Optimal Stopping Problems with A Random Time Horizon

Author:
Wu, Kris

Publication Date:
2023

DOI:
https://doi.org/10.26190/unsworks/24643

License:
https://creativecommons.org/licenses/by/4.0/
Link to license to see what you are allowed to do with this resource.

Downloaded from http://hdl.handle.net/1959.4/100936 in https://unswworks.unsw.edu.au on 2023-07-10
Optimal Stopping Problems with
A Random Time Horizon

by
Zhuoshu Wu

A dissertation
submitted to School of Mathematics and Statistics
of University of New South Wales
in partial fulfillment of the requirements
for the degree of
Doctor of Philosophy

October 2022
Inclusion of Publications Statement

UNSW is supportive of candidates publishing their research results during their candidature as detailed in the UNSW Thesis Examination Procedure.

Publications can be used in the candidate's thesis in lieu of a Chapter provided:

- The candidate contributed **greater than 50%** of the content in the publication and are the "primary author", i.e. they were responsible primarily for the planning, execution and preparation of the work for publication.
- The candidate has obtained approval to include the publication in their thesis in lieu of a Chapter from their Supervisor and Postgraduate Coordinator.
- The publication is not subject to any obligations or contractual agreements with a third party that would constrain its inclusion in the thesis.

The candidate has declared that some of the work described in their thesis has been published and has been documented in the relevant Chapters with acknowledgement.

A short statement on where this work appears in the thesis and how this work is acknowledged within chapter(s):

Mentioned at the beginning of the Chapter

Candidate's Declaration

I declare that I have complied with the Thesis Examination Procedure.
ORIGINALITY STATEMENT

‘I hereby declare that this submission is my own work and to the best of my knowledge it contains no materials previously published or written by another person, or substantial proportions of material which have been accepted for the award of any other degree or diploma at UNSW or any other educational institution, except where due acknowledgement is made in the thesis. Any contribution made to the research by others, with whom I have worked at UNSW or elsewhere, is explicitly acknowledged in the thesis. I also declare that the intellectual content of this thesis is the product of my own work, except to the extent that assistance from others in the project’s design and conception or in style, presentation and linguistic expression is acknowledged.’

Signed: ........................

Date: ............................
COPYRIGHT STATEMENT

‘I hereby grant the University of New South Wales or its agents a non-exclusive licence to archive and to make available (including to members of the public) my thesis or dissertation in whole or part in the University libraries in all forms of media, now or here after known. I acknowledge that I retain all intellectual property rights which subsist in my thesis or dissertation, such as copyright and patent rights, subject to applicable law. I also retain the right to use all or part of my thesis or dissertation in future works (such as articles or books).’

‘For any substantial portions of copyright material used in this thesis, written permission for use has been obtained, or the copyright material is removed from the final public version of the thesis.’

Signed: ........................
Date: ...........................

AUTHENTICITY STATEMENT

‘I certify that the Library deposit digital copy is a direct equivalent of the final officially approved version of my thesis.’

Signed: ........................
Date: .............................
He had come a long way to this blue lawn, and his dream must have seemed so close that he could hardly fail to grasp it. He did not know that it was already behind him ... It eluded us then, but that's no matter — tomorrow we will run faster, stretch out our arms further ... And one fine morning — So we beat on, boats against the current, borne back ceaselessly into the past.

F. Scott Fitzgerald, *The Great Gatsby*, 1926
Abstract

The theory of optimal stopping provides a powerful set of tools for the study of American contingent claim pricing problem in mathematical finance. We give a self-contained overview of the theory, including the complete proofs of existence and uniqueness theorems for the optimal stopping time in finite-time formulation. These theorems are developed with the goal of formulating the corresponding free-boundary problems for the valuation of atypical American contingent claims involving non-stopping times, which restrict the optimal stopping rules to be made before the last exit times of the price of the underlying assets at its running maximum or at any fixed level.

By exploiting the theory of enlargement of filtrations associated with random times, the original pricing problem with random times can be transformed into an equivalent optimal stopping problem with a semi-continuous, time-dependent gain function, whose partial derivative is singular at certain points. The difficulties in establishing the monotonicity of the optimal stopping boundary, the regularity of the value function and its differentiability in the boundary lie essentially in these somewhat unpleasant features of the gain function.

However, it turns out that a successful analysis of the continuation and stopping sets with proper assumptions can help us overcome these difficulties and further obtain the important properties that possessed by the free-boundary, which eventually leads us to the desired free-boundary problem. After this, we derive the nonlinear integral equations that characterise the free-boundary and the value function. The solutions to these equations are examined in details, and their financial justification is discussed briefly in the final chapter.
Acknowledgements

To my supervisors, Dr Libo Li and Professor Ben Goldys: a sincere thank-you for your guidance, advice over the years and not least, for reading various draft of this work in whole or in part, for helping me see ways of making it a better piece of work. For me, at least, the process has been a joy. I hope it has been for you as well.

To the staff of the School of Mathematics and Statistics who have given me so generously of their time and patience to help and support me: a heartfelt thank-you.

To University of New South Wales and the School of Mathematics and Statistics: a special thank-you for financially supporting me during my PhD journey, under the University International Postgraduate Award, Mathematics funded Top-ups and HDR Completion Scholarship.

To my friends from Room 4081 and the park run squad (you know who you are): thank you for sweetening those bitter days, for sharing your own struggles and most of all, for being there for me. All I wish is that, like Santiago the shepherd boy (from The Alchemist), we all have the courage to confront our personal calling, to believe that the Universe is conspiring in our favor, even though we may not understand how and that we deserve the miracle of life.

And to my parents: thank you. For everything.
Contents

1 Introduction 1

2 Optimal Stopping in Mathematical Finance: Background 7
  2.1 Probability Theory: Convergence and Others ......................... 8
  2.2 Stochastic Analysis: Basics ........................................ 10
  2.3 Martingales ....................................................... 15
  2.4 Returning To An Old Promise: Arbitrage ............................ 18

3 Optimal Stopping: The General Results 24
  3.1 The Continuation and Stopping Sets .................................. 24
  3.2 The Optimal Stopping Time: Existence and Uniqueness ............ 26
  3.3 Free-Boundary Problems In A Nutshell ............................. 33

4 Russian Option with Last Exit Time 37
  4.1 Infinite-time Horizon .............................................. 38
    4.1.1 Reformulation and Basics .................................... 38
    4.1.2 The Free-boundary Problem .................................. 41
  4.2 Finite-time Horizon ................................................ 47
    4.2.1 Reformulation and Basics .................................... 47
    4.2.2 The Free-boundary Problem .................................. 49
    4.2.3 The Continuation and Stopping Sets ............................ 52
    4.2.4 The Optimal Stopping Rule ................................... 57

5 American Put Option with Last Exit Time 67
  5.1 Infinite-time Horizon .............................................. 68
    5.1.1 Reformulation and Basics .................................... 68
    5.1.2 The Free-boundary Problems .................................. 70
  5.2 Finite-time Horizon ................................................ 76
<table>
<thead>
<tr>
<th>CONTENTS</th>
</tr>
</thead>
<tbody>
<tr>
<td>5.2.1 Reformulation and Basics .....................................</td>
</tr>
<tr>
<td>5.2.2 The Free-boundary Problem ......................................</td>
</tr>
<tr>
<td>5.2.3 The Continuation and Stopping Sets .............................</td>
</tr>
<tr>
<td>5.2.4 The Optimal Stopping Rule .......................................</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>6 Additional Remarks</th>
</tr>
</thead>
<tbody>
<tr>
<td>116</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>A Auxiliary Results</th>
</tr>
</thead>
<tbody>
<tr>
<td>121</td>
</tr>
<tr>
<td>A.1 Computations Associated with Azéma Supermartingale ........</td>
</tr>
<tr>
<td>A.1.1 For Chapter 4 ...................................................</td>
</tr>
<tr>
<td>A.1.2 For Chapter 5 ...................................................</td>
</tr>
<tr>
<td>A.2 Useful Results for the Proof of Smooth-fit Condition .......</td>
</tr>
<tr>
<td>A.3 Property of the Value Function .................................</td>
</tr>
<tr>
<td>A.4 The Uniform Integrability of $Y$ ...............................</td>
</tr>
<tr>
<td>A.5 Other Simple Results ..............................................</td>
</tr>
<tr>
<td>A.6 Algorithmic Remarks ...............................................</td>
</tr>
</tbody>
</table>
Notation and Symbols

\( \Omega \) The sample space
\( \omega \) An elementary event
\( \mathcal{F} \) The \( \sigma \)-algebra of events
\( \mathbb{R} \) The real numbers
\( \mathbb{N}_0 \) \( \mathbb{N}_0 = \{0, 1, 2, 3, \ldots \} \) is the set of non-negative integers
\( C^n \) space of functions with continuous derivatives up to order \( n \in \mathbb{N}_0 \)
\( C^{1,2}(\mathbb{R} \times \mathbb{R}^n) \) space of functions \( u = u(t, x) \)
with continuous first order derivative in the “time” variable \( t \in \mathbb{R} \) and
continuous second order derivatives in the “spatial” variable \( x \in \mathbb{R}^n \)
\( x^+ \) \( \max\{x, 0\} \)
\( x \wedge y \) \( \min\{x, y\} \)
\( x \vee y \) \( \max\{x, y\} \)
\( I\{A\} \) Indicator function of (the set) \( A \)
\( A^c \) Complement of \( A \)
\( \partial A \) The boundary of \( A \)
\( P(A) \) Probability of \( A \)
\( P(A|B) \) Conditional probability of \( A \) given \( B \)
\( \mathcal{L}_p^B \) space of functions locally integrable of order \( p \)
\( \Phi(x) \) Standard normal distribution function
\( \phi(x) \) Standard normal probability density function
\( X^x \) Stochastic process \( X \) starting at \( X_0 = x \)
\( X \equiv Y \) \( X \) and \( Y \) are equidistributed
\( X_n \overset{a.s.}{\rightarrow} X \) \( X_n \) converges almost surely to \( X \), \( P(\{\omega : X_n(\omega) \to X(\omega)\}) = 1 \)
\( X_n \overset{p}{\rightarrow} X \) \( X_n \) converges in probability to \( X \), \( P(|X_n - X| > \epsilon) \to 0 \)
\( X_n \overset{d}{\rightarrow} X \) \( X_n \) converges in distribution to \( X \)
a.s. Almost surely
i.e. \( Id est \), that is
iff If and only if
i.i.d Independent, identically distributed
i.o. Infinitely often
w.r.t With respect to
Chapter 1

Introduction

The present work is intended partly as a survey and partly as a forum for new results. Its subject is the application of optimal stopping theory in contingent claim valuation (option pricing), a highly specialised but nonetheless important topic in financial mathematics. For purposes of introduction, certain terms will be used in a temporary narrow sense and some of the mathematical definitions will be stated informally.

Our Aristotelian aitia

Consider an absurdly simple world, so that we can focus on what counts, with only one stock (risky asset), one bond (non-risky asset) in the market which is free of frictions (i.e. the market involves no transaction costs and no restriction in selling short). Then, there is this ticket (let us named it Ticket $A$ and its contents might vary) that entitles its holder to buy one share of stock for a specified $K$ dollars at any time before $T$, the expiration date of this ticket.

Now suppose an agent, hereafter called you, consider buying such a ticket: clearly, if at any time $t$ before $T$, the stock price $S_t$ has gone up above $K$ dollars, you can exercise your option to buy one share of the stock for $K$ dollars and turn around selling it for $S_t$ dollars, making an easy $(S_t - K)$-dollar profit. But of course, there is chance that the stock price has gone down below $K$ dollars, which makes the ticket worthless at time $t$ (note that in the near future, the price might go further up/down so current decision might not be a smart move). How much would you be willing to pay for Ticket $A$ at time $0$? The parallel question in optimal stopping is: you buy Ticket $A$ at time $0$ for its fair price (unknown) $V_0$ dollars, when will be the perfect timing (i.e. optimal stopping time) to exercise your option to buy?

At first glimpse, it seems perfectly reasonable to say that different agents shall have different answers, depending on their attitudes toward risk bearing or even their knowledge toward the future of the market. But, under the most fundamental hypothesis of efficient (rational) market in economics, it would be possible to find a unique rational value for such ticket. To be more specific, the efficiency means that the market responds rationally to new information so that the corrections of prices are instantaneous, leaving no room for arbitrage; the collectively rational investors correct their decisions instantaneously as the new information becomes available.

From the view of efficient market (it is no doubt a sweet dream of human kinds, even though
scientists start to rethink and reexamine its practicality, see [3] and [81]), we realise that the ultimate goal of financial mathematics is never to gain over the market or make risk-less profits but to help investors find rational strategies to a wide range of problems by taking account of the random factors in the economic environment (and eventually build an efficient market!). Among all these problems, the one attracts major attentions from researchers and professionals in the financial industry is the valuation of contingent claim (the formal name of Ticket A), a contract whose value depends on one or more securities or assets.

It is also due to the hypothesis of efficient market that the concept of arbitrage-free and the martingale property of the prices (if the fair price of Ticket A is a martingale, then the expected state of its future given the past history equals that of its present) are able to make their entrances and their very connections are revealed via the Fundamental Asset Pricing Theorems 1.

Therefore, the efficient market hypothesis provides a basic framework for us to further construct an arbitrage free market model so that our attempt to searching a unique rational solution mathematically to the contingent claim valuation problem will be possible.

**Historical Remarks**

From the outset of financial mathematics, there were attempts to describe the dynamics of the stock price mathematically. At first, Bachelier [4] proposed using the model based on random walk and their limited cases which was later what Einstein called Brownian motion, the irregular, random motion of small particles of dust suspended in a liquid, in his paper [26] written in 1905.

More than five decades later, the attempt was carried on by Kendall and Hill [50] via analysing the data of stock price, his intention was to detect the patterns (rhythms, trends or cycles) of the stock price but only found that “the demon of chance draw a random number and added it to the current price to determine the next price”, i.e. (the logarithms of the prices behave as) random walk. Although the use of random walk was not by then widely accepted, it did provide a natural (initial) home for martingale (a term originates in gambling theory) property.

After the findings of Kendall, the search continued and grossed increasing interests over years. But it was not until 1965 when Samuelson [71] developed the idea of geometric (economic) Brownian motion with positive drift and put forward the simplest argument that the stock price cannot go negative because of limited liability that the search eventually settled, leading to the era of using geometric Brownian motion as the model to describe the dynamics of stock price.

Under these models, researchers were able to obtain various yet somewhat unsatisfactory valuation formula (in the sense that they involved one or more arbitrary parameters) under different assumptions around 1960s. The highlight during this period was the idea of treating the valuation of American contingent claim as a question in optimal stopping and its associated free-boundary problem initiated by McKean [53] and Samuelson [71], which would be carried over further by Moerbeke [58], but their financial justification at the moment was rather murky.

The major breakthrough was made in 1973 when Black and Scholes [7] came along and postulated the “ideal conditions” in the market for the stock and the options, under which, a theoretical valuation

---

1See [63, Pages 23 and 31, Theorem 2.15 and Theorem 2.29] for discrete-time market, [76, Page 655, A List of Results] for continuous-time market.
formula for (European call) option was derived: suppose Ticket A has changed its content and it now only entitles you to exercise your option to buy on share of stock for \( K \) dollars at \( T \), the unique fair price of Ticket A under Black-Scholes Model is given as

\[
V_0 = e^{-rT} E_Q (S_T - K)^+ .
\]

In the late 70s and early 80s, their theory would further mature thanks to numerous researchers (just to name a few here): Merton \([56, 57]\) proposed alternative explanation to resolve some controversial argument in Black and Scholes methodology and made subsequent modification of the basic Black-Scholes model; Cox and Ross \([14, 15]\) employed Black-Scholes type analysis to examine the problem with discontinuous stock price dynamics; Harrison and Pliska \([41]\) considered continuous trading and unfolded more explicitly the connection between arbitrage and the existence and uniqueness of the equivalent martingale measure. The valuation formula during this period was generalised for every attainable contingent claim with payoff \( X \) (non-negative random variable):

\[
V_0 = e^{-rT} E_Q (X) ,
\]

and under continuous trading (see \([41, \text{Page 240}]\))

\[
V_t = E_Q \left( e^{-r(T-t)} X \mid F_t \right) .
\]

Fast forward, a rigorous theory of the valuation of American contingent claim in the complete market was then put forward by Bensoussan \([6]\), which was later simplified and extended by Karatzas \([46, 47]\): consider Ticket A now entitles you to sell one share of the stock for \( K \) dollars at any time before and including \( T \) (this is the well-known American put option), the unique fair price of this American put option at time \( t \in [0, T] \) is

\[
V_t = \text{ess sup}_{\tau \in [t, T]} E_Q \left( e^{-r(T-t)} (K - S_\tau)^+ \mid F_t \right) .
\]

Most importantly, they were able to resonate the proper financial justification lacked in the early work of free-boundary problem by using the hedging argument. Humanity’s deepest desire for knowledge is justification enough for our continuing quest and the story we are fond of telling about optimal stopping in contingent claim valuation starts really from here.

**A Short Outlook on Optimal Stopping**

From the theory of optimal stopping, the value process of the optimal stopping problem can be characterised as the smallest supermartingale majorant to the stopping reward. As such, the value of American contingent claim has a Riesz decomposition into martingale (the reward at the terminal date) and potential process (the early exercise premium representation), see \([59]\). The result for American put option (as the integral equation) can be derived from a simple consequence of McKean’s free boundary formulation.

The justification of the existence, uniqueness and regularity of the optimal stopping boundary from the integral equation were left open in McKean’s work, but they were later verified by Moerbeke.
and Jacka [44] under different settings. While reviewing the methodologies used in the American
contingent claim pricing problem, Myneni [59] pointed out that “the uniqueness and regularity of the
stopping boundary from this integral equation remains open”. Since then, a lot of efforts had been
made to answer these questions theoretically and numerically, among them, one of the most path-
breaking work is that of Peskir [65] which answered the the questions raised by Myneni affirmatively
based on his previous work [66] in 2002 on the change-of-variable formula.

From the perspectives of present work, the argument particularly consistent to the main result
presented here is that by Peskir and Shiryaev [64]. Their work is able to provide a rather systematic
methodology in a step-by-step fashion to tackle a wide range of optimal stopping problem associated
with the American contingent claim valuation (including American put option, Russian option and
Asian option). Let us briefly summarise their method here: (i) transform the targeted valuation
problem into optimal stopping problem; (ii) determine the existence of optimal stopping time; (iii)
analyse the continuation and stopping regions; (iv) derive the properties (including its monotonicity
and continuity) of the optimal stopping boundary; (v) prove the uniqueness of the solution from the
nonlinear integral equation.

An additional comment on such method is that it is (will still be?) highly dependent on the
monotonicity of the boundary, that is, if we are not able to accomplish step (iv), nothing further such
as the justification of the regularity of both the boundary and the value function can be proceeded
properly. Attempts are made in this direction, for instance, the probabilistic argument was proposed
by De Angelis and Stabile [18], De Angelis and Peskir [17].

A New Factor in Classic Pricing Problem

Readers might inevitably wonder: so what is new here?

A new factor that lies at the heart of current work is the presence of the last exit time in the classic
valuation of American contingent claim. For a solid introduction and background of last exit time
that is not stopping time, we refer to [61, Chapter 8.2]. The most interesting feature of the last exit
time is its involvement of “the knowledge of the future”, which suggests that this work is closely
related to the optimal prediction problem, see [10, 36, 38, 49] where this problem is also known as
the stopping rule problem, the secretary problem and the optimal selection problem; and for the more
recent work, see [24, 25]. For some financial applications of this factor that are remotely relevant to
this work, we refer to the monographs [27, 30, 8].

Because of this new factor, a significant amount of our attention is directed to the following
question: suppose that there are two options in that absurdly simple world, whose terminal date are
set as $T$ (which can be indefinitely long),

**Option B** entitles its holders to sell the stock at the highest price it has ever been traded during
the time frame between its purchase time and its exercise time or the last time the stock price
at its running maximum (whichever comes first); otherwise, it becomes worthless.

**Option C** entitles its holders to sell one share of stock for $K$ dollars at any time during the
time frame between its purchase time and its exercise time or the last time the stock price
price at any fixed level $L$ (whichever comes first); otherwise, it becomes worthless.

How much are you willing to pay for them?
With the aid of the classic pricing theory, the question posed above can be formulated into a precise mathematical form (clearly, any rational agent will want to maximise the stopping reward of Option B and C):

\[
V_B = \text{ess sup}_{\tau \in [0,T]} \mathbb{E}_Q \left( e^{-r\tau} e^{-\lambda\tau} \left( \max_{u \in [0,\tau]} S_u - LS_\tau \right)^+ I\{\tau < \theta\} \right),
\]

(1.0.1)

\[
V_C = \text{ess sup}_{\tau \in [0,T]} \mathbb{E}_Q \left( e^{-r\tau} (K - S_\tau)^+ I\{\tau < \theta\} \right),
\]

(1.0.2)

with \( \tau \) and \( \theta \) being the exercise time and the last exit time.

By invoking the theory of *enlargements of filtrations* studied by Mansuy and Yor [54] and Nikeghbali [61], both problems (1.0.1) and (1.0.2) can be further reduced to the equivalent optimal stopping problems without the presence of last exit time; for which, the classic method proposed in [64] can be partially applied. The resulting gain functions for finite terminal date are semi-continuous, time-dependent with their partial derivatives being singular at certain points. The difficulty in establishing the monotonicity of the optimal stopping boundary, the regularity of the value function and its differentiability in the boundary all comes from these essential yet somewhat unpleasant features of the gain functions.

As it turns out, these difficulties can all be resolved by a successful analysis of the continuation and stopping sets. Furthermore, under additional assumptions, we can establish the important properties that possessed by the optimal stopping boundaries, which in turn leads us to the desired free-boundary problem and eventually to the nonlinear integral equation that characterise the optimal stopping boundaries and the value functions.

Some useful references in this research area are [24, 25, 37]. In particular, the time-dependent reward problem was also studied in Du Toit, Peskir and Shiryaev [24, 25], Glover, Peskir and Samee [37]. A somewhat different approach to verify the regularity the optimal stopping boundary was developed by De Angelis [16]. A treatment of the falling-apart smooth-fit condition was given by Qiu [70], Detemple and Kitapbayev [21].

**Outline of the Dissertation**

Here we provide a quick overview of the contents of this dissertation.

Chapter 2 provides the relevant mathematical background on probability theory, stochastic analysis and the arbitrage pricing theory, which is necessary for a solid treatment of the material.

Chapter 3 reviews the optimal stopping theory in infinite time horizon and discusses its extension to the finite time formulation. We have found it necessary to detail the proof to see what hypotheses are really required. Classic arbitrage pricing examples are also provided with the hope that newcomers will be tempted to look further into the literature.

The heart of this work is Chapter 4 and 5, which contain the unique pricing formulas for Options B and C. We begin both chapters by discussing the fairly simple perpetual options pricing problem and their corresponding price are presented in Theorem 4.1.5 and Theorem 5.1.7. Both chapters then proceed by the finite-time formulated problems and provide the nonlinear integral equations that characterise the free-boundary and the value function in Theorem 4.2.19, Theorem 5.2.40 and Theorem 5.2.42.
Let us make some further comments on these two chapters:

In Chapter 4, with our choice of parameters in problem (1.0.1), one can avoid the singularity of the partial derivative and the establishment of the continuous value function is within easy reach, which in turn implies the existence of the optimal stopping rule. After this, we turn our attention to analysing the structure of the continuation and stopping sets and Assumption 4.2.10 is made principally to guarantee the monotonicity of the boundary in Lemma 4.2.13.

In Chapter 5, being confronted with the singularity of the partial derivative, it becomes challenging to justify the regularity of value function and the optimal stopping boundary. As a result of this, one can only derive the semi-continuity of the value function first; but fortunately, this is just enough to confirm the existence of the optimal stopping rule in Lemma 5.2.8. The next focal point of this chapter is once again to justify the monotonicity of the boundary, which shall pave the way for us to address the problems caused by the singularity and then show the regularity of the value function and the boundary. To achieve this, the first thing we have to do is to pinpoint the stopping set by two crucial observations in Lemma 5.2.12 and Lemma 5.2.23 respectively: (i) the relative position of the free-boundary for pricing the American put option (which we take for granted) and the fixed level $L$ (in Option C) and (ii) the maximum of the gain function for any fixed time. Based on these observations and Assumption 5.2.27 concerning the choice of parameters in problem (1.0.2), we are able to present the monotonicity of the boundary in Proposition 5.2.18 and Proposition 5.2.35.

Chapter 6 contains a number of additional results and remarks.

Lastly, Appendix A gathers together a host of computations needed throughout this work.
Chapter 2

Optimal Stopping in Mathematical Finance: Background

Stochastic analysis and probability theory are extensively used in the area of optimal stopping study, which will then be applied to the contingent claim valuation (Section 2.4 reveals how they clicked in the first place). The terminology sometimes differs slightly depending on the authors, this chapter makes clear our choice of terminology and in due course, reviews some general facts.

To distill such massive amount of information and determine what is important to the current work, it seems necessary to propose the basic skeleton from optimal stopping: namely, let $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \geq 0}, P)$ be a certain filter probability space and then consider the following optimal stopping problem for $T < \infty$:

$$V(0, x) = \sup_{\tau \in [0, T]} E_{0, x}(G(\tau, X_\tau)), \quad (2.0.1)$$

where the supremum of (2.0.1) is taken over all the stopping times of the continuous stochastic process $X$ ($X_t$ is $\mathcal{F}_t$-measurable) with $X_0 = x$, and the expectation is taken under measure $P_{0, x}$ and $G$ is the measurable function satisfying the following condition

$$E_{0, x} \left( \sup_{t \in [0, T]} |G(t, X_t)| \right) < \infty, \quad (2.0.2)$$

and in addition, for $T = \infty$, the infinite time formulated problem emerges:

$$V(x) = \sup_{\tau} E_x(G(\tau, X_\tau)), \quad (2.0.3)$$

in which case, let $G(\infty, X_\infty) = 0$.

The function $V$ is called the value function and $G$ the gain/reward function; in addition, if $G$ is non-negative, then it is a contingent claim in mathematical finance.

---

1. Hereafter, we write $E_{0, x}(A)$ and $E_{t, x}(A)$ as the conditional expectation of event $A$ relative to the conditional measure $P(A|X_0 = x)$ and $P(A|X_t = x)$, which are further simplified as $P_{0, x}(A)$ and $P_{t, x}(A)$.

2. We allow stopping times to assume the value $+\infty$ with positive probability, which corresponds to the situation the kind of stopping never occurs. Because of this, this condition is introduced to make certain that the optimal stopping time is finite.
2.1 Probability Theory: Convergence and Others

Not surprisingly, when we study the properties of the value function (or as a matter of fact, any functions), the first two things we are looking for are its continuity and differentiability, and almost inescapably, we run into the topic of convergence of moments. Another use of such topic is to relate local martingales to martingales, concepts we will introduce in the section 2.3. In this section, we collect some simple tools that enable us to conclude that if \( X_n \) converges to \( X \) almost surely as \( n \to \infty \), then \( E(X_n) \to E(X) \). For other modes of converges, readers are referred to [39, Chapter 5] for excellent treatment on such topic.

The definition of \textit{almost surely convergence} is a good place to start.

**Definition 2.1.1 (Almost Surely Convergence).** \( X_n \) converges almost surely (a.s.) to the random variable \( X \) as \( n \to \infty \) iff

\[
P\left( \{ \omega : X_n(\omega) \to X(\omega) \text{ as } n \to \infty \} \right) = 1.
\]

The first important tool to prove random variables converge almost surely is the Borel-Cantelli Lemmas.

**Theorem 2.1.2 (The First and Second Borel-Cantelli Lemmas).**

(i) Let \( \{A_n, n \geq 1\} \) be arbitrary events. Then

\[
\sum_{n=1}^{\infty} P(A_n) < \infty \Rightarrow P(A_n \text{ infinitely often}) = 0.
\]

(ii) Let \( \{A_n, n \geq 1\} \) be independent events. Then

\[
\sum_{n=1}^{\infty} P(A_n) = \infty \Rightarrow P(A_n \text{ infinitely often}) = 1.
\]

Theorem 2.1.2 then leads us to the following zero-one law, which in our context is used to conclude that independent random variables almost surely converge to a constant limiting variable. This is especially helpful in showing the first hitting time of a constant level is finite in the infinite time problem setting.

**Theorem 2.1.3 (A Zero-One Law).** If the events \( \{A_n, n \geq 1\} \) are independent, then

\[
P(A_n \text{ i.o.}) = \begin{cases} 1, & \text{when } \sum_{n=1}^{\infty} P(A_n) = \infty, \\ 0, & \text{when } \sum_{n=1}^{\infty} P(A_n) < \infty. \end{cases}
\]

**Corollary 2.1.4.** Suppose that \( X_1, X_2, \ldots \) are independent random variables. Then, for constant \( c \),

\[
X_n \xrightarrow{a.s.} c \text{ as } n \to \infty \iff \sum_{n=1}^{\infty} P(|X_n - c| > \epsilon) < \infty \text{ for all } \epsilon > 0.
\]

The result “always holds” is:
Theorem 2.1.5 (Fatou’s Lemma). (i) If \( \{X_n, n \geq 1\} \) are non-negative random variables, then
\[
E\left(\liminf_{n \to \infty} X_n\right) \leq \liminf_{n \to \infty} E(X_n).
\]

(ii) In addition, if \( X_n \leq Z \) with \( E|Z| < \infty \) (i.e. integrable), then
\[
E\left(\limsup_{n \to \infty} X_n\right) \geq \limsup_{n \to \infty} E(X_n).
\]

(iii) Furthermore, if \( Y \leq X_n \leq Z \) with integrable random variables \( Y \) and \( Z \) a.s. for all \( n \), then
\[
E\left(\liminf_{n \to \infty} X_n\right) \leq \liminf_{n \to \infty} E(X_n) \leq \limsup_{n \to \infty} E(X_n) \leq E\left(\limsup_{n \to \infty} X_n\right).
\]

Monotonically convergent random variables also provides us with a positive answer for moment convergence.

Theorem 2.1.6 (Monotone Convergence). (i) Let \( \{X_n, n \geq 1\} \) be non-negative random variables. If \( X_n \uparrow X \) as \( n \to \infty \), then
\[
E(X_n) \uparrow E(X), \quad \text{as } n \to \infty.
\]
where the limit might be infinite.

(ii) Let \( \{X_n, n \geq 1\} \) be non-negative random variables and suppose that \( X_1 \) is integrable. If \( X_n \downarrow X \) as \( n \to \infty \), then
\[
E(X_n) \downarrow E(X), \quad \text{as } n \to \infty.
\]
where the limit might be infinite.

Another adequate concept for convergence of moments is uniform integrability.

Definition 2.1.7 (Uniform Integrability). A sequence \( X_1, X_2, \ldots \) is called uniformly integrable iff
\[
E\left(|X_n|I\{|X_n| > a\}\right) \to 0, \quad \text{as } a \to \infty \text{ uniformly in } n.
\]

The following result is most frequently used to check the uniform integrability of the random variable:

Proposition 2.1.8. Suppose that \( X_1, X_2, \ldots \) are random variables such that
\[
|X_n| \leq Y \text{ a.s. for all } n,
\]
where \( Y \) is a positive integrable random variable. Then \( \{X_n, n \geq 1\} \) is uniformly integrable.

Here comes a convergence theorem that reassures us that uniform integrability is the right condition to guarantee moment convergence to happen.

Theorem 2.1.9 (The Lebesgue Dominated Convergence Theorem). Suppose that \( |X_n| \leq Y \) for all \( n \), where \( E(Y) < \infty \), and that \( X_n \overset{a.s.}{\longrightarrow} X \) as \( n \to \infty \). Then
\[
E(X_n) \to E(X), \quad \text{as } n \to \infty.
\]
Remark 2.1.10. Let $Y$ be a constant, Theorem 2.1.9 is sometimes called the bounded dominated convergence theorem.

These results can be easily extended to a function of random variables if we bear in mind the next theorem (not restricted to almost surely convergence):

Theorem 2.1.11 (The Continuous Mapping Theorem).

(i) If $X_n \xrightarrow{a.s.} X$ as $n \to \infty$, then $g(X_n) \xrightarrow{a.s.} g(X)$ as $n \to \infty$.

(ii) If $X_n \xrightarrow{p} X$ as $n \to \infty$, then $g(X_n) \xrightarrow{p} g(X)$ as $n \to \infty$.

(iii) If $X_n \xrightarrow{d} X$ as $n \to \infty$, then $g(X_n) \xrightarrow{d} g(X)$ as $n \to \infty$.

The above proposition remains true also if function $g$ is measurable with the set of discontinuity points denoted as $D_g$ so that $P(X \in D_g) = 0$.

Remark 2.1.12. The Borel measurable function of a random variable is itself a random variable, thus all the above result can be applied as long as the conditions are met.

We end this subsection with a fairly fundamental result.

Theorem 2.1.13 (Law of the Unconscious Statistician). If the probability density function of random variable $X$ is given as $f_X(x)$, and suppose that $G$ is a measurable function so that $G(X)$ is an integrable random variable, then

$$E(G(X)) = \int_{-\infty}^{\infty} G(x)f_X(x)dx.$$

2.2 Stochastic Analysis: Basics

A few words first about stochastic process and basic measurability. A stochastic process is a collection of random variables $X = (X_t)_{t \geq 0}$ on the sample space $(\Omega, \mathcal{F})$, which take value in the state space, the $d$-dimensional Euclidean space $\mathbb{R}^d$ equipped with the $\sigma$-field of Borel sets $\mathcal{B}(\mathbb{R}^d)$. In the beginning of Chapter 1, we mentioned that how the efficient market responds to new information so that the corrections of prices are instantaneous, so if we are to model the market using stochastic process, it is rather important to keep track of the information. To achieve this mathematically, we further equip the sample space with a filtration $\{\mathcal{F}_t : t \geq 0\}$ (where the filtration $\{\mathcal{F}_t : t \geq 0\}$ satisfying the conditions: (i) $\mathcal{F}_0$ contains all the null set of $P$ and (ii) it is right-continuous), which can then be interpreted as the information available up to time $t$ and chosen simply to be the one that generated by process $X$ itself, that is, $\mathcal{F}_t^X = \sigma(X_s : s \in [0, t])$ with respect to which, $X_s$ is measurable for every $s \in [0, t]$. The stochastic process $X$ is said to be adapted to the filtration $\{\mathcal{F}_t\}$ if for each $t \geq 0$, $X_t$ is an $\mathcal{F}_t$-measurable random variable. For more, see [48, Page 1-6].

The process of our interest is the continuous stochastic process $X = (X_t)_{t \geq 0}$ defined on a filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \geq 0}, P)$, which satisfies the stochastic differential equation (SDE) \footnote{We simply wrote $\mathcal{F}_t$ hereafter.}.

$$dX_t = \mu(t, X_t)dt + \sigma(t, X_t)dW_t, \quad \text{with } X_0 = x \text{ and } t \in [0, T], \quad (2.2.1)$$

\footnote{We write process $X$ starting from $X_0 = x$ for time $t$ as $X^{0,x}_t$ whenever it is needed.}
2.2 Stochastic Analysis: Basics

where $W = (W_t)_{t \geq 0}$ is the Brownian motion, $\mu$ is the drift coefficient and $\sigma$ is the diffusion coefficient; in addition, $\mu(t, x)$ and $\sigma(t, x)$ are measurable functions on $[0, T] \times \mathbb{R}$ that satisfy a space-variable Lipschitz condition with some constant $K$:

$$|\mu(t, x) - \mu(t, y)|^2 + |\sigma(t, x) - \sigma(t, y)|^2 \leq K|x - y|^2,$$

and the spatial growth condition:

$$|\mu(t, x)|^2 + |\sigma(t, x)|^2 \leq K(1 + |x|^2),$$

so that unique continuous solution for SDE (2.2.1) exists and $\sup_{t \in [0, T]} E(X^2_t) < \infty$.

Moreover, $X$ is time-homogeneous if SDE (2.2.1) is of the form:

$$dX_t = \mu(X_t)dt + \sigma(X_t)dW_t, \quad X_0 = x. \quad (2.2.2)$$

Shortly after the introduction of our main problem (2.0.1), we encounter a crucial concept so-called stopping times:

**Definition 2.2.1 (Stopping Time).** A nonnegative, possibly infinite random variable $\tau = \tau(\omega)$ is called a stopping time (with respect to $\{\mathcal{F}_t, t \geq 0\}$) if $\{\omega : \tau(\omega) \leq t\} \in \mathcal{F}_t$ for each $t \geq 0$.

From the definition, we see the important feature of stopping times is that they are measurable with respect to “what has happened so far”; thus, typical example of stopping times will be the first hitting time. We will also encounter some non-stopping times in the later chapters, that is the last exit time, which is only “measurable” by knowing the future, readers are referred to [54] and [61] for more information in this realm of research.

Inspecting (2.0.1), one might wonder if the starting time at 0 can be generalised to any time $t \in [0, T]$, the strong Markov property is the way to go.

