special function known as Tricomi's confluent hypergeometric function or the confluent hypergeometric function of the second kind, see [24], defined

$$U(a,b,z) = \frac{\Gamma(1-b)}{\Gamma(a-b+1)}M(a,b,z) + \frac{\Gamma(b-1)}{\Gamma(a)}z^{1-b}M(a-b,2-b,z).$$

It is straightforward to verify that with the prescribed relationship between C_0 and C_1,

$$C_0F_0(x) - C_1F_1(x) = CU\left(\frac{r}{2\alpha}, \frac{1}{2}, \frac{\alpha x^2}{\sigma^2}\right)$$

for some constant C. Tricomi's function has the asymptotic

$$U(a,b,z) \sim z^{-a}\left(1 + O\left(\frac{1}{z}\right)\right),$$

see [24] for details. Since $a = \frac{r}{2\alpha} > 0$ we have that

$$\lim_{x\to\infty} U\left(\frac{r}{2\alpha}, \frac{1}{2}, \frac{\alpha x^2}{\sigma^2}\right) = 0$$

as desired.

It is interesting to note that an analytic extension of the Hermite polynomials [6] is provided by

$$H(a,x) = 2^n xU\left(\frac{1}{2} - \frac{a}{2}, \frac{3}{2}, x^2\right) = 2^n U\left(-\frac{a}{2}, \frac{1}{2}, x^2\right)$$

where the second equality comes from a Kummer transformation $U(a,b,z) = z^{1-b}U(1+a-b, 2-b, z)$, see ([31], p. 505). This provides an elegant and straitforward way of writing the desired function

$$U\left(\frac{r}{2\alpha}, \frac{1}{2}, \frac{\alpha x^2}{\sigma^2}\right) = \kappa H\left(-\frac{r}{\alpha}, \frac{\sqrt{\alpha}x}{\sigma}\right)$$

for an appropriate constant κ. We now have a bounded solution to the ODEs given by

$$f(x) = \begin{cases} cH\left(-\frac{r}{\alpha}, \frac{\sqrt{\alpha}x}{\sigma}\right) & ; x > b \\ K - x & ; x \leq b \end{cases}$$

where c is an arbitrary constant. Again, we mandate that the solution be C^1 at the boundary b. Matching derivative and function value at the boundary point will completely determine both c and the boundary b.

All that remains is prove the optimality of the solution.

Theorem 5.2.1 (Optimality).

Suppose that the threshold $b < \frac{r}{r+\alpha}K$ has been found such that the unique solution to

$$-\alpha x f_x(x) + \frac{1}{2}\sigma^2 f_{xx}(x) = rf(x) \quad x > b$$

$$f(x) = K - x \quad x \leq b$$

is C^1 on its domain and bounded above b. Further, suppose that $f(x) \geq (K-x)^+$. Then the solution $f(x)$ and the stopping time $\tau_D = \{t \geq 0 : X_t \in D\}$ correspond to the value function

$$V(x) = \sup_{\tau} E^x[e^{-r\tau}(K - X_\tau)^+]$$

and its optimal stopping time, i.e.,

$$f(x) = V(x) = E^x[e^{-r\tau_D}(K - X_{\tau_D})^+].$$

Proof. To do this we will apply Ito's formula (2.1.1) to $e^{-rt}f(X_t)$ to get

$$e^{-rt}f(X_t) = f(X_0) + \int_0^t e^{-rs}\left(L_X[f](X_s) - rf(X_s)\right) ds + Martingale.$$

We need to show that

$$\begin{aligned}\Phi(x) &= (L_X[f](x) - rf(x))\, ds \\ &= -\alpha x f_x(x) + \frac{1}{2}\sigma^2 f_{xx}(x) - rf(x) \\ &\leq 0.\end{aligned}$$

From (5.2), we see that $\Phi(x) = 0$ for $x \in C$. When $x \in D$ we get $f(x) = (K - x)$ and

$$\begin{aligned}\Phi(x) &= \alpha x - r(K - x) \\ &= (r + \alpha)x - rK \\ &\leq (r + \alpha)b - rK \\ &\leq (r + \alpha)\frac{r}{r + \alpha}K - rK \\ &= 0.\end{aligned}$$

