August 2020

Analysis of the Continuity of the Value Function of an Optimal Stopping Problem

Samuel Morris Nehls
University of Wisconsin-Milwaukee

Analysis of the Continuity of the Value Function of an Optimal Stopping Problem

by

Samuel Morris Nehls

A Dissertation Submitted in
Partial Fulfillment of the
Requirements for the Degree of

Doctor of Philosophy
in Mathematics

at

The University of Wisconsin-Milwaukee
August 2020

ABSTRACT

Analysis of the Continuity of the Value Function of an Optimal Stopping Problem

by

Samuel Morris Nehls

The University of Wisconsin-Milwaukee, August 2020
Under the Supervision of Professor Richard Stockbridge

In order to study model uncertainty of an optimal stopping problem of a stochastic process with a given state dependent drift rate and volatility, we analyze the effects of perturbing the parameters of the problem. This is accomplished by translating the original problem into a semi-infinite linear program and its dual. We then approximate this dual linear program by a countably constrained sub-linear program as well as an infinite sequence of finitely constrained linear programs. We find that in this framework the value function will be lower semi-continuous with respect to the parameters. If in addition we restrict ourselves to a compact set of constraints and add smoothness conditions to the gain function, we have full continuity of the value function.

Table of Contents

List of Figures

List of Symbols

A	infinitesimal generator
$\mathscr{B}$	Borel σ-algebra
$\mathcal{D}$	dual linear program
$\mathbb{E}$	expectation
$\mathbb{E}_x$	conditional expectation, $\mathbb{E}[\,\cdot \mid X_0 = x]$
$\mathbb{F}$	filtration
$\mathbb{F}^X$	natural filtration generated by X
$\mathcal{F}$	σ-algebra on the sample space
$\mathcal{F}_t$	sub-σ-algebra of $\mathcal{F}$, part of a filtration
$\mathbb{L}_X$	infinitesimal generator
LP	linear program
$\mathcal{P}$	primal linear program
$\mathbb{P}$	probability measure
$\mathbb{R}_+$	the interval $[0, \infty)$
$\mathbb{R}^n_+$	$[0, \infty) \times \cdots \times [0, \infty)$
τ	stopping time
θ	a general parameter
$V(x)$	the value function
W_t	standard Brownian motion process, Wiener process
X_t	the stochastic process, also $X(t)$, $X(t, \omega)$, $X_t(\omega)$
x	initial position of the process, X_0

ACKNOWLEDGEMENTS

First of all I want to thank my advisor Richard Stockbridge for his years of invaluable advice and guidance and for convincing me to abandon the study of Algebra in favor of Probability and almost surely saving me from years of anguish.

I want to acknowledge my committee members, Vytaras Brazauskas, Istvan Lauko, Wei Wei, and Chao Zhu and thank them for the time they spent reading this document and for the advice they have given to improve it.

I want to thank some of my other professors, including Allen Bell, Rebecca Bourn, Craig Guilbault, and Jeb Willenbring for broadening my mathematical knowledge and for showing me how a great instructor should conduct themself.

I also want to thank my colleagues at University of Wisconsin - Milwaukee, including Durham Evans, Alison Becker, Martin Vieten, and Charles Beer for their collaboration and cooperation and for listening to my problems, both mathematical and personal.

Lastly, I want to thank my wife Brittany for her partnership and companionship, for playing games with me, for providing welcome distractions when the work became too great, and for parenting our wonderful children. I want to thank my son Simon for keeping me up late at night so I could find time to work. My daughter Charli requests that I not thank her for anything, but I will acknowledge her for being the most kind and caring daughter I could ever ask for.

Chapter 1

Introduction

Stochastic optimal control theory has many different flavors, and is a generally well understood area of probability. One of the most well studied implementations is optimal stopping. In this case a stochastic process is allowed to proceed undisturbed for either a finite or infinite length of time and the only control we have is to choose when to stop the process, if ever. The goal here is to implement a stopping rule that optimizes the expected value of some objective function, most commonly maximizing revenue or minimizing cost. This can be difficult to decide due to the inherent randomness in a stochastic process. That is, we may know how the process is expected to behave on average but we cannot predict its exact path.

The control theory currently used works under a few basic assumptions: that we have the correct model, and that the parameters in our model are accurate. Now let us suppose that these assumptions are not true. Perhaps the parameters in our model cannot be measured accurately, or it is expensive and time consuming to measure them accurately but a close approximation can be done cheaply. We would now be implementing a solution that is designed for different parameters and may not be optimal for the true situation at hand. In this case, how badly will this affect the outcome?

The first step to study the impact of the parameters on our optimal stopping problem is to analyze the parameters of the model that affect the optimal value function. We will study the continuity properties of the value function when we allow perturbations of these parameters. We will do this by leveraging the relatively well known structure of the linear program, both finite and infinite.

The motivating paper to this research was co-written by Kurt Helmes and my advisor

Richard Stockbridge (Helmes and Stockbridge, 2010). The major work of this paper was analyzing the optimal stopping time of a one-dimensional diffusion. This stopping time should maximize a given reward function which consists of a discounted running reward and a terminal reward which is obtained upon stopping the process. This sounds simple, but since we are working under a probability space we only know the dynamics of the process, which tells us the expected future behavior. Therefore this optimal stopping time should maximize the expected reward and we acknowledge that for certain individual outcomes it will not be optimal. To be precise our one dimensional diffusion must satisfy the following stochastic differential equation (SDE):

$$dX_t = \mu(X_t)\,dt + \sigma(X_t)\,dW_t, \qquad X_0 = x \tag{1.1}$$

and the goal of the optimal stopping problem is to find a stopping rule τ and a value function V which satisfies

$$\begin{aligned} V(x) &= \sup_{\tau\in\mathcal{T}} J(\tau;x) \\ &= \sup_{\tau\in\mathcal{T}} \mathbb{E}\left[\int_0^\tau e^{-\alpha s} r(X_s)\,ds + e^{-\alpha\tau} g(X_\tau)\mathbf{1}_{\tau<\infty}\right]. \end{aligned} \tag{1.2}$$

The general idea will be to formulate the value function as the objective function in a linear program which has probability measures as its variables. Then using strong duality, the optimal value of that primal linear program (LP) will be equal to the optimal value of the dual LP which is a semi-infinite linear program with three variables and uncountably many constraints. We will then argue that because of the nature of this dual LP we can achieve the same objective function by using a countable subset of the original constraints. Then we will only have to consider a finite number of constraints, since the continuity of the value function in finite linear programs is fairly well understood and we have several conditions that will grant us continuity of our value function. Finally, we will argue that because of the strong duality property of the original LP, these value functions of the finite LP and dual LP should converge to the value function of the LP on the countable constraint system, which should then be equal to our original value function. Additionally we will find that under some additional structure of the infinite dual program, we will have more continuity properties.

So, to summarize:

- Embed optimal stopping problem into a semi-infinite linear program with infinitely many variables

- Construct the semi-infinite dual linear program with infinitely many constraints
- Construct a sub-linear dual program with countably many constraints
- Estimate countable dual program with a sequence of finitely constrained dual linear programs
- Prove convergence of the sequence and determine continuity of the value function in the parameter space

Chapter 2

Literature Review

2.1 Solution Methods for Optimal Stopping Problems

In this paper we will work with a continuous-time stochastic process (one-dimensional diffusion to be precise) $X_t(\omega) : \mathbb{R}^+ \times \Omega \to \mathbb{R}$ defined on a filtered probability space $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P})$. If not specified we can take $\mathbb{F}$ to be the natural filtration $\mathbb{F}^X$ generated by the process X, that is, $\mathbb{F}^X = \sigma\{X_s^{-1}(A) \,|\, A \in \mathscr{B}(\mathbb{R}), 0 \leq s < \infty\}$. Alternatively, $\mathbb{F}^X = \sigma(\bigcup_{t \geq 0} \mathcal{F}_t^X)$ where $\mathcal{F}_t^X = \sigma\{X_s^{-1}(A) \,|\, A \in \mathscr{B}(\mathbb{R}), 0 \leq s \leq t\}$.

Here we must mention the differing notations for the stochastic process. Properly, the process is a function of time t and an outcome ω so it is common to see $X(t, \omega)$ or $X(t; \omega)$. However many probabilists downplay the role of ω and frequently use $X(t)$ or X_t and some use these interchangably. We will use X_t or $X_t(\omega)$ if we wish to emphasize the role of the sample space.

This process proceeds undisturbed until some stopping time τ. A stopping time is a random variable defined by some stopping rule. For example, this could be to stop at a fixed (deterministic) time, or to stop when the process reaches a certain state. The latter is commonly called the first hitting time and we denote it as $H_E = \inf\{t \,|\, X_t \in E\}$ to describe the first time the process enters any state in E. We denote the set of all stopping times as $\mathcal{T} = \{\tau : \Omega \to \mathbb{R}^+ \,|\, \{\tau(\omega) \leq t\} \in \mathcal{F}_t \text{ for all } t\}$ and allow for the possibility that $\tau = \infty$. In the case when $\tau = \infty$ we define $f(X_\tau) = \limsup_{t \to \infty} f(X_t)$ for any function $f(x)$ where the limit is defined. Some authors require that stopping times be finite and instead use the term Markov times to include infinite stopping times. At τ, the gain $g(X_\tau)$ is earned. We typically require that $g(X_\tau)$ be bounded from below and integrable (at a minimum). A common way to achieve integrability is to require that $\mathbb{E}_x[\sup_t g(X_t)] < \infty$.

The goal of the optimal stopping problem is to find the stopping time which maximizes the gain as well as find the value function corresponding to that stopping time. We define the value function to be the maximal expected gain for the process given that $X_0 = x$. That is,

$$V(x) := \sup_{\tau \in \mathcal{T}} \mathbb{E}_x \left[g(X_\tau) \right]. \tag{2.1}$$

Additionally, each Markov process has an associated infinitesimal generator which acts on a function and is a way to understand the expected change in a process, similar to a derivative. This is denoted $\mathbb{L}_X$ or A depending on the author. We will primarily use $\mathbb{L}_X$. This Chapter will detail three existing methods for solving the optimal stopping problem.

2.1.1 Majorizing Excessive Functions

The first of three methods for solving an optimal stopping problem was detailed by Shiryaev (1978). Here we will take X_t to be a homogeneous, nonterminating, standard Markov process. That means that $X = (X_t, \mathcal{F}, \mathbb{P})$ satisfies the following conditions.

1. For each $A \in \mathcal{F}$, $\mathbb{P}_x(A)$ is a $\mathscr{B} - \textit{measurable}$ function for x;
2. For all $x \in \mathbb{R}, B \in \mathscr{B}, s, t \geq 0, \quad \mathbb{P}_x(X_{t+s} \in B \,|\, \mathcal{F}_t) = \mathbb{P}_{X_t}(x_s \in B)$;
3. $\mathbb{P}_x(X_0 = x) = 1, x \in E$;
4. For each $t > 0$ and $\omega \in \Omega$ there will be a unique $\omega' \in \Omega$ such that $X_s(\omega') = X_{s+t}(\omega)$.

Shiryaev works with processes that take values in any semi-compact space but we will work with real valued processes.

As a motivating example we consider functions that for all x satisfy $\mathbb{E}_x \left[g(X_\tau) \right] \leq g(x)$ for any τ. For these functions, the supremum in (2.1) is achieved for $\tau \equiv 0$ so in fact the optimal stopping problem can be solved. In considering all other gain functions we introduce the following definitions.

Definition 2.1.1. A function f is said to be *regular* (for the process X_t) if for any x and τ, $\mathbb{E}_x[f(X_\tau)]$ is defined and for any other stopping time σ such that $\mathbb{P}_x(\sigma \leq \tau) = 1$, then $\mathbb{E}_x[f(X_\tau)] \leq \mathbb{E}_x[f(X_\sigma)]$.

Definition 2.1.2. A function f is said to be *excessive* (for the process X_t) if for any x and $t \geq 0$, $\mathbb{E}_x[f(X_t)] \leq f(x)$.

Notice that we can take $\sigma \equiv 0$ and τ the fixed stopping time t and see that a regular function satisfies the condition to be excessive. In addition, for f excessive and bounded from below, $f(x)$ is regular. Additionally we can define $g(x) := \mathbb{E}_x[f(X_t)]$ and see that $g(X_s)$ is a version of $\mathbb{E}[f(X_t) \,|\, X_s]$. Then since f is excessive, $g(x) \leq f(x)$ and substituting $x = X_s$ we have $g(X_s) \leq f(X_s)$ which implies that $\mathbb{E}[f(X_t) \,|\, X_s] \leq f(X_s)$ so $f(X_t)$ is a supermartingale.

Definition 2.1.3. The excessive function $f(x)$ is called the *excessive majorant* of the function g if $f(x) \geq g(x)$ for all $x \in E$. Moreover, we call $f(x)$ the *smallest excessive majorant* (of g) if for any other excessive majorant $h(x)$ of g, we have $h(x) \geq f(x) \geq g(x)$ for all x.

In the case where g is bounded from below Shiryaev showed the existence of a smallest excessive majorant by construction.
First, set $s_n(x) = \sup\limits_{\tau \in \mathcal{T}_n} \mathbb{E}_x[g(X_\tau)]$, where $\mathcal{T}_n$ is the set of stopping times taking values in $\mathbb{N} \cdot 2^{-n}$ and $s(x) = \lim\limits_{n \to \infty} s_n(x)$.
Then by Lemma 3.3 in Shiryaev (1978) $s(x)$ is the smallest excessive majorant of $g(x)$.

Now we have the following inequality,

$$\mathbb{E}_x[g(X_\tau)] \leq \mathbb{E}_x[s(X_\tau)] \leq s(x).$$

The first is true because $s(x)$ is a majorant of g and the second is true because $s(x)$ is an excessive (and therefore regular) function. Then,

$$\sup_{\tau \in \mathcal{T}_n} \mathbb{E}_x[g(X_\tau)] \leq \sup_{\tau \in \mathcal{T}} \mathbb{E}_x[g(X_\tau)] \leq s(x)$$
$$s_n(x) \leq V(x) \leq s(x)$$

Since $s(x) = \lim_{n \to \infty} s_n(x)$, we have the following result.

Theorem 2.1.4. *If X_t is a time-homogeneous Markov process and $g(x)$ is bounded from below then the value function $V(x) = \sup\limits_{\tau \in \mathcal{T}} \mathbb{E}_x[g(X_\tau)]$ is the smallest excessive majorant of $g(x)$.*

So to solve our optimal stopping problem we must find the smallest excessive majorant $V(x)$ of the gain function. Then our stopping rule is as follows: $\tau^* = \inf\{t \geq 0 \,|\, V(X_t) = g(X_t)\}$. We call the set $C = \{x \in E \,|\, V(x) > g(x)\}$ the continuation region and $D = \{x \in E \,|\, V(x) = g(x)\}$ the stopping region.

2.1.2 Free-Boundary PDE

Peskir and Shiryaev (2006) describe a method to solve the optimal stopping problem as a free boundary problem. Here we take X_t to be a strong Markov process (right-continuous and left-continuous over stopping times). The idea behind this approach is that the value function V should satisfy the free-boundary conditions:

$$\begin{aligned} &\mathbb{L}_X V \le 0 \quad (V \text{ minimal}), \\ &V \ge g \quad (V > g \text{ on } C \text{ and } V = g \text{ on } D) \end{aligned} \tag{2.2}$$

where $\mathbb{L}_X$ is the infinitesimal generator of the process X_t and is defined as:

$$\mathbb{L}_X f(x) = \lim_{t \searrow 0} \frac{\mathbb{E}_x[f(X_t)] - f(x)}{t} \tag{2.3}$$

on functions $f : E \to \mathbb{R}$ and is only defined for those functions for which the limit exists.

Notice that $\mathbb{L}_X f(x)$ has a similar form as a derivative. However, since X_t and $f(X_t)$ are random processes, $\mathbb{L}_X f(x)$ gives the *expected* forward change of $f(X_t)$. In other words $\mathbb{E}_x[f(X_t)] \approx f(x) + t\,\mathbb{L}_X f(x)$ for small values of t. In the previous section we argued that the value function should be excessive, or $\mathbb{E}_x[f(X_t)] \le f(x)$. This implies that $\mathbb{L}_X f(x) \le 0$ which is the motivation behind this approach.