**Theorem 2.2.2 (The Strong Markov Property for Itô Diffusions).** For any stopping time $\tau = \tau(\omega)$ with respect to $\{\mathcal{F}_t, t \geq 0\}$ and $P(\tau < \infty) = 1$, let $f$ be a bounded Borel function on $\mathbb{R}$. Then,

$$E_x(f(X_{\tau+h}) | \mathcal{F}_\tau) = E_{X_\tau}(f(X_h)), \quad \text{for all } h \geq 0. \quad (2.2.3)$$

This definition tells us that for Markov process, its future depends only on the fixed present and is independent of its past. Here are two examples to illuminate the above result.

**Example 2.2.3 (The Geometric Brownian Motion).** The frequent use of Geometric Brownian Motion (GBM) in finance and economic is second to none. To check its strong Markov property, we first note from its strong solution $X^{0,x}_t = xe^{(r-\frac{\sigma^2}{2})t + \sigma W_t}$ that the natural filtration of $X$ equals to the natural Brownian filtration, i.e. $\mathcal{F}_t = \sigma\{W_s, s \in [0, t]\}$.

From Theorem 2.2.2, we know that to show the strong Markov property, it is enough to show that for any stopping time $\tau < \infty$:

$$\text{Law}(X_{\tau+h}|\mathcal{F}_\tau, P) = \text{Law}(X_{\tau+h}|X_\tau, P), \quad (2.2.3)$$
where $P$ is the probability measure under which $W = (W_t)_{t \geq 0}$ is a standard Brownian motion started at 0. Observe that:

$$X_{\tau+s} = xe^{(r-\frac{\sigma^2}{2})(\tau+s)+\sigma W_{\tau+s}}$$

$$= xe^{(r-\frac{\sigma^2}{2})\tau+\sigma W_\tau} e^{(r-\frac{\sigma^2}{2})s+\sigma(W_{\tau+s}-W_\tau)}$$

$$= X_\tau e^{(r-\frac{\sigma^2}{2})s+\sigma W_s}.$$ 

where the third equality is due to the fact that $(W_{\tau+s}-W_\tau) \overset{d}{=} W_s$ and that $\{W_{\tau+s}-W_\tau\}_{s \geq 0}$ is independent of $\mathcal{F}_\tau$; hence, (2.2.3) follows.

Notice another pleasant property of GBM is being uniformly integrable for $t \in [0, T]$. Proposition 2.1.8 shows us the way to verify it:

$$X_t^{0,x} = xe^{(r-\frac{\sigma^2}{2})t+\sigma W_t} \leq xe^{rt+\sigma W_t} \leq xe^{rT+\sigma \max_{0\leq s \leq T} W_s},$$

and since the running maximum of Brownian motion $M_T = \max_{0 \leq s \leq T} W_s$ has the probability density function $f(y) = \sqrt{\frac{2}{\pi T}} e^{-\frac{y^2}{2T}},$

$$E\left(e^{\sigma M_T}\right) = \sqrt{\frac{2}{\pi T}} \int_{0}^{+\infty} e^{\sigma y - \frac{y^2}{2T}} dy = 2e^{\frac{\sigma^2 T}{2}} < \infty,$$

which means that $X_t^{0,x}$ is dominated by the positive integrable random variable, implying its uniform integrability.

**Example 2.2.4 (GBM and Its Running Maximum).** Define process $Y = (Y_t)_{t \geq 0} = \left(\frac{S_t}{X_t}\right)_{t \geq 0}$ where $X = (X_t)_{t \geq 0}$ is GBM defined in Example 2.2.3 with $X_0 = x = 1$ and $S = (S_t)_{t \geq 0}$ is the running maximum of GBM, i.e. $S_t = s \vee \max_{u \in [0,t]} X_u$ with $S_0 = s$; thereby, $Y_0 = y = s$. Once again, we lean on Theorem 2.2.2 to check its strong Markov property: for stopping time $\tau < \infty$,

$$Y_{\tau+s} = \frac{y \vee S_{\tau+s}}{X_{\tau+s}} = \frac{y \vee \max_{u \in [0,\tau]} X_u \vee \max_{u \in [\tau,\tau+s]} X_u}{X_{\tau+s}} = \frac{y \vee \max_{u \in [0,\tau]} X_u}{X_{\tau+s}} \vee \frac{\max_{u \in [\tau,\tau+s]} X_u}{X_{\tau+s}}$$

$$= X_\tau e^{(r-\frac{\sigma^2}{2})s+\sigma W_s} \vee \frac{X_\tau e^{(r-\frac{\sigma^2}{2})(u-\tau)+\sigma(W_u-W_\tau)}}{X_s}$$

$$= Y_\tau \vee \frac{e^{(r-\frac{\sigma^2}{2})(u-\tau)+\sigma(W_u-W_\tau)}}{X_s} \overset{d}{=} Y_\tau \vee S_s.$$

where the last distribution equality follows the same argument as that of Example 2.2.3 so that for stopping time $\tau < \infty$:

$$\text{Law} \left(Y_{\tau+s} | \mathcal{F}_\tau, P\right) = \text{Law} \left(Y_{\tau+s} | Y_\tau, P\right).$$
Since the optimal stopping problem is dealing with the function of process, we shall at the moment lay out some useful formulas.

**Theorem 2.2.5 (Itô’s Formula).** If \( f \in C^{1,2} (\mathbb{R}_+ \times \mathbb{R}) \) and the stochastic process \( X \) has the integral representation

\[
X_t = x + \int_0^t \mu(s, X_s) ds + \int_0^t \sigma(s, X_s) dW_s, \quad t \in [0, T], \tag{2.2.4}
\]

then we have

\[
f(t, X_t) = f(0, x) + \int_0^t \frac{\partial f}{\partial t}(s, X_s) ds + \int_0^t \frac{\partial f}{\partial x}(t, X_s) dX_s + \frac{1}{2} \int_0^t \sigma^2(s, X_s) \frac{\partial^2 f}{\partial x^2}(s, X_s) ds.
\]

The condition for Itô’s formula to be applicable is the function be appropriately smooth, but there are times when the smoothness fails for us to avail Itô’s formula, such as, the gain function of American call option \( (x - K)^+ \), so an advanced formula is required (see [64, Page 68]):

**Lemma 2.2.6 (The Extension in Itô-Tanaka-Meyer Formula).** If a concave (convex or the difference of the two) function \( f : \mathbb{R} \mapsto \mathbb{R} \) is not smooth at some constant point \( x = a \) and the process \( X \) has integral representation (2.2.4) and the process \( X \) has integral representation (2.2.4), then we have

\[
f(X_t) = f(x) + \int_0^t \frac{d}{dx} f(X_s) I\{X_s \neq a\} dX_s + \frac{1}{2} \int_0^t \frac{d^2}{dx^2} f(X_s) I\{X_s \neq a\} d\langle X, X \rangle_s
\]

\[+ \frac{1}{2} \int_0^t \left( \frac{d}{dx} f(a^+) - \frac{d}{dx} f(a^-) \right) d\mu_s(X),
\]

where the last term is the local time that process \( X \) spends at the level \( a \) with

\[
l^a_t(X) = P - \lim_{\epsilon \to 0} \frac{1}{2\epsilon} \int_0^t I\{a - \epsilon < X_u < a + \epsilon\} d\langle X, X \rangle_u.
\]

It is then only human to ask what if the function is not smooth on some curve(s)? The local time-space formula (also known as change-of-variable formula) can be our guide, (see [66] for detailed proofs and conditions to use; this result later is extended to multiple dimensions in [67])

**Theorem 2.2.7 (A Local Time-Space Formula).** Let the process \( X \) has the integral representation (2.2.4) and \( c : \mathbb{R}_+ \mapsto \mathbb{R} \) be a continuous function of bounded variation, and let \( f : \mathbb{R}_+ \times \mathbb{R} \mapsto \mathbb{R} \) be a continuous function satisfying \( f \) is \( C^{1,2} \) on \( \bar{C}_1 \cup \bar{C}_2 \), where \( \bar{C}_1 \) and \( \bar{C}_2 \) are given as follows:

\[
\bar{C}_1 = \{(t, x) \in \mathbb{R}_+ \times \mathbb{R} : x > c(t)\}, \quad \bar{C}_2 = \{(t, x) \in \mathbb{R}_+ \times \mathbb{R} : x < c(t)\}.
\]

Then, the following local time-space formula holds:

\[
f(t, X_t) = f(0, X_0) + \frac{1}{2} \int_0^t f_t(s, X_s+) + f_t(s, X_s-) ds + \frac{1}{2} \int_0^t f_x(s, X_s+) + f_x(s, X_s-) dX_s
\]

\[+ \frac{1}{2} \int_0^t f_{xx}(s, X_s) I\{X_s \neq c(s)\} d\langle X, X \rangle_s
\]

\[+ \frac{1}{2} \int_0^t \left( f_x(s, X_s+) - f_x(s, X_s-) \right) I\{X_s = c(s)\} dl^a_s(X),
\]

where \( l^a_t(X) \) is the local time of \( X \) at the curve \( c \) with

\[
l^a_t(X) = P - \lim_{\epsilon \to 0} \frac{1}{2\epsilon} \int_0^t I\{c(u) - \epsilon < X_u < c(u) + \epsilon\} d\langle X, X \rangle_u.
\]
Theorem 2.2.7 can be extended to the case involving finitely many functions (instead of one function) \( b_1, b_2, \ldots b_n \) which do not intersect, see [35, Page 7].

**Lemma 2.2.8** (Extension to Finitely Many Curves). If, in addition, the following conditions are fulfilled:

(i) \( b_i : \mathbb{R}_+ \to \mathbb{R} \) is continuous and bounded variation for \( 1 \leq i \leq n \);

(ii) \( f_i : \mathbb{R}_+ \to \mathbb{R} \) is \( C^{1,2} \) for \( 1 \leq i \leq n + 1 \);

(iii) 
\[
f(t, x) = \begin{cases} 
  f_1(t, x), & \text{if } x < b_1(t), \\
  f_i(t, x), & \text{if } b_{i-1}(t) < x < b_i(t) \text{ for } 2 \leq i \leq n, \\
  f_{n+1}(t, x), & \text{if } x > b_n(t), 
\end{cases}
\]
then Theorem (2.2.7) can be extended as follows
\[
f(t, X_t) = f(0, X_0) + \int_0^t f_i(s, X_{s-}) ds + \int_0^t f_x(s, X_{s-}) dX_s \\
+ \frac{1}{2} \int_0^t f_{xx}(s, X_{s-}) I\{X_s \notin \{b_1(s), \ldots, b_n(s)\}\} \langle X_s, X_s \rangle ds \\
+ \frac{1}{2} \sum_{i=1}^n \int_0^t (f_x(s, X_{s+}) - f_x(s, X_{s-})) I\{X_s = b_i(s)\} db_i(X).
\]

Can a stochastic process \( X \) with drift \( \mu \) also be viewed as a process without drift? This question is in fact closely connected to *arbitrage pricing theory* in mathematical finance, when we are required to find the *equivalent martingale measure* to make the *discounted stock price* a martingale, we shall soon return to this topic. And the answer to this question is positive, and the idea behind it is to interpret a single stochastic process by different probability measures. See [79, Page 222].

**Theorem 2.2.9** (Girsanov Theorem, Removing Non-Constant Drift). Suppose that \( \mu(t, \omega) \) is a bounded, adapted process on \([0, T]\), \( W_t \) is a \( P \)-Brownian motion, and the process \( X \) is given by
\[
X_t = W_t + \int_0^t \mu(s, \omega) ds.
\]

The process \( M_t \) defined by
\[
M_t = e^{-\int_0^t \mu(s, \omega) dW_s - \frac{1}{2} \int_0^t \mu^2(s, \omega) ds},
\]
is a \( P \)-martingale and the product \( X_t M_t \) is also a \( P \)-martingale. Finally, if \( Q \) denotes the measure defined by \( Q(A) = E(I\{A\} M_T) \) with the expectation taken under measure \( P \), then \( X_t \) is a \( Q \)-Brownian motion on \([0, T]\).

Inevitably, the uniqueness of such probability measure \( Q \) will be questioned. The next proposition provides the answer we need.

**Proposition 2.2.10** (The Uniqueness of \( Q \)). Suppose that \( \{W_t, t \in [0, T]\} \) is a \( Q \)-Brownian motion and \( \{F_t\} \) is the standard Brownian filtration for \( t \in [0, T] \). Further, suppose that \( \{M_t, \mathcal{F}_t\}_{t \in [0, T]} \) is a \( Q \)-martingale with an integral representation that is given by for \( t \in [0, T] \),
\[
M_t = M_0 + \int_0^t m(s, \omega) dB_s,
\]
where \( m(s, \omega) \neq 0 \), except possibly on a set with \( dt \times dQ \) measure zero. If \( Q' \) is a probability measure on \( \mathcal{F}_T \) that is equivalent to \( Q \) and if \( \{ M_t, \mathcal{F}_t \}_{t \in [0,T]} \) is also a \( Q' \)-martingale, then \( Q(A) = Q'(A) \) for all \( A \in \mathcal{F}_T \).

An immediate application Theorem 2.2.9 is to compute the following result, which was taken for granted from [76, Page 759 and 760]:

**Lemma 2.2.11** (The Distribution of the Running Maximum of Brownian motion with Drift). For \( x \geq 0, \mu \in \mathbb{R}, \) and \( \sigma > 0 \) we have (i) in finite time:

\[
P \left( \max_{s \leq t} (\mu s + \sigma W_s) \leq x \right) = \Phi \left( \frac{x - \mu t}{\sigma \sqrt{t}} \right) - e^{2\mu t} \Phi \left( \frac{-x - \mu t}{\sigma \sqrt{t}} \right),
\]

(2.2.5)

(ii) in infinite time: for \( \mu < 0 \), then

\[
P \left( \max_{t \geq 0} (\mu t + \sigma W_t) \leq x \right) = 1 - e^{2\mu x},
\]

(2.2.6)

as for \( \mu \geq 0 \), then

\[
P \left( \max_{t \geq 0} (\mu t + \sigma W_t) \leq x \right) = 0.
\]

(2.2.7)

Before bringing this section to a close, we introduce the concept of infinitesimal generator which is fundamental for many application to associate a second order differential operator to \( X \), denoted as \( L_X \); its relation with the coefficients \( \mu \) and \( \sigma \) in SDE (2.2.2) defining process \( X \), see [62, Page 117].

**Definition 2.2.12** (The Infinitesimal Generator). Let \( X \) be a time-homogeneous stochastic process defined by SDE (2.2.2) and \( D_A \) be the set of functions \( f : \mathbb{R} \mapsto \mathbb{R} \) so that the limit exists for all \( x \in \mathbb{R} \). The infinitesimal generator \( L_X \) of \( X \) is defined by for \( f \in D_A \)

\[
L_X f(x) = \lim_{t \to 0} \frac{E_x [f(X_t)] - f(x)}{t}.
\]

**Lemma 2.2.13.** Let \( X \) be a time-homogeneous stochastic process defined by SDE (2.2.2). If \( f \in C^2(\mathbb{R}) \) with compact support, then \( f \in D_A \) and

\[
L_X f(x) = \mu(x) \frac{df}{dx}(x) + \frac{1}{2} \sigma^2(x) \frac{d^2 f}{dx^2}(x).
\]

(2.2.8)

Loosely speaking, diffusion process can be regarded as a Markov process which has continuous sample paths and can be characterised in terms of its infinitesimal generator. (See [48, Page 282] for the formal definition of diffusion process.)

### 2.3 Martingales

“The martingale theory of arbitrage pricing is one of the greatest triumphs of probability theory!” [79, Page 233]

The first natural step is to acquaint ourselves with the basic definitions of martingale, alongside with submartingale and supermartingale:
Definition 2.3.1. An integrable \( \{F_t, t \geq 0\} \)-adapted stochastic process \( \{X_t, t \geq 0\} \) is called a martingale if

\[
E(X_t | F_s) = X_s \quad \text{a.s.} \quad \text{for all } 0 \leq s \leq t.
\]

It is called a submartingale if

\[
E(X_t | F_s) \geq X_s \quad \text{a.s.} \quad \text{for all } 0 \leq s \leq t,
\]

and a supermartingale if

\[
E(X_t | F_s) \leq X_s \quad \text{a.s.} \quad \text{for all } 0 \leq s \leq t.
\]

Definition 2.3.2. An \( \{F_t, t \geq 0\} \)-adapted process \( \{X_t, t \geq 0\} \) is called a local martingale provided that there is a nondecreasing sequence \( \{\tau_k\} \) of stopping times with property that \( \tau_k \to \infty \) with probability one as \( k \to \infty \) and such that for each \( k \) the process defined by \( X_t^{(k)} = X_{t \wedge \tau_k} - X_0 \) for \( t \geq 0 \) is a martingale with respect to the filtration \( \{F_t, t \geq 0\} \).

The connection between martingale and submartingale comes from the following result:

**Theorem 2.3.3 (The Doob-Meyer Decomposition).** Any submartingale \( \{(X_t, F_t), t \geq 0\} \) can be uniquely decomposed into the sum of a martingale \( \{(M_t, F_t), t \geq 0\} \) and an increasing predictable locally integrable process \( \{(A_t, F_t), t \geq 0\} \):

\[
X_t = X_0 + M_t + A_t, \quad t \geq 0. \tag{2.3.1}
\]

This can be certainly extended to the supermartingale by simply replacing \( A_t \) with \( -A_t \) in \( (2.3.1) \).

**Example 2.3.4 (Azéma Supermartingale).** Let \( \theta \) be an honest time that avoids \( F_t \)-stopping times \( \tau \) and let \( Z^\theta_t = P(\theta > t | F_t) \), which is the \( F_t \)-supermartingale chosen to be càdlàg, associated to \( \theta \).

By Theorem 2.3.3, we have (see [61, Page 385]):

\[
Z^\theta_t = M^\theta_t - A^\theta_t. \tag{2.3.2}
\]

where \( M^\theta \) is the càdlàg martingale and \( A^\theta \) is the dual predictable projection.

Their other connection comes through the convex function, thanks to the conditional Jensen inequality: \( g(E(X \mid G)) \leq E(g(X) \mid G) \) for \( g \) convex.

**Proposition 2.3.5.** If \( \{(X_t, F_t), t \geq 0\} \) is (i) a martingale and \( g \) a convex function, or (ii) a submartingale and \( g : \mathbb{R} \mapsto \mathbb{R} \) is a non-decreasing convex function, and moreover \( E(g(X_t)) < \infty \) for all \( t \geq 0 \), then \( \{(g(X_t), F_t), t \geq 0\} \) is a submartingale.

The concept of superharmonic function is rather fundamental in optimal stopping and closely related to supermartingale, we thus make room for its definition.

**Definition 2.3.6 (The Superharmonic Function).** A measurable function \( F : \mathbb{R}^d \mapsto \mathbb{R} \) is said to be superharmonic (w.r.t. \( X_t \)) if \( F(X_t) \) is uniformly integrable and

\[
E_x(F(X_\tau)) \leq F(x),
\]

for all stopping times \( \tau \) and \( x \in \mathbb{R}^d \).
**Definition 2.3.7** (The Smallest Superharmonic Function). Let \( h : \mathbb{R}^d \mapsto \mathbb{R} \) be a measurable function. If \( f : \mathbb{R}^d \mapsto \mathbb{R} \) is a superharmonic function and \( f \geq h \), then \( f \) is a superharmonic function that dominates \( h \). If \( F : \mathbb{R}^d \mapsto \mathbb{R} \) is any other superharmonic function dominating \( h \) and \( f \leq F \), then \( f \) is the smallest superharmonic function that dominates \( h \).

**Corollary 2.3.8.** A lower semicontinuous function \( F : \mathbb{R}^d \mapsto \mathbb{R} \) is superharmonic if and only if \( \text{uniformly integrable} \ (F(X_t))_{t \geq 0} \) is a right-continuous supermartingale under \( P_x \) for every \( x \in \mathbb{R}^d \).

**Definition 2.3.9.** The function \( F : \mathbb{R}^d \mapsto \mathbb{R} \) is lower semicontinuous (l.s.c) at \( x \in \mathbb{R}^d \) iff
\[
P_x \left( \liminf_{t \to 0} F(X_t) \geq F(x) \right) = 1,
\]
and upper semicontinuous (u.s.c) at \( x \in \mathbb{R}^d \) iff
\[
P_x \left( \limsup_{t \to 0} F(X_t) \leq F(x) \right) = 1,
\]
for \( X_t \xrightarrow{a.s.} x \) as \( t \to 0 \). (see [77, Page 113]).

One of the notions that lies at the heart of martingale theory is that of stopping times and remember that stopped martingale (submartingale) does not always remains a martingale (submartingale); but for the bounded stopping times, the martingale (submartingale) property is preserved.

**Theorem 2.3.10** (Doob’s Optional Sampling Theorem). Suppose that \( \{(X_t, \mathcal{F}_t), t \geq 0\} \) is a martingale of the form \( X_t = E(Z|\mathcal{F}_t) \) for some integrable random variable \( Z \). Then, the variables \( X_\tau \) and \( X_\sigma \) are integrable for any two finite stopping times \( \sigma, \tau \) and \( E(X_\tau|\mathcal{F}_\sigma) = X_\sigma \) \( P \)-a.s. in the set \( \{ \sigma \leq \tau \} \).

**Remark 2.3.11.** Notice one advantage of the finite time formulated problem is that the stopping times \( \tau \) are bounded by the terminal time \( T \), thus optional sampling theorem can therefore be applied freely. For infinite time formulated problem, the method of localization is required.

In later chapters, when we try to prove the martingale property, one particular result turns out to be handy in many situation.

**Lemma 2.3.12** (Smoothing Lemma). Suppose that \( \sigma \)-algebras \( \mathcal{F}_1 \subset \mathcal{F}_2 \) and \( X \) is a random variable with finite expectation. Then
\[
E(E(X|\mathcal{F}_2)|\mathcal{F}_1) = E(X|\mathcal{F}_1) = E(E(X|\mathcal{F}_1)|\mathcal{F}_2) \ a.s.
\]

A particular useful application of Theorem 2.3.10 and Lemma 2.3.12 is the following one, see [39, Page 496].

**Corollary 2.3.13.** Suppose that \( \{(X_t, \mathcal{F}_t), t \in [0, \infty)\} \) is a martingale, and that \( \tau \) is a stopping time. Then, \( \{(X_{\tau \wedge t}, \mathcal{F}_t), t \in [0, \infty)\} \) is a martingale and \( E(X_{\tau \wedge t}) = E(X_t) \).

**Proof.** We begin by asserting measurability:
\[
X_{\tau \wedge t} = X_\tau I\{\tau < t\} + X_t I\{\tau \geq t\} \in \mathcal{F}_t,
\]
since each term belongs to $\mathcal{F}_t$. Now observe that, for $s \in [0, \infty)$,

$$
E \left( X_{\tau \wedge (t+s)} | \mathcal{F}_t \right) = E \left( X_{\tau} \{ \tau < t+s \} + X_{t+s} I\{ \tau \geq t+s \} \right) | \mathcal{F}_t \\
= E \left( X_{\tau} \{ \tau < t+s \} + X_{t+s} I\{ \tau \leq t < t+s \} + X_{t+s} I\{ \tau \geq t + s \} \right) | \mathcal{F}_t \\
= X_{\tau} \{ \tau < t \} + E \left( X_{\tau \wedge (t+s)} | \mathcal{F}_t \right) I\{ \tau \geq t \} \\
= X_{\tau} \{ \tau < t \} + I\{ \tau \geq t \} E \left( X_{\tau \wedge (t+s)} | \mathcal{F}_t \right) \\
= X_{\tau} \{ \tau < t \} + X_t I\{ \tau \geq t \} = X_{\tau \wedge t}
$$

where the fifth equality follows immediately from Theorem 2.3.10 given that $\min\{ \tau, t+s \}$ is a stopping time bounded by $t + s$, and thereby, proving the martingale property; after which, by the martingale property of $\{(X_t, \mathcal{F}_t), t \in [0, \infty)\}$, $E \left( X_t | \mathcal{F}_{\tau \wedge t} \right) = X_{\tau \wedge t}$ and then we take the expectation again, Lemma 2.3.12 yields that $E \left( X_{\tau \wedge t} \right) = E \left( X_t \right)$. \hfill $\square$

### 2.4 Returning To An Old Promise: Arbitrage

The concept of arbitrage has made its numerous appearances up to this point and in Section 2.2, alongside with the introduction of Girsanov Theorem, we vaguely commented on its connection to some martingale measure. This section is designate to say more about the arbitrage-free pricing theory.

In what follows, we assume that the securities market under consideration operates in the condition of “uncertainty” that can be described in the probabilistic framework in terms of a filter probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \in [0, T]}, P)$, where we interpreted the filtration $(\mathcal{F}_t)_{t \in [0, T]}$ as the flow of incoming information and it satisfies conditions (i) $\mathcal{F}_0$ contains all the null sets of $P$ and (ii) $(\mathcal{F}_t)_{t \in [0, T]}$ is right-continuous, i.e. $\mathcal{F}_t = \cap_{s > t} \mathcal{F}_s$ for $t \in [0, T]$.

Then we consider an $(B, S)$-market formed by one bond (non-risky asset) and one stock (risky asset) with their prices following the standard stochastic processes $B = (B_t)_{t \geq 0}$ and $S = (S_t)_{t \geq 0}$, whose SDEs are given by

$$
dS_t = \mu(t, S_t) dt + \sigma(t, S_t) dW_t, \quad S_0 = s, \quad (2.4.1) \\
dB_t = r(t, S_t) B_t dt, \quad B_0 = 1, \quad (2.4.2)
$$

where the coefficients are fulfilling the Lipschitz and linear growth conditions and the processes $r$ and $\sigma$ are non-negative.

A few basic (yet sometimes hard to distinguish) concepts to start the ball rolling.

**Definition 2.4.1 (Self-Financing Strategy).** A strategy (or portfolio) is a stochastic process $(\alpha_t, \beta_t)$ for $t \in [0, T]$ where $\alpha \in L^2_{\text{loc}}$ and $\beta \in L^1_{\text{loc}}$ the value of the strategy $(\alpha_t, \beta_t)$ is the real value process

$$
V_t^{(\alpha, \beta)} = \alpha_t S_t + \beta_t B_t, \quad (2.4.3)
$$

which is said to be self-financing if

$$
dV_t^{(\alpha, \beta)} = \alpha_t dS_t + \beta_t dB_t, \quad (2.4.4)
$$
holds, that is
\begin{equation}
V_t^{(\alpha, \beta)} = V_0^{\alpha, \beta} + \int_0^t \alpha_u dS_u + \int_0^t \beta_u dB_u.
\end{equation}

(2.4.5)

It tells us that any change in the value of the self-financing portfolio must equal the profit or loss due to changes in the price of the stock or the price of the bond, see [79, Page 157].

**Definition 2.4.2 (Arbitrage).** An arbitrage is a self-financing strategy \((\alpha_t, \beta_t)\) for \(t \in [0, T]\) whose value \(V_t^{(\alpha, \beta)}\) is such that (i) \(V_0^{(\alpha, \beta)} = 0\) a.s. and there exists \(t_0 \in (0, T]\) such that (ii) \(V_{t_0}^{(\alpha, \beta)} \geq 0\) a.s. and (iii) \(P\left(V_{t_0}^{(\alpha, \beta)} > 0\right) > 0\).

**Definition 2.4.3 (Equivalent Martingale Measure, EMM).** An equivalent martingale measure \(Q\) is a probability measure on \((\Omega, \mathcal{F})\) so that (i) \(Q\) is equivalent to \(P\); (ii) The process of discounted stock price is an honest-\(Q\) martingale (not simply a local martingale), that is,
\begin{equation}
S_t = E_Q \left( e^{-\int_t^T r_u du} S_T | \mathcal{F}_t \right).
\end{equation}

**Definition 2.4.4 (Contingent Claim).** If there exists a unique equivalent martingale measure \(Q\), then a nonnegative random variable \(X \in \mathcal{F}_T\) is called a contingent claim provided that \(E_Q(X^2) < \infty\).

We can think of contingent claim as a contract pays an amount \(X\) at time \(t \in [0, T]\) or time \(T\).

**Remark 2.4.5.** In discrete-time market, these are all we need to determine the unique price of that ticket mentioned in the beginning of Chapter 1. Namely, there is no arbitrage opportunities in the market if and only if there exists at least one EMM. Then, if such EMM is unique, then the discrete market is said to be complete where every contingent claim can be priced uniquely by arbitrage.

In continuous-time market, the problem of existence of arbitrage opportunities has become rather delicate to the point that proper definition of an arbitrage opportunity and the ensuing mathematical development are extremely complex, see [40] and [41, Page 236]. Imposing the conditions of admissibility is the way out for the continuous-time theory of arbitrage pricing.

**Definition 2.4.6 (Admissible Strategy).** A self-financing strategy \((\alpha_t, \beta_t)\) whose value is positive (i.e. \(V_t^{\alpha, \beta} \geq 0\) for all \(t \in [0, T]\), is a admissible strategy if there exists an unique equivalent martingale measure \(Q\) so that \(\{V_t^{(\alpha, \beta)} B_t^{-1}\}\) is a \(Q\)-martingale.

**Remark 2.4.7.** An easier definition is provided in [63, Page 226]: a strategy \((\alpha_t, \beta_t)\) is admissible if it is bounded from below, that is there exists a constant such that \(V_t^{\alpha, \beta} \geq C\), for \(t \in [0, T]\) a.s..

We denote \(\mathcal{A}\) as the class of all admissible strategies, which contains no arbitrage strategies, then if \((\alpha_t, \beta_t) \in \mathcal{A}\) satisfies \(\alpha_T S_T + \beta_T B_T = X\) for a contingent claim \(X\), we say that \((\alpha_t, \beta_t)\) replicates \(X\).

The admissible strategy exists in the complete continuous-time market model, where every contingent claim is priced by arbitrage. In particular, the price is given by the expectation of the discounted contingent claim under the unique EMM \(Q\), i.e.
\begin{equation}
V_t = B_tE_Q \left( B_T^{-1} X | \mathcal{F}_t \right)
\end{equation}
Proposition 2.4.8 (Completeness of the Market). The \((B,S)\)-market given by SDEs (2.4.1) and (2.4.2) is complete if the following conditions are satisfied:

(i) \(E_P\left(e^{\int_0^T (\mu_t-r_t)^2 dt}\right) < \infty\);  
(ii) \(E_Q\left(e^{\int_0^T \sigma_t^2 dt}\right) < \infty\);

(iii) \(\sigma\left(\frac{S_t}{B_t}, t \in [0, T]\right) = \mathcal{F}_T\);  
(iv) \(P\left(\int_0^T \left(\frac{\mu_t}{\sigma_t}\right)^2 dt < \infty\right) = 1\).

Remark 2.4.9. Conditions (i) and (ii) are to ensure the existence and uniqueness of equivalent martingale measure \(Q\), conditions (iii) suggests all the sources of uncertainty can be explained by the price dynamics of the basic securities.

Finally it is time to enter our favourite example: the Black-Scholes Model to put all the pieces together, even though it can hardly count as an optimal stopping problem since the stopping time is fixed at \(T\), the spirit behind this is certainly priceless for the development of optimal stopping in mathematical finance.

Example 2.4.10 (The European Call Option Pricing Problem). Suppose that the time dynamics of processes (2.4.1) and (2.4.2) are now specified as follows

\[
dS_t = \mu S_t dt + \sigma S_t dW_t, \quad S_0 = s, \\
 dB_t = r B_t dt, \quad B_0 = 1,
\]

where the coefficients \(\mu\), \(\sigma\) and \(r\) are fulfilling the conditions on Proposition 2.4.8 so that the market is complete. The payoff of current contract at time \(T\) is \((S_T - K)^+\) where \(K\) is the strike price and \(T\) is the maturity time. Our mission is to find the arbitrage free price of this contract at any time \(t \in [0, T]\).

Aha!-type Solution

The market is complete so that there exists a unique equivalent \(Q\) and the arbitrage free price is directly given as \(V(t, S_t) = E_Q\left(e^{-r(T-t)}(S_T - K)^+|\mathcal{F}_t}\right)\) and by the Markov property of GMB (this is the Markovian setting), \(V(t, S_t) = E_Q\left(e^{-r(T-t)}(S_T - K)^+|S_t}\right)\).

The problem therefore is reduced to find EMM \(Q\) (also known as risk neutral measure in some context), under which, the discounted stock price can be interpreted as a martingale. This is where Theorem 2.2.9 shines.

To begin, via an appeal to the Itô formula, we have

\[
dS_t B_t^{-1} = B_t^{-1} S_t \left(\frac{\mu - r}{\sigma} dt + dW_t\right),
\]

and for simplicity, let \(m = \frac{\mu - r}{\sigma}\), which is often referred as the market price of risk, and if \(m = 0\), the model is the risk neutral model with \(P\) automatically being the risk neutral measure \(Q\). Then, by Theorem 2.2.9, the process defined by

\[
dW_t = \frac{\mu - r}{\sigma} dt + dW_t, \quad (2.4.6)
\]
A Laborious Solution

To the famous Black-Scholes partial differential equation (PDE): so that application of Itô’s formula yields such solution. The theory of option pricing come to being and we only confine ourselves to providing a sketch of paper written decades ago by Black and Scholes [7] is the perfect place for readers to understand its existence is automatic) that replicates payoff of the European call option at time T.

\[ \frac{\partial V_t}{\partial t} + \mu x \frac{\partial V_t}{\partial x} + \frac{1}{2} \sigma^2 x^2 \frac{\partial^2 V_t}{\partial x^2} = r V_t \quad \text{for } t \in [0, T], \]

which, coupling with \( V_T = \max(S_T - K, 0) \), leads us to the Black-Scholes formula, we omit further details.

\[ V_t = e^{-r(T-t)} \mathbb{E}_t \left[ (S_T - K)^+ \right] \]

A Laborious Solution

Another method to attack this problem is to find an admissible strategy (the market is complete, so its existence is automatic) that replicates payoff of the European call option at time T. The original paper written decades ago by Black and Scholes [7] is the perfect place for readers to understand the theory of option pricing come to being and we only confine ourselves to providing a sketch of such solution.

Let \( V_t^{\alpha, \beta} \) be the total value of the portfolio at time \( t \), and its construction be \( V_t^{\alpha, \beta} = \alpha_t S_t + \beta_t B_t \), which satisfies the self-financing condition:

\[ dV_t^{\alpha, \beta} = \alpha_t dS_t + \beta_t dB_t = (\mu \alpha_t S_t + r \beta_t B_t) dt + \sigma \alpha_t S_t dW_t, \tag{2.4.7} \]

and its payoff stream at \( T \) equals \( V_T^{\alpha, \beta} = \alpha_T S_T + b_T \beta_T = G(T, S_T) \).

Let us suppose that \( V_t^{\alpha, \beta} \) can also be written as a function in \( C^{1,2}([0, T] \times (0, \infty)) \) (this assumption is rather considered as the weak point of such method) i.e. \( V_t^{\alpha, \beta} = f(t, S_t) \) so that an application of Itô’s formula yields

\[ df(t, S_t) = f_t(t, S_t) dt + f_x(t, S_t) dS_t + \frac{1}{2} \sigma^2 S_t^2 f_{xx} dt \]

\[ = \left( f_t + f_x S_t + \frac{1}{2} \sigma^2 S_t^2 f_{xx} \right) (t, S_t) dt + \sigma S_t f_x(t, S_t) dW_t. \tag{2.4.8} \]

It is then only logical to match the coefficients on (2.4.7) and (2.4.8),

\[ \begin{cases} \left( f_t + \mu S_t f_x + \frac{1}{2} \sigma^2 S_t^2 f_{xx} \right) (t, S_t) = \mu \alpha_t S_t + r \beta_t B_t, \\ \sigma S_t f_x(t, S_t) = \sigma \alpha_t S_t, \end{cases} \tag{2.4.9} \]

so that \( \alpha_t = f_x(t, S_t) \) and \( \beta_t = f_t + \frac{1}{2} \sigma^2 S_t^2 f_{xx}(t, S_t) \). This, coupling with \( V_t^{\alpha, \beta} = f(t, S_t) \), leads us to the famous Black-Scholes partial differential equation (PDE):

\[ rf(t, x) - rx f_x(t, x) - f_t(t, x) - \frac{1}{2} \sigma^2 x^2 f_{xx}(t, x) = 0, \tag{2.4.10} \]

with the terminal boundary condition \( f(T, S_T) = (S_T - K)^+ \).
The only thing that does not add up is that the role of EMM \( Q \) is nowhere to be found. But once we invoke the Itô’s formula again on the discounted value function, together with (2.4.6) and (2.4.10),

\[
de^{-ru}f(t + u, S_{t+u}) = e^{-ru}
\left(-rf + ft + \mu S t f_x + \frac{1}{2} \sigma^2 S t^2 f_{xx}\right) (t + u, S_{t+u})du + \sigma S_{t+u} f_x (t, S_{t+u})dW_u
\]

\[
= e^{-ru}
\left(-rf + ft + \mu S t f_x + \frac{1}{2} \sigma^2 S t^2 f_{xx}\right) (t + u, S_{t+u})du
\]

\[
+ \sigma S_{t+u} f_x (t + u, S_{t+u}) \left(dW_u - \frac{\mu - r}{\sigma} du\right)
\]

\[
= e^{-ru} \sigma S_u f_x (t + u, S_{t+u}) dW_u,
\]

implying its martingale property under measure \( Q \), i.e. the precise definition of the admissible strategy (the local martingale property is fairly immediate, and its martingale property follows via checking \( E \left[ \int_0^T (e^{-ru} \sigma S_{t+u} f_x (t + u, S_{t+u}))^2 du \right] < \infty \)), which has captured the soul of arbitrage pricing theory and such fact will be confirmed over and over again in the remaining context.