Next, observe that

$$e^{-rt}(K - X_t)^+ \leq e^{-rt} f(X_t) \leq f(X_0) + \textit{Martingale}$$

and thus for any stopping time τ

$$E^x[e^{-r(\tau \wedge t)}(K - X_{\tau \wedge t})^+] \leq f(x)$$

and by Fatou's Lemma,

$$E^x[e^{-r\tau}(K - X_\tau)^+] \leq f(x).$$

Finally, we see that $e^{-r(\tau_D \wedge t)} f(X_{\tau_D \wedge t})$ is a martingale due to $\Phi(x) = 0$ when in region C. It is a bounded martingale since $f(x)$ is a bounded function on C. Using the optional sampling theorem (2.2.3) we get

$$E^x[e^{-r\tau_D} f(X_{\tau_D})] = E^x[e^{-r\tau_D}(K - X_{\tau_D})^+] = f(x)$$

and optimality is proven. □

5.3 The American Put in an Ornstein-Uhlenbeck Model with Non-Zero Mean

Next we will attempt to price the American put in model whose mean is non-zero. This is a much more realistic and general assumption, since no commodities are mean reverting to zero. The Ornstein-Uhlenbeck process with mean reversion to m is given by

$$dX_t = -\alpha(X_t - m)dt + \sigma dB_t$$

the solution of which is obtained by applying Ito's formula 2.1.1 to $e^{\alpha t}X_t$ to get

$$d(e^{\alpha t}X_t) = \alpha m e^{\alpha t} + e^{\alpha t}\sigma dB_t.$$

So,

$$X_t = X_0 e^{-\alpha t} + m(1 - e^{-\alpha t}) + \sigma \int_0^t e^{-\alpha(t-s)} dB_s.$$

The value of an American put is given by the optimal stopping problem

$$V(x) = \sup_{\tau} E^x[e^{-r\tau}(K - X_\tau)^+].$$

The same procedure from the previous section will be followed to verify an optimal stopping time of the form

$$\tau = \inf\{t \geq 0 : X_t \leq b\}$$

for some threshold b and to show that

$$L_X[V](x) = rV(x) \quad for \quad x > b$$
$$V(x) = (K - x)^+ \quad for \quad x \leq b.$$

This leads to the ODE

$$-\alpha(x - m)f_x(x) + \frac{1}{2}\sigma^2 f_{xx}(x) = rf(x) \quad x > b \tag{5.3}$$
$$f(x) = K - x \quad x \leq b.$$

Here we do a change of variables

$$y = x - m$$
$$g(y) = f(y + m)$$

arriving at

$$-\alpha y g_y(y) + \frac{1}{2}\sigma^2 g_{yy}(y) = rg(y) \quad y > b - m.$$

This is solved identically to (5.2), and we arrive at the solution

$$g(y) = cH\left(-\frac{r}{\alpha}, \frac{\sqrt{\alpha}y}{\sigma}\right) \quad for \quad y > b - m.$$

So,

$$f(x) = \begin{cases} cH\left(-\frac{r}{\alpha}, \frac{\sqrt{\alpha}(x-m)}{\sigma}\right) & ; x > b \\ K - x & ; x \le b \end{cases}$$

where c is an arbitrary constant. Again, we mandate that the solution be C^1 at the boundary b. Matching derivative and function value at the boundary point will completely determine both c and the boundary b.

All that remains is prove the optimality of the solution.

Theorem 5.3.1 (Optimality).