For functions $f \in \mathcal{C}^2$ it is known that the generator takes the form

$$\mathbb{L}_X f(x) = \lambda(x) f(x) + \mu(x) f'(x) + \frac{1}{2}\sigma^2(x) f''(x) + \int_{\mathbb{R}\setminus\{0\}} [f(y) - f(x) - (y-x) f'(x)] \nu(x, dy). \tag{2.4}$$

In this formulation λ is the killing (or creation if negative) rate of the process, μ is the drift rate, σ the diffusion (volatility), and ν is the compensator of the jump measure μ. In this paper we work with continuous processes without killing (Itô Processes) in which case the infinitesimal generator is simply

$$\mathbb{L}_X f(x) = \mu(x) f'(x) + \frac{1}{2}\sigma^2(x) f''(x). \tag{2.5}$$

Then in order to solve (2.2) we must find the appropriate function V and the region C. So long as condition 2 from (2.2) is satisfied, $E \setminus D = C$.

If, in addition we know that g is smooth and if X_t starts on the boundary of C it immediately enters $\text{int}(D)$ then we will have a smooth fit of V and g on the boundary of C. That is,

$$V(x)|_{\partial(C)} = g(x)|_{\partial(C)} \quad \text{and} \quad V'(x)|_{\partial(C)} = g'(x)|_{\partial(C)}.$$

If however we do not have that condition satisfied for X_t, then we will simply have continuous fit on the boundary. Knowing if V and g enjoy smooth fit can aid in the solution process. For example, if we know that V is in the class of degree-2 polynomials, we may be able to find the coefficients due to the smooth fit principle.

2.1.3 A Linear Programming Approach

As mentioned in the introduction, Helmes and Stockbridge (2010) describe how to model the optimal stopping program in a Linear Programming framework and then utilize the structure of the dual program in order to find a solution locally, rather than globally. They consider one-dimensional diffusions which satisfy the stochastic differential equation

$$dX_t = \mu(X_t)\,dt + \sigma(X_t)\,dW_t, \qquad X_0 = x \tag{2.6}$$

and their goal is to find a stopping time τ and a value function V which satisfies

$$\begin{aligned} V(x) &= \sup_{\tau\in\mathcal{T}} J(\tau;x) \\ &= \sup_{\tau\in\mathcal{T}} \mathbb{E}\left[\int_0^\tau e^{-\alpha\, s} r(X_s)\,ds + e^{-\alpha\,\tau} g(X_\tau)\mathbf{1}_{\tau<\infty}\right]. \end{aligned} \tag{2.7}$$

To accomplish this, they define an occupation measure $\mu_\tau(G) = \mathbb{E}[e^{-\alpha\tau}1_G(X_\tau)1_{\tau<\infty}]$ on $\mathscr{B}[x_l, x_r]$ to transform the optimal stopping problem into the following linear program.

$$\begin{aligned} \text{Maximize} \quad & \int g_r\,d\mu \\ \text{Subject to:} \quad & \int \phi\,d\mu = \phi(x), \\ & \int \psi\,d\mu = \psi(x), \\ & \int 1\,d\mu \leq 1, \\ & \mu \text{ a non-negative measure} \end{aligned}$$

where g_r is a function that captures both the total running reward accumulated and the terminal reward. From here they construct the following dual program using standard linear programming techniques.

$$
\begin{aligned}
\text{Minimize} \quad & \phi(x)\, c_1 + \psi(x)\, c_2 + c_3 \\
\text{Subject to:} \quad & \phi(y)\, c_1 + \psi(y)\, c_2 + c_3 \geq g_r(y) \qquad \forall y \in [x_l, x_r] \\
& c_1, c_2 \text{ free}, c_3 \geq 0
\end{aligned}
$$

In this dual program, ϕ, ψ are defined as follows (Rogers and Williams, 2000). Fix $q \in (x_l, x_r)$ arbitrarily and let $H_z = \inf\{t \geq 0 \,|\, X_t = z\}$ denote the first hitting time of $z \in [x_l, x_r]$, then

$$
\phi(y) = \begin{cases} \mathbb{E}_y[\exp(-\alpha H_q)] & y \geq q \\ 1/\mathbb{E}_q[\exp(-\alpha H_y)] & y \leq q \end{cases}
$$

and

$$
\psi(y) = \begin{cases} \mathbb{E}_y[\exp(-\alpha H_q)] & y \leq q \\ 1/\mathbb{E}_q[\exp(-\alpha H_y)] & y \geq q. \end{cases}
$$

Additionally they are continuously differentiable eigenfunctions. That is, they are solutions to $Af - \alpha f = 0$. The function ϕ is non-negative strictly decreasing and ψ is non-negative, strictly increasing. The following is also true (Borodin and Salminen, 2002):

$$
\mathbb{E}_x[e^{-\alpha H_z}] = \begin{cases} \dfrac{\psi(x)}{\psi(z)}, & z \geq x \\ \dfrac{\phi(x)}{\phi(z)}, & z \leq x \end{cases}
$$

Helmes and Stockbridge further show that their primal and dual programs enjoy strong duality, which then allows us to solve the optimal stopping problem locally using the dual. This is oftentimes done by appealing to the principle of smooth fit, depending on the particular gain function. More details to this approach will follow in Section 3.2.

One particularly interesting example was found in “Optimal stopping of oscillating Brownian motion” by Mordecki and Salminen (2019). They found that for a diffusion process with positive piecewise constant volatility changing at the point $x = 0$ and the gain function $((1 + x)^+)^2$ there are certain parameters which lead to a disconnected stopping region. This is of particular interest because for most of the possible parameters the stopping region is of the form $[c, \infty)$ which could lead to interesting results once we analyze what could happen if we allow the parameters to vary.

2.2 Stability of Semi-Infinite Programming

Further analysis of the structure of the semi-infinite linear programs led to a paper written by Jongen et al. (1992). Their work was focused on analyzing how the feasible set of a semi-infinite linear program changes under certain changes to the constraint system where the constraints were given by a family of functions. Historically, work has been using a similar formulation when there were finitely many constraints. However, they extended these results to a constraint system indexed by a continuum.

In their paper, they consider semi-infinite linear programs in the form:

$$\begin{aligned} &\text{minimize} && f \text{ on } M[H,G], \\ &\text{where} && M[H,G] = \{x \in \mathbb{R}^n \mid H(x) = 0,\ G(x,y) \geq 0, \text{ for all } y \in \mathcal{Y}\}, \\ & && \mathcal{Y} \subset \mathbb{R}^r \text{ compact.} \end{aligned}$$

In this framework the mappings $f : \mathbb{R}^n \to \mathbb{R}$, $H : (h_1, ..., h_m) : \mathbb{R}^n \to \mathbb{R}^m, \quad m < n$, and $G : \mathbb{R}^n \times \mathbb{R}^r \to \mathbb{R}$ are assumed to be continuously differentiable. The differential DH is defined as the matrix whose row vectors are $DH_i = (\frac{\partial h_i}{\partial x_1}, \frac{\partial h_i}{\partial x_2}, ..., \frac{\partial h_i}{\partial x_n})$ and $D_cG = (\frac{\partial G}{\partial c_1}, \frac{\partial G}{\partial c_2}, ..., \frac{\partial G}{\partial c_n})$.

In the result that we have used, the authors use the notion of lower and upper semi-continuity of a set valued mapping and define it as follows.

Definition 2.2.1. Let $\mathscr{M}$ be a mapping from a topological space T into the family $\mathscr{P}(\mathbb{R}^n)$ of all subsets of $\mathbb{R}^n$. We call $\mathscr{M}$ lower semi-continuous at $\bar{v} \in T$ if, for any open set $\mathscr{U} \subset \mathbb{R}^n$ with $\mathscr{M}(\bar{v}) \cap \mathscr{U} \neq \emptyset$, there exists a neighborhood $\mathscr{V}$ of $\bar{v}$ such that $\mathscr{M}(v) \cap \mathscr{U} \neq \emptyset$ whenever $v \in \mathscr{V}$. The mapping $\mathscr{M}$ is said to be upper semi-continuous at $\bar{v} \in T$ if, for any open set $\mathscr{U} \subset \mathbb{R}^n$ with $\mathscr{M}(\bar{v}) \subset \mathscr{U}$, there exists a neighborhood $\mathscr{V}$ of $\bar{v}$ such that $\mathscr{M}(v) \subset \mathscr{U}$ whenever $v \in \mathscr{V}$.

The conditions that need to be satisfied make up what are known as the extended Mangasarian-Fromovitz constraint qualification (EMFCQ). This is said to hold at $c \in M[H,G]$ if:

1) rank $DH(c) = m$

2) there exists a vector $\xi \in \mathbb{R}^n$ satisfying

$$DH(c) \cdot \xi = 0,$$

$$D_cG(c,y) \cdot \xi > 0, \text{ for all } y \text{ such that } G(c,y) = 0.$$

If these conditions are satisfied, they have the following result:

Theorem 2.2.2. *Let $H \in C^2(\mathbb{R}^n, \mathbb{R}^m)$, and suppose that EMFCQ is satisfied for all $x \in M[H, G]$. Then, there exists a C_s^1 - neighborhood $\mathcal{O} \subset C^2(\mathbb{R}^n, \mathbb{R}^m) \times C^1(\mathbb{R}^n \times \mathbb{R}^r, \mathbb{R})$ such that the set-valued mapping $\mathcal{M} : (\tilde{H}, \tilde{G}) \mapsto M[\tilde{H}, \tilde{G}]$ is both upper semi-continuous and lower semi-continuous at all $(\tilde{H}, \tilde{G}) \in \mathcal{O}$.*

They go on to conclude stronger results if compactness is assumed for various parts of the framework, however those results do not fit the work done here and are not listed.

2.3 Perturbations of Finite-Dimensional Linear Programming

The requirements to use the results above for the semi-infinite programs can be too strict for certain models. Therefore we then explored finite linear programs to see what results could be gained from looking at the continuity of a finite linear program and extending them to the semi-infinite case.

Robinson (1977) deals with finite dimensional linear programs and analyzes the primal and dual programs under arbitrary perturbations of the coefficients in the constraints and objective function. His main result is listed below, and states that under suitable conditions we have two results: the new perturbed system is still solvable and the optimal value of the perturbed system is close to that of the original system. First, his framework for the primal and dual programs is:

$$\begin{array}{llcll} min\ \langle c, x\rangle & & \qquad\qquad & max\ \langle u, b\rangle & \\ & Ax - b \in Q^* & & & c - uA \in P^* \\ & x \in P & & & u \in Q, \end{array}$$

where $P^* = \{z \in \mathbb{R}^n \,|\, \langle z, x\rangle \geq 0 \text{ for each } x \in P\}$ and Q^* is defined similarly. Additionally, he introduces the following definition.

Definition 2.3.1. These systems are regular if $b \in int\{A(P) - Q^*\}$ and $c \in int\{Q(A) + P^*\}$ where $P^* = \{z \in \mathbb{R}^n | \langle z, x\rangle \geq 0 \text{ for each } x \in P\}$ and $Q^* = \{z \in \mathbb{R}^m | \langle z, u\rangle \geq 0 \text{ for each } u \in Q\}$.

This then leads us to his main result.

Theorem 2.3.2. *The following are equivalent:*

(a) The constraints of (P) and (D) are regular.

(b) The sets of optimal solutions of (P) and (D) are nonempty and bounded.

(c) There exists an $\varepsilon_0 > 0$ *such that for any* A', b', *and* c' *with*

$$\varepsilon' \equiv \max\{||A' - A||, ||b' - b||, ||c' - c||\} < \varepsilon_0,$$

the two dual problems

$$\begin{array}{cc} (P') & (D') \\ \min\langle c', x\rangle & \max\langle u, b'\rangle \\ A'x - b' \in Q^*, x \in P & c' - uA' \in P^*, u \in Q \end{array}$$

are solvable.

If one of these conditions, and therefore all, are satisfied, then there exist constants $\varepsilon_1 \in (0, \varepsilon_0]$ *and* γ *such that for any* A', b', *and* c' *with* $\varepsilon' < \varepsilon_1$, *any* x' *solving* (P'), *and any* u' *solving* (D'), *one has* $d[(x', u'), S_P \times S_D] \leq \gamma\varepsilon'$, *where* S_P *and* S_D *are the sets of optimal solutions for* (P) *and* (D) *respectively.*

Chapter 3

Analysis of the Value Function of an Optimal Stopping Problem.

3.1 Outline of the Problem

We will work using the same basic framework as Helmes and Stockbridge (2010), that is we consider a one-dimensional diffusion process X which satisfies the stochastic differential equation (SDE) in (2.6).

Our ultimate goal is to show that the value function,

$$\begin{aligned} V(x,\theta) &= \sup_{\tau\in\mathcal{T}} J(\tau;x,\theta) \\ &= \sup_{\tau\in\mathcal{T}} \mathbb{E}\left[\int_0^{\tau} e^{-\alpha s} r(X_s)\,ds + e^{-\alpha\tau} g(X_\tau)\mathbf{1}_{\tau<\infty}\right] \end{aligned} \tag{3.1}$$

has certain continuity properties in terms of the parameters (θ) in the value function or those of the dynamics where τ represents a stopping time of the process (in some acceptable set of stopping times $\mathcal{T}$).

The obvious parameter in $V(x)$ is α, the discount rate of future value. In the objective function we have the functions $r(X)$, the running reward rate of the process, and $g(X)$, the terminal reward obtained upon stopping the process. These functions will oftentimes have parameters in them, such as the coefficients in a polynomial. We use θ to represent the parameters in general. This will be important later when we begin analyzing the value function in regards to the state space X or the parameter space Θ. One assumption we will have to make is that the terminal reward function be continuous rather than upper semi-continuous.

The optimal stopping problem comes back to the value function:

$$V(x, \theta) = \sup_{\tau \in \mathcal{T}} J(\tau; x, \theta).$$

The objective is to select a stopping time τ which achieves the supremum of $J(\tau; x, \theta)$ over all acceptable stopping times. In our case, such a stopping time exists. If that were not the case we could explore what are known as ε-*optimal* stopping times.

3.2 The Linear Program

To begin, we first must see how it is possible to take the optimal stopping problem and transform it into a linear program. The optimal stopping problem considers all stopping times, which, being random variables can be quite complicated. However, the solution to a linear program is simply an n-tuple of real numbers. To see this, we take this result from Helmes and Stockbridge (2010).

Proposition 3.2.1. *The optimal stopping problem is equivalent to solving the following primal LP (denoted $\mathcal{P}$):*

$$\begin{aligned} \textit{Maximize} \quad & \int g_r \, d\mu \\ \textit{Subject to:} \quad & \int \phi \, d\mu = \phi(x), \\ & \int \psi \, d\mu = \psi(x), \\ & \int 1 \, d\mu \leq 1, \\ & \mu \textit{ a non-negative measure} \end{aligned}$$

Or its dual LP, $\mathcal{D}$:

$$\begin{aligned} \textit{Minimize} \quad & \phi(x)\, c_1 + \psi(x)\, c_2 + c_3 \\ \textit{Subject to:} \quad & \phi(y)\, c_1 + \psi(y)\, c_2 + c_3 \geq g_r(y) \qquad \forall y \in [x_l, x_r] \\ & c_1, c_2 \textit{ free}, c_3 \geq 0 \end{aligned}$$

Also, this primal and dual LP enjoy strong duality. That is, the optimal values for the primal and dual programs are equal.

Proof. See Helmes and Stockbridge (2010) for full details. Their work will be summarized below.

First they show that

$$\begin{aligned} J(\tau;x) &= \mathbb{E}_x\left[\int_0^\tau e^{-\alpha\, s} r(X_s)\, ds + e^{-\alpha\, \tau} g(X_\tau)\mathbf{1}_{\tau<\infty}\right] \\ &= \mathbb{E}_x\left[e^{-\alpha\,\tau}(f_r+g)(X_\tau)\mathbf{1}_{\tau<\infty}\right] - f_r(x). \end{aligned} \tag{3.2}$$

Where $f_r(x) = -\mathbb{E}_x\left[\int_0^\infty e^{-\alpha\, s} r(X_s) ds\right]$.

Then by defining $g_r = f_r + g$ they see that the stopping problem is now to maximize

$$J_r(\tau;x) := \mathbb{E}[e^{-\alpha\tau} g_r(X_\tau)\, 1_{\tau<\infty}]$$

Now, if Itô's formula is applied to the function $g(X_t) = e^{-\alpha t} f(X_t)$ for some $f \in C^2(x_l, x_r)$, we get

$$e^{-\alpha t} f(X_t) = f(x) + \int_0^t e^{-\alpha s}[Af(X_s) - \alpha f(X_s)]\, ds + \int_0^t e^{-\alpha s} f'(X_s)\, dW_s.$$

Then the optional sampling theorem is used to replace t with $t \wedge \tau$ and upon rearrangement,

$$e^{-\alpha t\wedge\tau} f(X_{t\wedge\tau}) - f(x) - \int_0^{t\wedge\tau} e^{-\alpha s}[Af(X_s) - \alpha f(X_s)]\, ds = \int_0^{t\wedge\tau} e^{-\alpha s} f'(X_s)\, dW_s.$$

Notice that the right hand side is a stochastic integral, and therefore both sides are (mean zero) martingales.