Let us recast the example given in the beginning of Chapter 1 under the assumption that the arbitrage-free (continuous-time) market is complete. Any investors who are rational enough to enter this market need not to be a rocket scientist to know their goal is to maximise the payoff of that ticket (the prototype of American contingent claim) given that the exercise time is flexible, but the rule to decide when to exercise it must not anticipate the future (recall in the efficient market, you only correct the decision instantaneously as the new information becomes available), implying that it must be a stopping time.

**Definition 2.4.11 (American Contingent Claim).** An American Contingent Claim is a financial instrument consisting of (i) an expiration date \( T \in (0, \infty) \); (ii) the selection of an exercise time \( \tau \), which is a \( F_t \)-stopping time and (iv) a terminal payoff \( f_\tau \) at the exercise time.

The following theorem boils the arbitrage-free pricing problem down to its essence and puts the optimal stopping problem in the map. See [46, Page 50], [47, Page 223] and [76, Page 543].

**Theorem 2.4.12 (The Fair Price for the American Contingent Claim).** The fair price at time \( t = 0 \) for the American Contingent Claim is given by

\[
V = \sup_{\tau \in [0, T]} B_0 E_Q \left( \frac{f_\tau}{B_\tau} \right).
\]

Moreover, there exists a hedging strategy (inside \( A \) and satisfies the “balance” condition) whose value process \( X = (X_t)_{t \in [0, T]} \) can be described by formula

\[
X_t = \text{ess sup}_{\tau \in [t, T]} B_t E_Q \left( \frac{f_\tau}{B_\tau} \mid F_t \right).
\]
Remark 2.4.13. One might wonder why we care for the hedging strategies in the first place? Intuitively, investors can be divided into buyers and sellers of the contingent claims: buyer wants to maximise the payoff, while the seller must choose the strategies of value $X$ so that $X_\tau \geq f_\tau$ a.s. and thereby guaranteeing his ability to meet the contract terms for each stopping time $\tau \in [0, T]$ chosen by the buyers to exercise the contract. See [47, Page 221].
Chapter 3

Optimal Stopping: The General Results

As stated in the book [64] by Peskir and Shiryaev, the goal of studying the optimal stopping problem (3.0.1) is: “firstly, to exhibit an optimal stopping time, at which, the supremum is attained; secondly, to compute the value function as explicitly as possible ... Thus, the central problem arises as how to determine the stopping and the continuation sets.” Because of the important roles played by continuation and stopping sets in optimal stopping theory, we begin our investigation by studying the very existence of these sets under the umbrella of martingale property. We then review the basic existence and uniqueness theorems in [64] for the optimal stopping time in the infinite horizon case and as one would expect, we shall also extend these theorems within the finite-time framework.

In this chapter, we assume that $X$ is a strong Markov process, such that problem (2.0.1) can be generalised as follows:\footnote{Hereafter, we write $X_{t+s}^{t,x}$ simply as $X_{t+s}^{x}$ so that when $t = 0$, $X_{s}^{0,x}$ is just $X_{s}^{x}$.}

\begin{equation}
V(t, x) = \sup_{\tau \in [0, T-t]} E_{t, x}(G(t + \tau, X_{t+\tau})),
\end{equation}

with terminal boundary condition $V(T, x) = G(T, x)$ and as before, the measurable gain function fulfils the uniform integrable condition:

$$E_{t, x} \left( \sup_{s \in [0, T-t]} |G(t + s, X_{t+s})| \right) < \infty.$$

### 3.1 The Continuation and Stopping Sets

The optimal stopping problem is somewhat like gambling on a fair game, you cannot make money out of it (i.e. no arbitrage opportunity, see Chapter 2) with the rule that the game ends at certain time $T$ and the player who wants to maximise the outcome of the game can walk away from the game anytime $t \in [0, T]$ she prefers and obtains her reward $G(t, X_{t})$, or continue the game expecting for a better outcome in the near future with the knowledge of the present state of the game and the probability distribution of the future. Thus, it all boils down to the problem of when should the player stop/continue observing and playing the game; or mathematically speaking, finding the optimal stopping rules [77].
Lemma 3.1.1. The observation should always stop immediately if \( \{G(t, X_t), \mathcal{F}_t\}_{t \in [0, T]} \) is a supermartingale.

Proof. Immediate from its supermartingale property,

\[
E_{t,x}(G(t+s, X_{t+s})) \leq G(t, x),
\]
such that

\[
V(t, x) = \sup_{\tau \in [0, T-t]} E_{t,x}(G(t + \tau, X_{t+\tau})) \leq G(t, x).
\]
The fact that \( V(t, x) \geq G(t, x) \) then tells us that we must have equality down the line, and the conclusion follows.

Lemma 3.1.2. The observation should always continue till the terminal time \( T \) if \( \{G(t, X_t), \mathcal{F}_t\}_{t \in [0, T]} \) is a submartingale.

Proof. The submartingale property implies that \( E_{t,x}[G(t+s, X_{t+s})] \geq G(t, x) \) for \( t \in [0, T) \) so we similarly find that

\[
V(t, x) = \sup_{\tau \in [0, T-t]} E_{t,x}(G(t + \tau, X_{t+\tau})) > G(t, x),
\]
and thereby proving the claim as we hoped.

In the language of optimal stopping, the observation should stop immediately once the observed value is inside the stopping set, and continue while it is inside the continuation set; and as the last two corollaries hinted, we can begin to put the definition of sets into print.

Definition 3.1.3. The stopping set \( \mathcal{D} \) and the continuation set \( \mathcal{C} \) are given as:

\[
\mathcal{D} = \{(t, x) \in [0, T) \times \mathbb{R} : V(t, x) = G(t, x)\} \cup \{(T, x) : x \in \mathbb{R}\},
\]

\[
\mathcal{C} = \{(t, x) \in [0, T) \times \mathbb{R} : V(t, x) > G(t, x)\},
\]
and for \( t = T \), \( \mathcal{C} \) is an empty set.

The last two lemmas also help us establish the very existences of sets \( \mathcal{C}, \mathcal{D} \) and rule out those not so interesting cases, among them is an instructive example: pricing the American call option.

Example 3.1.4 (The American Call Option on Non-dividend-paying Stock). Let \( S = (S_t)_{t \in [0, T]} \) denote the price of a stock and \( B = (B_t)_{t \in [0, T]} \) the price of a bond at time \( t \) and then the time dynamics of these two processes be given by

**Bond Model:** \( dB_t = rB_t dt, \quad B_0 = 1 \)

**Stock Model:** \( dS_t = S_t(r dt + \sigma dW_t), \quad S_0 = s \)

that is, we assume that the stock price follows a geometric Brownian motion, and the bond price is a deterministic process with the exponential growth (notice that this is a risk neutral model given that...
the market price of risk is zero, the probability measure \( P \) is automatically the risk neutral measure \( Q \); and the gain function of a standard call option \( G(t, x) = (x - K)^+ \) for \( x \in (0, \infty) \). The value function thus equals:

\[
V(t, s) = \sup_{\tau \in [0, T-t]} E_{t,s} \left( e^{-r\tau}(S_{t+\tau} - K)^+ \right).
\] (3.1.1)

Then by invoking Lemma 2.2.6 and Theorem 2.3.10 on the gain function \( G \),

\[
E_{t,s}(G(t + \tau, S_{t+\tau})) = G(t, s) + E_{t,s} \left( rK \int_0^\tau e^{-ru} I\{S_u \geq K\} du + \frac{1}{2} \int_0^\tau e^{-ru} du K(S_u) \right)
\]

which after taking the supremum over all the stopping times \( \tau \), coupled with the proof of Lemma 3.1.2, shows that the observations should continue till the maturity time \( T \).

Tellingly, we have ended up with the arbitrage-free price of European call option in Example 2.4.10 considering it replicates the payoff stream of the American Call option, which implies that the value function \( V \) in (3.1.1) also satisfies the PDE (2.4.10) and exhibits the very same martingale property, and therefore reassures us that martingale is the key concept for solving the optimal stopping problem.

**Remark 3.1.5.** Another elegant way to see the submartingale property of the gain function in foregoing example is through the eye of Proposition 2.3.5! Just notice that \( g(x) = x^+ \) is a convex integrable function, and \( X_t = e^{-rt} (S_t - K) \) is submartingale in the sense that for \( 0 \leq s \leq t \leq T \)

\[
E(X_t|\mathcal{F}_s) = E(e^{-rt}S_t|\mathcal{F}_s) - e^{-rs}K + e^{-rs}K - e^{-rt}K = X_s + (e^{-rs} - e^{-rt}) K \geq X_s.
\]

The seed planted by the last example is how we can apply Theorem 2.2.5, Lemma 2.2.6 and Theorem 2.2.7 (depending on the smoothness of the gain function) and Theorem 2.3.10 to make our judgement call at first glance. Further exploration of the general properties of these sets will be possible after bringing the optimal stopping time into the picture.

### 3.2 The Optimal Stopping Time: Existence and Uniqueness

As usual, the first order of business is to define optimal stopping time.

**Definition 3.2.1** (Finite Time Formulation). A stopping time \( \tau_s \) with \( P_{t,x}(\tau_s \leq T-t) = 1 \) is said to be optimal in problem (3.0.1) if it satisfies \( V(t, x) = E_{t,x}(G(t + \tau_s, X_{t+\tau_s})) \).

**Definition 3.2.2** (Infinite Time Formulation). A stopping time \( \tau_s \in [0, \infty) \) is said to be optimal in problem (2.0.3) if it satisfies \( V(x) = E_{0,x}(G(\tau_s, X_{\tau_s})) \); otherwise, no optimal stopping time exists.

The existence in infinite-time formulation requires extra care as the stopping times are possibly infinite (just modify Example 3.1.4 with \( T = \infty \) and alas, \( \tau_s = \infty \)), the aftermath of this is an additional step of checking that the candidate optimal stopping times are finite. One natural candidate that grabs our attention in particular is the first entrance time of the stopping set \( \mathcal{D} \), denoted as \( \tau_D \) and its symbolic definition will be specified depending on the optimal stopping problem.
The next two results, borrowed from [64, Page 37, Theorem 2.4; Page 40, Theorem 2.7] establish the necessary and the sufficient conditions for the existence of optimal stopping time \( \tau_* \) in the infinite-time formulated problem:

\[
V(x) = \sup_{\tau} E_x(G(X_\tau)), \quad (3.2.1)
\]

and justify the optimality of \( \tau_D = \inf\{s \geq 0 : X_s \in \mathcal{D}\} \).

**Theorem 3.2.3 (Necessary Condition).** If there exists an optimal stopping time \( \tau_* \) in (3.2.1) for all \( x \in \mathbb{R} \), then, the value function \( V \) is the smallest superharmonic function that dominates the gain function \( G \) for \( x \in \mathbb{R} \).

In addition, if \( V \) is lower semicontinuous and \( G \) is upper semicontinuous, then the stopping time \( \tau_D \) satisfies \( \tau_D \leq \tau_* \) \( P_x \)-a.s. for all \( x \in \mathbb{R} \) and is optimal in (3.2.1). The stopped process \( (V(X_{t \wedge \tau_D}))_{t \geq 0} \) is right-continuous martingale under \( P_x \) for every \( x \in \mathbb{R} \).

**Theorem 3.2.4 (Sufficient Condition).** Consider the optimal stopping problem (3.2.1) and let \( \hat{V} \) be the smallest superharmonic function (if exists) that dominates the gain function \( G \) on \( x \in \mathbb{R} \) and that \( \hat{V} \) is lower semicontinuous and \( G \) is upper semicontinuous. Then, define \( \mathcal{D} = \{x \in \mathbb{R} : \hat{V}(x) = G(x)\} \).

(i) If \( \tau_D = \inf\{s \geq 0 : X_s \in \mathcal{D}\} \) and \( P_x(\tau_D < \infty) = 1 \) for all \( x \in \mathbb{R} \), then \( \hat{V} = V \) and \( \tau_D \) is optimal in (3.2.1);

(ii) If \( \tau_D = \inf\{s \geq 0 : X_s \in \mathcal{D}\} \) and \( P_x(\tau_D < \infty) < 1 \) for some \( x \in \mathbb{R} \), then there is no optimal stopping time in (3.2.1)

**Remark 3.2.5.** Under the weaker assumption that the gain function \( G \) is non-negative, bounded and semicontinuous, Theorem 3.2.4 still holds true. See [62, Page 203].

It is then only reasonable to ask if these results can be carried over to the finite-time formulated problem (3.0.1). In fact, they do and all details will be laid out in the upcoming text with \( \tau_D \) being reshuffled as

\[
\tau_D = \inf\{s \geq 0 : (t + s, X_{t+s}) \in \mathcal{D}\} \land (T - t).
\]

**Theorem 3.2.6 (Necessary Condition).** If there exists an optimal stopping time \( \tau_* \) in (3.0.1), then (i) the value function \( V \) is the smallest superharmonic function which dominates the gain function on \( (t, x) \in [0, T) \times \mathbb{R} \).

In addition, if \( V \) is lower semicontinuous and \( G \) is upper semicontinuous on \( [0, T] \times \mathbb{R} \), then (ii) the stopping time \( \tau_D \) satisfies \( \tau_D \leq \tau_* \) \( P_{t,x} \)-a.s. for all \( (t, x) \in [0, T] \times \mathbb{R} \) and is optimal in (3.0.1); (iii) the stopped process \( (V(t + s \wedge \tau_D, X_{t+s})_{s \in [0, T-t]} \) is a right-continuous martingale under \( P_{t,x} \) for all \( (t, x) \in [0, T] \times \mathbb{R} \).

**Proof.** (i) The first thing to prove is that \( V \) is the superharmonic function by simply using Definitions 2.3.6, 3.2.1 and the strong Markov property of \( X \) for each stopping times \( \sigma \in [0, T-t] \).

\[
E(V(t + \sigma, X_{t+\sigma})|F_t) = E\left(E(G(t + \sigma + \tau_*, X_{t+\sigma+\tau_*})|F_{t+\sigma})|F_t\right) = E_{t,x}(G(t + \sigma + \tau_*, X_{t+\sigma+\tau_*})) \leq \sup_{\tau \in [0, T-t]} E_{t,x}(G(t + \tau, X_{t+\tau})) = V(t, x),
\]
where the second equality follows from Lemma 2.3.12; which is precisely Definition 2.3.6.

Next in line is to show that $V$ is the smallest superharmonic function dominating $G$, which by Definition 2.3.7 is equivalently to prove that if there exists any other superharmonic function $F$ that dominates $G$, then $V \leq F$. Equipped with the definition, we thus have

$$F(t, x) \geq E_{t,x}(F(t + \tau, X_{t+\tau})) \geq E_{t,x}(G(t + \tau, X_{t+\tau})),$$

and by taking the supremum among all the stopping times $\tau \in [0, T - t]$ of the extreme member,

$$F(t, x) \geq \sup_{\tau \in [0, T - t]} E_{t,x}(G(t + \tau, X_{t+\tau})) = V(t, x),$$

and thereby proving statement (i).

(ii) The next point of the program is to show that $P_{t,x}(\tau_D \leq \tau_s) = 1$ and the optimality of $\tau_D$. Assume the contrary that $P_{t,x}(\tau_D > \tau_s) > 0$. Then

$$E_{t,x}(G(t + \tau_s, X_{t+\tau_s})) = E_{t,x}(G(t + \tau_s, X_{t+\tau_s}) I\{\tau_D > \tau_s\} + G(t + \tau_s, X_{t+\tau_s}) I\{\tau_D \leq \tau_s\})$$

$$< E_{t,x}(V(t + \tau_s, X_{t+\tau_s}) I\{\tau_D > \tau_s\} + V(t + \tau_s, X_{t+\tau_s}) I\{\tau_D \leq \tau_s\})$$

$$= E_{t,x}(V(t + \tau_s, X_{t+\tau_s})) \leq V(t, x),$$

where the strictly inequality is from the fact that $G(t + \tau_s, X_{t+\tau_s}) < V(t + \tau_s, X_{t+\tau_s})$ if $\tau_s > \tau_D$ and $G \leq V$ always; which is then in the violation of the optimality of $\tau_s$ and thereby, proving the claim as hoped.

It thus remains to verify the optimality of $\tau_D$. To begin, we choose $(t, x) \in C$, and since $V$ is the superharmonic function, we have

$$V(t, x) = E_{t,x}(G(t + \tau_s, X_{t+\tau_s})) \leq E_{t,x}(V(t + \tau_s, X_{t+\tau_s}))$$

$$\leq E_{t,x}(V(t + \tau_D, X_{t+\tau_D})) \leq V(t, x),$$

where the second inequality follows form the supermartingale property of $V$:

$$E(V(t + \tau_s, X_{t+\tau_s}) | F_t) = E(E(V(t + \tau_s, X_{t+\tau_s}) | F_{t+\tau_D}) | F_t)$$

$$\leq E_{t,x}(V(t + \tau_D, X_{t+\tau_D})).$$

Then, we choose an irregular boundary point $(t, x) \in \partial C$ so that $P(\tau_D > 0) = 1$, where $\partial C$ is the boundary of continuation set $C$. Let $\{\tau_k, k \geq 1\}$ be a sequence of stopping time such that $0 < \tau_k < \tau_D$ and $\tau_k \to 0$ P-a.s. as $k \to \infty$. It follows that $(t + \tau_k, X_{t+\tau_k}) \in C$, by Lemma 2.3.12:

$$E_{t,x}(G(t + \tau_D, X_{t+\tau_D})) = E(E(G(t + \tau_D, X_{t+\tau_D}) | F_{t+\tau_k}) | F_t)$$

$$= E_{t,x}(V(t + \tau_k, X_{t+\tau_k})). \quad (3.2.2)$$

By the l.s.c. of $V$, Lemma 2.1.5 and (3.2.2),

$$V(t, x) \leq E_{t,x}(\liminf_{k \to \infty} V(t + \tau_k, X_{t+\tau_k})) \leq \liminf_{k \to \infty} E_{t,x}(V(t + \tau_k, X_{t+\tau_k}))$$

$$= E_{t,x}(G(t + \tau_D, X_{t+\tau_D})) \leq V(t, x),$$
indicating the optimality of $\tau_D$.

To this end, if $(t, x) \in \partial C$ is regular so that $P(\tau_D = 0) = 1$ or $(t, x) \in D$, then

$$V(t, x) = G(t, x) = E_{t,x}(G(t + \tau_D, X_{t+\tau_D})),$$

after which, the assertion follows.

(iii) As usual, we lean on the definition of martingale property to prove the final statement,

$$E_{t,x}(V(t + s \wedge \tau_D, X_{t+s\wedge\tau_D})) = E\left(E(G(t + \tau_D, X_{t+\tau_D}) \mid \mathcal{F}_{t+s\wedge\tau_D}) \mid \mathcal{F}_t\right)$$

$$= E_{t,x}(G(t + \tau_D, X_{t+\tau_D})) = V(t, x),$$

where in the second equality Lemma 2.3.12 has managed to make its appearance in the third time throughout the proof! Finally, its right continuity follows from Corollary 2.3.8. \qed

**Remark 3.2.7.** Statement (ii) in Theorem 3.2.6 is also referred as the uniqueness theorem for optimal stopping in [62, Page 204].

In the infinite time formulated problem, the sufficient condition proposed in Theorem 3.2.4 will be fairly easy to check, because the candidate value function $\hat{V}$ can normally be derived as closed-form solution by solving the corresponding free-boundary problem in the form of ordinary differential equations (ODEs). However, in finite time setting, this is a different story mainly because of the involvement of PDEs.

Now, let us scrutinise what Theorem 3.2.4 tries to accomplish is to show that the value function $V$ in (3.0.1) is the smallest superharmonic function that dominates $G$, so if such fact can be confirmed directly from the definition of $V$, then we are on the right track, given that $P(\tau_D < \infty) = 1$.

**Theorem 3.2.8 (Sufficient Condition).** Consider the optimal stopping problem (3.0.1). If $\hat{V}$ is lower semicontinuous on $[0, T] \times \mathbb{R}$ and $G$ is upper semicontinuous on $[0, T] \times \mathbb{R}$. Then, (the optimal stopping time exists as) $\tau_D$ is optimal in (3.0.1).

Before we proceed to the proof, we pause for some useful properties of the superharmonic function (see [62, Page 197]):

**Lemma 3.2.9.** (a) If $f$ is superharmonic and $\alpha > 0$, then $\alpha f$ is superharmonic; 

(b) If $f_1$ and $f_2$ are superharmonic, then $f_1 + f_2$ is superharmonic; 

(c) If $f$ is superharmonic and $\sigma \leq \tau$ are stopping times, then $E_{0,x}(f(X_\sigma)) \geq E_{0,x}(f(X_\tau))$.

**Proof of Theorem 3.2.8.** (i) We wish to prove that $V$ is superharmonic; which, together with the proof of (i) in Theorem 3.2.6, presents the fact that $V$ is the smallest superharmonic function dominating $G$. The value function is l.s.c. on $[0, T] \times \mathbb{R}$ so it is measurable on $[0, T] \times \mathbb{R}$. Then, from the definition of the value function, we see that for $\sigma \in [0, T - t]$,

$$V(t + \sigma, X_{t+\sigma}) = \sup_{\tau \in [0, T-t-\sigma]} E_{t+\sigma,X_{t+\sigma}}(G(t + \sigma + \tau, X_{t+\sigma+\tau}))$$

$$= \sup_{\tau \in [0, T-t-\sigma]} E(G(t + \sigma + \tau, X_{t+\sigma+\tau}) \mid \mathcal{F}_{t+\sigma}),$$
and by the definition of essential supremum, it follows that
\[ V(t + \sigma, X_{t+\sigma}) = \operatorname{ess \sup}_{\tau \in [0, T-t-\sigma]} E(G(t + \sigma + \tau, X_{t+\sigma+\tau}) | \mathcal{F}_{t+\sigma}). \]

Next in line is to show that the family:
\[ \left\{ E(G(t + \sigma + \tau, X_{t+\sigma+\tau}) | \mathcal{F}_{t+\sigma}) : \tau \in [0, T-t-\sigma] \right\} \]
is upward directed in the sense that for any stopping times \( \tau_1, \tau_2 \in [0, T-t-\sigma] \), there exists stopping time \( \tau_3 \in [0, T-t-\sigma] \) so that with probability 1,
\[ E(G(t + \sigma + \tau_3, X_{t+\sigma+\tau_3}) | \mathcal{F}_{t+\sigma}) \geq E(G(t + \sigma + \tau_1, X_{t+\sigma+\tau_1}) | \mathcal{F}_{t+\sigma}) \vee E(G(t + \sigma + \tau_2, X_{t+\sigma+\tau_2}) | \mathcal{F}_{t+\sigma}). \]

To achieve this, let event
\[ B = \left\{ E(G(t + \sigma + \tau_1, X_{t+\sigma+\tau_1}) | \mathcal{F}_{t+\sigma}) \geq E(G(t + \sigma + \tau_2, X_{t+\sigma+\tau_2}) | \mathcal{F}_{t+\sigma}) \right\} \]
and that \( B \in \mathcal{F}_{t+\sigma} \). Let \( \tau_3 = \tau_1 I\{B\} + \tau_2 I\{B^c\} \), which is a stopping time as
\[ \{\tau_3 \leq t\} = \{\{\tau_1 \leq t\} \cap B\} \cup \{\{\tau_2 \leq t\} \cap B^c\} \in \mathcal{F}_t, \]
so that
\[
E(G(t + \sigma + \tau_3, X_{t+\sigma+\tau_3}) | \mathcal{F}_{t+\sigma}) \\
= E(G(t + \sigma + \tau_1, X_{t+\sigma+\tau_1}) I\{B\} | \mathcal{F}_{t+\sigma}) + E(G(t + \sigma + \tau_2, X_{t+\sigma+\tau_2}) I\{B^c\} | \mathcal{F}_{t+\sigma}) \\
= I\{B\} E(G(t + \sigma + \tau_1, X_{t+\sigma+\tau_1}) | \mathcal{F}_{t+\sigma}) + I\{B^c\} E(G(t + \sigma + \tau_2, X_{t+\sigma+\tau_2}) | \mathcal{F}_{t+\sigma}) \\
= E(G(t + \sigma + \tau_1, X_{t+\sigma+\tau_1}) | \mathcal{F}_{t+\sigma}) \vee E(G(t + \sigma + \tau_2, X_{t+\sigma+\tau_2}) | \mathcal{F}_{t+\sigma}),
\]
which is precisely the definition of being upward directed.

We note that being upward directed means a bounded sequence \( \{\tau_k, k \geq 1\} \) and \( \tau_k \leq T-t \) can be chosen so that
\[
V(t + \sigma, X_{t+\sigma}) = \lim_{k \to \infty} E(G(t + \sigma + \tau_k, X_{t+\sigma+\tau_k}) | \mathcal{F}_{t+\sigma}) \quad \text{P-a.s.} \quad (3.2.3)
\]
and that \( \left\{ E(G(t + \sigma + \tau_k, X_{t+\sigma+\tau_k}) | \mathcal{F}_{t+\sigma}), k \geq 1 \right\} \) is a non-decreasing sequence.

Toward this end, taking the expectation under measure \( P_{t,x} \) on both side of (3.2.3):
\[ E_{t,x}(V(t + \sigma, X_{t+\sigma})) = \left( \lim_{k \to \infty} E(G(t + \sigma + \tau_k, X_{t+\sigma+\tau_k}) | \mathcal{F}_{t+\sigma}) | \mathcal{F}_t \right), \]
after which, an appeal to the monotone convergence theorem yields
\[
E_{t,x}(V(t + \sigma, X_{t+\sigma})) = \lim_{k \to \infty} E\left( E(G(t + \sigma + \tau_k, X_{t+\sigma+\tau_k}) | \mathcal{F}_{t+\sigma}) | \mathcal{F}_t \right) \\
= \lim_{k \to \infty} E_{t,x}(G(t + \sigma + \tau_k, X_{t+\sigma+\tau_k})) \leq V(t,x),
\]
and thereby, proving that $V$ is a superharmonic function since it is also measurable on $[0, T] \times \mathbb{R}$.

(ii) In order to take care of the optimality of $\tau_D$, we first consider the sets, for $\lambda \in (0, 1)$:

$$
C_\lambda = \{(t, x) \in [0, T) \times \mathbb{R} : \lambda(V(t, x) - h(t, x)) > G(t, x) - h(t, x)\},
$$

$$
D_\lambda = \{(t, x) \in [0, T) \times \mathbb{R} : \lambda(V(t, x) - h(t, x)) \leq G(t, x) - h(t, x)\} \cup \{(T, x) : x \in \mathbb{R}\},
$$

where function $h$ is defined as follows for $(t, x) \in [0, T) \times \mathbb{R}$:

$$
h(t, x) = E_{t,x} \left( \inf_{s \in [0, T-t]} G(t + s, X_{t+s}) \right),
$$

from which, simply note that $-h(t, x)$ is a superharmonic function as for stopping time $\sigma \in [0, T-t]$, we have, for stopping time $\tau_D$ as follows:

$$
\tau_D = \inf \{s \geq 0 : (t + s, X_{t+s}) \in D_\lambda \} \wedge (T - t),
$$

so that

$$
G(t, x) - h(t, x) \leq \lambda(V(t, x) - h(t, x)) + (1 - \lambda)E_{t,x} \left( V\left(t + \tau_D, X_{t+\tau_D}\right) - h\left(t + \tau_D, X_{t+\tau_D}\right) \right),
$$

in the sense that if $(t, x) \in C_\lambda$, by the definition of the set and the nonnegative expectation term, the inequality is rather immediate; if $(t, x) \in D_\lambda$, $P(\tau_D = 0) = 1$, leading to the fact that $G \leq V$.

Next, by function $-h$ being superharmonic and statement (a) in Lemma 3.2.9,

$$
-(1 - \lambda)h(t, x) \geq -(1 - \lambda)E_{t,x} \left( h\left(t + \tau_D, X_{t+\tau_D}\right) \right).
$$

After combining (3.2.4) with (3.2.5), the following inequality emerges:

$$
G(t, x) \leq \lambda V(t, x) + (1 - \lambda)E_{t,x} \left( V\left(t + \tau_D, X_{t+\tau_D}\right) \right).
$$

Moreover, by the strong Markov property of $X$, we have, for stopping time $\sigma \in [0, T-t]$

$$
E_{t,x} \left( E_{t+\sigma, X_{t+\sigma}} \left( V\left(t + \sigma + \tau_D, X_{t+\sigma+\tau_D}\right) \right) \right)
= E \left( E \left( V\left(t + \sigma + \tau_D, X_{t+\sigma+\tau_D}\right) \left| \mathcal{F}_{t+\sigma}\right. \right) \left| \mathcal{F}_t \right. \right)
= E_{t,x} \left( V\left(t + \sigma + \tau_D, X_{t+\sigma+\tau_D}\right) \right) \leq E_{t,x} \left( V\left(t + \tau_D, X_{t+\tau_D}\right) \right),
$$
where the last inequality follows from the statement (c) in Lemma 3.2.9 and it follows that
\[ E_{t,x} \left( V \left( t + \tau_{D,}, X_{t+\tau_{D,}} \right) \right) \]
suggesting that \( E_{t,x} \left( V \left( t + \tau_{D,}, X_{t+\tau_{D,}} \right) \right) \) being the superharmonic function so that by statement (b) in Lemma 3.2.9
\[
\lambda V(t, x) + (1 - \lambda) E_{t,x} \left( V \left( t + \tau_{D,}, X_{t+\tau_{D,}} \right) \right),
\]
is itself a superharmonic function that dominates \( G \) because of (3.2.6) and since \( V \) is proven to be the smallest superharmonic function dominating \( G \) in statement (i), the following relation emerges:
\[
V(t, x) \leq \lambda V(t, x) + (1 - \lambda) E_{t,x} \left( V \left( t + \tau_{D,}, X_{t+\tau_{D,}} \right) \right),
\]
and consequently,
\[
V(t, x) - E_{t,x} \left( V \left( t + \tau_{D,}, X_{t+\tau_{D,}} \right) \right) \leq \lambda \left( V(t, x) - E_{t,x} \left( V \left( t + \tau_{D,}, X_{t+\tau_{D,}} \right) \right) \right),
\]
entailing that, in fact,
\[
V(t, x) = E_{t,x} \left( V \left( t + \tau_{D,}, X_{t+\tau_{D,}} \right) \right).
\]
By equation (3.2.8), the definition of \( \tau_{D,} \) and the definition of \( V \),
\[
V(t, x) = E_{t,x} \left( V \left( t + \tau_{D,}, X_{t+\tau_{D,}} \right) \right)
\leq \frac{1}{\lambda} E_{t,x} \left( G \left( t + \tau_{D,}, X_{t+\tau_{D,}} \right) + \left( 1 - \frac{1}{\lambda} \right) E_{t,x} \left( h \left( t + \tau_{D,}, X_{t+\tau_{D,}} \right) \right) \right).
\]
Finally, let \( \lambda \to 1 \), by invoking lemma 2.1.5 and using the upper semicontinuity of \( G \) in \( [0, T) \times \mathbb{R} \),
\[
V(t, x) \leq \limsup_{\lambda \to 1} E_{t,x} \left( G \left( t + \tau_{D,}, X_{t+\tau_{D,}} \right) \right)
\leq E_{t,x} \left( \limsup_{\lambda \to 1} G \left( t + \tau_{D,}, X_{t+\tau_{D,}} \right) \right)
\leq E_{t,x} \left( G \left( t + \tau_{D}, X_{t+\tau_{D}} \right) \right) \leq V(t, x),
\]
and the optimality of \( \tau_{D} \) follows.

**Remark 3.2.10.** The assumption towards functions \( G \) and \( V \) being upper semicontinuity (i.e. \( -G \) is l.s.c.) and lower semicontinuity (i.e. \( V - G \) is l.s.c.) implies that the stopping set \( \{ (t, x) : [0, T) \times \mathbb{R} : V(t, x) - G(t, x) \leq 0 \} \) is closed and the continuation set \( \{ (t, x) : [0, T) \times \mathbb{R} : V(t, x) - G(t, x) > 0 \} \) is open. See [16, Page 170] and [62, Page 203].

**Remark 3.2.11.** Despite the assumption of semicontinuity is useful in their own right, the reality is that if the gain function is non-negative, bounded measurable (together with other fairly common hypotheses), we normally cannot have them both unless one of them is continuous, see [5, Page 61] and [62, Page 135, Lemma 8.1.4]; namely, on the state space \( [0, T] \times \mathbb{R} \),
in both finite and infinite time horizon:
(a) if the gain function \( G \) is l.s.c., then the value function \( V \) is l.s.c;
(b) if the gain function \( G \) is u.s.c., then the value function \( V \) is u.s.c.;
(c) if the gain function \( G \) is continuous, then the value function \( V \) is continuous.
With the knowledge of optimal stopping time, we feel more than ready to pursue the solution of the optimal stopping problem.

Let us assume (and fall into the inventor’s paradox for a second) in this subsection that a bounded, nonnegative gain function \( G \) is \( C^{1,2}(\mathbb{R}) \) so that the Itô’s formula is available for us and by invoking Theorem 2.3.10 and then,

\[
E_t(x, G(t + s, X_{t+s})) = G(t, x) + E_t(x) \int_0^s H(t + u, X_{t+u}) du,
\]

where

\[
H(t, x) = G_t(t, x) + \mathbb{L}^X G(t, x),
\]

and by Remark 3.2.11 (let us bravely assuming those technical hypotheses are fulfilled hereafter), we have the value function being continuous on \([0, T] \times \mathbb{R}\).

**Remark 3.3.1.** Let us define set \( \mathcal{U} = \{(t, x) \in [0, T] \times \mathbb{R} : H(t, x) > 0\} \). Then, \( \mathcal{U} \subset \mathbb{C} \). Consequently, it is never optimal to stop before the process exits from \( \mathcal{U} \) (the process might even proceed for some time after leaving set \( \mathcal{U} \)). This is quite a sound method to determine that the continuation set is not empty.

Via an appeal to Theorem 3.2.8, we know that \( \tau_* \) exists and Theorem 3.2.6 is therefore at our service, which has given us the key we lacked before to initiate the free-boundary problem! If the value function is \( C^{1,2} \) except in the unknown boundary \( \partial \mathcal{D} \) (for \( \mathcal{D} \)), the very first statement (i) of Theorem 3.2.6 entails that

\[
\mathbb{L}^X V + V_t \leq 0 \quad \text{on} \quad \mathcal{C} \cup \mathcal{D} \setminus \partial \mathcal{D}
\]

and statement (iii) suggests that if the continuation set exists after all, the value function in \( \mathcal{C} \) is a right-continuous martingale under \( P_{t,x} \), namely,

\[
\mathbb{L}^X V + V_t = 0, \quad \text{in} \quad \mathcal{C},
\]

\[
V(t, x) = G(t, x), \quad \text{in} \quad \mathcal{D} \setminus \partial \mathcal{D}.
\]

The assumption of \( V \) being \( C^{1,2} \) inside \( \mathcal{C} \) can be well justified in [64, Page 131], which we shall state for easy reference.

**Lemma 3.3.2 (The \( C^{1,2} \)-property of Value Function Inside \( \mathcal{C} \)).** Let \( X \) be a strong Markov process characterised in terms of the infinitesimal generator defined as (2.2.8) with \( \mu \) and \( \sigma \) being sufficiently smooth and \( \partial \mathcal{D} \) sufficiently regular; meaning that each point \((t, x)\) from \( \partial \mathcal{D} \) gives us \( P_{t,x}(\tau_{\mathcal{D}} = 0) = 1 \) with \( \tau_{\mathcal{D}} = \inf\{s \geq 0 : (t + s, X_{t+s}) \in \mathcal{D}\} \wedge (T - t) \), then the value function \( V \in C^{1,2} \) in \( \mathcal{C} \) but not necessarily at boundary \( \partial \mathcal{D} \).

As one might suspect, some boundary conditions about \( \partial \mathcal{D} \) are missing for us to determine the solution of (3.3.2), but this can be soon resolved by extracting some pleasant features possessed by the boundary \( \partial \mathcal{D} \), i.e. continuous fit and smooth fit conditions.
The principle of continuous fit states that the optimal stopping boundary \( \partial D \) is selected so that the value function is continuous for \((t, x) \in \partial D\), i.e. \( V(t, x) = G(t, x) \); thus, with the assumption that the value function is continuous, our mission is accomplished rather immediately.