Suppose that the threshold $b < \frac{rK+\alpha m}{r+\alpha}$ has been found such that the unique solution to

$$-\alpha x f_x(x) + \frac{1}{2}\sigma^2 f_{xx}(x) = rf(x) \quad x > b$$
$$f(x) = K - x \quad x \le b$$

is C^1 on its domain. Further, suppose that $f(x) \ge (k-x)^+$. Then the solution $f(x)$ and the stopping time $\tau_D = \{t \ge 0 : X_t \in D\}$ correspond to the value function

$$V(x) = \sup_{\tau} E^x[e^{-r\tau}(K - X_\tau)^+]$$

and its optimal stopping time, i.e.,

$$f(x) = V(x) = E^x[e^{-r\tau_D}(K - X_{\tau_D})^+].$$

Proof. To do this we will apply Ito's formula (2.1.1) to $e^{-rt}f(X_t)$ to get

$$e^{-rt}f(X_t) = f(X_0) + \int_0^t e^{-rs}\left(L_X[f](X_s) - rf(X_s)\right) ds + Martingale.$$

We would like to show that

$$\begin{aligned}\Phi(x) &= (L_X[f](x) - rf(x))\, ds \\ &= -\alpha x f_x(x) + \frac{1}{2}\sigma^2 f_{xx}(x) - rf(x) \\ &\leq 0.\end{aligned}$$

From (5.3), we immediately see that $\Phi(x) = 0$ for $x > b$. When $x \leq b$ we get $f(x) = (K - x)$ and

$$\begin{aligned}\Phi(x) &= \alpha(x - m) - r(K - x) \\ &= (r + \alpha)x - rK \\ &\leq (r + \alpha)b - (rK + \alpha m) \\ &\leq (r + \alpha)\frac{rK + \alpha m}{r + \alpha} - (rK + \alpha m) \\ &= 0.\end{aligned}$$

Next, observe that

$$e^{-rt}(K - X_t)^+ \leq e^{-rt} f(X_t) \leq f(X_0) + Martingale$$

and thus for any stopping time τ

$$E^x[e^{-r(\tau \wedge t)}(K - X_{\tau \wedge t})^+] \leq f(x)$$

and applying Fatou's Lemma,

$$E^x[e^{-r\tau}(K - X_\tau)^+] \leq f(x).$$

Finally we see that $e^{-r(\tau_D \wedge t)} f(X_{\tau_D \wedge t})$ is a martingale since $\Phi(x) = 0$ when in region C. Further, it is bounded martingale since $f(x)$ is a bounded function on C. So using the optional sampling theorem (2.2.3) we get

$$E^x[e^{-r\tau_D} f(X_{\tau_D})] = E^x[e^{-r\tau_D}(K - X_{\tau_D})^+] = f(x)$$

and optimality is proven. □

5.4 The American Put in an Ornstein-Uhlenbeck Model with 2-State Markov Modulation

Here we modify the model Ornstein-Uhlenbeck model and incorporate Markov modulation

$$dX_t = -\alpha(\xi_t)X_t dt + \sigma(\xi_t)dB_t,$$

where $\xi : \Omega \times [0,\infty) \to \{1,2\}$ is a Markov chain defined by the infinitesimal generating matrix

$$Q = \begin{pmatrix} -\lambda_1 & \lambda_1 \\ \lambda_2 & -\lambda_2 \end{pmatrix}.$$

The value of an American put is given by the optimal stopping problem

$$V(x,i) = \sup_{\tau} E^{(x,i)}[e^{-\int_0^\tau r(\xi_s)ds}(K - X_\tau)^+].$$

The same procedure from section 5.2 will be followed to verify an optimal stopping time of the form

$$\tau^* = \inf\{t \geq 0 : X_t \leq b(\xi_t)\}$$

for two thresholds $b(1) = d_1$ and $b(2) = d_2$ and to show that

$$L_{X,\xi}[V](x,i) = r(i)V(x,i) \quad for \quad x > d_i$$

$$V(x) = (K-x)^+ \quad for \quad x \leq d_i.$$

This leads to the ODE system

$$-\alpha(i)xf_x(x,i) + \frac{1}{2}\sigma^2(i)f_{xx}(x,i) + \sum_{j=1}^{2} q_{ij}f(x,j) = r(i)f(x,i) \quad x > d_i$$

$$f(x,i) = K - x \quad x \leq d_i.$$

We will separate this into 3 regions supposing, with out loss of generality, that $d_1 > d_2$. The three regions to be considered are $(-\infty, d_2]$, $(d_2, d_1]$, and (d_1, ∞). When