Then, taking expectation and letting $t \to \infty$ yields

$$\mathbb{E}_x\left[e^{-\alpha\tau} f(X_\tau) 1_{\tau<\infty} - \int_0^\tau e^{-\alpha s}[Af(X_s) - \alpha f(X_s)]\, ds\right] = f(x).$$

If we replace f with the eigenfunctions ϕ and ψ and define the occupation measure $\mu_\tau(G) = \mathbb{E}[e^{-\alpha\tau} 1_G(X_\tau) 1_{\tau<\infty}]$ for any $G \in \mathscr{B}[x_l, x_r]$ we have

$$\mathbb{E}_x[e^{-\alpha\tau}\phi(X_\tau)] = \phi(x) \qquad \mathbb{E}_x[e^{-\alpha\tau}\psi(X_\tau)] = \psi(x)$$

$$\int \phi\, d\mu_\tau = \phi(x) \qquad \int \psi\, d\mu_\tau = \psi(x)$$

Then they argue that the original optimal stopping problem is embedded in the new LP ($\mathcal{P}$) and construct its dual LP naturally. Lastly, they prove (after considerable effort) that the primal and dual enjoy strong duality.

□

In order to ensure that the optimal stopping problem has a finite solution, that is, that the optimal strategy is not to simply allow the process to proceed ad infinitum the following restrictions are given to the new reward function g_r:

1. g_r is continuous
2. $\lim_{y \to x_l} \frac{g_r(y)}{\phi(y)} = 0$
3. $\lim_{y \to x_r} \frac{g_r(y)}{\psi(y)} = 0.$

We also observe that $g_r(y) > 0$ for some y. If not, then the optimal stopping strategy is to never stop so that we obtain a reward of zero. Also, g_r is not constant, otherwise the optimal stopping strategy is to stop immediately because we would never obtain a better reward.

If we consider the dual LP, we see that we have uncountably many constraints since our constraints are indexed by $y \in [x_l, x_r] \subset \mathbb{R}$. Note that one or both of these endpoints could be infinite so it may not be a compact interval. Also, in proving the strong duality of the primal and dual LPs, Helmes and Stockbridge prove the following result, which will be used in some of the results here and as such when we refer to the programs $\mathcal{P}$ and $\mathcal{D}$ we will use either form. Notice that the only difference is removing the third constraint from the primal which then removes a variable from the dual.

Corollary 3.2.2. *The programs $\mathcal{P}$ and $\mathcal{D}$ are equivalent to the following programs.*

$$\begin{aligned} &\textit{Maximize} && \int g_r \, d\mu \\ &\textit{Subject to:} && \int \phi \, d\mu = \phi(x), \\ & && \int \psi \, d\mu = \psi(x), \\ & && \mu \textit{ a non-negative measure} \end{aligned}$$

$$\begin{aligned} &\textit{Minimize} && \phi(x)\, c_1 + \psi(x)\, c_2 \\ &\textit{Subject to:} && \phi(y)\, c_1 + \psi(y)\, c_2 \geq g_r(y) \qquad \forall y \in [x_l, x_r] \\ & && c_1, c_2 \textit{ free} \end{aligned}$$

3.3 Countable and then Finite Sub-LP

Our next step will be to focus on a countably dense subset $\mathcal{Y}$ of $[x_l, x_r]$ and only use those countably many constraints. In fact we will take $\mathcal{Y}$ to be the dyadic rationals (those whose denominator is a power of two) that fall in $[x_l, x_r]$. We will need to verify that the objective function obtains the same value on $\mathcal{Y}$ as it does when the full constraint system is employed, however due to the continuity properties of the eigenfunctions and the reward function this is straightforward. Upon restricting ourselves to this smaller subset of constraints our dual LP is as follows:

$$
\begin{aligned}
\text{Minimize} \quad & \phi(x)\, c_1 + \psi(x)\, c_2 + c_3 \\
\text{Subject to:} \quad & \phi(y)\, c_1 + \psi(y)\, c_2 + c_3 \geq g_r(y) \qquad \forall y \in \mathcal{Y} \subset [x_l, x_r] \\
& c_1, c_2 \text{ free}, c_3 \geq 0
\end{aligned}
$$

And its resulting primal LP:

$$
\begin{aligned}
\text{Maximize} \quad & \sum_{y_i \in \mathcal{Y}} d_i g_r(y_i) \\
\text{Subject to:} \quad & \sum d_i\, \phi(y_i) = \phi(x), \\
& \sum d_i\, \psi(y_i) = \psi(x), \\
& \sum d_i \leq 1, \\
& d_i \geq 0
\end{aligned}
$$

which can equivalently be thought of as:

$$\begin{aligned}
\text{Maximize} \quad & \int g_r \, d\mu \\
\text{Subject to:} \quad & \int \phi \, d\mu = \phi(x), \\
& \int \psi \, d\mu = \psi(x), \\
& \int 1 \, d\mu \leq 1, \\
& \mu \text{ a non-negative measure on the set } \mathcal{Y}.
\end{aligned}$$

Now we have a semi-infinite linear program. The dual has countably many constraints and the primal has countably many variables. We must verify that this subsystem still enjoys strong duality.

From here forward I will focus on the dual LP. I will denote the countably infinite LP $\mathcal{P}^\infty$, and the finitely constrained LP (to be defined) $\mathcal{P}^N$. Similarly, their corresponding dual LPs will be $\mathcal{D}^\infty$ and $\mathcal{D}^N$ respectively. Lastly I will denote the optimal value of $\mathcal{P}$ by $V^{\mathcal{P}}(x;\theta)$ and likewise for the other primal and dual LPs.

Theorem 3.3.1. *Denote the feasible region of $\mathcal{D}$ by $\mathcal{F}(\mathcal{D})$ with the feasible regions of the other primal and dual LPs denoted similarly. Then $\mathcal{F}(\mathcal{D}) = \mathcal{F}(\mathcal{D}^\infty)$*

Proof. One direction is obvious. Since every constraint in $\mathcal{D}^\infty$ is also a constraint in $\mathcal{D}$, $\mathcal{F}(\mathcal{D}) \subseteq \mathcal{F}(\mathcal{D}^\infty)$.

Now let $(c_1, c_2, c_3) \in \mathcal{F}(\mathcal{D}^\infty)$ be fixed and $y \in [x_l, x_r]$ be chosen arbitrarily. We need to show that $\phi(y)\, c_1 + \psi(y)\, c_2 + c_3 \geq g_r(y)$.

Since $\mathcal{Y}$ is a dense subset of $[x_l, x_r]$ there must be a sequence (y_n) that converges to y. Moreover we know that $\phi(y_i)\, c_1 + \psi(y_i)\, c_2 + c_3 \geq g_r(y_i)$ for all $y_i \in (y_n)$ since $(c_1, c_2, c_3) \in \mathcal{F}(\mathcal{D}^\infty)$.

Then, since ϕ, ψ are continuous and g_r is continuous we have:

$$\begin{aligned}
\phi(y_i)\, c_1 + \psi(y_i)\, c_2 + c_3 &\geq g_r(y_i) \qquad \forall y_i \in (y_n) \\
\lim_{i\to\infty} [\phi(y_i)\, c_1 + \psi(y_i)\, c_2 + c_3] &\geq \lim_{i\to\infty} g_r(y_i) \\
\phi(y)\, c_1 + \psi(y)\, c_2 + c_3 &\geq g_r(y).
\end{aligned}$$

Since y was chosen arbitrarily, this inequality holds for all $y \in [x_l, x_r]$. Therefore (c_1, c_2, c_3) is feasible for $\mathcal{D}$ and $\mathcal{F}(\mathcal{D}^\infty) \subseteq \mathcal{F}(\mathcal{D})$.

□

Corollary 3.3.2. $V^{\mathcal{D}} = V^{\mathcal{D}^\infty}$

Proof. $\mathcal{D}$ and $\mathcal{D}^\infty$ have the same value functions. The theorem above gives us that they have the same feasible region. Therefore the optimal value must also be equal. □

Since $\mathcal{Y}$ is a countable subset of $[x_l, x_r]$, $\mathcal{Y} = \{y_1, y_2, y_3, ...\}$. Assume that $y_1 < x < y_2$. If not, reindex the set so that is true. We shall define the finite constraint system using the following:

Definition 3.3.3. $\mathcal{Y}^N := (\{-N, -N + \frac{1}{2^N}, -N + \frac{2}{2^N}, ..., N\} \cup \{x_l, x, x_r \,|\, x_l > -\infty, x_r < \infty\}) \cap [x_l, x_r]$.

Then we can see a natural definition of $\mathcal{D}^N$:

$$
\begin{aligned}
\text{Minimize} \quad & \phi(x)\, c_1 + \psi(x)\, c_2 + c_3 \\
\text{Subject to:} \quad & \phi(y)\, c_1 + \psi(y)\, c_2 + c_3 \geq g_r(y) \qquad \forall y \in \mathcal{Y}^N \subset [x_l, x_r] \\
& c_1, c_2 \text{ free}, c_3 \geq 0.
\end{aligned}
$$

Theorem 3.3.4. $(V^{\mathcal{D}^N}(\cdot))_{N=0}^\infty$ *is a monotonically increasing sequence of functions which converges pointwise to* $V^{\mathcal{D}^\infty}(\cdot)$.

Proof. First notice that because of how the finite and countably constrained dual LPs were formulated every single constraint in $\mathcal{D}^N$ is also a constraint in $\mathcal{D}^M$ for $M > N$, thus, $\mathcal{F}(\mathcal{D}^1) \supseteq \cdots \supseteq \mathcal{F}(\mathcal{D}^N) \supseteq \mathcal{F}(\mathcal{D}^{N+1}) \supseteq \cdots \supseteq \mathcal{F}(\mathcal{D}^\infty)$. Because of this and the fact that the objective functions are the same for all dual LPs, we have that $V^{\mathcal{D}^N}(\cdot) \leq V^{\mathcal{D}^M}(\cdot) \leq V^{\mathcal{D}^\infty}(\cdot)$ for all $M > N$.

Then choose some $z \in (x_l, x_r)$. The sequence $(V^{\mathcal{D}^N}(z))_{N=0}^\infty$ is monotonic as argued above, thus $\lim_{N\to\infty} V^{D^N}(z)$ exists and will be denoted $L(z)$. Moreover, $L(z) \leq V^{D^\infty}(z)$ by the inequality above.

Due to how $L(z)$ was defined, for any arbitrary $\varepsilon > 0$ there is some K so that $|L(z) - V^{D^K}(z)| < \varepsilon$ and $|y_i - y_{i+1}| = 1/2^K < \varepsilon$ for all $y_i \in \mathcal{Y}^K$. Let us take c_1^K, c_2^K, c_3^K to be the optimal coefficients from V^{D^K} and we can assume they are not feasible in D^∞, otherwise $V^{D^K}(z) \geq V^{D^\infty}(z)$ and we would be done.

So then if c_1^K, c_2^K, c_3^K are not feasible for D^∞ there must be some collection of $y_i \in \mathcal{Y}$ where $\phi(y_i)c_1^K + \psi(y_i)c_2^K + c_3^K \geq g_r(y_i)$ is violated for each y_i. Define

$$d_i^K := g_r(y_i) - [\phi(y_i)c_1^K + \psi(y_i)c_2^K + c_3^K]$$

and

$$d^K := \sup_i d_i^K.$$

This implies that $c_1^K, c_2^K, (c_3^K + d^K)$ is now feasible for D^∞ which means $V^{D^K}(z) + d^K = \phi(z)c_1^K + \psi(z)c_2^K + (c_3^K + d^K) > V^{D^\infty}(z)$ and then $d^K > V^{D^\infty}(z) - V^{D^K}(z)$.

We would like to utilize uniform continuity here, however the interval $[x_l, x_r]$ may not be compact. If the endpoints are finite, the functions ϕ and ψ have vertical asymptotes at one of the endpoints and so are not bounded nor uniformly continuous. Even so, the only thing that could give us trouble is if g_r is unbounded at an endpoint because otherwise we can argue that no such d^K would be created near the boundaries. We then must consider two cases.

Case 1. The boundary points x_l and x_r are both finite.

First define the following

$$\hat{c}_i := \inf_m \{c_i^m \,|\, m \geq K\},\ i = 1, 2, 3.$$

Notice that $\hat{c}_1, \hat{c}_2, \hat{c}_3 \geq 0$ since $\phi(y)c_1^K + \psi(y)c_2^K + c_3^K \geq g_r(y)$ must hold true at the boundaries and g_r is bounded from below.

We will first analyze the left boundary. If $\hat{c}_1 > 0$, the condition $\lim_{y\to x_l} \frac{g_r(y)}{\phi(y)} = 0$ on the gain function leads to the following:

$$\lim_{y\to x_l} \frac{g_r(y)}{\phi(y)} = 0$$
$$\lim_{y\to x_l} \frac{g_r(y)}{\hat{c}_1\phi(y)} = 0$$
$$\lim_{y\to x_l} \frac{g_r(y)}{\hat{c}_1\phi(y) + \hat{c}_2\psi(y) + \hat{c}_3} = 0$$

since $\hat{c_1}$, $\hat{c_2}$, $\hat{c_3}$, ϕ, and ψ are all non-negative. This implies that there is some $x_1 > x_l$ where for any $y < x_1$ we have $g_r(y) \leq \hat{c_1}\phi(y) + \hat{c_2}\psi(y) + \hat{c_3}$. We can find x_2 in the same manner by using $\lim_{y\to x_r} \frac{g_r(y)}{\psi(y)} = 0$.

Our functions g_r, ϕ, ψ are continuous and bounded on the compact interval $[x_1, x_2]$ and thus they are uniformly continuous. Then for our arbitrary $\varepsilon > 0$ there exist finite δ_g, δ_ϕ, and δ_ψ so that $|y_i - y_j| < \delta_g$ implies $|g_r(y_i) - g_r(y_j)| < \varepsilon$ and similarly for δ_ϕ and δ_ψ. If needed, increase K so that $1/2^K < \min\{\delta_g, \delta_\phi, \delta_\psi\}$. Also notice that $\hat{c_1}\phi(y) + \hat{c_2}\psi(y) + \hat{c_3} \leq c_1^m\phi(y) + c_2^m\psi(y) + c_3^m$ for all $y \in [x_l, x_r]$ and all $m \geq K$. Since we know that $\hat{c_1}\phi(y) + \hat{c_2}\psi(y) + \hat{c_3}$ majorizes g_r in the intervals $[x_l, x_1] \cup [x_2, x_r]$, no d_i^K will be created from any point outside of $[x_1, x_2]$.

Then for any y_i where $d_i^K > 0$ (i.e. where the constraint fails), there must be some $y_k, y_{k+1} \in \mathcal{Y}^K \cap [x_1, x_2]$ where $y_k < y_i < y_{k+1}$. Notice that $y_{k+1} - y_k < \varepsilon$ as well as $\phi(y_k)c_1^K + \psi(y_k)c_2^K + c_3^K \geq g_r(y_k)$ is true since our coefficients are feasible for D^K. In addition, since our functions g_r, ϕ, and ψ are continuous there is some $\hat{y} \in [y_k, y_i] \cap \mathcal{Y}$ so that $|g_r(\hat{y}) - [\phi(\hat{y})c_1^K + \psi(\hat{y})c_2^K + c_3^K]| < \varepsilon$ by the Intermediate Value Theorem.

Combining these and using a triangle inequality argument,

$$\begin{aligned}
d_i^K &= g_r(y_i) - [\phi(y_i)c_1^K + \psi(y_i)c_2^K + c_3^K] \\
&< |g_r(y_i) - g_r(\hat{y})| + |g_r(\hat{y}) - [\phi(y_i)c_1^K + \psi(y_i)c_2^K + c_3^K]| \\
&< \varepsilon + |g_r(\hat{y}) - [\phi(\hat{y})c_1^K + \psi(\hat{y})c_2^K + c_3^K]| \\
&\qquad + |[\phi(\hat{y})c_1^K + \psi(\hat{y})c_2^K + c_3^K] - [\phi(y_i)c_1^K + \psi(y_i)c_2^K + c_3^K]| \\
&< \varepsilon + \varepsilon + c_1^K\varepsilon + c_2^K\varepsilon \\
&= \varepsilon(2 + c_1^K + c_2^K).
\end{aligned}$$

Since $d_i^K < \varepsilon(2 + c_1^K + c_2^K)$ for each i, it must be true that $d^K \leq \varepsilon(2 + c_1^K + c_2^K)$. Also, $c_1^n\phi(z) + c_2^n\psi(z) \leq c_1^\infty\phi(z) + c_2^\infty\psi(z) + c_3^\infty < \infty$ for all n so we can assume $c_1^n + c_2^n$ is bounded by some M for all n.