The principle of smooth fit states that the value function is smooth for \((t, x) \in \partial D\) i.e.

\[
V_t(t, x) = G_t(t, x), \\
V_x(t, x) = G_x(t, x),
\]

which is a bit more delicate and thereby making its break-down easier and verification more difficult as we shall see in the later chapters. Along these lines, it is noteworthy that the justification of smooth-fit conditions depends highly on the monotonicity of the boundary in finite-time horizon (otherwise, assumption is needed. See [23, Page 1522]) and of course, the smoothness of \( G \) in the neighbourhood of \( \partial D \).

The real wisdom behind function \( H \) is that it nicely reveals the monotonicity of boundary \( \partial D \) and sometimes the number of boundary, which vaguely speaking is embedded on how many times function \( H \) changes its sign (from positive to negative, see for example [22]).

**Lemma 3.3.3.** If the map \( t \mapsto H(t, x) \) is decreasing, then the continuation set \( C \) is left-connected and the stopping set \( D \) is right-connected w.r.t time \( t \).

**Proof.** Take any \( t_1, t_2 \in [0, T) \) so that \( t_1 \leq t_2 \). Suppose that \( \tau_* \) is the optimal stopping time for \( V(t_2, x) \) so that \( 0 \leq \tau_* \leq T - t_2 \leq T - t_1 \), indicating the sub-optimality of \( \tau_* \) for \( V(t_1, x) \). Then,

\[
V(t_1, x) - V(t_2, x) \geq E_{t_1, x}(G(t_1 + \tau_*, X_{t_1+\tau_*})) - E_{t_2, x}(G(t_2 + \tau_*, X_{t_2+\tau_*})) \\
= G(t_1, x) - G(t_2, x) + E \left( \int_0^{\tau_*} H(t_1 + u, X^x_u) - H(t_2 + u, X^x_u) \, du \right) \\
\geq G(t_1, x) - G(t_2, x),
\]

where the first inequality is due to the definition of \( \tau_* \) and the last inequality is from the assumption that the map \( t \mapsto H(t, x) \) is decreasing; from which, we see that if \( V(t_1, x) - G(t_1, x) = 0 \), together with the fact that \( V \) dominates \( G \), then \( V(t_2, x) - G(t_2, x) = 0 \), i.e. \((t_1, x) \in D \implies (t_2, x) \in D\); if \( V(t_2, x) - G(t_2, x) > 0 \), then \( V(t_1, x) - G(t_1, x) > 0 \), i.e. \((t_2, x) \in C \implies (t_1, x) \in C \). This is precisely the conclusion.

If, in addition to Lemma 3.3.3, the continuation and stopping sets are up-down connected, that is, for \((t, x_1), (t, x_2) \in [0, T] \times \mathbb{R}, V(t, x_1) - G(t, x_1) \geq V(t, x_2) - G(t, x_2)\), then the boundary (or boundaries) \( \partial D \) is monotone.

To illuminate such idea, we recast the classic American put option pricing problem.

**Example 3.3.4 (The American Put Option).** Once again, we assume a risk neutral model as that of Example 3.1.4 s.t. its value function equals

\[
V(t, s) = \sup_{\tau \in [0,T-t]} E_{t,s}(e^{-r\tau}(K - S_{t+\tau})^+). 
\]

First order of business is to examine the existence of the optimal stopping time \( \tau_* \). Observe from the gain function \( G(x) = (K - x)^+ \) is non-negative, bounded, continuous and measurable, checking all the boxes in Remark 3.2.11 and hence, the value function is continuous, justifying the existence of \( \tau_* \).
Second of all is to check the connectedness of sets in time. The map \( t \mapsto V(t, x) \) is decreasing in the sense that for \( t_1 \leq t_2, [0, T - t_2] \subseteq [0, T - t_1] \) and the supremum taken over a set is greater than over its subset. Therefore, we have \( V(t_1, x) - G(x) \geq V(t_2, x) - G(x) \) and Lemma 3.3.3 is applicable.

Next is to check the connectedness of sets in space. Note that from the strong solution of GBM, \( X_t^{x_1} \leq X_t^{x_2} \) for \( x_1 \leq x_2 \) so that

\[
V(t, x_1) - V(t, x_2) \leq \sup_{\tau \in [0, T-t]} E \left( e^{-r\tau} \left( K - X^{x_1}_{\tau} \right)^+ - e^{-r\tau} \left( K - X^{x_2}_{\tau} \right)^+ \right) \\
\leq \sup_{\tau \in [0, T-t]} E \left( e^{-r\tau} \left( X^{x_2} - X^{x_1} \right)^+ \right) \\
= (x_2 - x_1) \sup_{\tau \in [0, T-t]} E \left( e^{-\frac{\sigma^2}{2}\tau + \sigma W_\tau} \right) \\
= G(x_1) - G(x_2),
\]

where the last equality follows from the martingale property of \( \{ (e^{-\frac{\sigma^2}{2}t + \sigma W_t}, F_t), t \geq 0 \} \) (see A.5) and thereby showing the connectedness in space, in particular, \( C \) is up-connected and \( D \) is down-connected. As a result, the boundary \( \partial D \) is monotone and it is increasing. (We know the boundary is unique because function \( H \) changes its sign once only!)

Examples 3.1.4 and 3.3.4 are not formatively matching our topic on problem concerning (starting) time dependent payoff \( G(t, x) \), so we are not going to settle until a real example is provided.

**Example 3.3.5** (The Asian Option). The arbitrage-free price of the Asian option is equivalent to the value function (after change of measure, details see [64, Page 417])

\[
V(t, x) = \sup_{\tau \in [0, T-t]} E_t,x \left( \left( 1 - \frac{X_{t+\tau}}{t+\tau} \right)^+ \right),
\]

where \( X = \frac{t_{t+s}}{S_{t+s}} = \frac{x + \int_0^s S_u du}{S_t} \) solves \( dX_{t+s} = (1 - r X_{t+s}) ds + \sigma X_{t+s} dB_s \) with \( X_t = x \), GBM \( (S_{t+s})_{s \geq 0} \) and \( S_t = 1 \). To begin, we notice that the gain function is once again non-negative, bounded, continuous on \( (0, T] \times [0, \infty) \), so is the value function. Hence, the optimal stopping time \( \tau_s \) exists on \( (0, T] \).

As before, establishing the connectedness of sets in time is next to go.

In this problem, invoking Lemma 3.3.3 is not a legitimate idea since we realise that the local-time term is of a curve which is a function of \( t \)! In later chapters, when we come up against the challenge of the local-time term of a constant level, the only obstacle needed to be overcome is the behaviour of its coefficients w.r.t. time \( t \). “Going back to definitions” adds another clues to this puzzle, in the lovely book [68, Page 91], George Pólya emphasises how one can benefit from such method, “Going back to definitions is important in inventing argument but it is also important in checking it”.

Let \( t_1, t_2 \in [0, T] \) so that \( t_1 \leq t_2 \). Suppose that \( \tau_s \) is the optimal stopping time of \( V(t_2, x) \) so that \( 0 \leq \tau_s \leq T - t_2 \leq T - t_1 \), upon noticing that

\[
V(t_2, x) = E_{t_2,x} \left( \left( 1 - \frac{X_{t_2+\tau_s}}{t_2+\tau_s} \right)^+ \right) = E \left( \left( 1 - \frac{X_{t_2+\tau_s}}{t_2+\tau_s} \right) I \{ X_{\tau_s}^x \leq t_2 + \tau_s \} \right)
\]
\[ P \left( X^x_{\tau_s} \leq t_2 + \tau_s \right) - E \left( \frac{X^x_{\tau_s} I \{ X^x_{\tau_s} \leq t_2 + \tau_s \} }{t_2 + \tau_s} \right) \geq G(t_2, x) = \left( 1 - \frac{x}{t_2} \right)^+ \geq 1 - \frac{x}{t_2}, \]

from which, the inequality \( E \left[ \frac{X^x_{\tau_s} I \{ X^x_{\tau_s} \leq t_2 + \tau_s \} }{t_2 + \tau_s} \right] - \frac{x}{t_2} \leq 0 \) holds. Then,

\[
V(t_2, x) - V(t_1, x) - (G(t_2, x) - G(t_1, x)) \\
\leq E \left( \left( 1 - \frac{X^x_{\tau_s}}{t_2 + \tau_s} \right)^+ - \left( 1 - \frac{X^x_{\tau_s}}{t_1 + \tau_s} \right)^+ \right) - \frac{x(t_2 - t_1)}{t_2 t_1} \\
\leq E \left( \frac{X^x_{\tau_s}}{t_1 + \tau_s} - \frac{X^x_{\tau_s}}{t_2 + \tau_s} \right) - \frac{x(t_2 - t_1)}{t_2 t_1} \\
\leq E \left( \left( \frac{X^x_{\tau_s}(t_2 - t_1)}{(t_1 + \tau_s)(t_2 + \tau_s)} \right) I \left\{ \frac{X^x_{\tau_s}}{t_2 + \tau_s} \leq 1 \right\} \right) - \frac{x(t_2 - t_1)}{t_2 t_1} \\
= E \left( \left( \frac{X^x_{\tau_s}(t_2 - t_1)}{(t_1 + \tau_s)(t_2 + \tau_s)} \right) I \left\{ \frac{X^x_{\tau_s}}{t_2 + \tau_s} \leq 1 \right\} \right) - \frac{x(t_2 - t_1)}{t_2 t_1} \\
\leq \frac{t_2 - t_1}{t_1} E \left( \frac{X^x_{\tau_s}}{t_2 + \tau_s} \right) I \left\{ \frac{X^x_{\tau_s}}{t_2 + \tau_s} \leq 1 \right\} - \frac{x}{t_2} \leq 0,
\]

and after arriving to the same inequality as (3.3.4), the rest of the claim is automatic.

Towards that end, let us look into the set connectedness in space once again via going back to definition: for \( 0 \leq x_1 \leq x_2 \) and \( t > 0 \), we have

\[
V(t, x_1) - V(t, x_2) \leq \sup_{\tau \in (0, T-t]} E \left( \left( 1 - \frac{x_1 + I_\tau}{(t + \tau)S_\tau} \right)^+ - \left( 1 - \frac{x_2 + I_\tau}{(t + \tau)S_\tau} \right)^+ \right) \\
\leq (x_2 - x_1) \sup_{\tau \in (0, T-t]} E \left( \frac{1}{(t + \tau)S_\tau} \right) \leq \frac{1}{t} (x_2 - x_1) = G(t, x_1) - G(t, x_2).
\]

From [64, Page 432, Equation 27.1.32], we see \( H(t, x) \) (the local time term is somewhat a by-product of the computation of \( H \) ) changes its sign once on \( x = \frac{1}{t+\tau} \), so the boundary is unique.

**Remark 3.3.6.** Single argument based on monotonicity of \( \partial \mathcal{D} \) to establish the smooth fit condition or the continuity of \( \partial \mathcal{D} \) will be easily outgrown by problems themselves because of our unlimited curiosity, we believe successful attempt will open the door for research of more atypical contingent claim pricing problems. Again we refer the interested readers to [17] for their probabilistic argument.

Finally, with every pieces of the puzzle in order, the free-boundary problem emerges.
Chapter 4

Russian Option with Last Exit Time

This chapter, whose content is largely based on the working paper [83] of the author, is devoted to solving the following optimal stopping problem in both infinite \((T = \infty)\) and finite \((T < \infty)\) time horizon\(^1\):

\[
V = \sup_{\tau \in [0,T]} E_{0,x,s} \left( e^{-(r+\lambda)\tau} (S_\tau - LX_\tau)^+ I\{\theta > \tau\} \right),
\]

with random time \(\theta = \sup\{t \in [0, T] : X_t = S_t\}\) and the supremum taken over all the stopping times \(\tau \in [0, T]\).

We assume the stock and bond price processes are given by the SDEs

\[
dX_t = rX_t dt + \sigma X_t dW_t, \quad X_0 = x, \tag{4.0.2}
\]

\[
dB_t = rB_t dt, \quad B_0 = 1, \tag{4.0.3}
\]

and \(S = (S_t)_{t \in [0, T]}\) is capturing the running maximum of the stock price process, i.e.

\[
S_t = s \lor \max_{0 \leq u \leq t} X_u, \quad S_0 = s,
\]

where \(W = (W_t)_{t \in [0, T]}\) is a standard Brownian motion process started at zero under measure \(P\), \(r > 0\) is the interest rate, \(\sigma > 0\) is the volatility coefficient and \(\lambda > 0\) is the discounting rate. The American option with payoff function \((S_t - LX_t)^+\) is called Russian option, where we let \(L > 0\) and \(L > 1\) in the infinite and finite time formulation respectively; such option belongs to the class of put option with aftereffect and discounting, see [77, Page 626]. It is worth mentioning that there is no restriction to assume that \(x = 1\) (and \(s \geq 1\)).

The indicator function in (4.0.1) imposes, of course, an additional layer to the original problem: what if our agent is not as rational as he supposed to be in this efficient market, in the sense that, instead of considering his optimal stopping strategy during the time being \([0, T]\), he actually restricts his decision to be made before the last time the underlying stock price is at its running maximum, reflecting his lack of risk appetite revealed in Chapter 6.

\(^1\)In this chapter, we write \(E_{0,x,s}(A), E_{t,x,s}(A)\) as the conditional expectation (of event A) relative to the conditional measure \(P(A|X_0 = x, S_0 = s)\) and \(P(A|X_t = x, S_t = s)\) respectively. And we write \(P(A|X_0 = x, S_0 = s) = P_{0,x,s}(A)\).
In Section 4.1, we first look into problem (4.0.1) by setting $T = \infty$ and introduce a new Markov process to reduce the dimension of the problem. Then, by Theorem 3.2.4 in Subsection 4.1.2, we examine the solution of the free-boundary problem and verify that it is the smallest superharmonic function dominating the gain function. Similar infinite-horizon problems have been studied in [31, 32, 74] without reducing the dimension of the problem.

In Section 4.2, analogously, one can reduce the dimension of the problem as $T < \infty$. The continuity of $V$ naturally leads to the existence of the optimal stopping time. Then, we investigate the features of the continuation and stopping sets in Subsection 4.2.3. Dealing with time-dependent gain function, a proper assumption is then made to guarantee the monotonicity of the free boundary; in so doing, we are able to arrive at the free-boundary problem. Subsection 4.2.4 covers the theoretical and numerical results concerning the price that the agent is willing to pay for this presuming contract.

### 4.1 Infinite-time Horizon

#### 4.1.1 Reformulation and Basics

The main optimal stopping problem can be reformulated as follows by the measurability of stopping time $\tau$ and Lemma 2.3.12:

$$
V = \sup_{\tau} E_{0,x,s} \left( e^{-(r+\lambda)\tau} (S_\tau - LX_\tau)^+ I\{\theta > \tau\} | \mathcal{F}_\tau \right)
= \sup_{\tau} E_{0,x,s} \left( e^{-(r+\lambda)\tau} (S_\tau - LX_\tau)^+ P(\theta > \tau | \mathcal{F}_\tau) \right). \tag{4.1.1}
$$

First of all, we investigate the property of Azéma supermartingale associated with $\theta$.

**Proposition 4.1.1.** In our setting, the following formula holds for $r - \frac{\sigma^2}{2} < 0$,

$$
P(\theta > t | \mathcal{F}_t) = \left( \frac{S_t}{X_t} \right)^\alpha,
$$

where $\alpha = \frac{2r}{\sigma^2} - 1$.

**Proof.** We note that $\{\theta > t\} = \{\max_{u \geq t} X_u \geq S_t\}$, and hence,

$$
P(\theta > t | \mathcal{F}_t) = P\left( \max_{u \geq t} X_u \geq S_t | \mathcal{F}_t \right) = P\left( \frac{\max_{u \geq t} X_u}{X_t} \geq \frac{S_t}{X_t} | \mathcal{F}_t \right)
= P\left( \max_{u \geq 0} \left( r - \frac{\sigma^2}{2} \right) u + \sigma W_u \geq \log \frac{S_t}{X_t} \right)
= 1 - P\left( \max_{u \geq 0} \left( r - \frac{\sigma^2}{2} \right) u + \sigma W_u \leq \log \frac{S_t}{X_t} \right)
= 1 - (1 - e^{\alpha \log \frac{S_t}{X_t}}) = \left( \frac{S_t}{X_t} \right)^\alpha,
$$

where the fifth equality follows from Lemma 2.2.11. \qed
Remark 4.1.2. Here one should note that by Lemma 2.2.11, \( P ( \theta > t | \mathcal{F}_t ) = 1 \) for \( r - \frac{\sigma^2}{2} > 0 \), problem (4.1.1) is therefore the classic Russian option pricing problem. For the detailed solution of pricing Russian option, see [64, Page 400].

Therefore, (4.1.1) gets the form

\[
V = \sup_{\tau} E_{0,x,s} \left( e^{-(r+\lambda)\tau} \frac{S_\tau}{X_\tau} \right). \tag{4.1.2}
\]

With some additional effort, we can reduce the above two-dimensional problem (4.1.2) into a one-dimensional problem. To do so, we introduce the probability measure \( \tilde{P} \), which satisfies

\[
d\tilde{P} = e^{\sigma W_t - \frac{\sigma^2}{2} t} dt \text{ and the process } Y = (Y_t)_{t \geq 0} = \left( \frac{S_t}{X_t} \right)_{t \geq 0}.
\]

It is a well-known fact that the strong solution of (4.0.2) is

\[
X_t = xe^{\left( r - \frac{\sigma^2}{2} \right) t + \sigma W_t} = xe^{\left( r + \frac{\sigma^2}{2} \right) t + \sigma \tilde{W}_t}, \tag{4.1.3}
\]

for \( t \geq 0 \), where \( W \) and \( \tilde{W} \) are standard Brownian motions under measure \( P \) and \( \tilde{P} \) respectively.

Corollary 4.1.3. The process \( Y = (Y_t)_{t \geq 0} \) is a (time-homogeneous) strong Markov process on the phase space \([1, \infty)\) with instantaneous reflection at the point \( \{1\} \), which satisfies the SDE

\[
dY_t = -rY_t dt - \sigma Y_t d\tilde{W}_t + dR_t, \quad Y_0 = y = \frac{s}{x}, \tag{4.1.4}
\]

where \( dR_t = I \{ Y_t = 1 \} \frac{dS_t}{X_t} \) and \( \tilde{W} \) is the standard Brownian motion under measure \( \tilde{P} \). See [76, Page 770].

Proof. (i) Its strong Markov property has already been verified in Example 2.2.4.

(ii) By exploiting Itô formula, we have

\[
dY_t = S_t d \left( \frac{1}{X_t} \right) + \left( \frac{1}{X_t} \right) dS_t
\]

\[
= - \frac{S_t}{X_t^2} (r X_t dt + \sigma X_t dW_t) + \frac{S_t}{X_t^3} \sigma^2 X_t^2 dt + dR_t
\]

\[
= (\sigma^2 - r) Y_t dt - \sigma Y_t dW_t + dR_t,
\]

where we set \( dR_t = \left( \frac{1}{X_t} \right) dS_t \) and it only changes value as \( (Y_t)_{t \geq 0} \) arrives at the boundary point \( \{1\} \), to stress this fact, we write \( (R_t)_{t \geq 0} = \int_0^t I \{ Y_s = 1 \} \frac{1}{X_s} dS_s \); after which, upon letting \( \tilde{W}_t = W_t - \sigma t \), SDE (4.1.4) follows.

(iii) Next in line is to show that \( \{1\} \) is the instantaneous reflection point (i.e. the process \( (Y_t)_{t \geq 0} \) spends zero time at \( \{1\} \) \( \tilde{P} \)-a.s.), that is,

\[
\int_0^t I \{ Y_u = 1 \} du = 0, \quad \tilde{P} \text{-a.s}
\]
for each \( t > 0 \). Via taking expectation under measure \( \tilde{P} \) and using Fubini’s theorem,

\[
\tilde{E}_y \int_0^t I\{Y_u = 1\} \, du = \int_0^t \tilde{E}_y (I\{Y_u = 1\}) \, du = \int_0^t \tilde{P}_y (Y_u = 1) \, du = 0,
\]

where the last equality holds via the fact that the probability density function of \( \{X_t, S_t\} \) exists, implying that \( Y \) is a continuous random variable, and that the probability of a continuous random variable equals a certain value is 0. The desired statement then follows from the simple fact that for non-negative random variables \( X \), if \( E(X) = 0 \), then \( X = 0 \) a.s.

(iv) The process \( Y \) has the property of being time-homogeneous, in the following sense:

\[
Y_{t+h}^y = y - \int_t^{t+h} rY_u^t \, du - \int_t^{t+h} \sigma Y_u^t \, d\tilde{W}_u + \int_t^{t+h} dR_u
\]

\[
= y - \int_0^h rY_{t+s}^y \, ds - \int_0^h \sigma Y_{t+s}^y \, d\tilde{W}_s + \int_0^h dR_{t+s}, \quad (u = t + s)
\]

where \( \tilde{W}_s = \tilde{W}_{t+s} - \tilde{W}_t, s \geq 0 \). On the other hand, of course,

\[
Y_{h}^0 = y - \int_0^h rY_s^0 \, ds - \int_0^h \sigma Y_s^0 \, d\tilde{W}_s + \int_0^h dR_s,
\]

Since \( \tilde{W}_s \overset{d}{=} \tilde{W}_s \) and

\[
dR_{t+s} = \frac{1}{X_{t+s}} dS_{t+s} = \frac{1}{X_{t+s}} \left( \frac{1}{r + \frac{\sigma^2}{2}} \right)_{s + \sigma \tilde{W}_s} d\left( \max_{0 \leq u \leq t+s} X_u \right)
\]

\[
= \frac{1}{X_{t+s}} \left( \frac{1}{r + \frac{\sigma^2}{2}} \right)_{s + \sigma \tilde{W}_s} d\left( \max_{0 \leq u \leq t} X_u \vee \max_{0 \leq u \leq t+s} X_u \right)
\]

\[
= \frac{1}{X_{t+s}} \left( \frac{1}{r + \frac{\sigma^2}{2}} \right)_{s + \sigma \tilde{W}_s} d\left( s \vee \max_{0 \leq u \leq s} X_u \right) = dR_s,
\]

under measure \( \tilde{P} \), it follows by weak uniqueness of the solution of the SDEs that \( \{Y_{t+h}^t \}_{h \geq 0} \overset{d}{=} \{Y_h^0 \}_{h \geq 0} \) that is, the process \( Y \) is time-homogeneous. \( \square \)

It then follows that \( (Y_t)_{t \geq 0} \) increases on the set \( \{ t : Y_t = 1 \} \), with the infinitesimal generator

\[
\mathbb{L}Y = -r y \frac{\partial}{\partial y} + \frac{1}{2} \sigma^2 y^2 \frac{\partial^2}{\partial y^2}, \quad \text{in } (1, \infty),
\]

and if function \( f \in C^2((1, \infty)) \) and its limit \( \lim_{y \downarrow 1} f(y) \) exists, then \( f'(1+) = 0 \).

Summarising our findings so far, we see that the problem (4.1.2), by using the process \( Y \) and change-of-measure, could be reduced to the following optimal stopping problem

\[
V = \sup_{\tau} E_{0,x,s} \left( e^{-(r+\lambda)\tau} X_\tau \left( \frac{S_\tau}{X_\tau} - L \right) \left( \frac{S_\tau}{X_\tau} \right)^\alpha \right)
\]

\[\footnote{We write \( Y_{t+h}^y \) simply as \( Y_{t+s}^y \) hereafter whenever needed, so that when \( t = 0 \), \( Y_s^0 \) is written as \( Y^y_s \).}
\[\footnote{For notational convenience, we simply write \( \tilde{E}_y(A) \) instead of \( \tilde{E}_{0,y}(A) \) hereafter till the end of Chapter 4.1.}
\[= \sup_{\tau} E_{0,y} \left( e^{-\lambda \tau} (Y_{\tau} - L)^{+} Y_{\tau}^{\alpha} \right). \tag{4.1.5} \]

### 4.1.2 The Free-boundary Problem

We are now ready to turn our attention to the free boundary problem but first for the sake of brevity, we define the gain function \( G(y) = (y - L)^{+} y^{\alpha} \).

To begin with, we invoke Theorem 2.2.7 to obtain

\[
e^{-\lambda t} G(Y_t) = G(y) + \int_0^t e^{-\lambda u} \left( -\lambda G - r Y_u^{y} G_y + \frac{1}{2} \sigma^2 Y_u^{yy} G_{yy} \right) (Y_u^{y}) I\{Y_u^{y} > L\} du \]

\[-\int_0^t e^{-\lambda u} \sigma Y_u^{y} G_y (Y_u^{y}) I\{Y_u^{y} \neq L\} d\tilde{W}_u \]

\[+ \int_0^t e^{-\lambda u} \sigma Y_u^{y} G_y (1+) dR_u \]

\[+ \frac{1}{2} \int_0^t e^{-\lambda u} (G_y(L) - G_y(L^{-})) dL_u(Y), \]

where \( L_u(Y) \) is the local time of \( Y \) at the level \( L \) given by

\[
l_u^L(Y) = \tilde{P}_y - \lim_{\epsilon \to 0} \frac{1}{2\epsilon} \int_0^u I\{|Y_r - L| < \epsilon\} d\langle Y, Y \rangle_r,\]

after which, the important observation follows, that is, the further away \( Y \) gets from \( \max(1, L) \), the less likely the gain function will increase upon continuing, which suggests that there exists a point \( b \in [\max(1, L), \infty) \) such that the stopping time

\[
\tau_b = \inf\{t \geq 0 : Y_t^{y} \geq b\} \tag{4.1.6} \]

should be optimal in the problem (4.1.5), such fact will soon be confirmed.

It is then not far-fetched to ask whether the stopping time \( \tau_b \) is finite?

**Corollary 4.1.4.** The stopping time \( \tau_b \) is finite.

**Proof.**\(^4\) To prove \( \tilde{P}_y(\tau_b < \infty) = 1 \) for \( y \in [1, b) \), it suffices to show that

\[
\tilde{P}_y \left( \max_{t \geq 0} Y_t \geq b \right) = 1, \text{ for } y = 1,
\]

To begin with, for \( n \geq 1, \)

\[
\tilde{P}_y \left( \max_{0 \leq t \leq n} \frac{S_t}{X_t} \geq b \right) = \tilde{P}_y \left( \max_{0 \leq u \leq t \leq n} \frac{X_u}{X_t} \geq b \right) = \tilde{P}_y \left( \max_{0 \leq u \leq t \leq n} \frac{X_u}{X_t} \geq b \right)
\]

\[= \tilde{P}_y \left( \max_{0 \leq u \leq t \leq n} e^{(r + \frac{a^2}{2})(u-t) + \sigma(\bar{W}_u - \bar{W}_t)} \geq b \right)\]

\(^4\)We essentially follows the proof from [75, Page 116], the original proof is missing the minus sign.
\[ \geq \tilde{P}_y \left( \max \left\{ \sigma \left( \tilde{W}_1 - \tilde{W}_0 \right), \ldots, \sigma \left( \tilde{W}_n - \tilde{W}_{n-1} \right) \right\} \geq \log b + r + \sigma^2 \right) . \]

Then, let \( C = \frac{\log b + r + \sigma^2}{\sigma} \), \( X_n = \max \left\{ \tilde{W}_1 - \tilde{W}_0, \ldots, \tilde{W}_n - \tilde{W}_{n-1} \right\} \) and set event

\[ A_k = \{ \tilde{W}_{k+1} - \tilde{W}_k \geq C \}, \quad \text{for } k \geq 0. \]

Note that the events \( \{ A_k, \ 0 \leq k \leq n \} \) are independent and the crucial observation is that

\[ \lim_{n \to \infty} X_n \geq C \iff \{ A_k \text{ infinitely often} \}. \]

Since \( 0 < \tilde{P}_y(A_k) < 1 \), it follows that \( \sum_{k=0}^\infty \tilde{P}_y(A_k) = +\infty \). An appeal to the second Borel-Cantelli lemma (Assertion (ii) in Theorem 2.1.2) asserts that \( \tilde{P}_y(A_k \text{ infinitely often}) = 1 \), and consequently that \( \tilde{P}_y \left( \max_{t \geq 0} Y_t \geq b \right) \geq \tilde{P}_y(A_k \text{ infinitely often}) = 1 \) as desired.

By reviewing Theorem 3.2.4, we know that the next task is to find the smallest superharmonic function \( \hat{V} \) that dominates \( G \) and the unknown point \( b \) by solving the corresponding free boundary problem:

\[
\begin{align*}
\mathbb{L}_Y \hat{V} &= \lambda \hat{V} \quad \text{for } y \in (1, b), \\
\hat{V}(y) &= G(y) \quad \text{for } y = b, \\
\hat{V}_y(y) &= G_y(y) \quad \text{for } y = b \quad \text{(smooth fit),} \\
\hat{V}_y(y) &= 0 \quad \text{for } y = 1 \quad \text{(normal reflection),} \\
\hat{V}(y) &> G(y) \quad \text{for } y \in [1, b), \\
\hat{V}(y) &= G(y) \quad \text{for } y \in (b, \infty). 
\end{align*}
\]

We proceed to solve the free-boundary problem, after which, we prove that its solution coincides with the value function in (4.1.5) and \( b \) is unique.

First, we apply the infinitesimal generator of \( Y \) to (4.1.7) and obtain the Cauchy-Euler equation as follows

\[
\frac{1}{2} \sigma^2 y^2 \frac{d^2 \hat{V}}{dy^2} - ry \frac{d\hat{V}}{dy} - \lambda \hat{V} = 0,
\]

from which, we know the solution gets form

\[ \hat{V}(y) = y^p. \]

By inserting (4.1.14) into (4.1.13), we obtain the following quadratic equation

\[
\frac{1}{2} \sigma^2 p^2 - \left( r + \frac{\sigma^2}{2} \right) p - \lambda = 0,
\]
whose roots are given by
\[ p_i = \frac{r + \sigma^2}{\sigma^2} \pm \sqrt{\left(\frac{r + \sigma^2}{\sigma^2}\right)^2 + 2\lambda \sigma^2}, \quad (i = 1, 2), \]
where \( p_1 > 1, p_2 < 0 \). The general solution of (4.1.7) therefore equals
\[ \hat{V}(y) = C_1 y^{p_1} + C_2 y^{p_2}, \quad (4.1.16) \]
where \( C_1 \) and \( C_2 \) are arbitrary constants. An exploitation of conditions (4.1.8) and (4.1.10) on (4.1.16) gives us
\[ C_1 = -\frac{p_2(b^{\alpha+1} - Lb^\alpha)}{p_1 b^{p_2} - p_2 b^{p_1}}, \quad C_2 = \frac{p_1(b^{\alpha+1} - Lb^\alpha)}{p_1 b^{p_2} - p_2 b^{p_1}}, \]
so that
\[ \hat{V}(y) = \begin{cases} \frac{(b^{\alpha+1} - Lb^\alpha)}{p_1 b^{p_2} - p_2 b^{p_1}} (p_1 y^{p_2} - p_2 y^{p_1}), & y \in [1, b], \\ (y - L)y^\alpha, & y \in [b, \infty), \end{cases} \]
and by using (4.1.9), we then know that \( b \in (L, \infty) \) satisfies the following transcendental equation
\[ (\alpha + 1)(p_1 b^{p_2} - p_2 b^{p_1}) - \alpha L (p_1 b^{p_2-1} - p_2 b^{p_1-1}) - p_1 p_2 (b - L) (b^{p_2-1} - b^{p_1-1}) = 0. \quad (4.1.17) \]

Thus, we have arrived at the following theorem.

**Theorem 4.1.5.** The value function \( V \) from (4.1.5) is given explicitly by
\[ V(y) = \begin{cases} \frac{(b^{\alpha+1} - Lb^\alpha)}{p_1 b^{p_2} - p_2 b^{p_1}} (p_1 y^{p_2} - p_2 y^{p_1}), & y \in [1, b], \\ (y - L)y^\alpha, & y \in [b, \infty), \end{cases} \quad (4.1.18) \]
that is, \( V = \hat{V} \). The stopping time \( \tau_b \) from (4.1.6) with \( b \) given as the unique solution to (4.1.17) above is optimal for the problem (4.1.5).

**Proof.** (i) To prove the first part of Theorem 4.1.5 is to verify that \( V = \hat{V} \). Notice that this is the same as showing that \( \hat{V} \) is the smallest superharmonic function dominating \( G \) from Theorem 3.2.4.

We begin by showing that \( \hat{V}(y) \geq G(y) \) for all \( y \in [1, \infty) \). Let \( h(y) = \hat{V}(y) - G(y) \), so that from (4.1.9), we have
\[ h'(b) = \hat{V}'(b) - G'(b) = 0. \]

We then wish to show that the stationary point \( b \) is the global minimum of \( h \) in \([1, b]\) so that \( h(y) \geq h(b) = 0 \) and thereby, proving that \( \hat{V}(y) \geq G(y) \) in this domain. For this, we take the second derivative of \( h \) and obtain
\[ h''(y) = \frac{b^\alpha (b - L)p_1 p_2}{p_1 b^{p_2} - p_2 b^{p_1}} \left( (p_2 - 1)y^{p_2-2} - (p_1 - 1)y^{p_1-2} \right) \]
where the first inequality follows from the fact that $p_1 > 1$, $p_2 < 0$, $b > L$ and $r - \frac{\sigma^2}{2} < 0$, indicating the convexity of function $h$.

An appeal to the second derivative test tells us that as $h'(b) = 0$ and $h''(y) > 0$ for every $y \in [1, b]$, the stationary point $b$ is the global minimum point and thus demonstrating $\hat{V}(y) \geq G(y)$ for all $y \in [1, \infty)$

Next we are ready to show that $V(y) \leq \hat{V}(y)$ for all $y \in [1, \infty)$. From (4.1.18), it is fairly obvious that $e^{-\lambda \hat{V}}(y)$ is $C^{1,2}$ on $\mathcal{C}$ and $\mathcal{D}$, where

$$\mathcal{C} = \{(t, y) \in [0, \infty) \times [1, \infty) : y < b\},$$

$$\mathcal{D} = \{(t, y) \in [0, \infty) \times [1, \infty) : y > b\},$$

and therefore, we exploit the local time-space formula to obtain

$$e^{-\lambda \hat{V}}(Y_t) = \hat{V}(y) + \int_0^t e^{-\lambda u} \left( L_Y \hat{V} - \lambda \hat{V} \right)(Y_u)du$$

$$+ \int_0^t e^{-\lambda u} \hat{V}_y(Y_u)dR_u - \int_0^t e^{-\lambda u} \sigma Y_u \hat{V}_y(Y_u)d\bar{W}_u$$

$$= \hat{V}(y) + \int_0^t e^{-\lambda u} \left( L_Y \hat{V} - \lambda \hat{V} \right)(Y_u)du$$

$$- \int_0^t e^{-\lambda u} \sigma Y_u \hat{V}_y(Y_u)d\bar{W}_u, \quad (4.1.19)$$

where the first equality follows via the smooth-fit condition (4.1.9) and the second equality holds as $V$ satisfies the normal reflection condition (4.1.10).

Since $\hat{V}(y) = G(y) = (y - L)g^\alpha$ in $\mathcal{D}$, we have

$$\left( L_Y G - \lambda G \right)(y) = -(\lambda + r)g^{\alpha + 1} + L(\lambda + 2r - \sigma^2)g^\alpha$$

$$= (\lambda + r)(L - y)g^\alpha + (r - \sigma^2)Lg^\alpha < 0,$$

where we have used the fact that $y > b > L$ to conclude that the first term is negative, and $r - \frac{\sigma^2}{2} < 0$ for the second one. This, together with (4.1.17), shows that

$$L_Y \hat{V} - \lambda \hat{V} \leq 0, \quad (4.1.20)$$

everywhere on $[1, \infty)$ but $b$. As $\tilde{P}_y(Y_u = b) = 0$, we thus have

$$e^{-\lambda \tilde{M}}(Y_t) \leq e^{-\lambda \hat{V}}(Y_t) \leq \hat{V}(y) + \tilde{M}_t, \quad (4.1.21)$$

where the first inequality follows from the first observation that $\hat{V} \geq G$, the second inequality holds via (4.1.19) and (4.1.20), and moreover, $\tilde{M}_t = -\int_0^t e^{-\lambda u} \sigma Y_u \hat{V}_y(Y_u)d\bar{W}_u$ is a continuous local martingale.

Remember that a stopped martingale does not always remain a martingale, but for bounded stopping times, it always preserves its martingale property. Therefore, let $\tau_n = \tau \wedge n$ be a bounded stopping time, for $n \geq 0$ so that by Corollary 2.3.13, $\tilde{E}_y(\tilde{M}_{\tau_n}) = \tilde{E}_y(\tilde{M}_{\tau_n}) = 0$. 

$$-\frac{y^{\alpha-2}}{\sigma^4} \left( (2r - \sigma^2) \left( 2ry + L(2\sigma^2 - 2r) \right) \right) > 0,$$
Then, taking the expectation under measure $\tilde{P}_y$ gives us
$$\tilde{E}_y \left( e^{-\lambda n \tau} G(Y_{\tau_n}) \right) \leq \hat{V}(y).$$

Now, let $n \to \infty$, so that $\tau_n \to \tau$. We invoke Fatou’s lemma to obtain
$$\tilde{E}_y \left( e^{-\lambda \tau} G(Y_{\tau}) \right) \leq \liminf_{n \to \infty} \tilde{E}_y \left( e^{-\lambda n \tau} G(Y_{\tau_n}) \right) \leq \hat{V}(y),$$
after which, we take the supremum over all stopping times $\tau$ of $Y$, together with (4.1.5), the desired assertion, that $V(y) \leq \hat{V}(y)$, suggests itself for all $y \in [1, \infty)$.