$x \leq d_2$ we get trivially that $f(x,i) = K - x$ for $i = 1, 2$. When $d_2 < x \leq d_1$ we get that $f(x,1) = K - x$ and that $f(x,2)$ satisfies the following non-homogeneous ODE:

$$-\alpha(2)xf_x(x,2) + \frac{1}{2}\sigma^2(2)f_{xx}(x,2) - (r(2) + \lambda_2)f(x,2) = -\lambda_2(K - x).$$

First, we solve the homogeneous equations exactly as in section 5.2 and arrive at the solution

$$c_0 M\left(\frac{r_2 + \lambda_2}{2\alpha_2}, \frac{1}{2}, \frac{\alpha_2 x^2}{\sigma_2^2}\right) + c_1 x M\left(\frac{r_2 + \lambda_2 + \alpha_2}{2\alpha_2}, \frac{3}{2}, \frac{\alpha_2 x^2}{\sigma_2^2}\right) \quad for \quad d_2 < x \leq d_1.$$

Next we find a particular solution of the form $\gamma_1 + \gamma_2 x$ arriving at the full solution of $f(x,2)$ below d_1

$$f(x,2) = \begin{cases} c_0 M\left(\frac{r_2+\lambda_2}{2\alpha_2}, \frac{1}{2}, \frac{\alpha_2 x^2}{\sigma_2^2}\right) + c_1 x M\left(\frac{r_2+\lambda_2+\alpha_2}{2\alpha_2}, \frac{3}{2}, \frac{\alpha_2 x^2}{\sigma_2^2}\right) + \gamma_1 + \gamma_2 x & for \quad d_2 < x \leq d_1 \\ K - x & for \quad x \leq d_2 \end{cases}.$$

For $x > d_1$, the problem becomes considerably harder. Here, both $f(x,1)$ and $f(x,2)$ are unknown and we have the following 2 dimensional ODE

$$\begin{cases} -\alpha(1)xf_x(x,1) + \frac{1}{2}\sigma^2(1)f_{xx}(x,1) - (r(1) + \lambda_1)f(x,1) + \lambda_1 f(x,2) = 0 \\ -\alpha(2)xf_x(x,2) + \frac{1}{2}\sigma^2(2)f_{xx}(x,2) - (r(2) + \lambda_2)f(x,2) + \lambda_2 f(x,1) = 0. \end{cases}$$

We will attempt to solve this ODE by the technique presented in section 5.2. Suppose $f(x,1)$ and $f(x,2)$ have the following expansion above d_1

$$f(x,1) = \sum_{k=0}^{\infty} a_k x^k$$
$$f(x,2) = \sum_{k=0}^{\infty} b_k x^k.$$

Using these expansions in the first ODE yields

$$
\begin{aligned}
&-\alpha_1 x \sum_{k=0}^{\infty}(k+1)a_{k+1}x^k + \frac{\sigma_1^2}{2}\sum_{k=0}^{\infty}(k+2)(k+1)a_{k+2}x^k - (r_1+\lambda_1)\sum_{k=0}^{\infty} a_k x^k + \sum_{k=0}^{\infty}\lambda_1 b_k x^k \\
&= \sum_{k=0}^{\infty}\left(-\alpha_1 k a_k x^k + \frac{\sigma_1^2}{2}(k+2)(k+1)a_{k+2}x^k - (r_1+\lambda_1)a_k x^k + \lambda_1 b_k x^k\right) \\
&= \sum_{k=0}^{\infty}\left(-\alpha_1 k a_k + \frac{\sigma_1^2}{2}(k+2)(k+1)a_{k+2} - (r_1+\lambda_1)a_k + \lambda_1 b_k\right)x^k \\
&= 0.
\end{aligned}
$$