Therefore, $0 < V^{D^\infty}(z) - V^{D^K}(z) < d^K < \varepsilon(2 + M)$. Let K go to infinity, and $V^{D^\infty}(z) - L(z) < \varepsilon(2 + M)$. Therefore $V^{D^\infty}(z) = L(z)$ at z and since z was arbitrary, $V^{D^\infty}(y) = L(y)$ as functions.

If $\hat{c_1} = 0$, then $\hat{c_1}\phi(y) + \hat{c_2}\psi(y) + \hat{c_3} \to \hat{c_3}$ as $y \to x_l$ so $\lim_{y\to x_l} g_r(y) \leq \hat{c_3}$ and there is some x_1 where $y < x_1$ implies $g_r(y) < \hat{c_3} + \varepsilon$.

From here we use an identical argument as above for any d_i^K created from $y_i > x_1$. However, there could be some d_i^K created from $y_i \in [x_l, x_1]$ but since $c_1^K\phi(y) + c_2^K\psi(y) + c_3^K \geq 0$

and $g_r(y) < \hat{c}_3 + \varepsilon$ we see that $d_i^k \leq \varepsilon$. Also, $\varepsilon < \varepsilon(2+M)$ so once again $0 < V^{D^\infty}(z) - V^{D^K}(z) < d^K < \varepsilon(2+M)$ and our result follows.

Case 2. Exactly one of x_l, x_r are infinite.
Without loss of generality, assume $x_l = -\infty$. Our argument will be similar to that in case 1, however we will use

$$\widetilde{c_1} = \liminf_{m\to\infty}\{c_1^m \,|\, m \geq K\}.$$

The need to use the limit inferior comes from the fact that when the boundary point is infinite, our finite collection of points is bounded and so some c_1^N, c_2^N could be negative.

Since $x_l = -\infty$ and we used the limit inferior rather than the infimum we know that there is some $K_l > K$ so that $|\inf_m\{c_1^m \,|\, m \geq K_l\} - \widetilde{c_1}| < \varepsilon$. Now choose $\hat{c}_i := \inf_m\{c_i^m \,|\, m \geq K_l\},\ i = 1,2,3$. If $\hat{c}_1 > 0$, we will again use the condition $\lim_{y\to x_l} \frac{g_r(y)}{\phi(y)} = 0$ to find x_1 so for any $y < x_1$ we have $g_r(y) \leq \hat{c}_1\phi(y) + \hat{c}_2\psi(y) + \hat{c}_3$.

This implies that for all $m \geq K_l$ and $y < x_1$:

$$\hat{c}_1\phi(y) + \hat{c}_2\psi(y) + \hat{c}_3 \leq c_1^m\phi(y) + c_2^m\psi(y) + c_3^m$$
$$g_r(y) \leq \hat{c}_1\phi(y) + \hat{c}_2\psi(y) + \hat{c}_3 \leq c_1^m\phi(y) + c_2^m\psi(y) + c_3^m$$

and so no d_i^m will be created.

If $\widetilde{c_1} \leq 0$ then $\lim_{y\to-\infty} g_r(y) \leq 0$ and there is some $x_1 \leq K_l$ so $g_r(y) < \varepsilon$ for all $y < x_1$ and then any d_i^K created by $y_i < x_1$ must be less than ε. Then using the same argument as case 1 we can show

$$d_i^K < \varepsilon(2 + c_1^K + c_2^k).$$

So once again, $V^{D^\infty}(y) = L(y)$.

Case 3. Both of the boundary points are infinite.
This case proceeds similarly to case 2 except that the limit inferior is used to find both x_1 and x_2.

□

Theorem 3.3.5. $V^{\mathcal{P}^\infty} = V^{\mathcal{P}}$

Proof. One of the major results from Helmes and Stockbridge (2010) is that the uncountably constrained primal and dual LPs have strong duality, so we know that $V^{\mathcal{P}} = V^{\mathcal{D}}$. Also $V^{\mathcal{P}^N} = V^{\mathcal{D}^N}$ since for each N, $\mathcal{P}^N$ and $\mathcal{D}^N$ are finite linear programs.

Even though we do not yet have strong duality for the countable linear programs, every primal and dual LP enjoy weak duality. That, combined with the inequalities and equalities above give:

$$V^{\mathcal{D}^N} = V^{\mathcal{P}^N} \leq V^{\mathcal{P}^\infty} \leq V^{\mathcal{D}^\infty}$$
$$V^{\mathcal{D}^N} \leq V^{\mathcal{P}^\infty} \leq V^{\mathcal{D}^\infty}.$$

Since these inequalities are true for every N, we can pass to the limit and use Corollary 3.3.2 to see

$$V^{\mathcal{D}^\infty} \leq V^{\mathcal{P}^\infty} \leq V^{\mathcal{D}^\infty}$$
$$V^{\mathcal{P}^\infty} = V^{\mathcal{D}^\infty}.$$

Corollary 3.3.2 gives us $V^{\mathcal{D}^\infty} = V^{\mathcal{D}}$ so we see $V^{\mathcal{P}^\infty} = V^{\mathcal{D}}$. Lastly use strong duality of the original primal and dual LP to get $V^{\mathcal{P}^\infty} = V^{\mathcal{P}}$, our desired result. □

Corollary 3.3.6. *Strong duality holds for $\mathcal{P}^\infty$ and $\mathcal{D}^\infty$.*

Proof. This is a direct result of the proof above. In particular, it is given by the equation $V^{\mathcal{D}^\infty} = V^{\mathcal{P}^\infty}$. □

3.4 Continuity of the Value Function in Finite LPs

At this point we would like to use results in the paper by Robinson (1977) to show continuity of $V(x;\theta)$ under small perturbations of the parameters (θ) of the model. Theorem 2.3.2 from his paper gives three conditions under which that would be true. However, this only applies to finite linear programs (finite variables and constraints). Theorem 2.3.2 (Robinson, 1977) states that for finite linear programs, if the constraints of the primal and dual systems are regular then the optimal solution sets are continuous under small perturbations of the parameters. For our problem, $V(x)$ is continuous in x so then we can conclude that $V^{\mathcal{P}^N}$ is continuous in the parameters provided our systems are regular.

Robinson works with primal linear programs of the form:

$$\begin{aligned} &min\ \langle c, x\rangle \\ &\quad Ax - b \in Q^* \\ &\quad x \in P \end{aligned}$$

and their duals

$$\begin{aligned} &max\ \langle u, b\rangle \\ &\quad c - uA \in P^* \\ &\quad u \in Q. \end{aligned}$$

Recalling Definition 2.3.1, these linear programs are regular if $b \in int\{A(P) - Q^*\}$ and $c \in int\{Q(A) + P^*\}$ where $P^* = \{z \in \mathbb{R}^n | \langle z, x\rangle \geq 0$ for each $x \in P\}$ and $Q^* = \{z \in \mathbb{R}^m | \langle z, u\rangle \geq 0$ for each $u \in Q\}$.

For our systems we have the following:

$$A = \begin{bmatrix} \phi(y_1) & \psi(y_1) & 1 \\ \vdots & \vdots & \vdots \\ \phi(y_n) & \psi(y_n) & 1 \end{bmatrix} \qquad b = \begin{bmatrix} g_r(y_1) \\ \vdots \\ g_r(y_n) \end{bmatrix} \qquad c = \begin{bmatrix} \phi(x) \\ \psi(x) \\ 1 \end{bmatrix}$$

$$P = \mathbb{R}^2 \times \mathbb{R}_+ \qquad P^* = \{z \in \mathbb{R}^3 \,|\, \langle z, x\rangle \geq 0 \text{ for each } x \in P\} = \vec{0} \times \vec{0} \times \mathbb{R}_+$$
$$Q = \mathbb{R}^n_+ \qquad Q^* = \{z \in \mathbb{R}^n \,|\, \langle z, u\rangle \geq 0 \text{ for each } u \in Q\} = \mathbb{R}^n_+$$

Theorem 3.4.1. *The constraints for the finite system $\mathcal{P}^N$ and its dual $\mathcal{D}^N$ are regular.*

Proof. To establish regularity we need to verify that $b \in int\{A(P) - Q^*\}$ and $c \in int\{(Q)A + P^*\}$. The first condition is straight forward. Consider $A(P) - Q^*$. Let $\vec{p} \in P$ and $\vec{q^*} \in Q^*$.

$$\text{Then } A\vec{p} - \vec{q^*} = \begin{bmatrix} \phi(y_1)p_1 + \psi(y_1)p_2 + p_3 - q_1* \\ \vdots \\ \phi(y_n)p_1 + \psi(y_n)p_2 + p_3 - q_n* \end{bmatrix}$$

where p_1, p_2 are free and p_3, q_i^* are non-negative. Let $p_1 = p_2 = 0$ and we see that $A(P) - Q^* = \mathbb{R}^n$, which is an open set and therefore the first regularity condition is satisfied.

The second condition requires more work. We must show that

$$\begin{bmatrix} \phi(x) \\ \psi(x) \\ 1 \end{bmatrix} \in int \left\{ q \begin{bmatrix} \phi(y_1) & \psi(y_1) & 1 \\ \vdots & \vdots & \vdots \\ \phi(y_n) & \psi(y_n) & 1 \end{bmatrix} + p^* \;\middle|\; q \in \mathbb{R}^n_+,\ p^* \in \vec{0} \times \vec{0} \times \mathbb{R}_+ \right\}.$$

Recall that $y_1 < x < y_2$ by construction. Now let us only consider q of the form $(q_1, q_2, 0, 0, ..., 0)$. I will show that the interior of this smaller set contains c and so therefore the interior of the full set does as well. In other words,

$$\begin{bmatrix}\phi(x) & \psi(x) & 1\end{bmatrix}$$

$$\in int \left\{ \begin{bmatrix} q_1 & q_2 & 0 & \cdots & 0 \end{bmatrix} \begin{bmatrix} \phi(y_1) & \psi(y_1) & 1 \\ \vdots & \vdots & \vdots \\ \phi(y_n) & \psi(y_n) & 1 \end{bmatrix} + p^* \;\middle|\; q_1, q_2 \in \mathbb{R}_+,\ p^* \in \vec{0} \times \vec{0} \times \mathbb{R}_+ \right\}.$$

Performing the matrix operations in this set we get (rewritten as a column vector for clarity in reading):

$$\begin{bmatrix}\phi(x) \\ \psi(x) \\ 1\end{bmatrix} \in int \left\{ \begin{bmatrix} \phi(y_1)q_1 + \phi(y_2)q_2 \\ \psi(y_1)q_1 + \psi(y_2)q_2 \\ q_1 + q_2 + p \end{bmatrix} \;\middle|\; q_1, q_2, p \geq 0 \right\}.$$

Since ϕ, ψ are strictly convex, we have $\phi(y) < \phi(y_1)q_1 + \phi(y_2)q_2$ and $\psi(y) < \psi(y_1)q_1 + \psi(y_2)q_2$ where $q_1 + q_2 = 1$ and $y_1 \leq y \leq y_2$. Recall that $x \in [y_1, y_2]$ by construction.

Now let $\delta = \min\{\phi(y_1)q_1 + \phi(y_2)q_2 - \phi(x),\ \psi(y_1)q_1 + \psi(y_2)q_2 - \psi(x)\}$.

Then we have that,

$$\begin{bmatrix}\phi(x) \\ \psi(x) \\ 1\end{bmatrix} \in \left\{ \begin{bmatrix} \phi(y_1)q_1 + \phi(y_2)q_2 \\ \psi(y_1)q_1 + \psi(y_2)q_2 \\ q_1 + q_2 + p \end{bmatrix} \text{ for } q_1 \in (0,1), q_2 \in (0, 1 - q_1), 0 < p < \delta + 1 \right\}$$

which is an open set, and is actually contained inside of the set $\{(Q)A + P^*\}$ and so is contained within its interior.

Therefore, the finite primal and dual constraints satisfy the regularity conditions. □

Theorem 3.4.2. *$V^{\mathcal{P}^N}(y, \cdot)$ and $V^{\mathcal{D}^N}(y, \cdot)$ are continuous.*

Proof. This is a direct result of Theorem 2.3.1 (Robinson, 1977). Since we have established that $\mathcal{P}^N$ and $\mathcal{D}^N$ are regular systems we get this result. □

Corollary 3.4.3. *$V^{\mathcal{D}^\infty}(y, \cdot)$ is lower semi-continuous.*

Proof. By the construction of the finite and countably infinite LPs as well as the results from Theorem 3.3.4 we have

$$V^{\mathcal{D}^\infty}(y,\theta) = \lim_{N\to\infty} V^{\mathcal{D}^N}(y,\theta) = \sup_N V^{\mathcal{D}^N}(y,\theta).$$

Theorem 3.4.2 gives the continuity of $V^{\mathcal{D}^N}(y,\cdot)$ and the supremum of a family of continuous functions is lower semi-continuous so we have our result. □

Theorem 3.4.4. *$V^{\mathcal{D}}(y,\cdot)$ is lower semi-continuous.*

Proof. This is a result of Corollaries 3.3.2 which gives $V^{\mathcal{D}}(y,\theta) = V^{\mathcal{D}^\infty}(y,\theta)$ and 3.4.3 which gives the lower semi-continuity of $V^{\mathcal{D}}(y,\cdot)$. □

3.5 Continuity of $V^{\mathcal{P}}(\cdot)$

At this point we have established the continuity (in the parameter space) of the finite value functions as well as lower semi-continuity of the value function for the full linear program. We would then like to introduce some ground work that will establish continuity of the infinite value function under certain conditions.

A paper by Jongen et al. (1992) explores the structure and stability of the feasible set of semi-infinite optimization. Recall that the optimization problems that fit this structure are of the form

$$\begin{aligned} \text{minimize} \quad & f \text{ on } M[H,G], \\ \text{where} \quad & M[H,G] = \{x \in \mathbb{R}^n \mid H(x) = 0,\ G(x,y) \geq 0,\ \text{for all } y \in \mathcal{Y}\}, \\ & \mathcal{Y} \subset \mathbb{R}^r \text{ compact.} \end{aligned}$$

In this formulation, the functions $f : \mathbb{R}^n \to \mathbb{R}$, $H = (h_1, ..., h_m) : \mathbb{R}^n \to \mathbb{R}^m$, $m < n$, and $G : \mathbb{R}^n \times \mathbb{R}^r \to \mathbb{R}$ are assumed to be continuously differentiable. We will now match $\mathcal{D}$ (our infinitely constrained dual problem) to this format.

Recall that $\mathcal{D}$ is given by:

$$\begin{aligned} \text{Minimize} \quad & \phi(x)\, c_1 + \psi(x)\, c_2 \\ \text{Subject to:} \quad & \phi(y)\, c_1 + \psi(y)\, c_2 \geq g_r(y) \qquad \forall y \in [x_l, x_r] \\ & c_1, c_2 \text{ free} \end{aligned}$$

and that x is not a variable but instead reserved for the starting point of the stochastic process. Also notice that c_3 has been removed from the dual problem, but this is because

Helmes and Stockbridge (2010) showed that the value of this sub-LP is in fact equal to the value of the full LP.

So then $f(c_1, c_2) = \phi(x)\, c_1 + \psi(x)\, c_2$, $G(c_1, c_2, y) = \phi(y)\, c_1 + \psi(y)\, c_2 - g_r(y)$ and $(n, m, r) = (2, 0, 1)$. Since our problem does not have any equality constraints, we define $H : \mathbb{R}^2 \to \mathbb{R}^0$ to be the zero function. Lastly we have $M[H, G] = \{c \in \mathbb{R}^2 \mid \phi(y)\, c_1 + \psi(y)\, c_2 - g_r(y) \geq 0\}$.

One of the main results of Jongen et al. (1992) states that under certain conditions on H and G the set-valued mapping $(\tilde{H}, \tilde{G}) \mapsto M[\tilde{H}, \tilde{G}]$ is both upper and lower semi-continuous in a neighborhood of H, G under a specific topology. This tells us that if we perturb the functions H, G a small amount, the feasible region $M[H, G]$ will undergo a small perturbation as well. This is due to lower semi-continuous part of definition 2.2.1 which tells us that this mapping is lower semi-continuous with respect to θ at $(0, \phi_\theta(y)\, c_1 + \psi_\theta(y)\, c_2 - g_{r,\theta}(y))$.

The conditions that need to be satisfied make up what are known as the extended Mangasarian-Fromovitz constraint qualification.