To finish off, we have to show that $V(y) = \hat{V}(y)$ for all $y \in [1, \infty)$. Let $\tau_n = \tau_b \wedge n$ for $n \geq 0$ and $\tau_b$ be the finite stopping time defined in (4.1.6). Then, set $t = \tau_n$ in (4.1.19) so that
$$e^{-\lambda \tau_n} \hat{V}(Y_{\tau_n}) = \hat{V}(y) + \tilde{M}_{\tau_n},$$
of which, taking the expectation under $\tilde{P}_y$, upon using the same argument as before yields $\tilde{E}_y \left( \tilde{M}_{\tau_n} \right) = \tilde{E}_y \left( \tilde{M}_n \right) = 0$ and letting $n \to \infty$,
$$\tilde{E}_y \left( e^{-\lambda \tau_b} \hat{V}(Y_{\tau_b}) \right) = \hat{V}(y).$$
Furthermore, we conclude that, in the view of the fact that $\hat{V}(Y_{\tau_b}) = G(Y_{\tau_b})$ and (4.1.5),
$$\hat{V}(y) = \tilde{E}_y \left( e^{-\lambda \tau_b} G(Y_{\tau_b}) \right) \leq \sup_{\tau} \tilde{E}_y \left( e^{-\lambda \tau} G(Y_{\tau}) \right) = V(y),$$
for all $y \in [1, \infty)$, which, joining with the fact that $\hat{V}(y) \geq V(y)$, proving that the equality holds true. With $\tau_b$ being finite, Theorem 3.2.4 then provides us with the positive answer for $\tau_b$ being optimal.

(ii) It thus remains to show the second part of the Theorem 4.1.5, that is $b$ is unique, i.e. the equation (4.1.17) has only one root. Let
$$g(y) = (\alpha + 1) (p_1 y^{p_2} - p_2 y^{p_1}) - \alpha L \left( p_1 y^{p_2-1} - p_2 y^{p_1-1} \right) - p_1 p_2 (y - L) \left( y^{p_2-1} - y^{p_1-1} \right),$$
for $y \in (\max\{1, L\}, \infty)$. Then, upon using $p_1 p_2 = \frac{-2\lambda}{\sigma^2}$, we have
$$g'(y) = \frac{2\lambda}{\sigma^2 y^2} (L - y) \left( (p_1 - 1) y^{p_1} + (1 - p_2) y^{p_2} \right) + \frac{\alpha}{y^2} \left( \frac{2\lambda}{\sigma^2} (y - L) (y^{p_1} - y^{p_2}) + L (p_1 y^{p_2} - p_2 y^{p_1}) \right) < 0,$$
where the strict inequality holds via $p_1 > 1, p_2 < 0$ and $\alpha < 0$, which implies that the map $y \mapsto g(y)$ is decreasing. To apply the intermediate value theorem, we need to further estimate the value of the endpoints:
$$g(1) = \left( \frac{2r}{\sigma^2} + L \left( 1 - \frac{2r}{\sigma^2} \right) \right) (p_1 - p_2) > 0, \quad g(L) = p_1 L^{p_2} - p_2 L^{p_1} > 0,$$
where the first strict inequality follows by $\frac{2r}{\sigma^2} - 1 < 0$.

Let $y \to \infty$, then as $p_2 < 0$, we have

$$\lim_{y \to \infty} g(y) = \lim_{y \to \infty} \frac{2y^{p_1}}{\sigma^2} (-rp_2 - \lambda) + \lim_{y \to \infty} Ly^{p_1-1} \left( \left( \frac{2r}{\sigma^2} - 1 \right) p_2 + \frac{2\lambda}{\sigma^2} \right). \tag{4.1.22}$$

With a little more effort, we could determine the sign of $-rp_2 - \lambda$,

$$-rp_2 - \lambda = -\sqrt{\frac{r^2 \left( r + \frac{\sigma^2}{2} \right)^2 + (\lambda \sigma^2)^2 + 2r^2 \lambda \sigma^2 + r^2 \sigma^4}{\sigma^2} + \sqrt{\frac{r^2 \left( r + \frac{\sigma^2}{2} \right)^2 + 2r^2 \lambda \sigma^2}{\sigma^2}} < 0,$$

from which, together with $p_1 > p_1 - 1$, we know that the first term of (4.1.22) heads towards $-\infty$ more rapidly than the second term heading towards $+\infty$ and thus, $\lim_{y \to \infty} g(y) \to -\infty$. Finally, we appeal to the intermediate value theorem, together with the fact that $y \mapsto g(y)$ is decreasing, establishing the uniqueness of $b$. \hfill \Box

**Figure 4.1:** This figure displays the maps $y \mapsto V(y)$ and $y \mapsto G(y)$ with chosen parameters $L = 4$, $r = 0.02$, $\sigma = 0.3$, $\lambda = 0.5$ and the optimal stopping point $b = 5.0845$. 
4.2 Finite-time Horizon

4.2.1 Reformulation and Basics

In this section, we consider problem (4.0.1) in the finite time horizon and by the same arguments as that in Section 4.1, it can be reformulated as

\[ V = \sup_{0 \leq \tau \leq T} E_{0,x,s}(e^{-(r+\lambda)\tau}(S_\tau - LX_\tau) + P(\theta > \tau | F_\tau)). \]  

Once again, some general facts of the Azéma supermartingale associated with random time \( \theta \) to start the ball rolling. For notational convenience, we set

\[ Z(t, y) = \Phi \left( \frac{-\log y + \left( r - \frac{\sigma^2}{2} \right)(T - t)}{\sigma \sqrt{T - t}} \right) + y^\alpha \Phi \left( \frac{-\log y - \left( r - \frac{\sigma^2}{2} \right)(T - t)}{\sigma \sqrt{T - t}} \right), \]

where \( \Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-z^2/2} dz \) and \( \alpha = \frac{2r}{\sigma^2} - 1 < 0 \).

**Proposition 4.2.1.** Let \( P(\theta > t|F_t) \) be the Azéma supermartingale of random time \( \theta \). Then \( P(\theta > t|F_t) = Z(t, Y_t) \), \quad (4.2.2)

where \( Y \) is the process introduced in Corollary 4.1.3.

**Proof.** From the set equality

\[ \{ \theta > t \} = \left\{ \max_{t \leq u \leq T} X_u > S_t \right\}, \]

we deduce that

\[ P(\theta > t|F_t) = P \left( \max_{t \leq u \leq T} X_u > S_t | F_t \right) \]

\[ = P \left( \frac{\max_{t \leq u \leq T} X_u}{X_t} > \frac{S_t}{X_t} | F_t \right) \]

\[ = P \left( \log \frac{S_t}{X_t} < \max_{0 \leq u \leq T-t} \left( r - \frac{\sigma^2}{2} \right) u + \sigma W_u \right) \]

\[ = P \left( \log Y_t < \max_{0 \leq u \leq T-t} \left( r - \frac{\sigma^2}{2} \right) u + \sigma W_u \right), \quad (4.2.3) \]

and by applying Lemma 2.2.11 to equation (4.2.3), the conclusion is immediate. \( \square \)

**Remark 4.2.2.** As the reader may have discovered, we have not yet defined \( \theta \) properly if the process \( X \) does not exceed level \( s \) at all on \([0, T]\), given that \( x < s \). However, it should now be clear that in such case, by Proposition 4.2.1, \( P(\theta > t|F_t) = 0 \) so that there is nothing to prove and thus the definition shall not concern us. Another more convenient way for this is to simply define the \( \sup \{ \emptyset \} = 0 \), see [55, Page 296], after which, the same result follows.

\[ ^5 \text{Not to confuse this } y \text{ with } Y_0 = y. \]
Proposition 4.2.3. The function $Z(t, y)$ satisfies the PDE for $(t, x) \in [0, T] \times (1, \infty)$.

$$
\frac{\partial Z(t, y)}{\partial t} + (\sigma^2 - r)y \frac{\partial Z(t, y)}{\partial y} + \frac{1}{2} \sigma^2 y^2 \frac{\partial^2 Z(t, y)}{\partial y^2} = 0. \quad (4.2.4)
$$

Proof. To prove (4.2.4), we first apply Itô’s formula and obtain

$$
dZ(t, Y_t) = \left( \frac{\partial Z}{\partial t} + (\sigma^2 - r)y \frac{\partial Z}{\partial y} + \frac{1}{2} \sigma^2 y^2 \frac{\partial^2 Z}{\partial y^2} \right) (t, Y_t) dt + \frac{\partial Z}{\partial y} (t, Y_t) dR_t - \sigma Y_t \frac{\partial Z}{\partial y} (t, Y_t) dW_t. \quad (4.2.5)
$$

Then, recall the Doob-Meyer decomposition of the Azéma supermartingale in Example 2.3.4, that is,

$$
Z(t, Y_t) = Z(0, Y_0) + M_t^\theta - A_t^\theta, \quad (4.2.6)
$$
made more, the measure $dA_t^\theta$ is carried by the set $\{t : X_t = S_t\}$.

The conclusion now follows from the uniqueness of the Doob-Meyer decomposition, upon comparing (4.2.5) and (4.2.6), more precisely, we have

$$
M_t^\theta = -\int_0^t Y_u \frac{\partial Z}{\partial y} (u, Y_u) dW_u,
$$

$$
A_t^\theta = -\int_0^{t \wedge T} \frac{\partial Z}{\partial y} (u, Y_u) dR_u.
$$

The following result is simple yet useful, which we state for easy reference.

Corollary 4.2.4. Both maps $t \mapsto Z(t, y)$ and $y \mapsto Z(t, y)$ are decreasing.

Proof. Immediate from Appendix A.1.1. \hfill \Box

For the sake of brevity, we set the gain function

$$
G(t, y) = (y - L)^+ Z(t, y).
$$

We are now in the position to reduce the original problem (4.2.1) to one-dimension\(^6\) by using the measure $\bar{P}$ and generalise it by the strong Markov property of the process $Y$, that is

$$
V(t, y) = \sup_{0 \leq \tau \leq T - t} E_{t,x,s} \left( e^{-(r + \lambda)\tau} X_{t+\tau} \left( \frac{S_{t+\tau}}{X_{t+\tau}} - L \right)^+ P(\theta > t + \tau | \mathcal{F}_{t+\tau}) \right)
$$

$$
= \sup_{0 \leq \tau \leq T - t} \bar{E}_{t,y} \left( e^{-\lambda \tau} (Y_{t+\tau} - L)^+ Z(t, \tau, Y_{t+\tau}) \right)
$$

$$
= \sup_{0 \leq \tau \leq T - t} \bar{E}_{t,y} \left( e^{-\lambda \tau} G(t, \tau, Y_{t+\tau}) \right), \quad (4.2.7)
$$

where $\tau$ is a stopping time of the diffusion process $Y$ with $Y_t = y$ under $\bar{P}_{t,y}$ for $(t, y) \in [0, T] \times [1, \infty)$ given and fixed.

\(^6\)From $(S_t/X_t)_{t \geq 0}$ to $(Y_t)_{t \geq 0}$, process-wise.
4.2 Finite-time Horizon

4.2.2 The Free-boundary Problem

In order to formulate and describe the free-boundary problem, it is convenient to introduce the following sets.

Definition 4.2.5. The continuation set and the stopping set are defined respectively as

\[ C = \{(t, y) \in [0, T) \times [1, \infty) : V(t, y) > G(t, y)\}, \]

\[ D = \{(t, y) \in [0, T) \times [1, \infty) : V(t, y) = G(t, y)\} \cup \{(T) \times [1, \infty)\}. \]

and we define the first entrance time of the stopping set \( D \), denoted as \( \tau_D \), as follows

\[ \tau_D = \inf\{s \geq 0 : Y_{t+s}^y \in D\} \wedge (T - t). \]

Remark 4.2.6. In finite-time case, we have \( \tau_D \leq T - t < \infty \) a.s, thus the condition \( \tilde{P}(\tau_D < \infty) = 1 \) is automatically satisfied for all \( y \in [1, \infty) \).

Having defined stopping time \( \tau_D \), another natural challenge is to determine the optimality of \( \tau_D \). To do so, we begin by showing that the value function \( V \) is continuous.

Lemma 4.2.7. The value function \( V \) is continuous on \([0, T) \times [1, \infty)\).

Proof. We first show that (i) \( t \mapsto V(t, y) \) is continuous on \([0, T)\) for each \( y \geq 1 \) given and fixed. Take any \( t_1 < t_2 \) in \([0, T)\), \( \epsilon > 0 \) and let \( \tau_1^y \) be a stopping time such that \( \tilde{P}_{t_1, y} (\tau_1^y = T - t_1) = 1 \) and that

\[ E_{t_1, y} \left( e^{-\lambda \tau_1^y} (Y_{t_1+\tau_1^y}^y - L) + Z \left( t_1 + \tau_1^y, Y_{t_1+\tau_1^y}^y \right) \right) \geq V(t_1, y) - \epsilon, \]

and set \( \tau_2^y = \tau_1^y \wedge (T - t_2) \), we see that

\[ E_{t_2, y} \left( e^{-\lambda \tau_2^y} (Y_{t_2+\tau_2^y}^y - L) + Z \left( t_2 + \tau_2^y, Y_{t_2+\tau_2^y}^y \right) \right) \leq V(t_2, y). \]

Noting that \( t \mapsto V(t, y) \) is decreasing (see A.3) and the time-homogeneous property of process \( Y \), we then obtain

\[ 0 \leq V(t_1, y) - V(t_2, y) \]

\[ \leq \tilde{E} \left( e^{-\lambda \tau_1^y} (Y_{t_1+\tau_1^y}^y - L) + Z \left( t_1 + \tau_1^y, Y_{t_1+\tau_1^y}^y \right) - e^{-\lambda \tau_2^y} (Y_{t_2+\tau_2^y}^y - L) + Z \left( t_2 + \tau_2^y, Y_{t_2+\tau_2^y}^y \right) \right) + \epsilon \]

\[ \leq \tilde{E} \left( e^{-\lambda \tau_2^y} \left( (Y_{t_1+\tau_1^y}^y - L) + Z \left( t_1 + \tau_1^y, Y_{t_1+\tau_1^y}^y \right) - (Y_{t_2+\tau_2^y}^y - L) + Z \left( t_2 + \tau_2^y, Y_{t_2+\tau_2^y}^y \right) \right) \right) + \epsilon \]

\[ \leq \tilde{E} \left( (Y_{t_1+\tau_1^y}^y - L) + Z \left( t_1 + \tau_1^y, Y_{t_1+\tau_1^y}^y \right) - (Y_{t_2+\tau_2^y}^y - L) + Z \left( t_2 + \tau_2^y, Y_{t_2+\tau_2^y}^y \right) \right) + \epsilon \]

\[ \leq \tilde{E} \left( (Y_{t_1+\tau_1^y}^y - Y_{t_2+\tau_2^y}^y) + Z \left( t_1 + \tau_1^y, Y_{t_1+\tau_1^y}^y \right) + (Y_{t_2+\tau_2^y}^y - L) + Z \left( t_1 + \tau_1^y, Y_{t_1+\tau_1^y}^y \right) - Z \left( t_2 + \tau_2^y, Y_{t_2+\tau_2^y}^y \right) \right) + \epsilon \]

\[ \leq \tilde{E} \left( (Y_{t_1+\tau_1^y}^y - Y_{t_2+\tau_2^y}^y) + Z \left( t_1 + \tau_1^y, Y_{t_1+\tau_1^y}^y \right) + (Y_{t_2+\tau_2^y}^y - L) + Z \left( t_1 + \tau_1^y, Y_{t_1+\tau_1^y}^y \right) - Z \left( t_2 + \tau_2^y, Y_{t_2+\tau_2^y}^y \right) \right) + \epsilon, \]
where the last inequality follows from
\[(x - z)^+ - (y - z)^+ \leq (x - y)^+ \quad \text{for } x, y, z \in \mathbb{R}. \tag{4.2.13}\]
Hence, by letting \(t_1 \to t_2\) and \(\epsilon \to 0\), \(\tau_1^\epsilon \to \tau_2^\epsilon\), Theorem 2.1.9 (see A.4 for the uniform integrability of \(Y\)) yields
\[V(t_1, y) - V(t_2, y) \to 0\]
which finishes the first part of the proof.

We then show that (ii) \(y \mapsto V(t, y)\) is continuous for all \(t \in [0, T)\). For this, we note that for all \(1 \leq y_1 < y_2 < \infty\) and \(t \in [0, T)\),
\[
0 \leq |V(t, y_2) - V(t, y_1)| \\
\leq \sup_{0 \leq \tau \leq T-t} \left| \tilde{E} \left( \left( \frac{y_2 \lor S_t}{X_\tau} - L \right)^+ Z \left( t + \tau, \frac{y_2 \lor S_t}{X_\tau} \right) - \left( \frac{y_1 \lor S_t}{X_\tau} - L \right)^+ Z \left( t + \tau, \frac{y_1 \lor S_t}{X_\tau} \right) \right) \right| \\
\leq \sup_{0 \leq \tau \leq T-t} \left| \tilde{E} \left( \left( \frac{y_2 \lor S_t}{X_\tau} - L \right)^+ Z \left( t + \tau, \frac{y_2 \lor S_t}{X_\tau} \right) - \left( \frac{y_1 \lor S_t}{X_\tau} - L \right)^+ Z \left( t + \tau, \frac{y_1 \lor S_t}{X_\tau} \right) \right) \right| \\
\leq \sup_{0 \leq \tau \leq T-t} \tilde{E} \left( \left( \frac{y_2 - S_t}{X_\tau} + S_t \left( \frac{y_1 \lor S_t}{X_\tau} - L \right)^+ Z \left( t + \tau, \frac{y_2 \lor S_t}{X_\tau} \right) - Z \left( t + \tau, \frac{y_1 \lor S_t}{X_\tau} \right) \right) \right| \\
\leq (y_2 - y_1) \sup_{0 \leq \tau \leq T-t} \tilde{E} \left( \frac{1}{X_\tau} \right) \\
\leq (y_2 - y_1) \sup_{0 \leq \tau \leq T-t} \tilde{E} \left( \frac{1}{X_\tau} \right)
\]
where the second inequality holds by
\[|\sup f - \sup g| \leq \sup |f - g|,\]
and the fourth inequality follows via the triangle inequality, the fifth inequality is immediate from (4.2.13) and the seventh one follows from
\[ \frac{1}{X_t} = e^{-\left(r + \frac{\sigma^2}{2}\right)t - \sigma \tilde{W}_t}, \]
where the last term is the martingale under probability measure \( \tilde{P} \), see A.5; and the ninth inequality is due to the mean value theorem for \( y_3 \in [y_1, y_2] \) so that \( \frac{y_1 \vee S_\tau}{X_\tau} > L \) implies \( \frac{y_3 \vee S_\tau}{X_\tau} > L \) a.s., while the last two steps are due to the fact that
\[ \frac{y_2 \vee S_\tau}{X_\tau} - \frac{y_1 \vee S_\tau}{X_\tau} = \left(\frac{y_2 - S_\tau}{X_\tau}\right)^+ + \frac{S_\tau - (y_1 - S_\tau)^+}{X_\tau} \leq \frac{(y_2 - y_1)^+}{X_\tau}, \]
and (4.1.3); in addition, note that function \( Z_y(t, y) \) is bounded by \( C_1 \) for \( y > L \) (see A.1.1) and the boundedness of the second term in the last inequality can be verified via the same fashion as in A.4.

By letting \( y_1 \to y_2 \),
\[ V(t, y_2) - V(t, y_1) \to 0, \]
after which statement (ii) follows. \( \square \)

**Lemma 4.2.8.** The stopping time \( \tau_D \) is optimal.

**Proof.** By Lemma 4.2.7 and Theorem 3.2.8, the stopping time defined as
\[ \bar{\tau}_D = \inf\{ s \in [0, T - t] : (t + s, Y_{t+s}^y) \in \bar{D} \} \]
is optimal (with the convention that the infimum of the empty set is infinite \(^7\)) for the following problem
\[ \bar{V}(t, y) = \sup_{\tau \in [0, T - t]} \tilde{E}_{t,y} \left( e^{-\lambda \tau} G(t + \tau, Y_{t+\tau}) \right), \]
so that \( \bar{V}(t, y) = \bar{E}_{t,y} \left( e^{-\lambda \bar{\tau}_D} G(t + \bar{\tau}_D, Y_{t+\bar{\tau}_D}) \right) \), where its corresponding stopping set is defined as \( \bar{D} = D \setminus \{(T, y) : y \in [1, \infty)\} \). Furthermore, observe that
\[
V(t, y) = \sup_{\tau \in [0, T - t]} \tilde{E}_{t,y} \left( e^{-\lambda \tau} G(t + \tau, Y_{t+\tau}) \right) \\
= \sup_{\tau \in [0, T - t]} \tilde{E}_{t,y} \left( e^{-\lambda \tau} G(t + \tau, Y_{t+\tau}) I\{\tau < T - t\} + e^{-\lambda (T - t)} G(T, Y_T) I\{\tau = T - t\} \right) \\
= \max \left\{ \sup_{\tau \in [0, T - t]} \tilde{E}_{t,y} \left( e^{-\lambda \tau} G(t + \tau, Y_{t+\tau}) \right), \tilde{E}_{t,y} \left( e^{-\lambda (T - t)} G(T, Y_T) \right) \right\} \\
= \max \left\{ \tilde{E}_{t,y} \left( e^{-\lambda \bar{\tau}_D} G(t + \bar{\tau}_D, Y_{t+\bar{\tau}_D}) \right), \tilde{E}_{t,y} \left( e^{-\lambda (T - t)} G(T, Y_T) \right) \right\},
\]

\(^7\)In Chapter 2, we say that if \( \tau = \infty \), \( G(\infty, X_\infty) = 0 \).
such that,

\[ V(t, y) = \begin{cases} 
\tilde{E}_{t,y}(e^{-\lambda \bar{\tau} G(t, Y_{t+\bar{\tau}})}), & \bar{\tau}_D < T - t, \\
\tilde{E}_{t,y}(e^{-\lambda(T-t)} G(T, Y_T)), & \bar{\tau}_D > T - t,
\end{cases} \]

which implies that \( \tau_D = \bar{\tau}_D \wedge (T-t) \), and the desired claim follows.

It will be shown in the next section that the continuation region and the stopping region could also be defined as

\[ C = \{(t, y) \in [0, T) \times [1, \infty) : y < b(t)\}, \] \hspace{1cm} (4.2.15)
\[ D = \{(t, y) \in [0, T) \times [1, \infty) : y \geq b(t)\} \cup \{(T, y) : y \geq b(T)\}, \] \hspace{1cm} (4.2.16)

where \( b : [0, T] \rightarrow \mathbb{R} \) is the unknown optimal stopping boundary, it then follows that \( \tau_D \) can be rewritten as

\[ \tau_D = \inf\{0 \leq s \leq T - t : Y_{t+s}^y \geq b(t + s)\} \wedge (T-t). \]

We are now ready to formulate the following free-boundary problem.

After establishing the optimality of \( \tau_D \), Theorem 3.2.6 and Lemma 3.3.2 suggest that the unknown value function \( V \) from (4.2.7), together with the unknown boundary \( b \), solve

\[ V_t + \mathbb{L}_Y V = \lambda V \] \hspace{1cm} in \( C \), \hspace{1cm} (4.2.17)
\[ V(t, y) = G(t, y) \] \hspace{1cm} for \( y = b(t) \) or \( t = T \), \hspace{1cm} (4.2.18)
\[ V_y(t, y) = G_y(t, y) \] \hspace{1cm} for \( y = b(t) \) (smooth fit), \hspace{1cm} (4.2.19)
\[ V_y(t, 1+) = 0 \] \hspace{1cm} (normal reflection), \hspace{1cm} (4.2.20)
\[ V(t, y) > G(t, y) \] \hspace{1cm} in \( C \), \hspace{1cm} (4.2.21)
\[ V(t, y) = G(t, y) \] \hspace{1cm} in \( D \). \hspace{1cm} (4.2.22)

Proofs of conditions (4.2.19)-(4.2.20) will be given in section 4.2.4.

### 4.2.3 The Continuation and Stopping Sets

We first need to make some preparations in order to prove (4.2.15) and (4.2.16).

Since the gain function \( G(t, y) \) is a continuous function whose derivative in \( y = L \) is not continuous, we apply Theorem 2.2.6 on \( e^{-\lambda s} G(t + s, Y_{t+s}) \) and obtain

\[
e^{-\lambda s} G(t + s, Y_{t+s}) = G(t, y) + \int_0^s e^{-\lambda u} H(t + u, Y_{t+u}) I\{Y_{t+u} > L\} du
+ \int_0^s e^{-\lambda u} G_y(t + u, Y_{t+u}) I\{Y_{t+u} = 1\} dR_{t+u} + M_s
+ \frac{1}{2} \int_0^s e^{-\lambda u} (G_y(t + u, L+) - G_y(t + u, L-)) dt Y_u, \] \hspace{1cm} (4.2.23)
where we set
\[ H(t, y) = \left( -\lambda G + \frac{\partial G}{\partial t} - ry \frac{\partial G}{\partial y} + \frac{1}{2} \sigma^2 y^2 \frac{\partial^2 G}{\partial y^2} \right)(t, y), \] (4.2.24)
and \( M_s = -\int_0^s \sigma Y_{t+s} G_y(t + s, Y_{t+s}) I\{Y_{t+s} \neq L\} d\bar{W}_s \) is a martingale.

To simplify (4.2.24), we see that
\[ \frac{\partial G}{\partial t}(t, y) = (y - L) \frac{\partial Z(t, y)}{\partial t}, \]
\[ \frac{\partial G}{\partial y}(t, y) = Z(t, y) + (y - L) \frac{\partial Z(t, y)}{\partial y}, \]
\[ \frac{\partial^2 G}{\partial y^2}(t, y) = 2 \frac{\partial Z(t, y)}{\partial y} + (y - L) \frac{\partial^2 Z(t, y)}{\partial y^2}, \]
for all \( y > L. \)

By combining the above computation and using the fact that \( Z(t, y) \) satisfying (4.2.4), we therefore have (4.2.24) rewritten as follows
\[ H(t, y) = \left( -\lambda(y - L)Z - ryZ + \sigma^2 Ly \frac{\partial Z}{\partial y} \right)(t, y). \]

Then, by taking the expectation under \( \bar{P}_{t,y} \) and applying the optional sampling theorem to get rid of the martingale part, we find that
\[ \bar{E}_{t,y} \left( e^{-\lambda s} G(t + s, Y_{t+s}) \right) = G(t, y) + \bar{E}_{t,y} \int_0^s e^{-\lambda u} H(t + u, Y_{t+u}) I\{Y_{t+u} > L\} du \]
\[ + \bar{E}_{t,y} \int_0^s e^{-\lambda u} G_y(t + u, 1+) I\{Y_{t+u} \neq L\} dR_{t+u} \]
\[ + \frac{1}{2} \bar{E}_{t,y} \int_0^s e^{-\lambda u} (G_y(t + u, L+) - G_y(t + u, L-)) dI_u^L(Y). \] (4.2.25)

We first investigate the map \( t \mapsto H(t, y) \) by a direct differentiation in \( t \), which yields
\[ \frac{\partial H}{\partial t}(t, y) = -(ry + \lambda y - \lambda L) \frac{\partial Z}{\partial t} + \sigma^2 Ly \frac{\partial^2 Z}{\partial y \partial t} \]
\[ = \frac{y^\alpha e^{-\frac{d^2}{2}}}{\sqrt{8\pi(T - t)^{\frac{3}{2}}}} \left( \frac{-2(ry + \lambda y - \lambda L) \log y + 2\sigma^2 L - \frac{2(\log y)^2 L}{T - t} + (2r - \sigma^2) L \log y}{\sigma} \right) \]
\[ < \frac{y^\alpha e^{-\frac{d^2}{2} L}}{\sqrt{8\pi(T - t)^{\frac{3}{2}}}} \left( \frac{-2r \log L + 2\sigma^2 - \frac{2(\log L)^2 L}{T} + (2r - \sigma^2) \log L}{\sigma} \right), \] (4.2.26)
from which one sees that there exists a wide range of parameters for \( t \mapsto H(t, y) \) to be decreasing.

**Remark 4.2.9.** The map \( t \mapsto H(t, y) \) being decreasing is rather crucial for showing the monotonicity of \( b \), the continuity of \( b \) and the smooth-fit condition.
For the purpose of proving the properties mentioned in Remark 4.2.9, from now on, we assume:

**Assumption 4.2.10.** Parameters \( r \) and \( \sigma \) can be chosen so that the map \( t \mapsto H(t, x) \) is decreasing.

**Lemma 4.2.11.** The continuation set \( C \) is left-connected and non-increasing with respect to time \( t \).

**Proof.** We shall show that for fixed \( 0 \leq t_1 < t_2 < T \) and \( y \in (L, \infty) \), \( (t_2, y) \in C \) implies \( (t_1, y) \in C \).

We begin by recalling that

\[
G_y(t, y) = \begin{cases} 0, & \text{for } y \leq L, \\ Z(t, y) + (y - L) \frac{\partial Z(t, y)}{\partial y}, & \text{for } y > L, \end{cases}
\]

and for \( L > 1 \), it follows that

\[
\begin{align*}
G_y(t, 1+) &= 0, \quad (4.2.27) \\
G_y(t, L+) &= Z(t, L+), \quad (4.2.28) \\
G_y(t, L-) &= 0. \quad (4.2.29)
\end{align*}
\]

Let \( \tau \) be the optimal stopping time for \( V(t_2, y) \) so that by (4.2.25) and (4.2.27)-(4.2.29),

\[
\begin{align*}
V(t_1, y) - V(t_2, y) &\geq \mathbb{E} \left( e^{-rT} G(t_1 + \tau, Y^y_{t_1+\tau}) \right) - \mathbb{E} \left( e^{-rT} G(t_2 + \tau, Y^y_{t_2+\tau}) \right) \\
&= \mathbb{E} \left( e^{-rT} G(t_1 + \tau, Y^y_{\tau}) \right) - \mathbb{E} \left( e^{-rT} G(t_2 + \tau, Y^y_{\tau}) \right) \\
&= G(t_1, y) - G(t_2, y) + \mathbb{E} \left( \int_0^\tau e^{-ru} \left( H(t_1 + u, Y^y_u) - H(t_2 + u, Y^y_u) \right) I\{Y^y_u > L\} du \right) \\
&\quad + \frac{1}{2} \mathbb{E} \left( \int_0^\tau e^{-ru} \left( G_y(t_1 + u, L+) - G_y(t_2 + u, L+) \right) d\ell_u(Y) \right) \\
&\geq G(t_1, y) - G(t_2, y),
\end{align*}
\]

where the first equality holds as \( Y^y_{t_1+\tau}, Y^y_{t_2+\tau} \) and \( Y^y_{\tau} \) are identically distributed and the last inequality holds by the fact that both \( t \mapsto H(t, y) \) and \( t \mapsto G_y(t, L+) \) are decreasing on \([0, T]\) (recall Assumption 4.2.10, Corollary 4.2.4 and equation (4.2.28)). Hence, we reach the following conclusion

\[
V(t_1, y) - G(t_1, y) \geq V(t_2, y) - G(t_2, y), \quad (4.2.30)
\]

indicating that if a point \( (t_2, y) \in C \), then we have \( V(t_1, y) - G(t_1, y) \geq V(t_2, y) - G(t_2, y) > 0 \), which implies that \( (t_1, y) \in C \), proving the initial claim.

From (4.2.30), the following statement does not come to us as a surprise, that is the map \( t \mapsto V(t, y) - G(t, y) \) is decreasing on \([0, T]\).

**Lemma 4.2.12.** The stopping set \( D \) is right-connected and up-connected.

**Proof.** The right-connectedness of the stopping set \( D \) is another direct consequence of (4.2.30), since

\[
0 = V(t_1, y) - G(t_1, y) \geq V(t_2, y) - G(t_2, y) \geq 0,
\]
that is, \((t_1, y) \in \mathcal{D}\) implies \((t_2, y) \in \mathcal{D}\).

To justify its up-connectedness, we first take \(t > 0\) and \(y_2 > y_1 > L\) such that \((t, y_1) \in \mathcal{D}\). Then, by the right-connectedness of the exercise region, we have \((t + s, y_1) \in \mathcal{D}\) for any \(s \in (0, T - t)\). If we now run the process \((t + s, \tilde{Y}_{t+s}^{y_2})\) from \((t, y_2)\), we cannot hit the level \(L\) to compensate the negative integrand \(H\) in (4.2.25) before exercise as \(y_2 > y_1\), which means that the local time term in (4.2.25) is 0 and the integrand \(H\) is negative. Therefore, it’s optimal to exercise at \((t, y_2)\) and we have established the up-connectedness of the exercise region \(\mathcal{D}\).

Lemma 4.2.13. The map \(t \mapsto b(t)\) is decreasing on \([0, T]\).

Proof. Combine Lemma 4.2.11 and Lemma 4.2.12.

Proposition 4.2.14. All points \(y < L\) for \(0 \leq t < T\) belongs to the continuation set \(\mathcal{C}\).

Proof. The proof follows, in fact, from the verbal statement, by exercising below \(L\), the option holder receive a null payoff, whereas by waiting would have a positive probability of collecting a strictly positive payoff in the future.

A more detailed yet simple proof of this is that, from (4.2.25), we know that for \(Y_t^{y} < L\), all integral terms on the right-hand side are nonnegative.

We also recall the solution to the infinite time horizon optimal stopping problem, where the stopping time \(\tau_0 = \inf\{t \geq 0 : Y_t > b_s\}\) is optimal and \(b_s > L\) is proved to be true, we therefore conclude that all points \((t, y)\) with \(y > b_s\) for \(0 < t < T\) belong to the stopping set \(\mathcal{D}\) and that as \(b_s\) is finite, so is \(b\).

By taking advantage of our findings so far, we can draw the conclusion that the continuation set \(\mathcal{C}\) and the stopping set \(\mathcal{D}\) indeed equal (4.2.15) and (4.2.16) respectively.

We close this section by constructing the continuity of the optimal stopping boundary \(b\).

Proposition 4.2.15. The optimal stopping boundary \(b\) is continuous on \([0, T]\) and \(b(T-) = L\).

Proof. We first show that (i) \(b\) is right-continuous. Let us fix \(t \in [0, T)\) and take a sequence \(t_n \downarrow t\) as \(n \to \infty\). Since \(b\) is decreasing on \([0, T]\), the right-limit \(b(t+)\) exists. Remember that \(\mathcal{D}\) is closed so that its limit point \(\lim_{n \to \infty} (t_n, b(t_n)) \to (t, b(t+))\) is contained in \(\mathcal{D}\), it then follows, together with (4.2.16), that \(b(t+) \geq b(t)\). However, the fact that \(b\) is decreasing suggests that \(b(t) \geq b(t+)\) and therefore, \(b\) is right-continuous as claimed.

We then show that (ii) \(b\) is left-continuous. Assume that, for contradiction, there exists \(t \in (0, T)\) such that \(b(t-) > b(t)\). Then, fix a point \(y_s \in (b(t), b(t-))\).

By (4.2.18) and (4.2.19), we have

\[
V(s, y_s) - G(s, y_s) = \int_{y_s}^{b(s)} \int_{x}^{b(s)} (V_{yy} - G_{yy})(s, z) dz dx, \tag{4.2.31}
\]

for each \(s \in (t - \delta, t)\) where \(\delta > 0\) and \(t - \delta > 0\). Knowing that the value function satisfies

\[
V_t + \mathbb{L}_Y V = \lambda V, \quad \text{in } \mathcal{C},
\]

we conclude that the optimal stopping boundary \(b\) is indeed continuous on \([0, T]\).
and that
\[ H(t, y) = (-\lambda G + G_t + L Y G) (t, y), \]
we have
\[ \frac{\sigma^2 y^2}{2} (V_{yy} - G_{yy}) (t, y) = (\lambda (V - G) - (V_t - G_t) + ry(V_y - G_y) - H) (t, y). \]

Now recall that \( t \mapsto V(t, y) - G(t, y) \) is decreasing and hence
\[ \frac{\sigma^2 y^2}{2} (V_{yy} - G_{yy}) (t, y) \geq (ry(V_y - G_y) - H) (t, y), \] (4.2.32)
in \( C \cup \{(t, y) \in [0, T] \times [1, \infty) : y = b(t)\} \).