From this we get the recursive relation

$$a_{k+2} = \frac{2}{\sigma_1^2}\frac{(r_1+\lambda_1+\alpha_1 k)a_k - \lambda_1 b_k}{(k+2)(k+1)}$$

and similarly

$$b_{k+2} = \frac{2}{\sigma_2^2}\frac{(r_2+\lambda_2+\alpha_2 k)b_k - \lambda_2 a_k}{(k+2)(k+1)}.$$

This can be restated in matrix form

$$C_{k+2} = A_k C_k$$

where

$$C_k = \begin{pmatrix} a_k \\ b_k \end{pmatrix} \qquad \text{and} \qquad A_k = \frac{2}{(k+2)(k+1)}\begin{pmatrix} \frac{r_1+\lambda_1+k\alpha_1}{\sigma_1^2} & -\frac{\lambda_1}{\sigma_1^2} \\ -\frac{\lambda_2}{\sigma_2^2} & \frac{r_2+\lambda_2+k\alpha_2}{\sigma_2^2} \end{pmatrix}.$$

From this we get

$$
\begin{aligned}
C_{2n} &= \left(\prod_{k=0}^{n-1} A_{2k}\right) C_0 \\
C_{2n+1} &= \left(\prod_{k=0}^{n-1} A_{2k+1}\right) C_1.
\end{aligned}
$$

If we let

$$A_k = \frac{2}{(k+2)(k+1)}A + kB$$

where

$$A = \begin{pmatrix} \frac{r_1+\lambda_1}{\sigma_1^2} & -\frac{\lambda_1}{\sigma_1^2} \\ -\frac{\lambda_2}{\sigma_2^2} & \frac{r_2+\lambda_2}{\sigma_2^2} \end{pmatrix} \qquad \text{and} \qquad B = \begin{pmatrix} \frac{\alpha_1}{\sigma_1^2} & 0 \\ 0 & \frac{\alpha_2}{\sigma_2^2} \end{pmatrix},$$

we have an elegant expression for C_k

$$C_{2n} = \frac{2^n}{(2n)!}\left(\prod_{k=0}^{n-1}(A + kB)\right)C_0$$

$$C_{2n+1} = \frac{2^n}{(2n+1)!}\left(\prod_{k=0}^{n-1}\left(\frac{2A+B}{2} + kB\right)\right)C_1.$$

Define

$$\mathcal{A}_{2n} = \frac{2^n}{(2n)!}\prod_{k=0}^{n-1}(A + kB)$$

$$\mathcal{A}_{2n+1} = \frac{2^n}{(2n+1)!}\prod_{k=0}^{n-1}\left(\frac{2A+B}{2} + kB\right).$$

Let $\mathcal{A}_n(i, j)$ represent the i, j element of the matrix $\mathcal{A}_n$. Now a compact form of $f(x, 1)$ for $x > d_1$ is presented

$$\begin{aligned} f(x,1) &= \sum_{k=0}^{\infty}\left(\mathcal{A}_{2k}(1,1)a_0 + \mathcal{A}_{2k}(1,2)b_0\right)x^{2k} + x\sum_{k=0}^{\infty}\left(\mathcal{A}_{2k+1}(1,1)a_1 + \mathcal{A}_{2k+1}(1,2)b_1\right)x^{2k} \\ &= a_0\sum_{k=0}^{\infty}\mathcal{A}_{2k}(1,1)x^{2k} + b_0\sum_{k=0}^{\infty}\mathcal{A}_{2k}(1,2)x^{2k} + a_1x\sum_{k=0}^{\infty}\mathcal{A}_{2k+1}(1,1)x^{2k} + b_1x\sum_{k=0}^{\infty}\mathcal{A}_{2k+1}(1,2)x^{2k} \end{aligned}$$

and similarly for $f(x, 2)$ when $x > d_1$

$$f(x,2) = a_0\sum_{k=0}^{\infty}\mathcal{A}_{2k}(2,1)x^{2k} + b_0\sum_{k=0}^{\infty}\mathcal{A}_{2k}(2,2)x^{2k} + a_1x\sum_{k=0}^{\infty}\mathcal{A}_{2k+1}(2,1)x^{2k} + b_1x\sum_{k=0}^{\infty}\mathcal{A}_{2k+1}(2,2)x^{2k}.$$

Conjecture 5.4.1.