Definition 3.5.1. A point c in the feasible set $M[H, G]$ of a semi-infinite linear program satisfies the extended Mangasarian-Fromovitz constraint qualification (EMFCQ) if:
1) rank $DH(c) = m$
2) there exists a vector $\xi \in \mathbb{R}^n$ satisfying

$DH(c) \cdot \xi = 0,$

$D_c G(c, y) \cdot \xi > 0$, for all y such that $G(c, y) = 0$.

Condition 1 is true since $m = 0$ and $DH(c) = 0$.
Condition 2 is satisfied by the vector $\xi = (1, 1)$. Since $D_c G(c, y) = (\phi(y), \psi(y))$ and $\phi(y) > 0, \psi(y) > 0$ for all y we have $D_c G(c, y) \cdot \xi > 0$ for all y not just for those in that small set. Now that we have established that our problem is of the right structure and satisfies this condition we can now introduce our result.

Definition 3.5.2. We say that a linear program $\mathcal{D}$ is **perturbable** if it satisfies the following conditions:

1. $\mathcal{D}$ has the structure of

$$\begin{aligned} \text{minimize} \quad & f \text{ on } M[H, G], \\ \text{where} \quad & M[H, G] = \{x \in \mathbb{R}^n \mid H(x) = 0,\ G(x, y) \geq 0,\ \text{for all } y \in \mathcal{Y}\}, \\ & \mathcal{Y} \subset \mathbb{R}^r \text{ compact}; \end{aligned}$$

2. The functions $f : \mathbb{R}^n \to \mathbb{R},\ H = (h_1, ..., h_m) : \mathbb{R}^n \to \mathbb{R}^m$, $m < n$, and $G : \mathbb{R}^n \times \mathbb{R}^r \to \mathbb{R}$ are continuously differentiable;

3. $\text{rank}(DH(c)) = m$;

4. There exists a vector $\xi \in \mathbb{R}^n$ satisfying

   $DH(c) \cdot \xi = 0$,

   $D_c G(c, y) \cdot \xi > 0$, for all y such that $G(c, y) = 0$.

5. The functions $\phi_\theta(y)$, $\psi_\theta(y)$, and $g_{r,\theta}(y)$ are continuous with respect to θ.

Theorem 3.5.3. *If $\mathcal{D}$ is perturbable, then $V^{\mathcal{D}}(x, \cdot)$ is continuous.*

Proof. Fix θ arbitrarily and let $\varepsilon > 0$. Assume for this proof that $\varepsilon < 1$.

For this proof we will be referring to three different linear programs. LP (1) is:

$$\begin{array}{lll} \text{Minimize} & \phi_\theta(x)\, c_1 + \psi_\theta(x)\, c_2 & \\ \text{Subject to:} & \phi_\theta(y)\, c_1 + \psi_\theta(y)\, c_2 \geq g_{r,\theta}(y) & \forall y \in [x_l, x_r] \\ & c_1, c_2 \text{ free.} & \end{array}$$

By the results in Jongen et al. (1992) we can choose some $\hat{\theta}$ near θ to perturb the constraint functions in our linear program. This new program will be LP (2):

$$\begin{array}{lll} \text{Minimize} & \phi_\theta(x)\, d_1 + \psi_\theta(x)\, d_2 & \\ \text{Subject to:} & \phi_{\hat{\theta}}(y)\, d_1 + \psi_{\hat{\theta}}(y)\, d_2 \geq g_{r,\hat{\theta}}(y) & \forall y \in [x_l, x_r] \\ & d_1, d_2 \text{ free.} & \end{array}$$

We must be careful about how $\hat{\theta}$ is chosen. So we will choose $\delta > 0$ so that the following are true if $|\theta - \hat{\theta}| < \delta$:

1. $\delta < \frac{\varepsilon}{2K}$ where $K = \phi_\theta(x) + \psi_\theta(x)$.

2. For any c in $\mathcal{F}(1)$ there is some $\tilde{d}$ in $\mathcal{F}(2)$ so $|c - \tilde{d}| < \frac{\varepsilon}{2K}$.

3. For any d in $\mathcal{F}(2)$ there is some $\tilde{c}$ in $\mathcal{F}(1)$ so $|d - \tilde{c}| < \frac{\varepsilon}{2K}$.

4. $|\phi_{\hat{\theta}}(x) - \phi_\theta(x)| + |\psi_{\hat{\theta}}(x) - \psi_\theta(x)| < \frac{\varepsilon}{8M}$ where $M = c_1^*(1 + \frac{\phi_\theta(x)}{\psi_\theta(x)}) + c_2^*(1 + \frac{\psi_\theta(x)}{\phi_\theta(x)}) + (1/\phi_\theta(x) + 1/\psi_\theta(x))$.

Conditions 2,3, and 4 above are accomplished by the lower semi-continuity on the set-valued mapping $\mathscr{M}$ as well as the assumption that $\phi_\theta(y)$, $\psi_\theta(y)$, and $g_{r,\theta}(y)$ are continuous with respect to θ.

Then lastly LP (3) is defined as:

$$\begin{aligned} \text{Minimize} \quad & \phi_{\hat{\theta}}(x)\, e_1 + \psi_{\hat{\theta}}(x)\, e_2 \\ \text{Subject to:} \quad & \phi_{\hat{\theta}}(y)\, e_1 + \psi_{\hat{\theta}}(y)\, e_2 \geq g_{r,\hat{\theta}}(y) \qquad \forall y \in [x_l, x_r] \\ & e_1, e_2 \text{ free.} \end{aligned}$$

Notice that (1) and (2) have the same objective functions but different feasible regions ($\mathcal{F}(1) \neq \mathcal{F}(2)$) because the constraints are different functions. Whereas $\mathcal{F}(2) = \mathcal{F}(3)$ but (2) and (3) have different objective functions. Denote the objective function for (1) and (2) as $f_\theta(\cdot)$ and that of (3) as $f_{\hat{\theta}}(\cdot)$ where these are now functions on the feasible sets.

The goal will be to show that the optimal value for (1) and (3) are close which would show that the new linear program we create by perturbing the parameters slightly will have an optimal value close to that of the original. This is in fact continuity in the parameter space.

Claim 1. The optimal values of (1) and (2) are within $\varepsilon/2$.

Proof of claim. Let c^* and d^* be the optimal solutions to (1) and (2) respectively and let c and d be as described so $d(c^*, d) < \frac{\varepsilon}{2K}$ and $|d^* - c| < \frac{\varepsilon}{2K}$. Then, since c^* and c are feasible in (1) and c^* is the optimal solution we know that $f_\theta(c^*) \leq f_\theta(c)$. Similarly, since d^* and d are feasible in (2), $f_\theta(d^*) \leq f_\theta(d)$.

Additionally,

$$\begin{aligned} |f_\theta(c^*) - f_\theta(d)| &= |[\phi_\theta(x)\, c_1^* + \psi_\theta(x)\, c_2^*] - [\phi_\theta(x)\, d_1 + \psi_\theta(x)\, d_2]| \\ &= |\phi_\theta(x)\, (c_1^* - d_1) + \psi_\theta(x)\, (c_2^* - d_2)| \\ &\leq \phi_\theta(x)|c_1^* - d_1| + \psi_\theta(x)|c_2^* - d_2| \\ &< (\phi_\theta(x) + \psi_\theta(x))\, \frac{\varepsilon}{2K}. \end{aligned}$$

Then $|f_\theta(c^*) - f_\theta(d)| < \varepsilon/2$ and likewise $|f_\theta(d^*) - f_\theta(c)| < \varepsilon/2$.

Suppose $f_\theta(c^*) \leq f_\theta(d^*)$.
Then $f_\theta(c^*) \leq f_\theta(d^*) \leq f_\theta(d)$ so in fact $|f_\theta(c^*) - f_\theta(d^*)| \leq |f_\theta(c^*) - f_\theta(d)| < \varepsilon/2$.

Suppose $f_\theta(d^*) < f_\theta(c^*)$.
Then $f_\theta(d^*) < f_\theta(c^*) \leq f_\theta(c)$ so in fact $|f_\theta(d^*) - f_\theta(c^*)| \leq |f_\theta(d^*) - f_\theta(c)| < \varepsilon/2$.

In either case, the optimal values $f_\theta(c^*)$ and $f_\theta(d^*)$ for (1) and (2) respectively are within $\varepsilon/2$.

Claim 2. The optimal values for (2) and (3) are within $\varepsilon/2$.

Proof of claim. Now we have two problems with the same feasible solutions. Recall d^* is the optimal solution for (2) and let e^* be that of (3).

By the results in claim 1,

$$\begin{aligned}
\phi_\theta(x)\, d_1^* + \psi_\theta(x)\, d_2^* &< \phi_\theta(x)\, c_1^* + \psi_\theta(x)\, c_2^* + \varepsilon/2 \\
\phi_\theta(x)\, d_1^* &< \phi_\theta(x)\, c_1^* + \psi_\theta(x)\, c_2^* + \varepsilon/2 \\
d_1^* &< c_1^* + \frac{\psi_\theta(x)}{\phi_\theta(x)}\, c_2^* + \frac{\varepsilon}{2\phi_\theta(x)}
\end{aligned}$$

and likewise

$$d_2^* < \frac{\phi_\theta(x)}{\psi_\theta(x)} c_1^* + c_2^* + \frac{\varepsilon}{2\psi_\theta(x)}.$$

So then we have $d_1^* + d_2^* < c_1^*(1 + \frac{\phi_\theta(x)}{\psi_\theta(x)}) + c_2^*(1 + \frac{\psi_\theta(x)}{\phi_\theta(x)}) + \frac{\varepsilon}{2}(1/\phi_\theta(x) + 1/\psi_\theta(x)) < M$.

We also know $f_\theta(d^*) \le f_\theta(e^*)$ and $f_{\hat\theta}(e^*) \le f_{\hat\theta}(d^*)$ since d^* and e^* are optimal for their respective linear programs. Then

$$\begin{aligned}
|f_\theta(d^*) - f_{\hat\theta}(d^*)| &= |[\phi_\theta(x)\, d_1^* + \psi_\theta(x)\, d_2^*] - [\phi_{\hat\theta}(x)\, d_1^* + \psi_{\hat\theta}(x)\, d_2^*]| \\
&= |[\phi_\theta(x) - \phi_{\hat\theta}(x)]d_1^* + [\psi_\theta(x) - \psi_{\hat\theta}(x)]d_2^*| \\
&\le |[\phi_\theta(x) - \phi_{\hat\theta}(x)]d_1^*| + |[\psi_\theta(x) - \psi_{\hat\theta}(x)]d_2^*| \\
&< \frac{\varepsilon(d_1^* + d_2^*)}{4M} \\
&< \frac{\varepsilon}{4} < \frac{\varepsilon}{2}.
\end{aligned}$$

Suppose $f_\theta(d^*) \le f_{\hat\theta}(e^*)$, then $f_\theta(d^*) \le f_{\hat\theta}(e^*) \le f_{\hat\theta}(d^*)$ and in fact $|f_\theta(d^*) - f_{\hat\theta}(e^*)| < \varepsilon/2$ is true.

Suppose $f_\theta(d^*) \ge f_{\hat\theta}(e^*)$. We will assume $\phi_{\hat\theta}(x) > \phi_\theta(x)/2$ and $\psi_{\hat\theta}(x) > \psi_\theta(x)/2$. By a similar argument as above we can show

$$\begin{aligned}
\phi_{\hat\theta}(x)\, e_1^* + \psi_{\hat\theta}(x)\, e_2^* &< \phi_\theta(x)\, d_1^* + \psi_\theta(x)\, d_2^* \\
\phi_{\hat\theta}(x)\, e_1^* + \psi_{\hat\theta}(x)\, e_2^* &< \phi_\theta(x)\, c_1^* + \psi_\theta(x)\, c_2^* + \varepsilon/2 \\
e_1^* &< \frac{\phi_\theta(x)\, c_1^* + \psi_\theta(x)\, c_2^* + \varepsilon/2}{\phi_{\hat\theta}(x)} \\
e_1^* &< 2\frac{\phi_\theta(x)\, c_1^* + \psi_\theta(x)\, c_2^* + \varepsilon/2}{\phi_\theta(x)}.
\end{aligned}$$

Then $e_1^* + e_2^* < 2c_1^*(1 + \frac{\phi_\theta(x)}{\psi_\theta(x)}) + 2c_2^*(1 + \frac{\psi_\theta(x)}{\phi_\theta(x)}) + 2\frac{\varepsilon}{2}(1/\phi_\theta(x) + 1/\psi_\theta(x)) < 2M$.

Additionally,

$$\begin{aligned}
|f_\theta(e^*) - f_{\hat{\theta}}(e^*)| &= |[\phi_\theta(x)\, e_1^* + \psi_\theta(x)\, e_2^*] - [\phi_{\hat{\theta}}(x)\, e_1^* + \psi_{\hat{\theta}}(x)\, e_2^*]| \\
&= |[\phi_\theta(x) - \phi_{\hat{\theta}}(x)]e_1^* + [\psi_\theta(x) - \psi_{\hat{\theta}}(x)]e_2^*| \\
&\le |[\phi_\theta(x) - \phi_{\hat{\theta}}(x)]e_1^*| + |[\psi_\theta(x) - \psi_{\hat{\theta}}(x)]e_2^*| \\
&< \frac{\varepsilon(e_1^* + e_2^*)}{4M} \\
&< \frac{\varepsilon}{2}.
\end{aligned}$$

Then $f_{\hat{\theta}}(e^*) \le f_\theta(d^*) \le f_\theta(e^*)$ and $|f_{\hat{\theta}}(e^*) - f_\theta(d^*)| \le |f_\theta(e^*) - f_{\hat{\theta}}(e^*)| < \frac{\varepsilon}{2}$.

In either case, the optimal values for (2) and (3) are within $\varepsilon/2$.

Combining claims 1 and 2 we have that the optimal value of LPs (1) and (3) are within ε which finishes our proof. □

Chapter 4

Examples

4.1 Forest Harvest with Carbon Credits

This problem was analyzed as example 6.2 in Helmes and Stockbridge (2010). They found the closed form expression of the value function, we go further to examine the effects of allowing the parameters to vary. This example illustrates that if we are able to find a closed form expression of the value function with respect to the parameters of the model, we are able to analyze continuity directly without using the results of this paper. In this example we are examining a geometric Brownian motion process X_t which satisfies the stochastic differential equation

$$dX_t = \mu X_t \, dt + \sigma X_t \, dW_t, \qquad X_0 = x > 0, \quad \mu, \sigma > 0. \tag{4.1}$$

In this example X_t represents the quantity of lumber in a forest. While the forest stands, it earns carbon credits for the owner proportional to the size of the forest. So the running reward is $r(y) = Ry^{\beta}$ with $\beta > 0$. Upon harvesting the lumber, the owner receives a terminal reward of $g(y) = k_1 y^{\beta} - k_2$ where $k_1, k_2 > 0$. Here we assume that all of the lumber is harvested at the terminal time and the process X_t has reached zero.

The objective then, is to choose a stopping time τ that would maximize the combined earnings of the carbon credits along with the profit from harvesting. Each reward is discounted at rate α, otherwise the optimal solution would be simply collect the running reward forever. The value function is then

$$\begin{aligned} V(x) &= \sup_{\tau} \mathbb{E}_x \left[\int_0^{\tau} e^{-\alpha t} r(X_t) \, dt + e^{-\alpha \tau} g(X_\tau) \right] \\ &= \sup_{\tau} \mathbb{E}_x \left[\int_0^{\tau} e^{-\alpha t} R X_t^{\beta} \, dt + e^{-\alpha \tau} (k_1 X_\tau^{\beta} - k_2) \right]. \end{aligned} \tag{4.2}$$

Through the method developed by Helmes and Stockbridge, they showed that the optimal stopping time is $\tau_{b^*} = \inf\{t \geq 0 \mid X_t \geq b^*\}$ which then leads to a closed form expression of the value function, namely:

$$V(x) = \begin{cases} \left(\dfrac{k_3 (b^*)^\beta - k_2}{(b^*)^{\gamma_2}} - \dfrac{R}{(\sigma^2/2)\beta(\beta-1) + \mu\beta - \alpha} \right) x^{\gamma_2} & \text{for } x \leq b^* \\ k_1 x^\beta - k_2 & \text{for } x \geq b^* \end{cases} \tag{4.3}$$

Where $\gamma_2 = \dfrac{1}{2} - \dfrac{\mu}{\sigma^2} + \sqrt{\left(\dfrac{1}{2} - \dfrac{\mu}{\sigma^2}\right)^2 + \dfrac{2\alpha}{\sigma^2}} < 0$ (there is a γ_1 in general, but in this example it does not appear in the value function because it is a one-sided stopping region), $b^* = \left(\dfrac{k_2 \gamma_2}{k_3(\gamma_2 - \beta)}\right)^{\frac{1}{\beta}}$, and $k_3 = k_1 + \dfrac{R}{\frac{1}{2}\sigma^2\beta(\beta-1) + \mu\beta - \alpha}$.