Using (4.2.31) and (4.2.32), we find that
\[ V(s, y_*) - G(s, y_*) \geq \int_{y_*}^{b(s)} \int_x^{b(s)} \left( \frac{2r}{\sigma^2 z^2} (V_z - G_z) - \frac{2}{\sigma^2 z^2} H \right)(s, z) dz dx, \]
and since \( H \) is strictly negative in \( \{(s, y) \in [t - \delta, t] \times [y_*, b(t - \delta)] \} \) (as \( y_* > b(t) \geq L \)), we set
\[ m = \inf \left\{ -\frac{2}{\sigma^2 y^2} H(s, y) : (s, y) \in [t - \delta, t] \times [y_*, b(t - \delta)] \right\} > 0, \]
so that
\[ V(s, y_*) - G(s, y_*) \geq \frac{2r}{\sigma^2} \int_{y_*}^{b(s)} \int_x^{b(s)} \frac{1}{z} (V_z - G_z) (s, z) dz dx + m \int_{y_*}^{b(s)} (b(s) - x) dx. \]

Via an integration by parts, we have
\[
\int_{y_*}^{b(s)} \int_x^{b(s)} \frac{1}{z} d(V - G)(s, z) dx
\]
\[ = \int_{y_*}^{b(s)} \left( \frac{1}{z} (V - G)(s, z) \big|_x^{b(s)} \right) \int_x^{b(s)} \frac{1}{z} (V - G)(s, z) dz \] \( dx \)
\[ = \int_{y_*}^{b(s)} \left( \frac{1}{x} (V - G)(s, x) dx + \int_{y_*}^{b(s)} \int_x^{b(s)} \frac{1}{z} (V - G)(s, z) dz dx, \]
where the second equality follows from (4.2.18).

From Lemma 4.2.7 and the fact that the gain function is continuous on \((1, \infty) \times [0, T]\), we know that
\[ V(t, y) = V(t-, y), \]
\[ G(t, y) = G(t-, y), \]
and from (4.2.22), it follows that
\[ V(t, y) = G(t, y) \text{ for all } y \in [y_*, b(t-)).\]
Then, we observe that

\[ |V - G|(s, y) \leq V(s, y) \leq \sup_{0 \leq \tau \leq T-s} \tilde{E}_{s,y} (Y^{y}_{s+	au}) \leq ye^{\left(r + \frac{\sigma^2}{2}\right) T} \tilde{E} \left(e^{2\sigma \sqrt{T}X}\right) < \infty, \]

where the third inequality holds via (A.4.1). Finally, let \( s \uparrow t \) and an application of dominated convergence theorem shows

\[
V(t, y_\ast) - G(t, y_\ast) \geq \frac{2r}{\sigma^2} \left( \int_{y_\ast}^{b(t)} \frac{1}{x} (V - G)(t, x) dx + \int_{y_\ast}^{b(t)} \int_{x}^{b(t)} \frac{1}{z^2} (V - G)(t, z) dz dx \right) + m \int_{y_\ast}^{b(t)} (b(t) - x) dx = m \int_{y_\ast}^{b(t)} (b(t) - x) dx = \frac{m}{2} (b(t) - y_\ast)^2 > 0,
\]

where the first equality is due to the fact that \( \{t\} \times [y_\ast, b(t)] \subset \mathcal{D} \), which is a contradiction, as \((t, y_\ast)\) belongs to the stopping set \( \mathcal{D} \), indicating that such point cannot exist. Therefore, by combining statements (i) and (ii), we have established the continuity of \( b \) on \([0, T]\).

(iii) It remains to take care of the final piece of the claim, that is, \( b(T\!-) = L \).

According to Proposition 4.2.14, we must have \( b(T\!-) \geq L \). Assume that, \( b(T\!-) > L \) such that a point \( y_\ast \) exists on \( (L, b(T\!-)) \) and let \( s \) be an arbitrarily fixed point in the interval \( (T - \delta, T) \) with \( 0 < \delta < T \). Rerunning the proof of (ii) with the above modifications, upon letting \( s \uparrow T \), we thus have arrived at the contradiction, that is

\[ V(T, y_\ast) - G(T, y_\ast) > 0, \]

but the definition of the stopping set \( \mathcal{D} \) tells us that \( V(T, y_\ast) = G(T, y_\ast) \), which may allow us to conclude that \( b(T\!-) = L \).

4.2.4 The Optimal Stopping Rule

Before we turn to presenting the main result, let us first complete the promise given in section 4.2.2, namely, to verify conditions (4.2.19)-(4.2.20) for the value function.

**Lemma 4.2.16.** (Smooth-fit Condition) The value function \( V(t, y) \) is differentiable at the optimal stopping boundary \( b \) and

\[
\frac{\partial}{\partial y} V(t, y) = \frac{\partial}{\partial y} G(t, y),
\]

whenever \( V(t, y) = G(t, y) \) for \( y = b(t) \).

**Proof.** Let \( t \in [0, T) \) be given and fixed and set \( y = b(t) \). Knowing that \( b(t) > L \), let \( \epsilon > 0 \) so that \( y - \epsilon > L \). Since

\[ V(t, y) = G(t, y), \]
we have
\[
\frac{V(t, y - \epsilon) - V(t, y) - V(t, y - \epsilon)}{\epsilon} \leq \frac{G(t, y) - G(t, y - \epsilon)}{\epsilon},
\]
and taking the limit as \(\epsilon \to 0\) shows that
\[
\lim_{\epsilon \to 0} \frac{V(t, y) - V(t, y - \epsilon)}{\epsilon} = \frac{\partial}{\partial y} G(t, y).
\]

To prove the reverse inequality, let \(\tau\) be the optimal stopping time for \(V(t, y - \epsilon)\), so that
\[
\frac{V(t, y) - V(t, y - \epsilon)}{\epsilon} \geq \frac{1}{\epsilon} \mathbb{E} \left( e^{-\lambda \tau} \left( G \left( t + \tau, Y_{t+\tau}^y \right) - G \left( t + \tau, Y_{t+\tau}^{y-\epsilon} \right) \right) \right)
\]
\[
= \frac{1}{\epsilon} \mathbb{E} \left( e^{-\lambda \tau} \left( (Y_{t+\tau}^y - L)^+ - (Y_{t+\tau}^{y-\epsilon} - L)^+ \right) Z \left( t + \tau, Y_{t+\tau}^y \right) \right)
\]
\[
+ \frac{1}{\epsilon} \mathbb{E} \left( e^{-\lambda \tau} \left( Z \left( t + \tau, Y_{t+\tau}^y \right) - Z \left( t + \tau, Y_{t+\tau}^{y-\epsilon} \right) \right) \left( Y_{t+\tau}^{y-\epsilon} - L \right)^+ \right)
\]
\[
\geq \frac{1}{\epsilon} \mathbb{E} \left( e^{-\lambda \tau} I \{ Y_{t+\tau}^y - L \} I \{ y - \epsilon \geq S_{\tau} \} \frac{1}{X_{\tau}} Z \left( t + \tau, Y_{t+\tau}^y \right) \right)
\]
\[
+ \frac{1}{\epsilon} \mathbb{E} \left( e^{-\lambda \tau} \left( Z \left( t + \tau, Y_{t+\tau}^y \right) - Z \left( t + \tau, Y_{t+\tau}^{y-\epsilon} \right) \right) \left( Y_{t+\tau}^{y-\epsilon} - L \right)^+ \right)
\]
\[
= \mathbb{E} \left( e^{-\lambda \tau} I \{ Y_{t+\tau}^y - L \} I \{ y - \epsilon \geq S_{\tau} \} \frac{1}{X_{\tau}} Z \left( t + \tau, Y_{t+\tau}^y \right) \right)
\]
\[
+ \mathbb{E} \left( e^{-\lambda \tau} \frac{Y_{t+\tau}^y}{X_{\tau}} \left( Y_{t+\tau}^{y-\epsilon} - L \right) I \{ Y_{t+\tau}^{y-\epsilon} - L \} I \{ y - \epsilon \geq S_{\tau} \} \right),
\]
where the second inequality holds via the facts that \(Y_{t+\tau}^y \geq Y_{t+\tau}^{y-\epsilon}\) (\(\geq L\) because of the definition of the optimal stopping time) and that
\[
(Y_{t+\tau}^y - L)^+ - (Y_{t+\tau}^{y-\epsilon} - L)^+ = I \{ Y_{t+\tau}^{y-\epsilon} - L \} \left( Y_{t+\tau}^y - Y_{t+\tau}^{y-\epsilon} \right) + (Y_{t+\tau}^y - L)^+
\]
\[
\geq I \{ Y_{t+\tau}^{y-\epsilon} \} \left( Y_{t+\tau}^y - Y_{t+\tau}^{y-\epsilon} \right),
\]
and the third inequality follows from
\[
Y_{t+\tau}^y - Y_{t+\tau}^{y-\epsilon} = \frac{(y - S_{\tau})^+ + S_{\tau} - (y - \epsilon - S_{\tau})^+}{X_{\tau}} - S_{\tau}
\]
\[
= \frac{I \{ y - \epsilon \geq S_{\tau} \} \epsilon + (y - S_{\tau})^+}{X_{\tau}} \geq \frac{I \{ y - \epsilon \geq S_{\tau} \} \epsilon}{X_{\tau}},
\]
and by mean value theory for \( x \in [y - \epsilon, y] \), the last equality follows. Note that \( I \{ Y_t^{y-\epsilon} \geq L \} = 1 \) implies \( I \{ Y_t^{y+} \geq L \} = 1 \) so that \( Z_t = \tau_e, Y_t^{x+\epsilon} \) is bounded by \( C_1 \).

Finally, let \( \epsilon \to 0 \) so that \( \tau_e \to 0 \) a.s (see A.2.1) and the dominated convergence theorem once again yields \( S_{\tau_e} \to 1, X_{\tau_e} \to y, Y_{t+\tau^-} \to y, Y_{t+\tau^\epsilon} \to y, Y_{t+\tau^-} \to y \). Then, recalling that \( y = b(t) \geq L > 1 \), we obtain

\[
\lim_{\epsilon \to 0} \frac{V(t,y) - V(t,y-\epsilon)}{\epsilon} = \frac{\partial}{\partial y} G(t,y),
\]

which, joining with (4.2.33), proves the claim.

Lemma 4.2.17. The normal reflection condition holds, that is

\[
\frac{\partial}{\partial y} V(t,1+) = 0.
\]

Proof. We begin by noticing that from (4.2.7) and the construction of the process \( Y \)

\[
V(t,y) = \sup_{0 \leq \tau \leq T - t} \tilde{E} \left( e^{-\lambda \tau} G \left( t + \tau, (y - S_\tau)^+ + S_\tau \right) \right),
\]

and that, in fact, there exists \( y_* \in (1, L] \) such that the map \( y \mapsto G(t,y) \) is increasing on \([1, y_*]\) in the sense that \( \frac{\partial}{\partial y} G(t,y) = \left( Z + (y - L) \frac{\partial}{\partial y} Z \right) (t,y) > 0 \).

It then follows, from the view of (4.2.35), that \( y \mapsto V(t,y) \) is increasing on \([1, y_*]\), meaning that \( V_y(t,1+) \geq 0 \) for all \( t \in [0, T] \), and moreover, since the value function \( V \) is \( C^{1,2} \) on the continuation set, we have \( t \mapsto V_y(t,1+) \) is continuous on \([0, T]\).

Assume that, for contradiction, there exists \( t \in [0, T] \) so that \( V_y(t,1+) > 0 \). Let \( \tau_D \) be the optimal stopping time for \( V(t,1+) \) and let \( \tilde{\tau}_D = \tau_D \wedge s \) for \( s \in [0, T - t] \). Then, we apply Itô’s formula and the optional sampling theorem to obtain

\[
\tilde{E}_{t,1} \left( e^{-\lambda \tilde{\tau}_D} V \left( t + \tilde{\tau}_D, Y_{t+\tilde{\tau}_D} \right) \right) = V(t,1) + \tilde{E}_{t,1} \left( \int_0^{\tilde{\tau}_D} e^{-\lambda u} V_y(t + u, Y_{t+u}) dR_{t+u} \right).
\]

Now in order to complete the proof, we need to further establish the martingale property of

\[
\left\{ e^{-\lambda \tilde{\tau}_D} V \left( t + \tilde{\tau}_D, Y_{t+\tilde{\tau}_D} \right), \mathcal{F}_{t+\tilde{\tau}_D} \right\} _{0 \leq \tilde{\tau}_D \leq T - t}.
\]

(4.2.36)

Let us check the martingale property claimed in (4.2.36) for all \( u \leq \tilde{\tau}_D \):}

\[
\tilde{E} \left( e^{-\lambda \tilde{\tau}_D} V \left( t + \tilde{\tau}_D, Y_{t+\tilde{\tau}_D} \right) | \mathcal{F}_{t+u} \right) = \tilde{E} \left( e^{-\lambda \tilde{\tau}_D} \tilde{E}_{t+u} \left( e^{-\lambda (\tau_D - \tilde{\tau}_D)} G(t + \tau_D, Y_{t+\tau_D}) \right) | \mathcal{F}_{t+u} \right) = \tilde{E} \left( e^{-\lambda \tilde{\tau}_D} \tilde{E} \left( e^{-\lambda (\tau_D - \tilde{\tau}_D)} G(t + \tau_D, Y_{t+\tau_D}) | \mathcal{F}_{t+\tau_D} \right) | \mathcal{F}_{t+u} \right).
\]

\footnote{Theorem 3.2.6 says it all, but let us prove it here again.}
\[
\begin{align*}
&= \mathbb{E} \left( \mathbb{E} \left( e^{-\lambda \tau_D} G (t + \tau_D, Y_{t+\tau_D}) \mid \mathcal{F}_{t+\tau_D} \right) \mid \mathcal{F}_{t+u} \right) \\
&= \mathbb{E} \left( e^{-\lambda \tau_D} G (t + \tau_D, Y_{t+\tau_D}) \mid \mathcal{F}_{t+u} \right) \\
&= e^{-ru} \mathbb{E}_{t+u,Y_{t+u}^1} \left( e^{-\tau(t+\tau_D)} G (t + \tau_D, Y_{t+\tau_D}) \right) = e^{-ru} V \left( t + u, Y_{t+u}^1 \right),
\end{align*}
\]

where the first equality follows from the definitions of \( \tau_D \) and the value function \( V \), the second and the fifth equalities are immediate from the strong Markov property of the process \( Y \) whereas the fourth equality holds by Lemma 2.3.12 as \( \mathcal{F}_{t+u} \subseteq \mathcal{F}_{t+\hat{\tau}_D} \), and thereby, proving the martingale property.

With the aid of the above observation, we then obtain
\[
\mathbb{E}_{t,1} \left( e^{-\lambda \hat{\tau}_D} V \left( t + \hat{\tau}_D, Y_{t+\hat{\tau}_D} \right) \right) = V(t, 1),
\]

which implies that
\[
\mathbb{E}_{t,1} \left( \int_0^{\hat{\tau}_D} e^{-\lambda u} V_y(t + u, Y_{t+u}) dR_{t+u} \right) = 0.
\]

In particular, \( V_y(t + u, Y_{t+u}) dR_{t+u} = V_y(t + u, 1+) dR_{t+u} \), which, together with the assumption that \( V_y(t + u, 1+) > 0 \) for all \( u \in [0, \hat{\tau}_D] \), tells us that
\[
\mathbb{E}_{t,1} \left( \int_0^{\hat{\tau}_D} dR_{t+u} \right) = 0.
\]

By the solution of (4.1.4), equation (4.2.37) and the optional sampling theorem, we see that
\[
\mathbb{E}_{t,1} \left( Y_{t+\hat{\tau}_D} \right) - 1 + r \mathbb{E}_{t,1} \int_0^{\hat{\tau}_D} Y_{t+u} du = 0. \quad (4.2.38)
\]

As \( Y_{t+u} > 1 \) for all \( u \in [0, T - t] \), (4.2.38) implies that \( \mathbb{P}_{t,1}(\hat{\tau}_D = 0) = 1 \), which is not possible in the sense that the optimal stopping boundary \( b > 1 \) and \( \hat{\tau}_D > 0 \), proving the desired assertion.

With a bit more effort, we can verify the following result to apply Theorem 2.2.7 freely later.

**Corollary 4.2.18.** Let the function \( F(t, y) = e^{-\lambda t} V(t, y) \). Then

\[
\begin{align*}
F(t, y) &\text{ is } C^{1,2} \text{ on } \mathcal{C} \cup \overline{D}, \quad (4.2.39) \\
F_t + \mathbb{E}_Y F &\text{ is locally bounded } \mathcal{C} \cup \overline{D}, \quad (4.2.40) \\
t &\mapsto F_y(t, b(t) \pm) &\text{ is continuous,} \quad (4.2.41) \\
F_{yy} &\text{ is non-negative and } F_{2} \text{ is continuous on } \mathcal{C} \text{ and } \overline{D}, \quad (4.2.42)
\end{align*}
\]

where \( \overline{D} = D \setminus \{(t, b(t))\} \).
Proof. First of all, (4.2.39) is immediate from the fact that

\[
F(t,y) = \begin{cases} 
  e^{-\lambda t}V(t,y), & \text{for } (t,y) \in \mathcal{C}, \\
  e^{-\lambda t}G(t,y), & \text{for } (t,y) \in \mathcal{D}, 
\end{cases}
\]

and that \(V(t,y)\) is \(C^{1,2}\) in \(\mathcal{C}\), so is \(G(t,y)\) in \(\mathcal{D}\).

From [64, Page 409], we know that to verify (4.2.40) is to show that \(F_t + \mathbb{L}_Y F\) is locally bounded on \(\mathcal{K} \cap (\mathcal{C} \cup \mathcal{D})\) for each compact set \(\mathcal{K}\) in \([0, T] \times [1, \infty)\). In \(\mathcal{C}\), we have \(F_t + \mathbb{L}_Y F = 0\) in the sense that \(V\) satisfies (4.2.17). As in \(\mathcal{D}\), \(F_t + \mathbb{L}_Y F = e^{-\lambda t}H(t,y)\), which is continuous on the compact set \(\mathcal{K} \cap \mathcal{D}\) and hence, the range of \(H\) is bounded as claimed.

Next, (4.2.41) follows via (4.2.19) and Lemma 4.2.15.

Moving down the list, (4.2.42) follows from

\[
F_{yy}(t,y) = \begin{cases} 
  \frac{2e^{-\lambda t}}{\sigma^2 y^2} (\lambda V - V_t + ryV_y)(t,y), & \text{for } (t,y) \in \mathcal{C}, \\
  \frac{2e^{-\lambda t}}{\sigma^2 y^2} (\lambda G - G_t + ryG_y + H)(t,y), & \text{for } (t,y) \in \mathcal{D}, 
\end{cases}
\]

where the first two terms on both regions are nonnegative and since the continuous function \(V\) is \(C^{1,2}\) in \(\mathcal{C}\) because of Lemma 3.3.2 the partial derivatives \(V_t, V_y\) and \(V_{yy}\) exist and are continuous, whereas as \(y = b(t)\),

\[
V_y(t,b(t)) = G_y(t,b(t)) = Z(t,b(t)) + (b(t) - L) \frac{\partial}{\partial y} Z(t,b(t)),
\]

whose continuity is immediate, so are the latter two terms in \(\mathcal{D}\).

Here is, finally, the result we have been waiting for.

**Theorem 4.2.19.** The optimal stopping boundary in problem (4.2.7) can be characterised as the the unique decreasing solution \(b : [0, T] \mapsto \mathbb{R}\) of the following non-linear integral equation

\[
(b(t) - L)Z(t,b(t)) = -\int_0^{T-t} e^{-\lambda u} \mathbb{E}_{t,b(t)} (H(t + u, Y_{t+u})I\{Y_{t+u} \geq b(t + u))} du, \tag{4.2.43}
\]

satisfying \(b(t) > L\) for all \(0 < t < T\). The solution \(b\) also satisfies \(b(T-) = L\). The value function in problem (4.2.7) has the following representation

\[
V(t,y) = -\int_0^{T-t} e^{-\lambda u} \mathbb{E}_{t,y} (H(t + u, Y_{t+u})I\{Y_{t+u} \geq b(t + u))} du, \tag{4.2.44}
\]

for all \((t,y) \in [0, T] \times [1, \infty)\).

**Proof.** We follow, essentially, [64].

(i) We prove that the unknown optimal stopping boundary \(b\) and value function \(V\) indeed satisfy (4.2.43) and (4.2.44) respectively.

First of all, via an application of the local space-time formula on \(e^{-\lambda s}V(t+s,Y_{t+s})\), we obtain

\[
e^{-\lambda s}V(t+s,Y_{t+s}) = V(t,y) + \tilde{M}_s \tag{4.2.45}
\]
Russian Option

\[ + \int_0^s e^{-\lambda u} (V_t + \mathbb{L}YV - \lambda V) (t + u, Y_{t+u}^y) I\{Y_{t+u}^y \neq b(t + u)\} du \]

\[ + \int_0^s e^{-\lambda u} V_y (t + u, 1+) I\{Y_{t+u}^y \neq b(t + u)\} dR_{t+u} \]

\[ + \frac{1}{2} \int_0^s e^{-\lambda u} (V_y (t + u, b(t + u) +) - V_y (t + u, b(t + u) -)) dI_u^b (Y) \]

\[ = V(t, y) + \tilde{M}_s \]

\[ (4.2.46) \]

\[ + \int_0^s e^{-\lambda u} (G_t + \mathbb{L}YG - \lambda G) (t + u, Y_{t+u}^y) I\{Y_{t+u}^y \geq b(t + u)\} du \]

\[ = V(t, y) + \tilde{M}_s + \int_0^s e^{-\lambda u} H (t + u, Y_{t+u}^y) I\{Y_{t+u}^y \geq b(t + u)\} du, \]

where \( \left( \frac{I_u^b (Y)}{u} \right)_{u \geq 0} \) is the local time process of \( Y \) on the curve \( b \) given as follows

\[ l_u^b (Y) := \tilde{P}_{t,y} - \lim_{\epsilon \rightarrow 0} \frac{1}{2\epsilon} \int_0^u I\{b(t + s) - \epsilon < Y_u < b(t + s) + \epsilon\} \frac{\sigma^2}{2} (Y_s)^2 du, \]

and the second equality holds by lemma 4.2.16 and lemma 4.2.17 and moreover,

\[ \tilde{M}_s = -\int_0^s e^{-\lambda u} \sigma Y_{t+u} V_y (t + u, Y_{t+u}) d\tilde{W}_u, \]

is a martingale.

By setting \( s = T - t \), using that \( V(T, Y_T) = G(T, Y_T) = 0 \) and taking the \( \tilde{P}_{t,y} \)-expectation, we obtain by the optional sampling theorem

\[ V(t, y) = -\int_0^{T-t} e^{-\lambda u} \tilde{E}_{t,y} (H (t + u, Y_{t+u}) I\{Y_{t+u} \geq b(t + u)\}) du, \]

which is the desired result \((4.2.44)\).

Next, recall \((4.2.18)\) and set \( y = b(t) \) in \((4.2.44)\) so that

\[ G(t, b(t)) = -\int_0^{T-t} e^{-\lambda u} \tilde{E}_{t,b(t)} (H (t + u, Y_{t+u}) I\{Y_{t+u} \geq b(t + u)\}) du, \]

after which, \((4.2.43)\) follows.

(ii) We now establish the uniqueness of the optimal stopping boundary \( b \).

First, assume that there exists a decreasing function \( c : [0, T) \rightarrow \mathbb{R} \), which solves \((4.2.43)\) and satisfies \( c(t) \geq L \) for all \( 0 \leq t < T \) and let

\[ U^c(t, y) = -\int_0^{T-t} e^{-\lambda u} \tilde{E}_{t,y} (H (t + u, Y_{t+u}) I\{Y_{t+u} \geq c(t + u)\}) du \]

\[ (4.2.47) \]

for \( (t, y) \in [0, T) \times [1, \infty) \). The following equalities then emerge by \((4.2.47)\), the strong Markov property of \( Y \) and Lemma 2.3.12:

\[ \tilde{E}_{t,y} \left( e^{-\lambda s} U^c (t + s, Y_{t+s}) - \int_0^s e^{-\lambda u} H (t + u, Y_{t+u}) I\{Y_{t+u} \geq c(t + u)\} du \right) \]
With the verification being similar as Corollary 4.2.18, we manage to apply Theorem 2.2.7 to obtain
\[
\tilde{E}(e^{-\lambda s} \tilde{E}(- \int_t^T e^{-\lambda(u-t-s)} H(u,Y_u) I\{Y_u \geq c(u)\} du | \mathcal{F}_{t+s}) | \mathcal{F}_t)
\] 
\[ - \tilde{E}_{t,y} \left( \int_0^s e^{-\lambda u} H(t + u, Y_{t+u}) I\{Y_{t+u} \geq c(t + u)\} du \right)
\] 
\[ = \tilde{E}_{t,y} \left( - \int_t^T e^{-\lambda(u-t)} H(u,Y_u) I\{Y_u \geq c(u)\} du - \int_{t+s}^T e^{-\lambda(u-t)} H(u,Y_u) I\{Y_u \geq c(u)\} du \right)
\] 
\[ = \tilde{E}_{t,y} \left( \int_t^T e^{-\lambda(u-t)} H(u,Y_u) I\{Y_u \geq c(u)\} du \right)
\] 
\[ = \tilde{E}_{t,y} \left( \int_0^{T-t} e^{-\lambda u} H(t + u, Y_{t+u}) I\{Y_{t+u} \geq c(t + u)\} du \right) = U^c(t, y),
\]
which can be easily reshuffled into
\[
\tilde{E}_{t,y} \left( e^{-\lambda s} U^c (t + s, Y_{t+s}) \right) - U^c(t, y)
\] 
\[ = \tilde{E}_{t,y} \left( \int_0^s e^{-\lambda u} H(t + u, Y_{t+u}) I\{Y_{t+u} \geq c(t + u)\} du \right).
\] (4.2.48)

Then, define the following function
\[
V^c(t, y) = \begin{cases} 
U^c(t, y), & \text{for } y < c(t), \\
G(t, y), & \text{for } y \geq c(t), 
\end{cases}
\]
and notice that since \(c\) solves (4.2.43), it follows that \(U^c(t, c(t)) = G(t, c(t))\) and thus \(V^c(t, y)\) is continuous on \([0, T) \times [1, \infty)\).

Next in line, we reconstruct the continuation region and the stopping region by the means of \(c\)
\[
\mathcal{C} = \{(t, y) \in [0, T) \times [1, \infty) : y < c(t)\},
\]
\[
\mathcal{D} = \{(t, y) \in [0, T] \times [1, \infty) : y \geq c(t)\}.
\]

Since \(V^c(t, y)\) equals \(U^c(t, y)\) in the continuation set \(\mathcal{C}\),
\[
V^c_t + \mathbb{L}_Y V^c = \lambda V^c \quad \text{in } \mathcal{C},
\]
\[
V^c_y(t, 1+) = 0 \quad \text{for all } t \in [0, T).
\]

With the verification being similar as Corollary 4.2.18, we manage to apply Theorem 2.2.7 to obtain
\[
e^{-\lambda s} V^c (t + s, Y^u_{t+s}) = V^c(t, y) + \tilde{M}^c_s
\] 
\[ + \int_0^s e^{-\lambda u} \left( V^c_t + \mathbb{L}_Y V^c - \lambda V^c \right) (t + u, Y^u_{t+u}) I\{Y^u_{t+u} \neq c(t + u)\} du
\] 
\[ + \frac{1}{2} \int_0^s e^{-\lambda u} \left( V^c_y(t + u, c(t + u) \text{+}) - V^c_y(t + u, c(t + u) \text{−}) \right) dl^c_u (Y),
\]
where \(\tilde{M}^c_s\) is the martingale under measure \(\tilde{P}\).
Inspiring by (4.2.46), it seems only natural to investigate whether \( V^c(t, y) \) is differentiable at \( c(t) \) for each \( t \in [0, T) \). And if we had known that
\[
U^c(t, y) = G(t, y) \quad \text{for all } y \geq c(t), \tag{4.2.49}
\]
we would have directly obtained
\[
V^c_y(t, c(t)+) - V^c_y(t, c(t)-) = U^c_y(t, c(t)+) - U^c_y(t, c(t)-) = 0,
\]
and then consequently,
\[
e^{-\lambda s} V^c(t + s, Y^{y}_{t+s}) = V^c(t, y) + \tilde{M}^c_s + \int_0^s e^{-\lambda u} H(t + u, Y^{y}_{t+u}) I\{Y^{y}_{t+u} \geq c(t + u)\} du. \tag{4.2.50}
\]

To derive (4.2.49), we consider the stopping time
\[
\sigma_c = \inf\{0 \leq s \leq T - t : Y_{t+s} \leq c(t + s)\},
\]
and by setting \( s = \sigma_c \) in (4.2.48), we find that
\[
U^c(t, y) = \tilde{E}_{t,y} \left( e^{-\lambda \sigma_c} U^c(t + \sigma_c, Y_{t+\sigma_c}) \right) - \tilde{E}_{t,y} \left( \int_0^{\sigma_c} e^{-\lambda u} H(t + u, Y_{t+u}) I\{Y_{t+u} \geq c(t + u)\} du \right)
\]
\[
= \tilde{E}_{t,y} \left( e^{-\lambda \sigma_c} G(t + \sigma_c, Y_{t+\sigma_c}) \right) - \tilde{E}_{t,y} \left( \int_0^{\sigma_c} e^{-\lambda u} H(t + u, Y_{t+u}) I\{Y_{t+u} \geq c(t + u)\} du \right)
\]
\[
= G(t, y) + \tilde{E}_{t,y} \left( \int_0^{\sigma_c} e^{-\lambda u} H(t + u, Y_{t+u}) du \right) - \tilde{E}_{t,y} \left( \int_0^{\sigma_c} e^{-\lambda u} H(t + u, Y_{t+u}) I\{Y_{t+u} < c(t + u)\} du \right)
\]
\[
= G(t, y),
\]
where the second equality holds as \( U^c(t, c(t)) = G(t, c(t)) \) and the third equality follows from (4.2.23) in the sense that \( c(t) > L > 1 \) for all \( 0 \leq t < T \), and the fifth equality holds by the definition of the stopping time \( \sigma_c \), after which, (4.2.49) and (4.2.50) are fairly immediate.

By using (4.2.50) and considering the stopping time
\[
\tau_c = \inf\{0 \leq s \leq T - t : Y_{t+s} \geq c(t + s)\},
\]
and taking the expectation under \( \tilde{P}_{t,y} \), together with the optional sampling theorem, we have
\[
V^c(t, y) = \tilde{E}_{t,y} \left( e^{-\lambda \tau_c} G(\tau_c, Y_{t+\tau_c}) \right),
\]
and then recalling the value function $V$ in (4.2.7), we find that

$$V^c(t, y) \leq V(t, y),$$

(4.2.51)

for all $(t, y) \in [0, T) \times [1, \infty)$.

With the above observation in mind, we are now ready to prove the first relation between $b$ and $c$, that is, $b \geq c$ on $[0, T)$.

Let $y > b(t) \lor c(t)$ for $t \in [0, T)$ and consider the stopping time

$$\sigma_b = \inf\{0 \leq s \leq T - t \mid Y_{t+s} \leq b(t + s)\}.$$

By replacing $s$ in (4.2.46) and (4.2.50) with $\sigma_b$ and taking the expectation under $\tilde{P}_{t,y}$, we have

$$\tilde{E}_{t,y} \left( e^{-\sigma_b \lambda} V(t + \sigma_b, Y_{t+\sigma_b}) \right) = G(t, y) + \tilde{E}_{t,y} \left( \int_0^{\sigma_b} e^{-\lambda u} H(t + u, Y_{t+u}) \, du \right),$$

$$\tilde{E}_{t,y} \left( e^{-\sigma_b \lambda} V^c(t + \sigma_b, Y_{t+\sigma_b}) \right) = G(t, y) + \tilde{E}_{t,y} \left( \int_0^{\sigma_b} e^{-\lambda u} H(t + u, Y_{t+u}) I\{Y_{t+u} \geq c(t + u)\} \, du \right),$$

and from (4.2.51), we know that

$$\tilde{E}_{t,y} \left( \int_0^{\sigma_b} e^{-\lambda u} H(t + u, Y_{t+u}) I\{Y_{t+u} < c(t + u)\} \, du \right) \geq 0,$$

that is

$$\tilde{E}_{t,y} \left( \int_0^{\sigma_b} e^{-\lambda u} H(t + u, Y_{t+u}) I\{Y_{t+u} \geq c(t + u)\} \, du \right) \geq 0,$$

and since $H(t, y) < 0$ for all $(t, y) \in [0, T) \times [1, \infty)$, it follows that $\tilde{P}_{t,y} (Y_{t+u} \geq c(t + u)) = 1$, proving that $b(t) \geq c(t)$ for all $t \in [0, T]$.

In order to complete the proof, we must show that $c$ equals $b$.

Assume that, for contradiction, there exists $t \in (0, T)$ such that $b(t) > c(t)$ and pick a point $y \in (c(t), b(t))$.

Consider the stopping time

$$\tau_b = \inf\{0 \leq s \leq T - t \mid Y_{t+s} \geq b(t + s)\}.$$

By replacing $s$ in (4.2.46) and (4.2.50) with $\tau_b$ and taking the expectation under $\tilde{P}_{t,y}$, we have

$$\tilde{E}_{t,y} \left( e^{-\tau_b \lambda} G(t + \tau_b, Y_{t+\tau_b}) \right) = V(t, y),$$

$$\tilde{E}_{t,y} \left( e^{-\tau_b \lambda} G(t + \tau_b, Y_{t+\tau_b}) \right) = G(t, y) + \tilde{E}_{t,y} \left( \int_0^{\tau_b} e^{-\lambda u} H(t + u, Y_{t+u}) I\{Y_{t+u} \geq c(t + u)\} \, du \right).$$

Knowing that $V(t, y) > G(t, y)$ as $y < b(t)$, we find that

$$\tilde{E}_{t,y} \left( \int_0^{\tau_b} e^{-\lambda u} H(t + u, Y_{t+u}) I\{Y_{t+u} \geq c(t + u)\} \, du \right) > 0,$$

but $H(t, y) \leq 0$ in $[0, T) \times [1, \infty)$ suggests otherwise, we have reach a contradiction and thereby establishing the uniqueness of $b$.

The theorem is completely proved. \qed
Remark 4.2.20. The mainstream method of obtaining the optimal stopping boundary $b$ and the value function $V$ is to solve (4.2.43) and (4.2.44) numerically by using the terminal point $b(T)$, for detailed description of this approach, we refer to [64, Page 432] or A.6.

Figure 4.2: This figure displays the maps $t \mapsto b(t)$ with chosen parameters $r = 0.02$, $\sigma = 0.3$, $\lambda = 0.4$, $L = 5$ and $T = 10$.

Remark 4.2.21. Figure 4.2 seemingly has a flat tail for $t \in [9.5, 10]$, this is because the plot function in Matlab has automatically rounded the number to 4 decimal places; but if vpa function is used to recover their values, we see that

\[
\begin{align*}
    b(9.5) &\approx 5.0011848117586090722852532053366, \\
    b(9.6) &\approx 5.00010594413864506435629666433819, \\
    b(9.7) &\approx 5.0000008520428087521736415510532, \\
    b(9.8) &\approx 5.0000000000049764636855798016768.
\end{align*}
\]
Chapter 5

American Put Option with Last Exit Time

This chapter, whose content is primarily based on the working paper [82] of the author, is designated to solve the following optimal stopping problem in both infinite \((T = \infty)\) and finite \((T < \infty)\) horizon:

\[
V = \sup_{\tau \in [0,T]} E_0, x \left( e^{-r\tau} (K - X_\tau)^+ I\{\tau < \theta\} \right),
\]

with random time \(\theta = \sup\{t \in [0, T] : X_t \geq L\}\) and the supremum is taken over all the stopping times \(\tau \in [0, T]\). As in Chapter 4, we assume that the stock and bond price process follow the SDEs:

\[
dX_t = r X_t dt + \sigma X_t dW_t, \quad X_0 = x > 0, \\
\]

\[
 dbs = r B_t dt, \quad B_0 = 1,
\]

where \(W = (W_t)_{t \geq 0}\) is a standard Brownian motion started at zero under measure \(P\), \(r > 0\) is the interest rate and \(\sigma > 0\) is the volatility coefficient. Recall from Example 3.1.4 that this is a risk neutral model so that the discounted stock price is a martingale under measure \(P\). We assume \(L < K\) where \(K > 0\) is the strike price. The indicator function in (5.0.1) can be interpreted as the risk preference of the agent, see [60]⁴.

Section 5.1 is based on the published paper [33] of the author and the result presented here contains extension by considering the payoff of conventional contingent claim. This section covers the solution for problem (5.0.1) as \(T = \infty\). We begin by formulating the standard optimal stopping problem and its associated free-boundary problem in the form of ODE. After solving the ODE, we also prove its solution to be the smallest superharmonic function that dominates the gain function in Subsection 5.2.2.

Section 5.2 begins our discussion on problem (5.0.1) as \(T < \infty\). We once again reformulate the problem and acquaint ourselves with the properties possessed by the Azéma supermartingale. Then, Theorem 3.2.8 tells us that the existence of the optimal stopping rule lies in the semi-continuity of

⁴This is only one of the many possible financial interpretations to this problem, readers are referred to [33] for other financial justification.
the gain function and the value function. Before turning to the free-boundary problem, we investigate
the features of continuation and stopping sets by breaking down the stopping set piece by piece
so that assumption can be made to establish the monotonicity of the boundary in Subsection 5.2.3.
After this, we shall have sufficient mathematics at our disposal to avoid singular point of the partial
derivative of the gain function to further justify the regularity of \( V \). The free-boundary problem is
then within easy reach. The price of such contract is given with proof in the Subsection 5.2.4.