There exists an appropriate choice of the parameters $\{a_0, a_1, b_0, b_1\}$ *such that*

$$a_0\sum_{k=0}^{\infty}\mathcal{A}_{2k}(1,1)x^{2k} + b_0\sum_{k=0}^{\infty}\mathcal{A}_{2k}(1,2)x^{2k} + a_1x\sum_{k=0}^{\infty}\mathcal{A}_{2k+1}(1,1)x^{2k} + b_1x\sum_{k=0}^{\infty}\mathcal{A}_{2k+1}(1,2)x^{2k}$$

$$a_0\sum_{k=0}^{\infty}\mathcal{A}_{2k}(2,1)x^{2k} + b_0\sum_{k=0}^{\infty}\mathcal{A}_{2k}(2,2)x^{2k} + a_1x\sum_{k=0}^{\infty}\mathcal{A}_{2k+1}(2,1)x^{2k} + b_1x\sum_{k=0}^{\infty}\mathcal{A}_{2k+1}(2,2)x^{2k}$$

are bounded functions in $\mathbb{R}$. With the appropriate choice of $\{a_0, a_1, b_0, b_1\}$, we will rename the above functions $\psi_1(x)$ and $\psi_2(x)$ respectively.

Relying on this conjecture, we have a solution bounded above d_1

$$f(x,1) = \begin{cases} a\psi_1(x) & x > d_1 \\ K - x & x \leq d_1 \end{cases}$$

and

$$f(x,2) = \begin{cases} b\psi_2(x) & x > d_1 \\ c_0 M\left(\frac{r_2+\lambda_2}{2\alpha_2}, \frac{1}{2}, \frac{\alpha_2 x^2}{\sigma_2^2}\right) + c_1 x M\left(\frac{r_2+\lambda_2+\alpha_2}{2\alpha_2}, \frac{3}{2}, \frac{\alpha_2 x^2}{\sigma_2^2}\right) + \gamma_1 + \gamma_2 x & d_2 < x \leq d_1 \\ K - x & x \leq d_2 \end{cases}.$$

We mandate that the solution be C^1 at the boundaries, and in so doing determine all unknown parameters $\{a, b, c_0, c_1\}$ and unknown boundaries d_1 and d_2. We now prove that such a solution is optimal.

Theorem 5.4.1 (Optimality).

Suppose that thresholds $d_2 < d_1 < K$ have been found such that the unique solution to

$$-\alpha(i)xf_x(x,i) + \frac{1}{2}\sigma^2(i)f_{xx}(x,i) + \sum_{j=1}^{2} q_{ij}f(x,j) = rf(x,i) \quad x > d_i \tag{5.4}$$

$$f(x,i) = K - x \quad x \leq d_i$$

is C^1 on its domain. Further, suppose that the following assumptions hold

1. *$f(x,i) \geq (k-x)^+$*
2. *$f(x,i) \leq M$ for $i = 1, 2$.*
3. *$d_i \leq \frac{r_i K - \lambda_i M}{\alpha_i + r_i}$.*

Then the solution $f(x,i)$ and the stopping time $\tau^* = \{t \geq 0 : X_t \leq b(\xi_t)\}$ correspond to the value function

$$V(x,i) = \sup_{\tau} E^{(x,i)}[e^{-\int_0^\tau r(\xi_s)ds}(K - X_\tau)^+]$$

and its optimal stopping time, i.e.,

$$f(x,i) = V(x,i) = E^{(x,i)}[e^{-\int_0^{\tau^*} r(\xi_s)ds}(K - X_{\tau^*})^+].$$

Proof. To do this we will apply Ito's formula (2.1.1) to $e^{-\int_0^{\tau} r(\xi_s)ds} f(X_t)$ to get

$$e^{-\int_0^{\tau} r(\xi_s)ds} f(X_t) = f(X_0) + \int_0^t e^{-\int_0^s r(\xi_u)du}\left(L_{X,\xi}[f](X_s,\xi_s) - r(\xi_s)f(X_s,\xi_s)\right)ds + Martingale.$$