This closed form expression better allows us to analyze the value function in terms of the parameter space. That is we can think of $V(x)$ as $V(x, \theta)$ where $\theta = (\mu, \sigma, \alpha, R, \beta, k_1, k_2)$. To summarize, the parameters are:

μ - the drift rate of the process,

σ - the volatility of the process, a way to measure its randomness,

α - discount rate of the reward,

R - carbon credit reward multiplier,

β - power of the size of the forest for running and terminal reward ($\beta < \gamma_2$),

k_1 - terminal reward multiplier,

k_2 - cost of harvest.

Unless stated otherwise, each parameter is assumed to be positive. In addition, the restriction that ($\beta < \gamma_2$) is required to insure a finite maximum, otherwise any arbitrarily large reward could be attained by waiting until the forest is large enough (which is not necessarily realistic). Any one of these parameters could change so we are interested in the continuity of the value function with respect to each of them.

There is a lot to unravel to show the continuity that we are looking for. We will examine three cases. First, let $\varepsilon > 0$ and recall that throughout this we are assuming that $x = X_0$ is a fixed positive quantity.

Case 1: $x > b^* + \varepsilon$

In this region $V(x, \theta) = k_1 x^\beta - k_2$ and is clearly continuous with respect to k_1, k_2, and β.

Case 2: $x < b^* - \varepsilon$

In this region $V(x,\theta) = [(k_3(b^*)^\beta - k_2)/(b^*)^{\gamma_2} - R/((\sigma^2/2)\beta(\beta-1) + \mu\beta - \alpha)]x^{\gamma_2}$ and we have lots of different layers to unravel before we can claim continuity.

First we want to show that the optimal stopping level, b^* is continuous under small perturbations of θ. Since $b^* = (k_2\gamma_2/k_3(\gamma_2 - \beta))^{1/\beta}$, showing this will essentially involve everything we need to show continuity of the value function for this case.

We will begin with the continuity of γ_2. We have $\gamma_2 = \frac{1}{2} - \frac{\mu}{\sigma^2} + \sqrt{\left(\frac{1}{2} - \frac{\mu}{\sigma^2}\right)^2 + \frac{2\alpha}{\sigma^2}}$. The only possible issues with γ_2 are if $\sigma = 0$ or if the term inside the root is negative. By assumption $\sigma > 0$, and also $(\frac{1}{2} - \frac{\mu}{\sigma^2})^2 + \frac{2\alpha}{\sigma^2} > \frac{2\alpha}{\sigma^2} > 0$ since $\alpha > 0$. So γ_2 is continuous with respect to the parameters.

Now onto $k_3 = k_1 + R/(\frac{1}{2}\sigma^2\beta(\beta-1) + \mu\beta - \alpha)$. The only potential problem here is the possibility of $\frac{1}{2}\sigma^2\beta(\beta-1) + \mu\beta - \alpha = 0$. This however is not an issue due to our assumption that $\beta < \gamma_2$:

$$\beta < \gamma_2 = \frac{1}{2} - \frac{\mu}{\sigma^2} + \sqrt{\left(\frac{1}{2} - \frac{\mu}{\sigma^2}\right)^2 + \frac{2\alpha}{\sigma^2}}$$

$$\left(\beta - \left(\frac{1}{2} - \frac{\mu}{\sigma^2}\right)\right)^2 < \left(\frac{1}{2} - \frac{\mu}{\sigma^2}\right)^2 + \frac{2\alpha}{\sigma^2}$$

$$\beta^2 - 2\beta\left(\frac{1}{2} - \frac{\mu}{\sigma^2}\right) + \left(\frac{1}{2} - \frac{\mu}{\sigma^2}\right)^2 < \left(\frac{1}{2} - \frac{\mu}{\sigma^2}\right)^2 + \frac{2\alpha}{\sigma^2}$$

$$\beta^2 - 2\beta\left(\frac{1}{2} - \frac{\mu}{\sigma^2}\right) < \frac{2\alpha}{\sigma^2}$$

$$\sigma^2\beta(\beta-1) + 2\mu\beta - 2\alpha < 0$$

The last part that could cause concern for b^* is if $k_3(\gamma_2 - \beta) = 0$. By assumption $\gamma_2 - \beta \neq 0$ and to ensure the existence of a finite stopping time it's assumed $k_1[\alpha - \beta\mu - (\sigma^2/2)\beta(\beta-1)] > R$ so $k_3 = k_1 + R/(\frac{1}{2}\sigma^2\beta(\beta-1) + \mu\beta - \alpha) > 0$. Thus we have continuity of b^* with respect to the parameters.

Back to the value function $V(x,\theta) = \left(\frac{k_3(b^*)^\beta - k_2}{(b^*)^{\gamma_2}} - \frac{R}{(\sigma^2/2)\beta(\beta-1) + \mu\beta - \alpha}\right)x^{\gamma_2}$ we notice that we have taken care of all aspects of this function that could lead to discontinuities save for one, if $b^* = 0$. This could occur if either k_2 or γ_2 are zero. However, $\gamma_2 > 0$ and k_2 corresponds to the cost of harvesting so realistically it will never be zero. Therefore $V(x,\theta)$ is continuous with respect to θ in this case.

Case 3: $x \in (b^* - \varepsilon,\, b^* + \varepsilon)$

We have shown continuity for each region individually, now the concern is if the parameter space begins to vary and causes the value function to shift from one region to the next.

To simplify notation, denote $V(x,\theta)$ in the region of case 1 as V_1 and that of case 2 as V_2. Now for any particular θ_1 and θ_2, $|V_1(x,\theta_1) - V_2(x,\theta_1)|$ and $|V_1(x,\theta_2) - V_2(x,\theta_2)|$ can be made arbitrarily small for any $x \in (b^* - \varepsilon, b^* + \varepsilon)$ by making ε smaller. This is due to the construction of $V(x)$ as a continuous pasting of V_1 and V_2 in Helmes and Stockbridge (2010). Then so long as θ_1 and θ_2 are close enough then $|V_1(x,\theta_1) - V_1(x,\theta_2)|$ and $|V_2(x,\theta_1) - V_2(x,\theta_2)|$ will be small as shown in cases 1 and 2 above. Then a simple triangularization argument shows that $|V_1(x,\theta_1) - V_2(x,\theta_2)|$ and $|V_2(x,\theta_1) - V_1(x,\theta_2)|$ will also be small. That is:

$$|V_1(x,\theta_1) - V_2(x,\theta_2)| < |V_1(x,\theta_1) - V_1(x,\theta_2)| + |V_1(x,\theta_2) - V_2(x,\theta_2)| \text{ and}$$
$$|V_2(x,\theta_1) - V_1(x,\theta_2)| < |V_2(x,\theta_1) - V_1(x,\theta_1)| + |V_1(x,\theta_1) - V_1(x,\theta_2)|$$

Therefore for this example the value function is in fact continuous when taking into account perturbations in the parameters.

4.2 Oscillating Brownian Motion

The examples from this section were taken from Mordecki and Salminen (2019) in which they explored the stopping regions of the oscillating Brownian motion using two similar gain functions. They were able to partially solve these problems in general using methods from analysis. We will show how to solve them completely for any particular problem using our linear programming method and a graphical approach.

The process X_t is the strong solution to the SDE:

$$dX_t = [\sigma_1\, 1_{\{X_t<0\}} + \sigma_2\, 1_{\{X\geq 0\}}]\, dW_t,$$

where $0 < \sigma_1 \leq \sigma_2$.

The infinitesimal generator of the process is

$$Af(x) = \begin{cases} 0.5\, \sigma_2^2\, f''(x), & x \geq 0 \\ 0.5\, \sigma_1^2\, f''(x), & x < 0 \end{cases}$$

and the eigenfunctions ϕ and ψ that solve $Af - \alpha f = 0$ are:

$$\phi(x) = \begin{cases} \frac{\sigma_2+\sigma_1}{2\sigma_2}\exp\left(\frac{-\sqrt{2r}}{\sigma_1}x\right) + \frac{\sigma_2-\sigma_1}{2\sigma_2}\exp\left(\frac{\sqrt{2r}}{\sigma_1}x\right) & x < 0 \\ \exp\left(\frac{-\sqrt{2r}}{\sigma_2}x\right) & x \geq 0 \end{cases}$$

$$\psi(x) = \begin{cases} \exp\left(\frac{\sqrt{2r}}{\sigma_1}x\right) & x < 0 \\ \frac{\sigma_1+\sigma_2}{2\sigma_1}\exp\left(\frac{\sqrt{2r}}{\sigma_2}x\right) + \frac{\sigma_1-\sigma_2}{2\sigma_1}\exp\left(\frac{-\sqrt{2r}}{\sigma_2}x\right) & x \geq 0. \end{cases}$$

4.2.1 Piecewise Linear Gain Function

We will analyze two different gain functions for the oscillating Brownian motion process. The first of which is $g_1(x) = (1+x)^+$. Transferring this problem into the Linear Program framework we have the primal LP:

$$\begin{aligned} \text{maximize} \quad & \int (1+x)^+ \, d\mu \\ \text{subject to} \quad & \int \phi \, d\mu = \phi(x) \\ & \int \psi \, d\mu = \psi(x) \\ & \int d\mu \leq 1 \\ & \mu \text{ a non-negative measure} \end{aligned}$$

as well as the Dual LP:

$$\begin{aligned} \text{minimize} \quad & c_1\phi(x) + c_2\psi(x) \\ \text{subject to} \quad & c_1\phi(y) + c_2\psi(y) \geq (1+y)^+ \text{ for all } y \in \mathbb{R} \\ & c_1, c_2 \text{ free.} \end{aligned}$$

First recall that ϕ and ψ are unbounded, monotone functions and $g_r(y) = (1+y)^+$ is bounded from below so c_1 and c_2 must be non-negative in order to satisfy the constraints.

We know that $(-\infty, -1]$ must be in the continuation region because waiting until we receive some positive reward is always better than stopping and receiving a reward of zero. One approach to find the stopping region is to appeal to the principle of smooth fit because for any x in the stopping region we expect to find c_1, c_2 so that the constraints are satisfied and that $c_1\phi(x) + c_2\psi(x) = g_r(x)$ and $c_1\phi'(x) + c_2\psi'(x) = g_r'(x)$. Solving this system of equations we find:

$$c_1 = \frac{(1+x)\psi'(x) - \psi(x)}{\phi(x)\psi'(x) - \phi'(x)\psi(x)}, \qquad c_2 = \frac{\phi(x) - (1+x)\phi'(x)}{\phi(x)\psi'(x) - \phi'(x)\psi(x)}. \tag{4.4}$$

Substituting ϕ and ψ we have:

$$c_1 = \begin{cases} \frac{\sigma_2}{\sqrt{2r}(\sigma_2+\sigma_1)}\, e^{\frac{\sqrt{2r}}{\sigma_1}x}[(1+x)\sqrt{2r} - \sigma_1], & x < 0 \\ \frac{1}{\sqrt{2r}(\sigma_2+\sigma_1)}\left(e^{\frac{\sqrt{2r}}{\sigma_2}x}((1+x)\sqrt{2r} - \sigma_2)(\sigma_1+\sigma_2)\right. & \\ \left. +e^{\frac{-\sqrt{2r}}{\sigma_2}x}((1+x)\sqrt{2r} + \sigma_2)(\sigma_2 - \sigma_1)\right), & x \geq 0. \end{cases}$$

and

$$c_2 = \begin{cases} \frac{1}{\sqrt{2r}(\sigma_2+\sigma_1)}\left(e^{\frac{-\sqrt{2r}}{\sigma_1}x}((1+x)\sqrt{2r} + \sigma_1)(\sigma_1+\sigma_2)\right. & \\ \left. -e^{\frac{\sqrt{2r}}{\sigma_1}x}((1+x)\sqrt{2r} - \sigma_1)(\sigma_2 - \sigma_1)\right), & x < 0 \\ \frac{\sigma_1}{\sqrt{2r}(\sigma_2+\sigma_1)}\, e^{\frac{-\sqrt{2r}}{\sigma_2}x}[(1+x)\sqrt{2r} + \sigma_2], & x \geq 0 \end{cases}$$

For $x < 0$ we see that $c_1 < 0$ if $x < \frac{\sigma_1}{\sqrt{2r}} - 1$. This tells us that if $\frac{\sigma_1}{\sqrt{2r}} - 1 > 0$ then our continuation region contains $(-\infty, 0)$ and if $\frac{\sigma_1}{\sqrt{2r}} - 1 < 0$ then the continuation region contains $(-\infty, \frac{\sigma_1}{\sqrt{2r}} - 1)$.

The biggest downfall of this method is that we might expect x to be in the stopping region if c_1, c_2 are both non-negative. However, that does not guarantee that $c_1\phi(y) + c_2\psi(y) \geq (1+y)^+$ for all y. What we have used so far is that if either c_1 or c_2 are negative, then that inequality clearly fails. Showing in generality that $c_1\phi(y) + c_2\psi(y)$ majorizes our gain function for any particular c_1 and c_2 has proven to be difficult in general. However, if we choose the parameters r, σ_1, σ_2 then we can use a graphing utility to find the value function for any particular x.

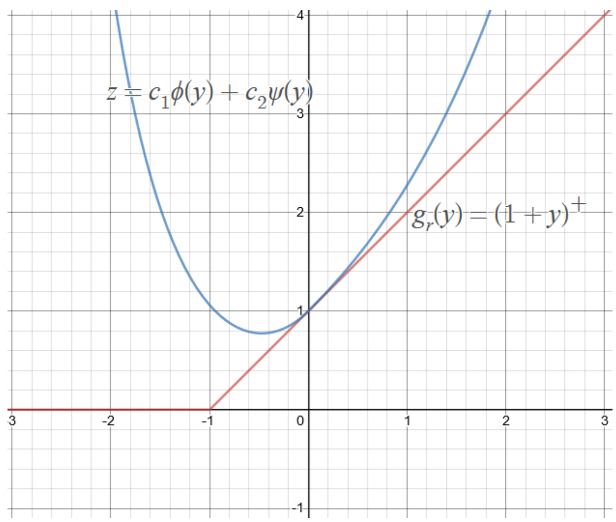

Figure 4.1: $X_0 = 0$

Consider Figure 4.1 above. The parameters chosen are $r = 0.2, \sigma_1 = 0.4$, and $\sigma_2 = 1$. If we solve (4.4) with a starting value of $x = 0$ we see that the function $c_1\phi(y) + c_2\psi(y)$ is smoothly fit to $g_r(y)$ at $(0, 1)$ and majorizes $g_r(y)$ everywhere. This implies that 0 is in the stopping region and $V(0) = g_r(0)$. However, the figure below illustrates what happens when we try the same approach for a starting value of $x = -0.5$. Immediately we see that even though we have smooth fit, $c_1\phi(y) + c_2\psi(y) \geq g_r(y)$ is violated for some y. Thus -0.5 is in the continuation region and $V(-0.5) > g_r(-0.5)$.

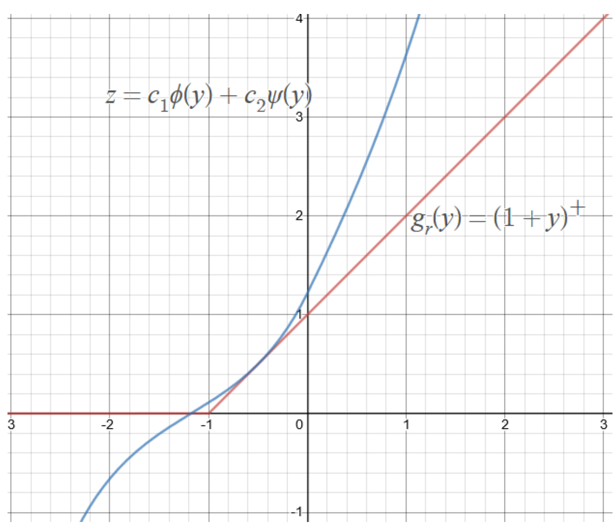

Figure 4.2: $X_0 = -0.5$

So then the question is, where does the stopping region begin? The stopping region begins at the smallest x value where $c_1 \geq 0$. For these particular parameters that occurs when $x = \frac{\sigma_1}{\sqrt{2r}} - 1 \approx -0.3675$. As shown below, for $x = \frac{\sigma_1}{\sqrt{2r}} - 1$ we can achieve smooth fit while keeping $c_1\phi(y) + c_2\psi(y) \geq g_r(y)$ true for all y. In fact we can achieve smooth fit and satisfy the constraints so long as $x \geq \frac{\sigma_1}{\sqrt{2r}} - 1$. However, if $x < \frac{\sigma_1}{\sqrt{2r}} - 1$ then we cannot.