5.1 Infinite-time Horizon

5.1.1 Reformulation and Basics

In this subsection, we reformulate our main optimal stopping problem\(^2\) and prove some basic facts.
By Lemma 2.3.12 and the measurability of stopping time \( \tau \),

\[
V(x) = \sup_\tau E_x \left( e^{-r\tau}(K - X_\tau)^+ I\{\tau < \theta\}|F_\tau \right),
\]

\[
= \sup_\tau E_x \left( e^{-r\tau}(K - X_\tau)^+ P(\tau < \theta|F_\tau) \right). \tag{5.1.1}
\]

Next in line is the result of the Azéma supermartingale but first for notational convenience, we set

\[
Z(x) = \left( \frac{L}{x} \right)^\alpha \wedge 1 \text{ and } \alpha = \frac{2r}{\sigma^2} - 1 < 0.
\]

**Proposition 5.1.1.** Let \( P(\theta > t|F_t) \) be the Azéma supermartingales associated with the random
times \( \theta \). Then,

\[
P(\theta > t|F_t) = Z(X_t).
\]

**Proof.** As before, the set equality suggests that

\[
\{\theta > t\} = \left\{ \max_{u \geq t} X_u \geq L \right\},
\]

so that

\[
P(\theta > t|F_t) = P\left( \max_{u \geq t} X_{t+u} \geq L \right)
= P\left( \max_{u \geq 0} X_t e^{(r-\frac{\sigma^2}{2})u+\sigma W_u} \geq L \right)
= P\left( \max_{u \geq 0} \left( r - \frac{\sigma^2}{2} \right) u + \sigma W_u \geq \log \frac{L}{X_t} \right)
= \left( \frac{L}{X_t} \right)^\alpha,
\]

and the desired result follows. \( \Box \)

\(^2\)For notational convenience, we simply write \( E_x(A) \) instead of \( E_{0,x}(A) \) hereafter till the end of Chapter 5.1.
Combining the above computation, the original problems (5.1.1) can be rewritten as

\[ V(x) = \sup_{\tau} E_x \left( e^{-r\tau} (K - X_\tau)^+ Z(X_\tau) \right), \]

and for the sake of brevity, we denote the gain function \( G(x) = (K - x)^+ Z(x) \).

Let us end this subsection by a useful property possessed by function \( Z \).

**Proposition 5.1.2.** The function \( Z(x) \) satisfies the following ordinary differential equation, for \( x \in (0, L) \cup (L, \infty) \),

\[ rx \frac{d}{dx} Z(x) + \frac{1}{2} \sigma^2 x^2 \frac{d^2}{dx^2} Z(x) = 0. \]

**Proof.** Remember that the Doob-Meyer decomposition of the Azéma supermartingale in Example 2.3.4 is given by

\[ Z(X_t) = Z(X_0) + M_\theta^t - A_\theta^t \]

and the measure \( dA_\theta^t \) is carried by the set \( \{ t : X_t = L \} \).

Now, by Theorem 2.2.7 \(^3\), we have

\[
dZ(X_t) = \left( rX_t Z_x(X_t) + \frac{1}{2} \sigma^2 X_t^2 Z_{xx}(X_t) \right) I\{X_t \neq L\} dt + \sigma X_t Z_x(X_t) I\{X_t \neq L\} dW_t + \frac{1}{2} (Z_x(L^-) - Z_x(L^+)) dl^t_L(X),
\]

where \( l^t_L \) is the local time of \( X \) at the level \( L \) given by

\[ l^t_L(X) = P - \lim_{\epsilon \to 0} \frac{1}{2\epsilon} \int_0^t I\{|X_s - L| < \epsilon\} d\langle X, X \rangle_s, \]

which, joining with (5.1.4), shows that

\[
M_\theta^t = \int_0^t \left( \sigma - \frac{2r}{\sigma} \right) \left( \frac{L}{X_s} \right)^\alpha I\{X_s \neq L\} dW_s
\]

\[
A_\theta^t = \int_0^t \left( \frac{1}{2} - \frac{r}{\sigma^2} \right) \frac{1}{L} dl^t_L(X),
\]

and thereby, proving that (5.1.3) holds true. \( \square \)

**Remark 5.1.3.** An alternative is to compute the first and second derivative of the function \( Z \) directly. Admittedly, the above proof is more clumsy than direct computation, but latter, as we shall see in the setting of finite time, the method of using Doob-Meyer decomposition is far more convenient.

\(^3\)The corresponding conditions to apply this Theorem can be checked in a fairly easy fashion, given that \( Z \) and later \( G \) are the deterministic functions. We omit further details.
5.1.2 The Free-boundary Problems

This subsection is devoted to formulating the free-boundary problem, whose solution will also be provided and verified.

The point of departure is to appeal to Theorem 2.2.8 as $G$ is of class $C^2$ on $(0, \infty) \setminus \{L, K\}$ and obtain

$$e^{-rt}G(X_t) = G(x) + \int_0^t e^{-ru} \left(-rG + rX_uG_x + \frac{1}{2} \sigma^2 X_u^2 G_{xx} \right)(X_u) I\{X_u \neq L, X_u \neq K\} du$$

$$+ \int_0^t e^{-ru} \sigma X_u G_x(X_u) I\{X_u \neq L, X_u \neq K\} dW_u$$

$$+ \frac{1}{2} \int_0^t e^{-ru} (G_x(K+) - G_x(K-)) d\ell^K_u(X)$$

$$+ \frac{1}{2} \int_0^t e^{-ru} (G_x(L+) - G_x(L-)) d\ell^L_u(X)$$

(5.1.5)

where $\ell^L_s$ and $\ell^K_s$ are the local time of $X$ at the levels $L, K$ respectively, and for brevity, we denote the martingale term as $M_t = \int_0^t e^{-ru} \sigma X_u G_x(X_u) I\{X_u \neq L, X_u \neq K\} dW_u$.

Now, with a little bit of effort, we can compute:

$$\left( -rG + rX_u G_x + \frac{1}{2} \sigma^2 X_u^2 G_{xx} \right) = \begin{cases} -rKZ - \sigma^2 x^2 Z_x & \text{for } x < L, \\ -rK & \text{for } L < x < K, \\ 0 & \text{for } x > K, \end{cases}$$

and

$$G_x(L-) = -1 - \alpha(K - L)L^{-1}, \quad G_x(L+) = -1,$$

$$G_x(K+) = 0, \quad G_x(K-) = -1,$$

so that (5.1.5) is reshuffled accordingly as

$$e^{-rt}G(X_t) = G(x) + M_t - rK \int_0^t e^{-ru} I\{L < X_u < K\} du$$

$$- \int_0^t e^{-ru} \left( rKZ + \sigma^2 X_u^2 Z_x \right)(X_u) I\{X_u < L\} du$$

$$+ \int_0^t e^{-ru} \alpha(K - L)L^{-1} d\ell^L_u(X) + \frac{1}{2} \int_0^t e^{-ru} d\ell^K_u(X),$$

of which, we take the expectation under measure $P_x$,

$$E_x \left( e^{-rt}G(X_t) \right) = G(x) + E \left( -rK \int_0^t e^{-ru} I\{L < X_u < K\} du \right)$$

$$- \int_0^t e^{-ru} \left( rK - \alpha \sigma^2 X_u \right) Z(X_u) I\{X_u < L\} du$$

$$+ \int_0^t e^{-ru} \alpha(K - L)L^{-1} d\ell^L_u(X) + \frac{1}{2} \int_0^t e^{-ru} d\ell^K_u(X).$$

(5.1.6)
Immediate from (5.1.6), we see that the further \( X \) gets away from \( K \) the less likely that the gain will increase upon continuing, and this suggests that there exists a point \( b \in (0, K) \) such that the stopping time

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

(5.1.7)
is optimal for problem (5.1.2), with the convention that the infimum of the empty set is infinite. The optimality of \( \tau_b \) will be verified in section 5.1.3.

The next logical task is, of course, to determine whether or not such stopping time \( \tau_b \) is finite.

**Corollary 5.1.4.** The stopping time \( \tau_b \) is finite, i.e. \( P_x(\tau_b < \infty) = 1 \).

**Proof.** To verify \( P_x(\tau_b < \infty) = 1 \) for all \( x > 0 \), it suffices to show that \( P_x(\min_{t \geq 0} X_t \leq b) = 1 \).

With appeal to Lemma 2.2.11, we find that

\[
P_x\left(\min_{t \geq 0} xe^{(r - \frac{\sigma^2}{2})t + \sigma W_t} \leq b\right) = P_x\left(\min_{t \geq 0} \left( r - \frac{\sigma^2}{2} \right) t + \sigma W_t \leq \log \frac{b}{x} \right)
\]

\[
= P_x\left( - \max_{t \geq 0} \left( r - \frac{\sigma^2}{2} \right) t - \sigma W_t \leq \log \frac{b}{x} \right)
\]

\[
= P_x\left( \max_{t \geq 0} \left( r - \frac{\sigma^2}{2} \right) t + \sigma W_t \geq - \log \frac{b}{x} \right)
\]

\[
= 1 - P_x\left( \max_{t \geq 0} \left( r - \frac{\sigma^2}{2} \right) t + \sigma W_t \leq - \log \frac{b}{x} \right)
\]

where the third equality holds true as \((W_t)_{t \geq 0} \overset{d}{=} (-W_t)_{t \geq 0} \). Hence, for \( x \geq b \) (i.e. \(- \log \frac{b}{x} \geq 0 \)), by the assumption that \( \alpha = \frac{2r}{\sigma^2} - 1 < 0 \) and statement (ii) in Lemma 2.2.11, we have

\[
P_x\left(\min_{t \geq 0} X_t \leq b\right) = 1,
\]

which, joining with the fact that \( x < b \) implies \( \tau_b = 0 \), shows that \( P_x(\tau_b < \infty) = 1 \) for all \( x > 0 \). \( \square \)

With the aid of the above observations and the standard arguments based on the strong Markov property, we are now ready to formulate the free boundary problem to find the smallest superharmonic function \( \hat{V} \) dominating \( G \) and the unknown point \( b \),

\[
\mathbb{L}_X \hat{V} - r \hat{V} = 0 \quad \text{for} \quad x \in (b, \infty), \quad (5.1.8)
\]

\[
\hat{V}(x) = G(x) \quad \text{for} \quad x = b, \quad (5.1.9)
\]

\[
\hat{V}_x(x) = G_x(x) \quad \text{for} \quad x = b, \quad (5.1.10)
\]

\[
\hat{V}(x) = G(x) \quad \text{for} \quad x \in (0, b), \quad (5.1.11)
\]

\[
\hat{V}(x) > G(x) \quad \text{for} \quad x \in (b, \infty), \quad (5.1.12)
\]

where \( \mathbb{L}_X = r x \frac{\partial}{\partial x} + \frac{1}{2} \sigma^2 x^2 \frac{\partial^2}{\partial x^2} \) is the infinitesimal generator of the strong Markov diffusion process \( X \).
**Remark 5.1.5.** Another message of the free-boundary problem is that the optimal stopping level \( b \) depends on the relative positions of the level \( L \).

We now proceed to solve the free-boundary problem. Equation (5.1.8) is the well-known Cauchy-Euler equation, whose solution gets form as follows

\[
\hat{V}(x) = x^p,
\]

(5.1.13)

Then, by putting (5.1.13) into (5.1.8), we have the following quadratic equation

\[
\frac{1}{2} \sigma^2 p^2 + \left( r - \frac{\sigma^2}{2} \right) p - r = 0,
\]

whose roots equal

\[
p_1 = 1, \quad p_2 = -\frac{2r}{\sigma^2},
\]

from which, we know that the general solution of (5.1.8) is

\[
\hat{V}(x) = C_1 x + C_2 x^{-\frac{2r}{\sigma^2}},
\]

where \( C_1 \) and \( C_2 \) are undetermined constants and by noticing that, in fact, for all \( x > 0 \), function \( \hat{V}(x) \leq K \), implying that \( C_1 \) must be zero.

By inspecting (5.1.9) and (5.1.10), we see that they are of different forms as \( b > L \) and \( b < L \),

\[
\hat{V}(b) = C_2 b^{-\frac{2r}{\sigma^2}} = \begin{cases} (K - b) \left( \frac{L}{b} \right)^\alpha & \text{for } b > L, \\ (K - b) \left( \frac{L}{b} \right)^\alpha & \text{for } b < L, \end{cases}
\]

(5.1.14)

and

\[
\hat{V}_x(b) = -\frac{C_2}{\sigma^2} b^{-\frac{2r}{\sigma^2} - 1} = \begin{cases} -1 \left( \frac{L}{b} \right)^\alpha & \text{for } b > L, \\ -1 - \frac{\alpha(K-b)}{b} \left( \frac{L}{b} \right)^\alpha & \text{for } b < L, \end{cases}
\]

(5.1.15)

and we have rediscovered Remark 5.1.5, from which, determining \( C_2 \) is within easy reach:

\[
C_2 = \begin{cases} (K - b) b^{\frac{2r}{\sigma^2}} & \text{for } b > L, \\ (K - b) L^{\frac{2r}{\sigma^2} - 1} & \text{for } b < L. \end{cases}
\]

With a little additional effort, we work out the math for \( b \) by using (5.1.10):

\[
b = \begin{cases} \frac{2rK}{2r + \sigma^2}, & \text{for } \frac{2rK}{2r + \sigma^2} > L, \\ \frac{K}{2}, & \text{for } \frac{K}{2} < L. \end{cases}
\]

(5.1.16)

**Remark 5.1.6.** Notice that, if \( \frac{2r}{\sigma^2} + 2r = \frac{1}{2} \), i.e. \( \alpha = 0 \), then the result of standard perpetual American put option emerges.

The conclusion runs as follows:
Theorem 5.1.7. If $K < 2L$, that is $b < L$, then the value function $V$ from (5.1.1) is given explicitly by

$$V(x) = \begin{cases} G(x), & x \in (0, b], \\ \frac{L^2 K^2}{4} x^{-\frac{2\alpha}{\sigma^2}} - \frac{2\alpha}{\sigma^2}, & x \in (b, \infty); \end{cases} \quad (5.1.17)$$

if, on the other hand, $2rK > L(\sigma^2 + 2r)$, that is $b > L$, then

$$V(x) = \begin{cases} G(x), & x \in (0, b), \\ \left( K - \frac{2rK}{2r + \sigma^2} \right) \left( \frac{2rK}{2r + \sigma^2} \right)^{\frac{2\alpha}{\sigma^2}} x^{-\frac{2\alpha}{\sigma^2}}, & x \in [b, \infty). \end{cases} \quad (5.1.18)$$

The stopping time $\tau_0$ from (5.1.7) are optimal in problem (5.1.1).

Proof of Solution (5.1.17). We wish to prove that $V(x)$ in (5.1.1) equals $\hat{V}(x)$ in (5.1.14) for all $x \in (0, \infty)$, to achieve which, we proceed via three steps (proving three statements):

(i) $\hat{V}(x) \geq G(x)$ for all $x \in (0, \infty)$. To prove this, we set

$$h(x) = \hat{V}(x) - G(x) = \begin{cases} 0, & x \in (0, K], \\ \frac{L^2 K^2}{4} x^{-\frac{2\alpha}{\sigma^2}} - (K - x) \left( \frac{L}{x} \right)^{\alpha}, & x \in [K, L], \\ \frac{L^2 K^2}{4} x^{-\frac{2\alpha}{\sigma^2}} - (K - x), & x \in [L, \infty). \end{cases}$$

Knowing that $\hat{V}(x) = G(x)$ for all $x \in (0, K]$ reduces our task to verifying that $h(x) \geq 0$ in the domain $[K, \infty)$. By rearrangement, we have

$$h(x) = \begin{cases} \left( \frac{L}{x} \right)^{\alpha} x^{-1} \left( \frac{K}{2} - x \right)^2, & x \in [K, L], \\ \left( \frac{L}{x} \right)^{\alpha} K^2 x^{-1} - K + x \geq K^2 x^{-1} - K + x = x^{-1} \left( \frac{K}{2} - x \right)^2, & x \in [L, \infty), \end{cases}$$

which yields that $h(x) \geq 0$, and thereby proving statement (i).

(ii) $V(x) \leq \hat{V}(x)$ for all $x \in (0, \infty)$. By invoking Theorem 2.2.7, we obtain

$$e^{-rt} \hat{V}(X_t) = \hat{V}(x) + \int_0^t e^{-rs} \left( -r \hat{V} + \mathbb{L}_X \hat{V} \right)(X_s) I\{X_s \neq b\} ds$$

$$+ \int_0^t e^{-rs} \sigma X_s \hat{V}_x(X_s) I\{X_s \neq b\} dW_s + \frac{1}{2} \int_0^t e^{-rs} \left( \hat{V}_x(b+) - \hat{V}_x(b-) \right) dl^b_s (X^s)$$

$$= \hat{V}(x) + M_t + \int_0^t e^{-rs} \left( -r \hat{V} + \mathbb{L}_X \hat{V} \right)(X_s) ds, \quad (5.1.19)$$

where the second equality holds via the smooth-fit condition (5.1.10). Since $\hat{V}(x) = G(x)$ for all $0 < x < b$ such that

$$-r \hat{V} + \mathbb{L}_X \hat{V} = -r G + \mathbb{L}_X G = -r K Z - \sigma^2 x^2 Z_x,$$

which, together with (5.1.8), tells us that

$$-r \hat{V} + \mathbb{L}_X \hat{V} \leq 0, \quad (5.1.20)$$
and then, immediate from statement (i) and (5.1.20) for all \( x \in \mathbb{R}_+ \), we see that
\[
e^{-rt}G(X_t) \leq e^{-rt}\hat{V}(X_t) \leq \hat{V}(x) + M_t. \tag{5.1.21}
\]

Let \( \tau_n = \tau \wedge n \) for \( n \geq 0 \) and then by Corollary 2.3.13, for every bounded stopping time \( \tau_n \) of \( X \), taking the expectation under measure \( P_x \) yields \( E_x (M_{\tau_n}) = E_x (M_n) = 0 \) and
\[
E_x (e^{-r\tau_n}G(X_{\tau_n})) \leq \hat{V}(x).
\]
Next, let \( n \rightarrow \infty \), by Fatou’s lemma, \( E_x (e^{-r\tau}G(X_{\tau})) \leq \hat{V}(x) \). After this, by taking the supremum over all the stopping times of \( X \),
\[
V(x) = \sup_{\tau} E_x (e^{-r\tau}G(X_{\tau})) \leq \hat{V}(x), \tag{5.1.22}
\]
and the desired statement (ii) follows.

(iii) \( V(x) \geq \hat{V}(x) \) for all \( x \in (0, \infty) \). Put \( t = \tau_0 \wedge n \) for \( n \geq 0 \) in (5.1.19), with \( \tau_0 \) being the finite stopping time defined in (5.1.7),
\[
e^{-r(\tau_0 \wedge n)}\hat{V}(X_{\tau_0 \wedge n}) = \hat{V}(x) + M_{\tau_0 \wedge n},
\]
and then by taking its expectation under measure \( P_x \), we have \( E_x (M_{\tau_0 \wedge n}) = E_x (M_n) = 0 \) so that
\[
E_x (e^{-r(\tau_0 \wedge n)}\hat{V}(X_{\tau_0 \wedge n})) = \hat{V}(x).
\]
Let \( n \rightarrow \infty \), via an appeal to Theorem 2.1.9 and the fact that \( V(X_{\tau_0}) = G(X_{\tau_0}) \), we see that
\[
\hat{V}(x) = E_x (e^{-r\tau_0}\hat{V}(X_{\tau_0})) = E_x (e^{-r\tau_0}G(X_{\tau_0})) \leq \sup_{\tau} E_x (e^{-r\tau}G(X_{\tau})) = V(x), \tag{5.1.23}
\]
which yields statement (iii). Finally, combining (5.1.22) and (5.1.23) completes the proof.

**Remark 5.1.8.** The above proof is the same as showing that \( \hat{V} \) is the smallest superharmonic function that dominates \( G \) in Theorem 3.2.4: step (i) shows the domination, step (ii) we have \( \hat{V} \) being superharmonic function because of inequality (5.1.21) and (5.1.23) in step (iii) tells us that \( \hat{V} \) is the smallest superharmonic function that dominates \( G \). Therefore, \( \tau_0 \) is indeed optimal given that it is finite almost surely.

**Proof of Solution (5.1.18).** Similar argument works for this solution, but since solution (5.1.18) is not differentiable on \( L \), which should require extra care in the following part of step (ii):
\[
e^{-rt}\hat{V}(X_t) = \hat{V}(x) + \int_0^t e^{-r s} \left( -r \hat{V} + \frac{r}{2} X_t \right) I\{X_s \neq b, X_s \neq L\} ds
\]
\[
+ \int_0^t e^{-r s} \sigma X_s \hat{V}_x(X_s) I\{X_s \neq b, X_s \neq L\} dW_s
\]
\[
+ \frac{1}{2} \int_0^t e^{-r s} \left( \hat{V}_x(b+) - \hat{V}_x(b-) \right) dl_s^b(X^x)
\]
\[
+ \frac{1}{2} \int_0^t e^{-r s} \left( \hat{V}_x(L+) - \hat{V}_x(L-) \right) dl_s^L(X^x)
\]
\[ \hat{V}(x) + M_t + \int_0^t e^{-rs} \left( -r\hat{V} + \mathbb{L}_X \hat{V} \right) (X_s) ds \]
\[ + \frac{1}{2} \int_0^t e^{-rs} (G_x(L^+) - G_x(L^-)) d\tilde{F}_t(X^x) \]
\[ \leq \hat{V}(x) + M_t + \int_0^t e^{-rs} \left( -r\hat{V} + \mathbb{L}_X \hat{V} \right) (X_s) ds, \quad (5.1.24) \]

where the second equality holds via the smooth-fit condition and the inequality follows via \( G_x(L^+) - G_x(L^-) < 0 \). Since \( \hat{V}(x) = G(x) \) for all \( 0 < x < b \) such that
\[ -r\hat{V} + \mathbb{L}_X \hat{V} = -rG + \mathbb{L}_X G = -rKZ - \sigma^2 x^2 Z_x, \]

which, together with (5.1.8), tells us that \(-r\hat{V} + \mathbb{L}_X \hat{V} \leq 0\). This, combining with statement (i) yields \( e^{-rt}G(X_t) \leq e^{-rt} \hat{V}(X_t) \leq \hat{V}(x) + M_t \). The rest of the proof follows similarly as the proof of (5.1.17). \( \square \)

**Figure 5.1:** This figure displays the maps \( x \mapsto V(x) \) and \( x \mapsto G(x) \) with chosen parameters \( r = 0.05, \sigma = 0.4 \), and the optimal stopping points \( b = 2.6923 \) for \( L = 2 \) and \( b = 3.5 \) for \( L = 4 \).

**Remark 5.1.9.** The figure shows that the value of perpetual contract is a decreasing function of \( L \), and is bounded above by the global maximum of \( G \) and below by 0.
5.2 Finite-time Horizon

5.2.1 Reformulation and Basics

In this subsection, we reformulate the main optimal stopping problem and collect some elementary facts that will be used later.

We begin by restating the following optimal stopping problems:

\[
V = \sup_{0 \leq \tau \leq T} E_{0, x} \left( e^{-r\tau} (K - X_{\tau})^+ I\{\theta > \tau\} \right) \\
\sup_{0 \leq \tau \leq T} E_{0, x} \left( e^{-r\tau} (K - X_{\tau})^+ P(\theta > \tau | F_\tau) \right). 
\] (5.2.1)

Remark 5.2.1. The question is then raised when the geometric Brownian motion starts below \( L \), i.e. \( x < L \) and never hits \( L \) at all on \([0, T]\), which value shall be assigned to \( \theta \). However, as it turns out, the value of \( \theta \) in such case is irrelevant. (See Proposition 5.2.2 and [24, Page 230].)

The process \( X \) is a strong Markov (diffusion) process with the infinitesimal generator given by

\[
\mathbb{L}_X = rx \frac{\partial}{\partial x} + \frac{\sigma^2}{2} x^2 \frac{\partial^2}{\partial x^2}.
\]

In the view of (5.2.1), the first thing is, of course, to calculate \( P(\theta > \tau | F_\tau) \) and let us introduce the notation

\[
Z(t, x) = \left( \Phi\left( \frac{-\log \frac{L}{x} + (r - \frac{\sigma^2}{2}) (T - t)}{\sigma \sqrt{T - t}} \right) + \left( \frac{L}{x} \right)^\alpha \Phi\left( \frac{-\log \frac{L}{x} - (r - \frac{\sigma^2}{2}) (T - t)}{\sigma \sqrt{T - t}} \right) \right) \land 1,
\]

where \( \Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x e^{-\frac{z^2}{2}} dz \) and \( \alpha = \frac{2r}{\sigma^2} - 1 < 0 \).

Proposition 5.2.2. Let \( P(\theta > t | F_t) \) be the Azéma supermartingale associated with the random time \( \theta \). Then, for \( r - \frac{\sigma^2}{2} \in \mathbb{R} \) and \( \sigma > 0 \), for

\[
P(\theta > t | F_t) = Z(t, X_t).
\]

Proof. First of all, suppose that

\[
d_t = \inf\{ u \geq t : X_u \geq L \} = t + \inf\{ u \geq 0 : \bar{X}_0 = X_t, \bar{X}_u \geq L \},
\]

where the second equality follows from the Markov property of \( X \) and for easy reference, let

\[
\bar{h}_L = \inf\{ u \geq 0 : \bar{X}_0 = X_t, \bar{X}_u \geq L \},
\]

in plain language, \( \bar{h}_L \) is the first time when a geometric Brownian motion \( \bar{X} \), starting from \( X_t \), hits the level \( L \).

\( ^4 \)Not to confuse this \( x \) with \( X_0 = x \).
Then, it follows that
\[
P(\theta > t | \mathcal{F}_t) = 1 - P(\theta \leq t | \mathcal{F}_t) \\
= 1 - P(d_t > T | \mathcal{F}_t) \\
= 1 - P(\bar{h}_L > T - t | \mathcal{F}_t).
\]
Upon observing that, in fact,
\[
\{\bar{h}_L \leq T - t | \mathcal{F}_t\} = \left\{ \max_{0 \leq u \leq T-t} \bar{X}_u \geq L | \mathcal{F}_t \right\}
\]
(5.2.2) after which, an application of Lemma 2.2.11 proves that, for 
\[
\log \frac{L}{\bar{X}_t} \geq 0, \quad r - \frac{\sigma^2}{2} \in \mathbb{R} \quad \text{and} \quad \sigma > 0,
\]
\[
P(\theta > t | \mathcal{F}_t) = \Phi \left( \frac{-\log \frac{L}{\bar{X}_t} + \left( r - \frac{\sigma^2}{2} \right)(T - t)}{\sigma \sqrt{T - t}} \right)
\]
while for \( \log \frac{L}{\bar{X}_t} < 0 \), observe that event (5.2.2) happens almost surely.

As for \( t = T \), from the set equality (5.2.2), it follows that
\[
P(\theta > t | \mathcal{F}_t) = \begin{cases} 
1, & \log \frac{L}{\bar{X}_t} \leq 0, \\
0, & \log \frac{L}{\bar{X}_t} > 0.
\end{cases}
\]

Having introduced the Azéma supermartingale, the following facts are just around the corner. 

**Proposition 5.2.3.** The function \( Z(t, x) \) satisfies the following partial differential equation, for \( (t, x) \in [0, T] \times \{(0, L) \cup (L, \infty)\} \),
\[
\frac{\partial}{\partial t} Z(t, x) + rx \frac{\partial}{\partial x} Z(t, x) + \frac{1}{2} \sigma^2 x^2 \frac{\partial^2}{\partial x^2} Z(t, x) = 0.
\]
(5.2.3)

**Proof.** Notice that direct computation is certainly an option, but making use of the Doob-Meyer decomposition shall be a much efficient way.

Let us apply the change-of-variable formula to obtain
\[
dZ(t, X_t) = \left( \frac{\partial}{\partial t} Z(t, x) + rx \frac{\partial}{\partial x} Z(t, x) + \frac{1}{2} \sigma^2 x^2 \frac{\partial^2}{\partial x^2} Z(t, x) \right) (t, x) I\{X_t \neq L\} dt \\
+ \sigma X_t \frac{\partial}{\partial x} Z(t, x) I\{X_t \neq L\} dW_t + \frac{1}{2} \left( \frac{\partial}{\partial x} Z(t, L^+) - \frac{\partial}{\partial x} Z(t, L^-) \right) dt^L.
\]
where $l_t^L$ is the local time of $X$ at the level $L$ given by

$$l_t^L = P - \lim_{\epsilon \to 0} \frac{1}{2\epsilon} \int_0^t I\{|X_s - L| < \epsilon\} d\langle X, X \rangle_s,$$

after which, we can lean on the Doob-Meyer decomposition $Z(t, X_t) = Z(0, x) + M_t - A_t$ to conclude that

$$M_t = \int_0^t \sigma X_s \frac{\partial}{\partial x} Z(t, X_s) I\{X_s \neq L\} dW_s,$$

$$A_t = \frac{1}{2} \int_0^t \frac{\partial}{\partial x} Z(s, L-) d\langle L \rangle_s,$$

where $M_t$ is the càdlàg martingale and $A_t$ is the predictable projection whose measure $dA_t$ is carried by the set $\{t : X_t = L\}$, and consequently, (5.2.3) follows.

**Corollary 5.2.4.** (i) The map $t \mapsto Z(t, x)$ is decreasing on $[0, T]$; (ii) the map $x \mapsto Z(t, x)$ is increasing on $(0, \infty)$.

**Proof.** These two statements follow directly from the computation in A.1.2.

Summarising our findings so far, by the strong Markov property of $X$, we can rewrite and generalise the original problem (5.2.1) as

$$V(t, x) = \sup_{0 \leq \tau \leq T - t} E_{t,x} \left( e^{-r\tau} (K - X_{t+\tau})^+ Z(\tau, X_{t+\tau}) \right), \quad (5.2.5)$$

and for simplicity, let the gain function $G(t, x) = (K - x)^+ Z(t, x)$.

### 5.2.2 The Free-boundary Problem

Before we turn to formulating the free-boundary problem, we squeeze the important concepts of continuation and stopping sets into the text.

**Definition 5.2.5.** The continuation set $C$ and the stopping set $D$ are defined as follows:

$$C = \{(t, x) \in [0, T) \times (0, \infty) : V(t, x) > G(t, x)\}, \quad (5.2.6)$$

$$D = \{(t, x) \in [0, T) \times (0, \infty) : V(t, x) = G(t, x)\} \cup \{(T, x) : x \in (0, \infty)\}. \quad (5.2.7)$$

Based on the above definition, the first entry time $\tau_D$ of $X$ into $D$ is then defined as

$$\tau_D = \inf\{0 \leq s \leq T - t : (t + s, X_{t+s}^x) \in D\} \wedge (T - t). \quad (5.2.8)$$

It is then not far-fetched to ask for the optimality of $\tau_D$, which, according to Theorem 3.2.8, is the same as showing the (semi)continuity of $V$ and $G$.

**Remark 5.2.6.** It is worth mentioning that $Z_{x}(t, x)$ (see A.1.2) has a singularity at point $(L, T)$, so establishing the continuity of the value function can only be pursued later.
Lemma 5.2.7. The value function $V$ is l.s.c on $[0, T] \times (0, \infty)$.

Proof. As was noticed before, the supremum of an l.s.c function defined an l.s.c function, so we only have to establish the continuity (thereby l.s.c) of $E_{t,x} \left( e^{-rt} G(t + \tau, X_{t+\tau}) \right)$ on $[0, T] \times (0, \infty)$ and the assertion will follow immediately.

To achieve this, we begin by showing that (i) $t \mapsto E_{t,x} \left( e^{-rt} G(t + \tau, X_{t+\tau}) \right)$ is continuous on $[0, T)$ for each $x \in (0, \infty)$. Take any $t_1 < t_2$ in $[0, T)$ so that

$$0 \leq E_{t_1,x} \left( e^{-rt} G(t_1 + \tau, X_{t_1+\tau}) \right) - E_{t_2,x} \left( e^{-rt} G(t_2 + \tau, X_{t_2+\tau}) \right) = E \left( e^{-rt} (K - X^x_\tau)^+ (Z(t_1 + \tau, X^x_\tau) - Z(t_2 + \tau, X^x_\tau)) \right) \leq KE \left( Z(t_1 + \tau, X^x_\tau) - Z(t_2 + \tau, X^x_\tau) \right),$$

where the first inequality is based on the map $t \mapsto G(t, x)$ is decreasing, and first equality follows via the time-homogeneous property of GBM; after which, let $t_1 \to t_2$, we see that by Theorem 2.1.6 (function $Z : [0, T] \times (0, \infty) \mapsto [0, 1]$),

$$E_{t_1,x} \left( e^{-rt} G(t_1 + \tau, X_{t_1+\tau}) \right) - E_{t_2,x} \left( e^{-rt} G(t_2 + \tau, X_{t_2+\tau}) \right) \to 0,$$

and assertion (i) follows.

It remains to show that (ii) $t \mapsto E_{t,x} \left( e^{-rt} G(t + \tau, X_{t+\tau}) \right)$ is continuous on $(0, \infty)$ for each $t \in [0, T)$. We simply observe that

$$0 \leq \left| E_{t,x_1} \left( e^{-rt} G(t + \tau, X_{t+\tau}) \right) - E_{t,x_2} \left( e^{-rt} G(t + \tau, X_{t+\tau}) \right) \right| \leq E \left| G(t + \tau, X^x_\tau) - G(t + \tau, X^x_\tau) \right| \leq E \left| Z(t + \tau, X^x_\tau) - Z(t + \tau, X^x_\tau) \right| \leq E \left| X^x_\tau - X^x_\tau \right| + KE \left| Z(t + \tau, X^x_\tau) - Z(t + \tau, X^x_\tau) \right|,$$

where the third inequality follows from $(K - y)^+ - (K - z)^+ \leq (z - y)^+$ for $y, z \in \mathbb{R}$; after which, let $x_1 \to x_2$, Theorem 2.1.6 (it is applicable because of Corollary 5.2.4) suggests

$$E_{t,x_1} \left( e^{-rt} G(t + \tau, X_{t+\tau}) \right) - E_{t,x_2} \left( e^{-rt} G(t + \tau, X_{t+\tau}) \right) \to 0.$$

By combining assertions (i) and (ii), the conclusion follows.

With the aid of Lemma 5.2.7, we can then establish the optimality of $\tau_D$.

Lemma 5.2.8. The stopping time $\tau_D$ is optimal in (5.2.5).

Proof. Note that function $G$ is continuous (thereby u.s.c) on $[0, T) \times (0, \infty)$, which, together with Lemma 5.2.7 and Theorem 3.2.8, yield that

$$\bar{\tau}_D = \inf \{ s \in [0, T-t) : (t + s, X_{t+s}^x) \in \bar{D} \},$$
is optimal (with the convention that the infimum of the empty set is infinite\(^5\)) and \(\bar{D} = D \setminus \{(T,x) : x \in (0,\infty)\}\) in the following problem

\[
\bar{V}(t,x) = \sup_{\tau \in [0,T-t]} E_{t,x} \left( e^{-r\tau} G(t+\tau, X_{t+\tau}) \right),
\]

so that \(\bar{V}(t,x) = E_{t,x} \left( e^{-r\tau_D} G(t+\tau_D, X_{t+\tau_D}) \right)\). The rest of the proof is similar to that of Lemma 4.2.8.

It will be shown in the next section that the continuation and stopping sets can be defined as

\[
\mathcal{C} = \{(t,x) \in [0,T) \times (0,\infty) : x > b(t)\}, \tag{5.2.9}
\]

\[
\mathcal{D} = \{(t,x) \in [0,T) \times (0,\infty) : x \leq b(t)\} \cup \{(T,x) : x \leq b(T)\}, \tag{5.2.10}
\]

where \(b : [0,T] \to \mathbb{R}\) is the unknown optimal stopping boundary such that \(\tau_D\) can be rewritten as

\[
\tau_D = \inf\{0 \leq s \leq T-t : X^x_{t+s} \leq b(t+s)\} \wedge (T-t), \tag{5.2.11}
\]

and by Theorem 3.2.6, Lemma 3.3.2 and Corollary 5.2.8, the free-boundary problem can be formulated as:

\[
V_t + \mathbb{L}_x V - rV = 0 \quad \text{for} \quad (t,x) \in \mathcal{C}, \tag{5.2.12}
\]

\[
V(t,x) = G(t,x) \quad \text{for} \quad x = b(t), \text{instantaneous stopping}, \tag{5.2.13}
\]

\[
V_x(t,x) = G_x(t,x) \quad \text{for} \quad x = b(t) \neq L, \text{smooth-fit}, \tag{5.2.14}
\]

\[
V(t,x) > G(t,x) \quad \text{for} \quad (t,x) \in \mathcal{C}, \tag{5.2.15}
\]

\[
V(t,x) = G(t,x) \quad \text{for} \quad (t,x) \in \mathcal{D}. \tag{5.2.16}
\]

### 5.2.3 The Continuation and Stopping Sets

As a further preparation for proceeding the detailed analysis, we lean on Theorem 2.2.8\(^6\), as the function \(G\) is not differentiable \(K\) and \(L\), to obtain

\[
e^{-rs} G(t+s, X^x_{t+s}) = G(t,x) \\
+ \int_0^s e^{-ru} (-rG + G_t + \mathbb{L}_x G)(t+u, X^x_{t+u}) I\{X_{t+u} \neq K, X_{t+u} \neq L\} du \\
+ \int_0^s e^{-ru} \sigma X^x_{t+u} G_x(t+u, X^x_{t+u}) I\{X_{t+u} \neq K, X_{t+u} \neq L\} dW_u \\
+ \frac{1}{2} \int_0^s e^{-ru} (G_x(t+u, L+) - G_x(t+u, L-)) dl^L_u(X^x) \\
+ \frac{1}{2} \int_0^s e^{-ru} (G_x(t+u, K+) - G_x(t+u, K-)) dl^K_u(X^x), \tag{5.2.17}
\]

\(^5\)That is, \(G(\infty, X_\infty) = 0\) as assumed in Chapter 2.