We need to show that

$$\begin{aligned}\Phi(x,i) &= (L_{X,\xi}[f](x,i) - r(i)f(x,i))\,ds\\ &= -\alpha(i)xf_x(x,i) + \frac{1}{2}\sigma(i)^2 f_{xx}(x,i) - r(i)f(x,i) + \sum_{j=1}^{2} q_{ij}f(x,j)\\ &\leq 0.\end{aligned}$$

From (5.2), we see that $\Phi(x) = 0$ for $x \in C = \{(x,i) : x > b_i\}$. When $x \in D = \{(x,i) : x \leq b_i\}$ we get $f(x) = (K - x)$ and

$$\begin{aligned}\Phi(x,i) &= \alpha_i x - r_i(K - x) + \lambda_i(f(x,i) - (K - x))\\ &\leq (r_i + \alpha_i)x - r_iK + \lambda_i M\\ &\leq (r_i + \alpha_i)d_i - r_iK + \lambda_i M\\ &\leq (r_i + \alpha_i)\frac{r_iK - \lambda_i M}{\alpha_i + r_i} - r_iK + \lambda_i M\\ &= 0.\end{aligned}$$

Next, observe that

$$e^{-\int_0^t r(\xi_s)ds}(K - X_t)^+ \leq e^{-\int_0^t r(\xi_s)ds} f(X_t, i) \leq f(X_0,\xi_0) + Martingale$$

and thus for any stopping time τ

$$E^{(x,i)}[e^{-\int_0^{\tau\wedge t} r(\xi_s)ds}(K - X_{\tau\wedge t})^+] \leq f(x)$$

whence by Fatou's Lemma,

$$E^{(x,i)}[e^{-\int_0^{\tau} r(\xi_s)ds}(K - X_\tau)^+] \leq f(x).$$

Finally, we see that $e^{-\int_0^{\tau^* \wedge t} r(\xi_s)ds} f(X_{\tau^* \wedge t})$ is a martingale due to $\Phi(x) = 0$ in the region C. Further, it is bounded martingale since $f(x, i)$ is a bounded function on C. By the optional sampling theorem (2.2.3) we get

$$E^{(x,i)}[e^{-\int_0^{\tau^*} r(\xi_s)ds} f(X_{\tau^*})] = E^{(x,i)}[e^{-\int_0^{\tau^*} r(\xi_s)ds}(K - X_{\tau^*})^+] = f(x, i)$$

which concludes the proof of optimality. □

REFERENCES

[1] Martin Baxter and Andrew Rennie. *An Introduction to Derivative Pricing*. University Press, Cambridge, New York, 1996.

[2] Fischer Black and Myron Scholes. The pricing of options and corporate liabilities. *Journal of Political Economy*, 81(3):637–654, 1973.

[3] John Buffington and Robert J. Elliott. American options with regime switching. *International Journal of Theoretical and Applied Finance*, 5(5):497–514, 2002.

[4] David S. Bunch and Herb Johnson. The american put option and its critical stock price. *The Journal of Finance*, 55(5):2333–2356, 2000.

[5] Arun Chockalingam and Kumar Muthuraman. American options under stochastic volatility. *Operations Research*, 59(4):793–809, 2011.

[6] NIST Digital Library of Mathematical Functions. http://dlmf.nist.gov/13.6, Release 1.0.5 of 2012-10-01. Online companion to [25].

[7] Robert J. Elliott and P. Ekkehard Kopp. *Mathematics of Financial Markets*. Springer, New York, 2005.

[8] Robert J. Elliottand, Lakhdar Aggoun, and John B. Moore. *Hidden Markov Models Estimation and Control*. Springer-Verlag, New York, 1995.

[9] Stewart N. Ethier and Thomas G. Kurtz. *Markov Processes Characterization and Convergence*. John Wiley & Sons, New York, 1986.

[10] X. Guo. An explicit solution to an optimal stopping problem with regime switching. *Journal of Applied Probability*, 38:464–481, 2001.