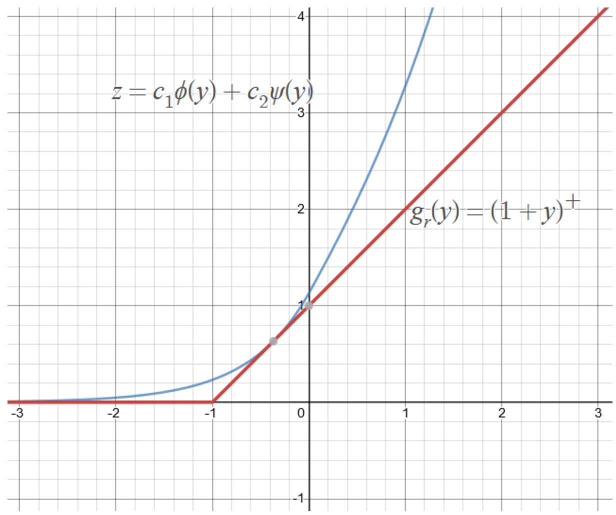

Figure 4.3: $X_0 = \frac{\sigma_1}{\sqrt{2r}} - 1$

Therefore our value function is:

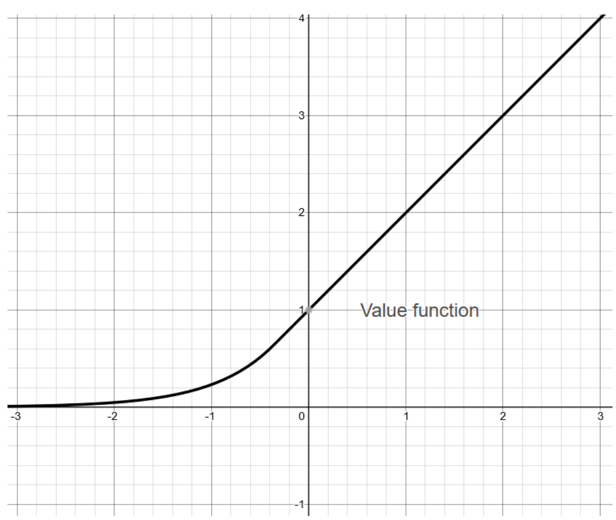

Figure 4.4: Value Function

4.2.2 Piecewise Quadratic Gain Function

Once again we appeal to the principle of smooth fit and are able to assume that c_1 and c_2 must be non-negative. Solving a similar system of equations as the linear gain function we find:

$$c_1 = \frac{(1+x)^2\psi'(x) - 2(1+x)\psi(x)}{\phi(x)\psi'(x) - \phi'(x)\psi(x)}, \qquad c_2 = \frac{2(1+x)\phi(x) - (1+x)^2\phi'(x)}{\phi(x)\psi'(x) - \phi'(x)\psi(x)}.$$

Substituting ϕ and ψ we have:

$$c_1 = \begin{cases} \frac{(1+x)\sigma_2}{\sqrt{2r}(\sigma_2+\sigma_1)} e^{\frac{\sqrt{2r}}{\sigma_1}x}[(1+x)\sqrt{2r} - 2\sigma_1], & x < 0 \\ \frac{(1+x)}{2\sqrt{2r}(\sigma_2+\sigma_1)} \bigg(e^{\frac{\sqrt{2r}}{\sigma_2}x}((1+x)\sqrt{2r} - 2\sigma_2)(\sigma_1+\sigma_2) & \\ +e^{\frac{-\sqrt{2r}}{\sigma_2}x}((1+x)\sqrt{2r} + 2\sigma_2)(\sigma_2-\sigma_1) \bigg), & x \geq 0 \end{cases}$$

and

$$c_2 = \begin{cases} \frac{(1+x)}{2\sqrt{2r}(\sigma_2+\sigma_1)} \bigg(e^{\frac{-\sqrt{2r}}{\sigma_1}x}((1+x)\sqrt{2r} + 2\sigma_1)(\sigma_1+\sigma_2) & \\ -e^{\frac{\sqrt{2r}}{\sigma_1}x}((1+x)\sqrt{2r} - 2\sigma_1)(\sigma_2-\sigma_1) \bigg), & x < 0 \\ \frac{(1+x)\sigma_1}{\sqrt{2r}(\sigma_2+\sigma_1)} e^{\frac{-\sqrt{2r}}{\sigma_2}x}[(1+x)\sqrt{2r} + 2\sigma_2], & x \geq 0. \end{cases}$$

First we will analyze c_1 to identify the areas that are known to be in the continuation region. When $x < 0$ we see that c_1 is negative when $(1+x)\sqrt{2r} - 2\sigma_1 < 0$ or $x < \sigma_1\sqrt{2/r} - 1$. When $r > 2\sigma_1^2$ this value is negative and the interval $(-\infty, \sigma_1\sqrt{2/r} - 1)$ is contained in the continuation region. When $r \leq 2\sigma_1^2$ the value is non-negative and due to the piecewise nature of c_1 all we can assure is that $(-\infty, 0)$ is in the continuation region.

Now when $x = 0$, $c_1 = \sigma_2(\sqrt{2r} - \sigma_1)/(2\sqrt{2r}(\sigma_2+\sigma_1))$ and we see that this value is negative if $r < 2\sigma_1^2$ so the continuation region extends (at least) until the zero of $(1+x)\psi'(x) - 2\psi(x)$.

However, the most interesting case happens when we have $2\sigma_1^2 < r < \sigma_2^2$. Mordecki and Salminen (2019) refer to this as "the bubble".

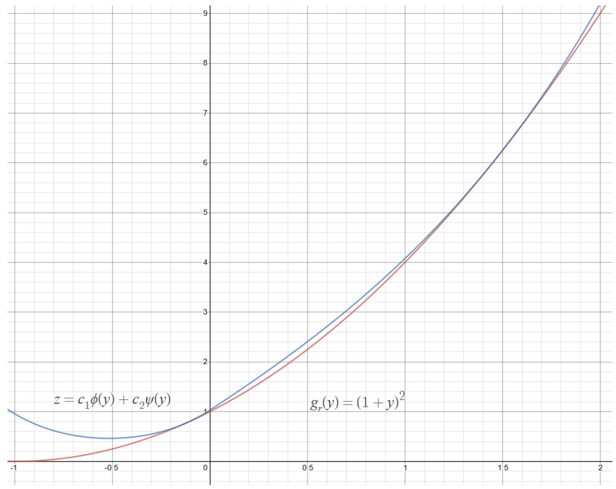

Figure 4.5: The Bubble

Here we solve the system of equations using parameters $r = 3.9$, $\sigma_1 = 1$ and $\sigma_2 = 4$. We are able to achieve smooth fit at two different values simultaneously, $x \approx -0.1136$ and $x \approx 1.4869$. These values are found by setting up a graphing utility as follows. We graph the gain function $g_r(y)$ and define $c_1(t), c_2(t)$ as above but now they are functions where t is the value for which we would like to see smooth fit. Then we graph the following function $z = c_1(t)\phi(y) + c_2(t)\psi(y)$ and use a slider on t.

In this particular case we begin with $t = 0$ are are able to observe that $c_1(0)\phi(y) + c_2(0)\psi(y) \geq g_r(y)$ are violated for some y so then $x = 0$ is in the continuation region. Then we can slide t towards -1 and observe the first time which the constraints are satisfied for all y, and then do the same in the positive direction. Perhaps unsurprisingly based on the graph above, we notice that $c_1(-0.1136) = c_1(1.4869)$ and $c_2(-0.1136) = c_2(1.4869)$.

In short, unlike most cases where we have a stopping region of the form $[c, \infty)$, here we have a stopping region with disjoint intervals $[a, b] \cup [c, \infty)$ where $a, b < 0$ and $c > 0$. This method gives a quick and dirty way to find the stopping and continuation regions for particular starting values of the process. For a more exact method, we can use numerical solvers. For the case where we have a bubble, we use the fact that there are some $a < 0$ and $b > 0$ where $c_1(a) = c_1(b)$ and $c_2(a) = c_2(b)$ and have the following system of equations to solve:

$$\begin{aligned}
c_1\phi(a) + c_2\psi(a) &= g_r(a)\\
c_1\phi'(a) + c_2\psi'(a) &= g_r'(a)\\
c_1\phi(b) + c_2\psi(b) &= g_r(b)\\
c_1\phi'(b) + c_2\psi'(b) &= g_r'(b)\\
-1 < a < 0, &\quad b > 0.
\end{aligned}$$

This system can be easily solved using a numerical solver such as Mathematica, the code for which is in Appendix A.1. Solving this system using the parameters listed above verifies the results that are found using the graphical method.

The most fascinating aspect of this example is that if we allow the parameters r, σ_1, and σ_2 to vary not only does the value function change but the form of the stopping rule can change. Based on the results in Mordecki and Salminen (2019) there is some $r_0 > 2\sigma_1^2$ so that as long as $r \in [r_0, \sigma_2^2)$ we have this bubble and the stopping region is of the form $[a, b] \cup [c, \infty)$ but once r leaves that interval, the stopping region is (c, ∞). However, if $r = r_0$ then we do have a bubble but $a = b$ and the stopping region is $\{a\} \cup [c, \infty)$.

Using the results of Theorem 3.4.3 we know that the value function is lower semi-continuous with respect to the parameters. If we focus on r, we can look at the behavior of the value function when we allow r to leave the interval $[r_0, \sigma_2^2)$. Since $V(x, \theta)$ is lower semi-continuous with respect to $\theta = (r, \sigma_1, \sigma_2)$ we know that for fixed σ_1 and σ_2 we have

$$\lim_{r \nearrow \sigma_2^2} V(x, r, \sigma_1, \sigma_2) \geq V(x, \sigma_2^2, \sigma_1, \sigma_2).$$

As r approaches σ_2^2 we expect there to be a bubble in the value function, and once $r = \sigma_2^2$ the bubble is gone. We could expect the value function to jump up but the inequality above suggests that does not happen so we suspect full continuity at $r = \sigma_2^2$. We would like to make a similar argument for $r = r_0$ but at that point, there is a bubble so we cannot. However, for the values $\sigma_1 = 1$ and $\sigma_2 = 4$ we can simulate the finitely constrained and compactly supported linear programs by using the python lp solver in Appendix A.2. Now we must understand that due to our previous results, this finitely contrained and compactly supported value function is continuous with respect to the parameters.

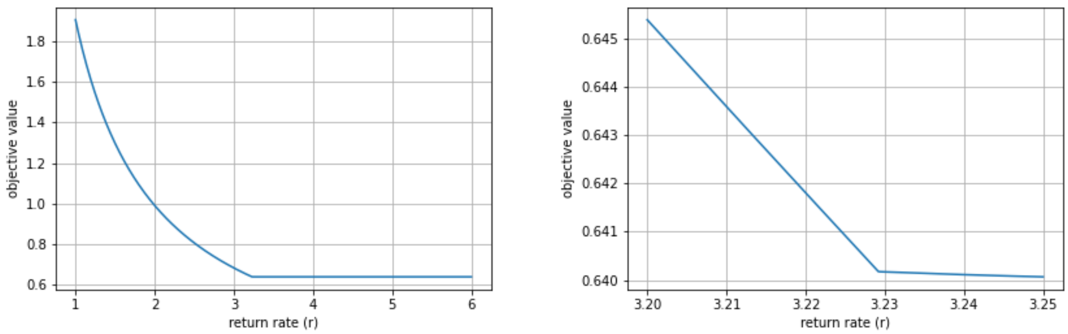

Figure 4.6: Continuity of Value function with respect to r when $x = -0.2$.

These graphs suggest that $r_0 \approx 3.228$ and that we do in fact have continuity when $r = r_0$.

4.2.3 Long Term Average Criterion with Linear Quadratic Gaussian Control

This example shows that we can extend this work to more classes of stochastic control. Here we have a long term average rather than a discounted gain function and our control is linear quadratic Gaussian instead of optimal stopping.

Let $C : E \times U \to \mathbb{R}$ be a cost rate function and let (X, Λ) be a relaxed weak solution of the martingale problem for (A, ν). We say (X, Λ) is a relaxed weak solution for (A, ν) if:

$$f(X_t) - f(x) - \int_0^t \int_U Af(X_s, u)\Lambda_s(du)ds \tag{4.5}$$

is a martingale. Note: Λ_s is a $\mathcal{P}(U)$-valued process.

Define the long term average criterion to be

$$J(X, \Lambda) = \limsup_{t\to\infty} \frac{1}{t}\,\mathbb{E}_x \left[\int_0^t \int_U C(X_s, u)\Lambda_s(du)ds\right] \tag{4.6}$$

and the lta (long term average) problem is to minimize $J(X, \Lambda)$ over relaxed solutions (X, Λ) of the martingale problem for A.

For this particular LQG (Linear Quadratic Gaussian) control problem the process X_t satisfies the SDE $dX_t = u_t dt + \sigma dW_t$, $X_0 = x$, so the generator of the process is $Af(x, u) = uf'(x) + 0.5\sigma^2 f''(x)$ and the cost rate function is $C(x, u) = mx^2 + nu^2$.

Now, for a given (X, Λ) we define the average occupation measure μ_t by

$$\mu_t(G_1 \times G_2) = \frac{1}{t}\mathbb{E}_x\left[\int_0^t \int_U I_{G_1 \times G_2}(X_s, u)\Lambda_s(du)ds\right]$$

This measure allows us to rewrite (4.6) as

$$\begin{aligned} J(X,\Lambda) &= \limsup_{t\to\infty} \int_{E\times U} C(x,u)\mu_t(dx \times du) \\ &= \int_{E\times U} C(x,u)\mu_\infty(dx \times du) \end{aligned}$$

where μ_∞ is defined to be a weak limit of (μ_t). Next, if we take equation 4.5 and rewrite it using this new measure, take expectations, and let $t \to \infty$ we obtain

$$\int_{E\times U} Af(x,u)\mu_\infty(dx \times du) = 0 \qquad \forall f \in dom(A)$$

Finally we can rewrite our lta problem as a linear program,

$$\begin{aligned} \text{minimize} &\int C(x,u)\mu(dx \times du) \\ \text{subject to} &\int Af(x,u)\mu(dx \times du) = 0 \qquad \forall f \in dom(A) \\ &\mu \in \mathcal{P}(E \times U) \end{aligned}$$

So then for our particular LQG we have:

$$\begin{aligned} \text{minimize} &\int [mx^2 + nu^2]\mu(dx \times du) \\ \text{subject to} &\int [uf'(x) + 0.5\sigma^2 f''(x)]\,\mu(dx \times du) = 0 \qquad \forall f \in dom(A) \\ &\mu \in \mathcal{P}(E \times U) \end{aligned} \tag{4.7}$$

One way to solve this problem is through exploiting the constraints and using carefully chosen functions. First for any measure μ on $E \times U$ we factor it into its marginal and conditional probabilities.

$$\begin{aligned} \int [uf'(x) + 0.5\sigma^2 f''(x)]\,\mu(dx \times du) &= 0 \\ \int\int [uf'(x) + 0.5\sigma^2 f''(x)]\,\eta(x, du)\,\mu_E(dx) &= 0 \\ \int [\bar{u}f'(x) + 0.5\sigma^2 f''(x)]\,\mu_E(dx) &= 0 \\ \int [\bar{u}f'(x) + 0.5\sigma^2 f''(x)]\,p(x)dx &= 0 \end{aligned}$$

where $\bar{u} = \int u\,\eta(du)$ and $\mu_E(dx) = p(x)dx$ (we assume μ_E is absolutely continuous with respect to Lebesgue measure).

Then for functions $f(x)$ where $f(\infty) = f(-\infty) = 0$ (e.g. Schwartz functions, functions of compact support, etc) we have

$$\int [\bar{u}f'(x) + 0.5\sigma^2 f''(x)]\, p(x)dx = 0$$
$$\int f(x)[0.5\sigma^2 p''(x) - (\bar{u}(x)p(x))']\, dx = 0.$$

If that is to be true for all such functions f, then $[0.5\sigma^2 p''(x) - (\bar{u}(x)p(x))'] = 0$ must be true. The solution to this differential equation is $p(x) = c\exp\left(\int_0^x \frac{2\bar{u}(v)}{\sigma^2}\,dv\right)$ where c is the constant so that $\int p(x)\,dx = 1$. Now we can revisit the objective function:

$$\begin{aligned}
&\int_{E\times U} [mx^2 + nu^2]\,\mu(dx \times du) \\
&= \int_{-\infty}^{\infty}\int_U [mx^2 + nu^2]\,\eta(x, du)\,\mu_E(dx) \\
&= \int_{-\infty}^{\infty} [mx^2 + n\bar{u}^2]\,\mu(dx) \\
&= \int_{-\infty}^{\infty} [mx^2 + n\bar{u}^2]\,p(x)dx \\
&= \int_{-\infty}^{\infty} [mx^2 + n\bar{u}^2]\,c\exp\left(\int_0^x \frac{2\bar{u}(v)}{\sigma^2}\,dv\right)dx
\end{aligned}$$

Recall that our goal is to find the optimal control $\bar{u}(x)$ to minimize the objective function. For now let us assume that $\bar{u}(x) = -kx$ for some constant k. Then $p(x)$ is a normal distribution and c can be calculated easily:

$$\begin{aligned}
p(x) =& c\exp\left(\int_0^x \frac{-2kv}{\sigma^2}\,dv\right) \\
=& c\exp\left(\frac{-kx^2}{\sigma^2}\right) \\
=& c\exp\left(\frac{-x^2}{2\left(\frac{\sigma}{\sqrt{2k}}\right)^2}\right).
\end{aligned}$$

This shows that $p(x)$ is the distribution function for a normal distribution with mean zero and standard deviation $\sigma/\sqrt{2k}$, which implies that $c = \sqrt{k/(\pi\sigma^2)}$ and the objective function is:

$$\begin{aligned}
&\int_{-\infty}^{\infty} [mx^2 + n(-kx)^2] \sqrt{\frac{k}{\pi\sigma^2}} \exp\left(\frac{-x^2}{2\left(\frac{\sigma}{\sqrt{2k}}\right)^2}\right) dx \\
=&(m + nk^2) \int_{-\infty}^{\infty} x^2 \sqrt{\frac{k}{\pi\sigma^2}} \exp\left(\frac{-x^2}{2\left(\frac{\sigma}{\sqrt{2k}}\right)^2}\right) dx \\
=&(m + nk^2)\left(\frac{\sigma^2}{2k}\right)
\end{aligned}$$

the last equality holds because the integral in line two is calculating the second moment of this normal distribution and since it has mean zero, the second moment is equal to the variance. Once again we wish to minimize this objective function. In this case our choice of control is simply to choose the value for k. The minimal value for k is easily found to be $\sqrt{m/n}$ and the value of the objective function is $\sigma^2\sqrt{mn}$. If we think of the value as a function of the parameters, i.e. $J(x, \theta) = \sigma^2\sqrt{mn}$ we see that J is clearly continuous with respect to the parameters σ, m, and n.

This solution seems unsatisfying because we ignore the possibility of having a different optimal control, such as quadratic. However, we can show that a linear control is in fact optimal by solving the problem a different way. First consider the Itô equation for the process X, for any suitable function f we have

$$f(X_t) = f(X_0) + \int_0^t \bar{u}_s f'(X_s) + 0.5\sigma^2 f''(X_s)\, ds + \int_0^t \sigma f'(X_s)\, dW_s.$$

Then take expectation and divide each term by t,

$$\begin{aligned}
\frac{1}{t}\mathbb{E}[f(X_t)] &= \frac{1}{t}\mathbb{E}[f(X_0)] + \frac{1}{t}\mathbb{E}\left[\int_0^t \bar{u}_s f'(X_s) + 0.5\sigma^2 f''(X_s)\, ds\right] + 0 \\
\frac{1}{t}\mathbb{E}[f(X_t)] &= \frac{1}{t}\mathbb{E}[f(X_0)] + \frac{1}{t}\mathbb{E}\left[\int_0^t Af(x, \bar{u})\, ds\right]
\end{aligned}$$

where the last integral in Itô's equation is a mean zero martingale and we used the definition of the generator in the last line. Then the Hamilton-Jacobi-Bellman equation for the long term average criterion is $\min_u\{AV(x, u) + mx^2 + nu^2\} = \lambda$ where λ is some constant. So then for any control u we have:

$$\begin{aligned}
\frac{1}{t}\mathbb{E}[V(X_t)] &\geq \frac{1}{t}\mathbb{E}[V(X_0)] + \frac{1}{t}\mathbb{E}\left[\int_0^t \lambda - (mX_s^2 + nu_s^2)\,ds\right] \\
\lim_{t\to\infty}\frac{1}{t}\mathbb{E}[V(X_t)] &\geq \lim_{t\to\infty}\frac{1}{t}\mathbb{E}[V(X_0)] + \lim_{t\to\infty}\frac{1}{t}\mathbb{E}\left[\int_0^t \lambda - (mX_s^2 + nu_s^2)\,ds\right] \\
0 &\geq \lambda - \lim_{t\to\infty}\frac{1}{t}\mathbb{E}\left[\int_0^t (mX_s^2 + nu_s^2)\,ds\right]
\end{aligned}$$

Notice that for the optimal control u^*, we would have equality throughout, and the last line would show that λ is in fact the long term average value. Returning to the HJB equation, $\min_u\{0.5\sigma^2 V''(x) + uV'(x) + mx^2 + nu^2\} = \lambda$ we have an equation that is quadratic in u so the minimum is easily found to be $u^* = -V'(x)/(2n)$. Then from the HJB equation,

$$\begin{aligned}
\frac{1}{2}\sigma^2 V''(x) - \frac{(V'(x))^2}{2n} + mx^2 + \frac{n(V'(x))^2}{4n^2} &= \lambda \\
\frac{1}{2}\sigma^2 V''(x) - \frac{(V'(x))^2}{4n} + mx^2 &= \lambda
\end{aligned}$$

and it is clear that $V(x)$ must be of the form ax^2+bx+c. Inserting this into the HJB equation yields $a\sigma^2 - a^2x^2/n - abx/n - b^2/(4n) + mx^2 = \lambda$. From here we find that $V(x) = \sqrt{mn}x^2 + c$ which gives $u^* = -x\sqrt{m/n}$ and $\lambda = \sigma^2\sqrt{mn}$ which agree with our previous solution. We can rule out $a = -\sqrt{mn}$ as a possible solution because that would give a negative value for λ, and the cost function makes this impossible.

Chapter 5

Future Directions and Conclusion

5.1 Future Directions

5.1.1 General Optimal Control

Our work so far has focused on optimal stopping as the implemented control. However, optimal control theory allows for a wide variety of more complicated controls. This work can be extended into analyzing continuity of the value function when we allow for other controls. One such example was analyzed previously in Section 4.2.3 where the cost function used a long term average instead of a discount, and the controls were linear quadratic Gaussian instead of choosing a stopping rule.

5.1.2 Risk Measures

Risk Measures are a well studied tool from the field of actuarial sciences. They are typically used as a way to quantify risk using historical statistical data. Future research would involve finding a distribution on our parameter space, and then we could use the same tools to quantify how badly the value function could change under perturbations of the model parameters.

5.2 Conclusion

To summarize, we have analyzed the value function for the optimal stopping of a one dimensional diffusion process. The main results from Helmes and Stockbridge (2010) show how to embed this optimal stopping problem into a semi-infinite linear program and they also show how to construct the dual program. Additionally, they prove that the primal and dual programs enjoy strong duality. In this paper, we first constructed an approximation to the semi-infinite dual program by only using countably many constraints. One of our results was showing that the optimal value for this approximation agreed with the optimal value of the semi-infinite dual program.

We then constructed a sequence of finitely constrained dual programs. We proved that the sequence of value functions for their respective dual programs converged pointwise to the value function of the semi-infinite dual program. We also proved that the value function for any finite dual program was continuous in the parameter space, which then led to our first continuity result. The value function for the semi-infinite dual program is lower semi-continuous in the parameters.

Our last result required us to instill further structure to the dual program. If our constraint system was indexed over a compact set and we had some additional smoothness of the gain functions then we showed that the value function enjoyed full continuity in the parameter space.

Bibliography

Borodin, A. N. and Salminen, P. (2002). *Handbook of Brownian motion—facts and formulae.* Probability and its Applications. Birkhäuser Verlag, Basel, second edition.

Helmes, K. and Stockbridge, R. H. (2010). Construction of the value function and optimal rules in optimal stopping of one-dimensional diffusions. *Adv. in Appl. Probab.*, 42(1):158–182.

Jongen, H. T., Twilt, F., and Weber, G.-W. (1992). Semi-infinite optimization: structure and stability of the feasible set. *J. Optim. Theory Appl.*, 72(3):529–552.

Mordecki, E. and Salminen, P. (2019). Optimal stopping of oscillating Brownian motion. *arXiv e-prints*, page arXiv:1903.01457.

Peskir, G. and Shiryaev, A. (2006). *Optimal stopping and free-boundary problems.* Lectures in Mathematics ETH Zürich. Birkhäuser Verlag, Basel.

Robinson, S. M. (1977). A characterization of stability in linear programming. *Operations Res.*, 25(3):435–447.

Rogers, L. C. G. and Williams, D. (2000). *Diffusions, Markov processes, and martingales. Vol. 2.* Cambridge Mathematical Library. Cambridge University Press, Cambridge. Itô calculus, Reprint of the second (1994) edition.

Shiryaev, A. N. (1978). *Optimal stopping rules.* Springer-Verlag, New York-Heidelberg. Translated from the Russian by A. B. Aries, Applications of Mathematics, Vol. 8.

Appendix A

Code

A.1 Mathematica

```
{r, s1, s2} = {2.49, 1, 2.5};
l1 = Sqrt[2 * r] / s1;
l2 = -Sqrt[2 * r] / s1;
l3 = Sqrt[2 * r] / s2;
l4 = -Sqrt[2 * r] / s2;
A1 = (1 + s1 / s2) / 2;
A2 = (1 - s1 / s2) / 2;
B1 = (1 + s2 / s1) / 2;
B2 = (1 - s2 / s1) / 2;
phi[x_] := Piecewise[{{A1 * Exp[l2 * x] + A2 * Exp[l1 * x], x < 0}, {Exp[l4 * x], x ≥ 0}}];
psi[x_] := Piecewise[{{Exp[l1 * x], x < 0}, {B1 * Exp[l3 * x] + B2 * Exp[l4 * x], x ≥ 0}}];
g[x_] := Piecewise[{{(1 + x)^2, x ≥ -1}, {0, x < -1}}];
t = FindRoot[{c * phi[a] + d * psi[a] - g[a], c * phi'[a] + d * psi'[a] - g'[a],
     c * phi[b] + d * psi[b] - g[b], c * phi'[b] + d * psi'[b] - g'[b]},
    {a, -0.1, -1, 0}, {b, 0.8, 0, 100}, {c, 1}, {d, 1}, WorkingPrecision → 30];
{c, d, e1, e2} = t[[All, 2]];
Plot[{e1 * phi[x] + e2 * psi[x], g[x]},
   {x, -1, 1}, PlotRange → {-0.5, 3}, PlotLegends -> "Expressions"];
```

A.2 Python

```
import numpy as np
import math
from pulp import *
import matplotlib.pyplot as plt

#Define Parameters
#r=3.3
s1=1
s2=4
z=-0.114


#Gain function
def g(y):
    if y<= -1:
        return 0
    if y > -1:
        return (1+y)**2

#Phi function
def phi(y,r):
    if y < 0:
        return
((s2+s1)/(2*s2))*math.exp(-y*math.sqrt(2*r)/s1)+((s2-s1)/(2*s2))*math.exp(y*math.sqr
t(2*r)/s1)
    if y >= 0:
        return math.exp(-y*math.sqrt(2*r)/s2)

#Psi function
def psi(y,r):
    if y < 0:
        return math.exp(y*math.sqrt(2*r)/s1)
    if y >= 0:
        return
((s2+s1)/(2*s1))*math.exp(y*math.sqrt(2*r)/s2)+((s1-s2)/(2*s1))*math.exp(-y*math.sqr
t(2*r)/s2)

#LP Solver the LP
def linsolve(r,lb,ub,meshsize):
    c1 = LpVariable("c1",lowBound=0)
    c2 = LpVariable("c2",lowBound=0)
    prob = LpProblem("myProblem", LpMinimize)
    prob+= phi(z,r)*c1 + psi(z,r)*c2
    A = np.linspace(lb,ub,meshsize).tolist()
    for a in A:
        prob+= LpConstraint(e=c1*phi(a,r) + c2*psi(a,r), sense=LpConstraintGE,
name=None, rhs=g(a))
    status=prob.solve()
    return value(prob.objective)
```

Appendix B

Other Results

The following theorem was proved when we thought the primal problem would be the key to our main results and were toying with using the finitely supported measures.

Theorem B.0.1. *Given any measure $\mu \in \mathcal{F}(\mathcal{P})$ there exists a sequence of measures on $\mathcal{Y}^\infty$ which converge strongly to μ.*

Proof. Let $\mu \in \mathcal{F}(\mathcal{P})$ be arbitrary. Assume without loss of generality that $\mathcal{Y}^N$ is ordered. Define a measure μ_N on $\mathcal{Y}^N$ as follows:

$$\mu_N(\{y_i\}) = \mu((y_{i-1}, y_i]), \quad i \geq 1$$
$$\mu_N(\{y_0\}) = \mu(\{y_0\})$$

We can extend μ_N to a measure on $\mathscr{B}([x_l, x_r])$ by:

$$\mu_N(A) = \sum_{i=0}^{N} \{\mu_N(\{y_i\}) |\, y_i \in A\}$$

Define $\hat{\mu}$ as the limit of the μ_N measures. That is,

$$\hat{\mu}(A) = \lim_{N \to \infty} \mu_N(A)$$

Claim. $\hat{\mu}(A) = \mu(A) \qquad \forall A \in \mathscr{B}([x_l, x_r])$.

Proof of claim. Since $A \in \mathscr{B}([x_l, x_r])$ it suffices to prove this for sets of the form (x_1, x_2) which generate the Borel σ-algebra.

Then

$$
\begin{aligned}
\hat{\mu}((x_1,x_2)) &= \lim_{N\to\infty} \mu_N((x_1,x_2)) \\
&= \lim_{N\to\infty} \sum_{i=0}^{N} \{\mu_N(\{y_i\}) \,|\, x_1 < y_i < x_2\} \\
&= \lim_{N\to\infty} \sum_{i=0}^{N} \{\mu((y_{i-1},y_i]) \,|\, x_1 < y_i < x_2\} \\
&= \lim_{N\to\infty} \{\mu((y_{i-1},y_j]) \,|\, y_{i-1} < x_1 < y_i \quad and \quad y_j < x_2 < y_{j+1}\} \\
&= \mu((x_1,x_2))
\end{aligned}
$$

□

CURRICULUM VITAE

Samuel Nehls

Place of Birth: Park Falls, WI

Education

2015–2020	PhD	Mathematics	University of Wisconsin - Milwaukee
2013–2015	MS	Mathematics	University of Wisconsin - Milwaukee
2008–2013	BS	Mathematics	University of Wisconsin - Stevens Point

Dissertation Adviser: Richard Stockbridge

Dissertation Title: Analysis of the Continuity of the Value Function of an Optimal Stopping Problem

Employment History

- Graduate Teaching Assistant at University of Wisconsin - Milwaukee: September 2013 - present
- Senior Medical NCO in US Army Reserve: June 2015 - present
- Math Tutor at University of Wisconsin - Stevens Point: January 2011 - May 2013

Honors and Awards

Ernst Schwandt Teaching Award from the Department of Mathematics, May 2016. This is awarded to recognize teaching excellence among graduate teaching assistants.

William Lowell Putnam Competition Award from UW - Stevens Point, May 2013. This award recognizes the highest score achieved on the Putnam Exam for this year.

Distinguished Achievement Award from UW - Stevens Point, April 2012. This award is given to two students in each major discipline.

Teaching Experience

- Taught the three courses in the calculus sequence, fall 2018 - fall 2019.

- Taught business calculus both as lecturer and TA in discussion sections,
- Taught 5-hour remedial math course,
- Taught algebra, pre-calculus, and trigonometry course,
- Substitute taught 300 level probability course, fall 2019.

Other Professional Experiences & Skills

1. Developed working knowledge of Python, Mathematica, and AMPL.
2. Implemented Combat Life Saver course as lead instructor at a chemical battalion in Chicago IL in July 2016.
3. Proctored exams in calculus testing center which sees over 1500 visitors a semester from fall 2018 to fall 2019.
4. Minored in Spanish at UW - Stevens Point and passed the foreign language exam for the PhD requirement demonstrating an intermediate knowledge of Spanish.

Major areas of research interest

My dissertation in is the area of probability, specifically optimal control theory. This research focuses heavily on linear programming methods in which I have a strong background from undergraduate courses. My work analyzes the effects of perturbations of the parameters of the problem on the value function. In addition, I have taken several additional algebra and topology courses beyond the PhD requirements to broaden my mathematical knowledge.