[11] X. Guo and Q. Zhang. Closed-form solutions for american put options with regime switching. *SIAM J. Appl. Math*, 64(6):2034–2049, 2004.

[12] X. Guo and Q. Zhang. Optimal selling rules in a regime switching model. *Transactions on Automatic Control*, 50(9):1450–1455, 2005.

[13] A. Jobert and L. C. G. Rogers. Option pricing with markov-modulated dynamics. *SIAM J. Control Optim.*, 44(6):2063–2078, 2006.

[14] Edward P. C. Kao. *An Introduction to Stochastic Processes*. Duxbury Press, Boston, 1997.

[15] Ioannis Karatzas and Steven Shreve. *Methods of Mathematical Finance*. Springer, New York, 1998.

[16] Samuel Karlin and Howard Taylor. *A First Course in Stochastic Processes*. Academic Press, New York, 1975.

[17] Samuel Karlin and Howard Taylor. *A Second Course in Stochastic Processes*. John Wiley & Sons, New York, 1981.

[18] Fima C Klebaner. *Introduction to Stochastic Calculus with Applications*. Imperial College Press, London, 2012.

[19] Y. J. Liu, G. Yin, and Q. Zhang. Computational methods for pricing american put options. *Journal of Optimization Theory and Applications*, 127(2):389–410, 2005.

[20] Henry P. McKean. Appendix: A free boundary problem for the heat equation arising from a problem of mathematical economics. *Industrial Management Review*, 6(2):32–39, 1965.

[21] Robert C. Merton. Theory of rational option pricing. *Bell Journal of Economics and Management Science*, 4(1):141–183, 1973.

[22] Marek Musiela and Marek Rutkowski. *Martingale Methods in Finanacial Modelling*. Springer, New York, 2005.

[23] Bernt Øksendal. *Stochastic Differential Equations*. Springer-Verlag, New York, 2003.

[24] Keith Oldham, Jan Myland, and Jerome Spanier. *An Atlas of Functions*. Springer, New York, 2009.

[25] F. W. J. Olver, D. W. Lozier, R. F. Boisvert, and C. W. Clark, editors. *NIST Handbook of Mathematical Functions*. Cambridge University Press, New York, NY, 2010. Print companion to [6].

[26] Goran Peskir and Albert Shiryaev. *Optimal Stopping and Free-Boundary Problems*. Birkhauser Verlag, Boston, 2006.

[27] Philip Protter. *Stochastic Integration and Differential Equations*. Springer Verlag, New York, 1992.

[28] L. C. G. Rogers and D. Williams. *Diffusions, Markov Processes, and Martingales Volume One: Foundations*. John Wiley & Sons, New York, 1994.

[29] L. C. G. Rogers and D. Williams. *Diffusions, Markov Processes, and Martingales Volume 2: Ito Calculus*. Cambridge University Press, New York, 2000.

[30] Sheldon M. Ross. *Stochastic Processes*. John Wiley & Sons, New York, 1995.

[31] Lucy Joan Slater. *Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables*, chapter Confluent Hypergeometric Functions. John Wiley & Sons, New York, 1972.

[32] G. Yin and C. Zhu. *Hybrid Switching Diffusions: Properties and Applications*. Springer, New York, 2010.

[33] Q. Zhang. Stock trading: An optimal selling rule. *SIAM Journal of Optimal Control*, 40(1):64–87, 2001.

[34] Z. Zhang, G. Yin, and Z. Liang. A stochastic algorithm for american lookback put options. *Stochastic Analysis and Applications*, 29(2):332–351, 2011.

BIOGRAPHICAL STATEMENT

Thomas Seaquist holds a B.S. degree in mathematics from the University of Texas at Arlington. He has completed research in the modeling of an invasive brain cancer, Glioblastoma-Multiforme at Worcester Polytechnic Institute. Subsequently, he researched the vector transmitted Chagas disease at Arizona State University which was continued in his final year as an undergraduate at the University of Texas of Arlington. His current research interest lies in optimal stopping and Markov modulation.