\(^6\)The technical conditions for its application shall be fairly easy to check as \(G\) is known.
where $t^K_s$ and $t^L_s$ are the local times of $X$ at the level $K$ and $L$ respectively, and for simplicity, we denote the martingale term as

$$M_s = \int_0^s e^{-ru} \sigma X_{t+u}^x(t + u, X_{t+u}^x) I \{ X_{t+u} \neq K, X_{t+u} \neq L \} dW_u,$$

so that $E_{t,x}(M_s) = 0$ under measure $P_{t,x}$ for each $s \in [0, T-t]$.

A fairly easy calculation then shows that

$$G_x(t, L+) = -1,$$
$$G_x(t, L-) = -1 + (K - L)Z_x(t, L-),$$
$$G_x(t, K+) = 0,$$
$$G_x(t, K-) = -1,$$

and that, upon putting $H(t, x) = (-rG + G_t + \mathbb{1}_x G)(t, x)$ and using Proposition 5.2.3,

$$H(t, x) = \begin{cases} (-rKZ - \sigma^2 x^2 Z_x)(t, x) & \text{for } x < L < K, \\ -rK & \text{for } L \leq x < K, \\ 0 & \text{for } x > K, \end{cases}$$

(5.2.18)

after which, (5.2.17) is reshuffled accordingly as

$$e^{-rs}G(t + s, X_{t+s}^x) = G(t, x) + M_s + \int_0^s e^{-ru} H(t + u, X_{t+u}^x) I \{ X_{t+u} \neq L, X_{t+u} \neq K \} du$$
$$+ \frac{1}{2} \int_0^s e^{-ru} dt^K_u(X^x) - \frac{1}{2} \int_0^s e^{-ru}(K - L)Z_x(t + u, L-) du \mathbb{1}_u(X) + (X^x).$$

**Remark 5.2.9.** Note how very nicely the property of function $Z$ in Proposition 5.2.3 has saved us from massive computations in the finite-time problem setup!

Using this, we find that

$$E_{t,x}(e^{-rs}G(t + s, X_{t+s}^x)) = G(t, x) + E_{t,x}\left( -rK \int_0^s e^{-ru} I \{ L \leq X_{t+u} < K \} du \right)$$
$$+ E_{t,x}\left( \int_0^s e^{-ru} \left( -rKZ - \sigma^2 X_{t+u}^2 Z_x \right)(t + u, X_{t+u}) I \{ X_{t+u} < L \} du \right)$$
$$- \frac{1}{2} E_{t,x}\left( \int_0^s e^{-ru}(K - L)Z_x(t + u, L-) du \mathbb{1}_u(X) \right) + \frac{1}{2} E_{t,x}\left( \int_0^s e^{-ru} dt^K_u(X) \right),$$

(5.2.19)

from which, what intuition would suggest is the following:

**Lemma 5.2.10.** All points $(t, x) \in [0, T) \times (K, \infty)$ belong to the continuation set $C$.

**Proof.** The proof follows, in fact, from the verbal statement: since stopping at $[0, T) \times (K, \infty)$ gives one null payoff, while waiting, there is a positive probability to collect strictly positive payoff.

**Remark 5.2.11.** Another message, as we shall see in a minute, of equation (5.2.19) is that the optimal stopping sets over $[0, T]$ depends on the relative positions of the level $L$ and of the optimal stopping boundary $B(t)$ of the standard American put option with payoff $G^A(x) = (K - x)^+$. (See [11] and [21].)
For notational convenience, we let \( V^A(t, x) \), \( G^A(x) = (K - x)^+ \) and \( B(t) \) denote the value function, the gain function and the optimal stopping boundary for the standard American put option with the same maturity and strike price as the current contract. Then, we know from [64] that

\[
V^A(t, x) > G^A(x) \quad \text{for } x > B(t),
V^A(t, x) = G^A(x) \quad \text{for } x \leq B(t),
\]

where \( B(t) \) is unique and \( B(T^-) = K \); and thus, its stopping set \( D^A \) and continuation set \( C^A \) equal:

\[
D^A = \{(t, x) \in [0, T] \times (0, \infty) : x \leq B(t)\},
C^A = \{(t, x) \in [0, T] \times (0, \infty) : x > B(t)\}.
\]

Also we define \( t_\ast \in [0, T) \) as follows:

\[
\begin{cases}
  t_\ast > 0, & \text{if } B(t_\ast) = L, \\
  t_\ast = 0, & \text{if } B(0) > L.
\end{cases}
\]

and since the map \( t \mapsto B(t) \) is strictly increasing, \( t_\ast \) is unique. Moreover, let us emphasise that \( t_\ast < T \), as \( L < K \).

**Lemma 5.2.12.** All points \((t, x) \in [t_\ast, T] \times [L, B(t)]\) belongs to the stopping set \( D \).

**Proof.** Since \((t, x) \in [t_\ast, T] \times [L, B(t)]\), it follows from the construction of the stopping set for the standard American put option that

\[
V^A(t, x) = G^A(x),
\]

and notice that in this set, \( G^A(x) = G(t, x) \).

Now, knowing that

\[
V(t, x) = \sup_{0 \leq \tau \leq T-t} E_{t,x} \left( e^{-\tau r} Z(t, t+\tau, X(t+\tau))(K - X(t+\tau))^+ \right) \\
\leq \sup_{0 \leq \tau \leq T-t} E_{t,x} \left( e^{-\tau r} (K - X(t+\tau))^+ \right) = V^A(t, x),
\]

where the inequality is due to function \( Z \) being a probability and consequently,

\[
0 = V^A(t, x) - G^A(x) \geq V(t, x) - G(t, x),
\]

on the set \( \{(t, x) \in [t_\ast, T] \times [L, B(t)]\} \), which in the sense of \( V \geq G \) forces the equation \( V = G \) and thereby, proving the desired assertion.

With a little additional work, we can extend Lemma 5.2.12 to the following result:

**Lemma 5.2.13.** All points \((t, x) \in [t_\ast, T] \times (0, B(t)]\) belongs to the stopping set \( D \).
Proof. The proof amounts to showing that $[t_s, T] \times [0, L]$ also belongs to the stopping set $D$.

Lemma 5.2.12 tells us that $[t_s, T] \times \{L\} \subset D$.

Suppose that we run the process $X$ starting at point $(t, x) \in [t_s, T] \times (0, L)$, and we see from (5.2.19) that in order to compensate the negative term inside the expectation, the process shall hit level $K$ as the local time term is positive, but $L < K$, indicating the process cannot hit level $K$ before exercising at level $L$. Therefore, it is optimal to stop immediately on this set and the conclusion follows. 

\[ \begin{array}{c}
\includegraphics[width=0.8\textwidth]{Figure52.png} \\
\text{Figure 5.2: Two sketches to illuminate Lemmas 5.2.12 and 5.2.13 for } t_s = 0. \text{ The area coloured in yellow is the subset of } D \text{ and that } D^A \subset D. 
\end{array} \]

\[ \begin{array}{c}
\includegraphics[width=0.8\textwidth]{Figure53.png} \\
\text{Figure 5.3: Two sketches to illuminate Lemmas 5.2.12 and 5.2.13 for } t_s > 0. \text{ The area coloured in green is the subset of } D. 
\end{array} \]

Remark 5.2.14. Lemmas 5.2.12 and 5.2.13 show that the stopping set is not empty and that if $t_s = 0$, then all points $(t, x) \in [0, T] \times (0, L]$ belongs to the stopping set $D$, in the sense that

\[ [0, T] \times [L, B(t)] \subset [0, T] \times (0, B(t)], \]
so that in \([0, T] \times [L, B(t)]\), the relation \(V^A(t, x) = G^A(x)\) and \(G^A(x) = G(t, x)\) once again entails \(V = G\), after which, the similar argument as Lemma 5.2.13 shows \((t, x) \in [0, T] \times (0, L]\) belongs to \(D\).

With Remark 5.2.14 in mind, we now enter the discussions of the properties for sets \(C\) and \(D\) with \(t_0 = 0\) and \(t_0 > 0\).

**In the Case of \(t_0 = 0\)**

**Proposition 5.2.15.** The stopping set \(D\) is right-connected and the continuation set \(C\) is left-connected.

**Proof.** Let \(0 \leq t_1 \leq t_2 \leq T\) and consider the stopping time \(\tau_L = \inf\{0 \leq u \leq T-t_2 : X_{t_2}^{X_{t_1}^u+u} \leq L\}\) and suppose that \(\tau_2\) is the optimal stopping time for \(V(t_2, x)\) so that \(\tau_2 < \tau_L\), which is a consequence of Remark 5.2.14. Then, by the time-homogeneous property of GBM,

\[
V(t_1, x) - V(t_2, x) \geq E\left(e^{-r\tau_2} G\left(t_1 + \tau_2, X_{t_1}^{X_{t_1}^u+u}\right) - e^{-r\tau_2} G\left(t_2 + \tau_2, X_{t_2}^{X_{t_2}^u+u}\right)\right)
\]

\[
= G(t_1, x) - G(t_2, x) + E\left(\int_0^{\tau_2} e^{-ru} (H(t_1 + u, X_u^x) - H(t_2 + u, X_u^x)) I\{L < X_u^K\} du\right)
\]

where the second equality follows from \(H(t_1 + u, X_u^x) = H(t_2 + u, X_u^x) = -rK\). This establishes the following relation:

\[
V(t_1, x) - G(t_1, x) \geq V(t_2, x) - G(t_2, x),
\]

from which, we see that \((t_1, x) \in D\) implies \((t_2, x) \in D\) and that \((t_2, x) \in C\) implies \((t_1, x) \in C\).

**Proposition 5.2.16.** The stopping set \(D\) is down-connected and the continuation set \(C\) is up-connected.

**Proof.** Immediate from Proposition 5.2.15 and the similar argument of the proof in Lemma 5.2.13.

We now have sufficient mathematics at our disposal to go around the singularity of \(Z_x(t, x)\) for the establishment of the continuity of the value function for \(t_0 = 0\).

**Lemma 5.2.17.** The value function \(V\) is continuous on \([0, T] \times (0, \infty)\).

**Proof.** The proof involves proving two assertions: (i) The map \(t \mapsto V(t, x)\) is continuous on \([0, T]\) for each \(x \in (0, \infty)\) given and fixed. Take any \(t_1 < t_2\) in \([0, T]\), let \(\epsilon > 0\) and \(\tau_1^t\) be the stopping time such that

\[
E_{t_1, x}\left(e^{-r\tau_1^t} G\left(t_1 + \tau_1^t, X_{t_1+\tau_1^t}\right)\right) \geq V(t_1, x) - \epsilon.
\]

\(^7\)That is, \(L < B(0)\).
Then, by setting \( \tau_2^* = \tau_1^* \land (T - t_2) \), we have

\[
E_{t_2,x} \left( e^{-rt_2^*} G \left( t_2 + \tau_2^*, X_{t_2 + \tau_2^*} \right) \right) \leq V(t_2, x),
\]

and that

\[
0 \leq V(t_1, x) - V(t_2, x) \\
\leq E \left( e^{-r\tau_1^*} G \left( t_1 + \tau_1^*, X_{t_1 + \tau_1^*} \right) \right) - E \left( e^{-r\tau_2^*} G \left( t_2 + \tau_2^*, X_{t_2 + \tau_2^*} \right) \right) + \epsilon \\
\leq E \left( e^{-r\tau_2^*} \left( G \left( t_1 + \tau_1^*, X_{\tau_1^*} \right) - G \left( t_2 + \tau_2^*, X_{\tau_2^*} \right) \right) \right) + \epsilon \\
\leq E \left( \left( X_{\tau_1^*} - X_{\tau_2^*} \right)^+ \left( Z \left( t_1 + \tau_1^*, X_{\tau_1^*} \right) - Z \left( t_2 + \tau_2^*, X_{\tau_2^*} \right) \right) \right) + \epsilon,
\]

(5.2.20)

where the first inequality is because of the map \( t \mapsto V(t, x) \) being decreasing and the last inequality holds via

\[
(K - y)^+ - (K - z)^+ \leq (z - y)^+ \quad \text{for } y, z \in \mathbb{R}.
\]

(5.2.21)

By letting \( t_1 \to t_2 \), using \( \tau_1^* \to \tau_2^* \) and then \( \epsilon \to 0 \) on (5.2.20), Theorem 2.1.9 allows us to conclude that (see Example 2.2.3 for its uniform integrability)

\[
V(t_1, x) - V(t_2, x) \to 0,
\]

and assertion (i) follows.

(ii) The map \( x \mapsto V(t, x) \) is continuous on \((0, \infty)\) for each \( t \in [0, T) \) given and fixed.

Note that by the up-down connectedness of \( \mathcal{C} \) and \( \mathcal{D} \), for any \( x_1 < x_2 \) in \((0, \infty)\),

\[
G(t, x_2) - G(t, x_1) \leq V(t, x_2) - V(t, x_1),
\]

and let \( \tau_2 \) be optimal for \( V(t, x_2) \) such that

\[
G(t, x_2) - G(t, x_1) \leq V(t, x_2) - V(t, x_1) \\
\leq E \left( e^{-r\tau_2} G(t + \tau_2, X_{\tau_2}^x) \right) - E \left( e^{-r\tau_2} G(t + \tau_2, X_{\tau_2}^x) \right),
\]

(5.2.22)

after which, by letting \( x_1 \to x_2 \), the dominated convergence theorem leads us to

\[
V(t, x_2) - V(t, x_1) \to 0
\]

and thereby, proving the left-continuity of \( V \).

And, needless to say, if \( x_2 \leq B(t) \), Lemma 5.2.13 tells us that \( \tau_2 = 0 \) and the continuity of \( V \) follows directly from that of the gain function \( G \).

On the other hand, if \( x_2 > B(t) \), we know there exists \( \delta > 0 \) so that \( B(t) > L + \delta > L \) (see figure 5.2) and that for \( 0 < x_2 - x_1 < \frac{x_2 \delta}{2L} \), immediately from the monotonicity of \( B \), we have

\[
X_{\tau_2}^x \geq B(t + \tau_2) > B(t) > L + \delta.
\]
which implies \( X_{\tau_2}^{x_1} > L \) a.s.; more precisely, by the strong solution of GBM and the optimality of \( \tau_2 \), \( X_{\tau_2}^{x_2} > L + \delta \) entails
\[
e^{\left( r - \frac{\sigma^2}{2}\right) \tau_2 + \sigma W_{\tau_2}} > \frac{L + \delta}{x_2},
\]
and that for \( x_2 \in (0, x_1 (1 + \delta/(2L))) \),
\[
X_{\tau_2}^{x_1} = x_1 e^{\left( r - \frac{\sigma^2}{2}\right) \tau_2 + \sigma W_{\tau_2}} > \frac{x_1 (L + \delta)}{x_2} > \frac{x_1 (L + \delta)}{x_1 (1 + \delta/(2L))} = L + \frac{\delta L}{\delta + 2L} > L,
\]
after which, it follows that as \( Z(t, x) = 1 \) for \( x > L \),
\[
G(t, x_2) - G(t, x_1) \leq V(t, x_2) - V(t, x_1)
\]
\[
\leq E \left( e^{-r \tau_2} G(t + \tau_2, X_{\tau_2}^{x_2}) \right) - E \left( e^{-r \tau_2} G(t + \tau_2, X_{\tau_2}^{x_1}) \right)
\]
\[
\leq (x_1 - x_2)^+ E \left( e^{-\frac{\sigma^2}{2} \tau_2 + \sigma W_{\tau_2}} \right) = (x_1 - x_2)^+,
\]
and then by letting \( x_2 \to x_1 \), we have \( V(t, x_2) - V(t, x_1) \to 0 \), that is, the right-continuity of \( V \). By combining the claims that \( V \) is left and right-continuous for \( x \in (0, \infty) \) for each \( t \in [0, T) \) given and fixed, statement (ii) follows.

By taking advantage of the above analysis, we obtain the following result:

**Proposition 5.2.18 (The Free-boundary Problem).** The stopping set and the continuation set are of the forms:

\[
C = \{(t, x) \in [0, T) \times (0, \infty) : x > b(t)\},
\]
\[
D = \{(t, x) \in [0, T) \times (0, \infty) : x \leq b(t)\} \cup \{(T, x) : x \leq b(T)\},
\]

where the map \( t \mapsto b(t) \) is increasing. The free-boundary problem is rearranged as follows:

\[
V_t + L_x V - r V = 0 \quad \text{for } (t, x) \in C, \quad (5.2.23)
\]
\[
V(t, x) = K - x \quad \text{for } x = b(t), \text{ instantaneous stopping}, \quad (5.2.24)
\]
\[
V_x(t, x) = -1 \quad \text{for } x = b(t), \text{ smooth-fit}, \quad (5.2.25)
\]
\[
V(t, x) > G(t, x) \quad \text{for } (t, x) \in C, \quad (5.2.26)
\]
\[
V(t, x) = G(t, x) \quad \text{for } (t, x) \in D. \quad (5.2.27)
\]

Just to complete the picture of the free-boundary problem, we now justify equation (5.2.25).

**Proof of the Smooth-fit Condition.** Let a point \( (t, x) \in (0, T) \times (0, \infty) \) lying on the boundary \( b \) be fixed, i.e. \( x = b(t) \). Then, \( L < x < K \) and for all \( \epsilon > 0 \) such that \( L < x + \epsilon < K \), we have
\[
\frac{V(t, x + \epsilon) - V(t, x)}{\epsilon} \geq \frac{G(t, x + \epsilon) - G(t, x)}{\epsilon} = \frac{(K - x - \epsilon) - (K - x)}{\epsilon} = -1,
\]
of which, taking the limit as \( \epsilon \to 0 \), we have \( \lim_{\epsilon \to 0} \frac{V(t, x + \epsilon) - V(t, x)}{\epsilon} \geq -1 \).
In order to prove the converse inequality, we fix a sufficiently small $\epsilon$ so that $0 < \epsilon < \frac{x-L}{2}$ and that $L < x + \epsilon < K$, and consider the optimal stopping time $\tau_{\epsilon}$ for $V(t, x + \epsilon)$ so that the monotonicity of $b$ tells us that $X_{\tau_{\epsilon}}^{x+\epsilon} = b(t + \tau_{\epsilon}) > b(t) \geq L$ implies $X_{\tau_{\epsilon}}^{x} > L$ almost surely, as before, by the strong solution of GBM,

$$e^{\left(-\frac{\sigma}{2} \tau_{\epsilon}\right)} \frac{x}{x + \epsilon} > \frac{b(t)}{x + \epsilon} = \frac{x}{x + \epsilon},$$

so that, for $x > L$,

$$x e^{\left(-\frac{\sigma}{2} \tau_{\epsilon}\right)} \frac{x^2}{x + \epsilon} = \frac{x^2 - \epsilon^2 + \epsilon^2}{x + \epsilon} > x - \epsilon > x - \frac{x-L}{2} = \frac{x+L}{2} > L.$$  

Then, we have

$$V(t, x + \epsilon) - V(t, x) \leq \frac{1}{\epsilon} E\left(e^{-\tau_{\epsilon}} G(t + \tau_{\epsilon}, X_{\tau_{\epsilon}}^{x+\epsilon}) - E\left(e^{-\tau_{\epsilon}} G(t + \tau_{\epsilon}, X_{\tau_{\epsilon}}^{x})\right)\right) \leq \frac{1}{\epsilon} E\left(e^{-\tau_{\epsilon}} \left((K - X_{\tau_{\epsilon}}^{x+\epsilon})^+ - (K - X_{\tau_{\epsilon}}^{x})^+\right) Z(t + \tau_{\epsilon}, X_{\tau_{\epsilon}}^{x+\epsilon})\right) + e^{-\tau_{\epsilon}} \left((K - X_{\tau_{\epsilon}}^{x+\epsilon})^+ \left(Z(t + \tau_{\epsilon}, X_{\tau_{\epsilon}}^{x+\epsilon}) - Z(t + \tau_{\epsilon}, X_{\tau_{\epsilon}}^{x})\right)\right) \leq \frac{1}{\epsilon} E\left(e^{-\tau_{\epsilon}} \left(X_{\tau_{\epsilon}}^{x} - X_{\tau_{\epsilon}}^{x+\epsilon}\right) Z(t + \tau_{\epsilon}, X_{\tau_{\epsilon}}^{x+\epsilon}) I\{X_{\tau_{\epsilon}}^{x+\epsilon} < K\} + e^{-\tau_{\epsilon}} \left(K - X_{\tau_{\epsilon}}^{x}\right)^+ \left(Z(t + \tau_{\epsilon}, X_{\tau_{\epsilon}}^{x+\epsilon}) - Z(t + \tau_{\epsilon}, X_{\tau_{\epsilon}}^{x})\right)\right) \leq -E\left(e^{-\frac{\sigma^2}{2} \tau_{\epsilon}} + \sigma W_{\tau_{\epsilon}} Z(t + \tau_{\epsilon}, X_{\tau_{\epsilon}}^{x+\epsilon}) I\{X_{\tau_{\epsilon}}^{x+\epsilon} < K\}\right) + \frac{1}{\epsilon} E\left(e^{-\tau_{\epsilon}} \left(K - X_{\tau_{\epsilon}}^{x}\right)^+ \left(Z(t + \tau_{\epsilon}, X_{\tau_{\epsilon}}^{x+\epsilon}) - Z(t + \tau_{\epsilon}, X_{\tau_{\epsilon}}^{x})\right)\right) \leq -E\left(e^{-\frac{\sigma^2}{2} \tau_{\epsilon}} + \sigma W_{\tau_{\epsilon}} Z(t + \tau_{\epsilon}, X_{\tau_{\epsilon}}^{x+\epsilon}) I\{X_{\tau_{\epsilon}}^{x+\epsilon} < K\}\right) + \frac{1}{\epsilon} E\left((K - X_{\tau_{\epsilon}}^{x})^+ \left(Z(t + \tau_{\epsilon}, X_{\tau_{\epsilon}}^{x+\epsilon}) - Z(t + \tau_{\epsilon}, X_{\tau_{\epsilon}}^{x})\right)\right) \leq -E\left(e^{-\frac{\sigma^2}{2} \tau_{\epsilon}} + \sigma W_{\tau_{\epsilon}} Z(t + \tau_{\epsilon}, X_{\tau_{\epsilon}}^{x+\epsilon}) I\{X_{\tau_{\epsilon}}^{x+\epsilon} < K\}\right) + \frac{K}{\epsilon} E\left(I\{X_{\tau_{\epsilon}}^{x} < L\} \left(Z(t + \tau_{\epsilon}, X_{\tau_{\epsilon}}^{x+\epsilon}) - Z(t + \tau_{\epsilon}, X_{\tau_{\epsilon}}^{x})\right)\right) \leq -E\left(e^{-\frac{\sigma^2}{2} \tau_{\epsilon}} + \sigma W_{\tau_{\epsilon}} Z(t + \tau_{\epsilon}, X_{\tau_{\epsilon}}^{x+\epsilon}) I\{X_{\tau_{\epsilon}}^{x+\epsilon} < K\}\right)$$

where the second inequality follows from

$$\left(K - X_{\tau_{\epsilon}}^{x+\epsilon}\right)^+ - (K - X_{\tau_{\epsilon}}^{x})^+ = \left(X_{\tau_{\epsilon}}^{x} - X_{\tau_{\epsilon}}^{x+\epsilon}\right) I\{X_{\tau_{\epsilon}}^{x+\epsilon} < K\} - (K - X_{\tau_{\epsilon}}^{x})^+ I\{X_{\tau_{\epsilon}}^{x+\epsilon} \geq K\} \leq \left(X_{\tau_{\epsilon}}^{x} - X_{\tau_{\epsilon}}^{x+\epsilon}\right) I\{X_{\tau_{\epsilon}}^{x+\epsilon} < K\}, \quad (5.2.28)$$

and the second equality is from the strong solution of $X$ and linearity; moreover, the fourth inequality is due to the fact that $I\{X_{\tau_{\epsilon}}^{x} \geq L\} = 1$ entails $I\{X_{\tau_{\epsilon}}^{x+\epsilon} \geq L\} = 1$ and that $Z(t + \epsilon, X_{\tau_{\epsilon}}^{x}) = Z(t + \epsilon, X_{\tau_{\epsilon}}^{x+\epsilon}) = 1$, while we used $I\{X_{\tau_{\epsilon}}^{x} < L\} = 0$ for chosen $\epsilon$ in the last step. Via an appeal
to the dominated convergence theorem, upon letting $\epsilon \to 0$, we have $\tau_\epsilon \to 0$ (see A.2.2 for its justification) so that
\[
E \left( e^{-\frac{\sigma^2 \epsilon^2 + \sigma W_{\tau_\epsilon}}{2}} Z(t + \tau_\epsilon, X^{x+\epsilon}_{\tau_\epsilon}) I\{X^{x+\epsilon}_{\tau_\epsilon} < K\} \right) \to 1.
\]
Thus,
\[
\lim_{\epsilon \to 0} \frac{V(t, x + \epsilon) - V(t, x)}{\epsilon} \leq -1,
\]
after which, the conclusion follows via the fact that for $(t, x)$ and $(t, x - \epsilon)$ belong to $\mathcal{D}$,
\[
\frac{V(t, x) - V(t, x - \epsilon)}{\epsilon} = \frac{G(t, x) - G(t, x - \epsilon)}{\epsilon} = -1.
\]

\[\text{Remark 5.2.19. The noteworthy feature of the value function is that it is not smooth in the stopping set $\mathcal{D}$, to be more specific, at level $L$. This fact implies the presence of a local time term in the contract pricing formula. (see [21, Page 7] and [70] for similar results)}\]

To prepare for the main theorem presented in Subsection 5.2.4, we further justify the continuity of the optimal stopping boundary.

**Lemma 5.2.20.** The optimal stopping boundary $b$ is continuous on $[0, T]$ and $b(T^-) = K$.

**Proof.** The proof proceeds in three steps, which follows essentially from [64, Page 383].

(i) The boundary $b$ is right-continuous on $[0, T]$.

Let us fix $t \in [0, T)$ and take a sequence $t_n \downarrow t$ as $n \to \infty$. Since $b$ is increasing, the right-limit $b(t^+)$ exists. Remember that $\mathcal{D}$ is a closed set, meaning that its limit point $\lim_{n \to \infty} (t_n, b(t_n)) \to (t, b(t^+))$ is contained in $\mathcal{D}$, it then follows, together with the construction of the stopping set that $b(t^+) \leq b(t)$. However, the fact that $b$ is increasing suggests that $b(t^+) \geq b(t)$ and thereby, forcing the equation $b(t^+) = b(t)$ and proving statement (i).

(ii) The boundary $b$ is left-continuous on $[0, T)$.

Assume that, for contradiction, at some point $t \in (0, T)$, the function makes a jump, that is, $b(t) > b(t^-)$. Then, fix a point $y_s \in (b(t^-), b(t))$. By (5.2.24) and (5.2.25), we have
\[
V(s, y_s) - G(s, y_s) = \int_{b(s)}^{y_s} \int_{b(s)}^{x} (V_{xx} - G_{xx}) (s, z)dzdx,
\]
for each $s \in (t-\delta, t)$, where $\delta > 0$ and $t-\delta > 0$. Despite the fact that function $G$ is not differentiable at $L$, the fact that $b(t) > L$ once again comes to our rescue.

Knowing that the value function satisfies (5.2.23), we see that, for $(t, x) \in \mathcal{R}$, where $\mathcal{R}$ is a curved trapezoid formed by the vertices $(s, b(s))$, $(s, y_s)$, $(t, b(t^-))$ and $(t, y_s)$,
\[
V_{xx}(t, x) = \frac{2}{\sigma^2 x^2} \left( rV - V_t - rxV_x \right) (t, x) > \frac{2}{\sigma^2 x^2} rV(t, x) > \frac{2}{\sigma^2 x^2} rG(t, x) \geq c > 0,
\]
where the first inequality follows via the fact that both maps \( t \mapsto V(t, x) \) and \( x \mapsto V(t, x) \) are decreasing on \( \mathcal{R} \) and the last one is due to the fact that \( x < K \) on \( \mathcal{R} \). Moreover, \( G_{xx}(t, x) = 0 \) on \( \mathcal{R} \).

Then, from the continuity of the value function and the gain function, upon letting \( s \uparrow t \) on (5.2.29) and using dominated convergence theorem,

\[
V(t-, y_*) - G(t-, y_*) = V(t, y_*) - G(t, y) \geq \int_{b(t-)}^{y_*} \int_{b(t-)}^{x} cdzdx \\
= \frac{c(y_* - b(t-))^2}{2} > 0,
\]

which is a contradiction, since \((t, y_*) \in \mathcal{D}\), indicating that such point cannot exist and \( b(t-) = b(t) \) must hold true. Immediate from statements (i) and (ii), the continuity of \( b \) follows.

(iii) \( b(T-) = K \).

In Lemma 5.2.10, we learned that \( b(T-) \in [L, K] \).

Assume that, \( b(T-) < K \) so that there exists a point \( y_* \in (b(T-), K) \) and let \( s \in [T - \delta, T) \) with \( 0 < \delta < T \). Then, by rerunning the proof of statement (ii) with the above modifications and letting \( s \uparrow T \), we find that from the continuity of \( V \) and \( G \) on \([0, T] \times [L, \infty)\),

\[
V(T, y_*) - G(T, y_*) > 0,
\]

which contradicts the fact that \( \{(T, x) : x < b(T)\} \subset \mathcal{D} \) and thus implies that \( b(T-) = K \) (because of its continuity, \( b(T) = K \)).

**In the Case of** \( 0 < t_* < T \)

**Proposition 5.2.21.** On the state space \([0, T] \times [L, \infty)\), the stopping set \( \mathcal{D} \) is right-connected and the continuation set \( \mathcal{C} \) is left-connected.

**Proof.** Since on the state space \([0, T] \times [L, \infty)\), the map \( t \mapsto V(t, x) \) is decreasing and that \( G(t, x) = (K-x)^+ \), the map \( t \mapsto V(t, x) - G(t, x) \) is decreasing also. That is, for any \( t_1, t_2 \in [0, T] \), if \( t_1 < t_2 \), then

\[
V(t_1, x) - G(t_1, x) \geq V(t_2, x) - G(t_2, x),
\]

such that \((t_1, x) \in \mathcal{D}\) implies \((t_1, x) \in \mathcal{D}\) and \((t_2, x) \in \mathcal{C}\) implies \((t_1, x) \in \mathcal{C}\), which is precisely the claim. \( \square \)

Proposition 5.2.21 immediately tells us the following result, which we state without proof:

**Corollary 5.2.22.** On the state space \([0, T] \times [L, \infty)\), \( \mathcal{D}^A \subseteq \mathcal{D} \).

To fill the last slot of the diagram, we shall show that the stopping set \( \mathcal{D} \) is not empty on state space \([0, t_*] \times (0, L]\), but first we present some important facts:

---

\(^8\)That is, \( L > B(0) \).
Lemma 5.2.23. (i) The gain function $G$ is uniformly continuous on the state space $[0, t_s] \times (0, L]$:

(ii) The map $t \mapsto \max_{x \in (0, L]} G(t, x)$ is continuous on $[0, t_s]$.

(iii) Given that for each $t \in [0, t_s]$, $\max_{x \in (0, L]} G(t, x)$ is reached at a unique point $x^*(t)$, in particular,

$$\max_{x \in (0, L]} G(t, x) = G(t, x^*(t)),$$

(5.2.30)

the map $t \mapsto x^*(t)$ is continuous on $[0, t_s]$.

Remark 5.2.24. Knowing from derivative computation that for each fixed $t$, the maximum of $G$ achieved in $(0, L]$ is the global maximum over the domain $(0, \infty)$, we can write

$$\max_{x \in (0, \infty)} G(t, x) = \max_{x \in (0, L]} G(t, x).$$

As a matter of fact, $t \mapsto x^*(t)$ is non-decreasing, see Figure 5.4.

Proof of Lemma 5.2.23.

(i) To show that $G$ is uniformly continuous on $[0, t_s] \times (0, L]$, is to show that for every $\epsilon > 0$, there exists a $\delta > 0$ so that for all $(t_1, x_1), (t_2, x_2) \in [0, t_s] \times (0, L]$, $|t_1 - t_2| + |x_1 - x_2| < \delta$ implies $|G(t_1, x_1) - G(t_2, x_2)| < \epsilon$.

Notice that

$$|G(t_1, x_1) - G(t_2, x_2)| \leq |x_2 - x_1| Z(t_1, x_1) + |K - x_2| Z(t_1, x_1) - Z(t_2, x_2)|$$

$$< |x_2 - x_1| + K|Z(t_1, x_1) - Z(t_2, x_2)|.$$

Observe that on $[0, t_s] \times (0, L]$,

$$|Z(t_1, x_1) - Z(t_2, x_2)| = |Z(t_1, x_1) - Z(t_2, x_1) + Z(t_2, x_1) - Z(t_2, x_2)|$$

$$\leq |Z(t_1, x_1) - Z(t_2, x_1)| + |Z(t_2, x_1) - Z(t_2, x_2)|$$

$$\leq M|t_1 - t_2| + N|x_1 - x_2|,$$

where the second inequality holds true by the mean value theorem and the fact that the partial derivatives of $Z_x(t, x)$ and $Z_t(t, x)$ are bounded, say by constants $N$ and $M$, on the set $[0, t_s] \times (0, L]$ (see A.1.2).

Hence,

$$|G(t_1, x_1) - G(t_2, x_2)| < KM|t_1 - t_2| + (KN + 1)|x_1 - x_2|,$$

and we can now choose $\delta = \frac{\epsilon}{\max\{KM, KN + 1\}}$ and verify that if $(t_1, x_1), (t_2, x_2) \in [0, t_s] \times (0, L]$ satisfy $|t_1 - t_2| + |x_1 - x_2| < \delta$, then

$$|G(t_1, x_1) - G(t_2, x_2)| < \epsilon,$$

thus proving statement (i) as desired.
(ii) The uniform continuity of $G$ means that for every $\epsilon > 0$, there exists a $\delta > 0$ such that for all $t_1, t_2 \in [0, t_\star]$ and $x \in (0, L]$,

$$|t_1 - t_2| < \delta \implies |G(t_1, x) - G(t_2, x)| < \epsilon,$$

which, after rearranging, equals

$$G(t_2, x) - \epsilon \leq G(t_1, x) \leq G(t_2, x) + \epsilon,$$

so we must have,

$$\max_{x \in (0, L]} G(t_1, x) \geq \max_{x \in (0, L]} G(t_2, x) - \epsilon,$$

$$\max_{x \in (0, L]} G(t_1, x) \leq \max_{x \in (0, L]} G(t_2, x) + \epsilon,$$

after which, the conclusion follows from

$$|\max_{x \in (0, L]} G(t_1, x) - \max_{x \in (0, L]} G(t_2, x)| < \epsilon.$$

(iii) Assume that, for contradiction, statement (iii) is false. To say that, $t \mapsto x^\star(t)$ is not continuous on $[0, t_\star]$ means that there exists a sequence $(t_n) \subseteq [0, t_\star]$ where $t_n \to t$ (Since the the set is compact, the limit point $t$ is always contained on $[0, t_\star]$,) such that $x^\star(t_n)$ does not converge $x^\star(t)$.

According to statement (ii) and the definition of $x^\star(t)$,

$$\lim_{t_n \to t} \max_{x \in (0, L]} G(t_n, x) \to \max_{x \in (0, L]} G(t, x) = G(t, x^\star(t)),$$

but the assumption is

$$\lim_{t_n \to t} \max_{x \in (0, L]} G(t_n, x) \to \max_{x \in (0, L]} G(t, x) = G\left(t, \lim_{t_n \to t} x^\star(t_n)\right) \neq G(t, x^\star(t)),$$

which is a contradiction, we may therefore conclude that statement (iii) holds true.

What Lemma 5.2.23 tells us is that:

**Lemma 5.2.25.** All points $(t, x) \in [0, t_\star] \times (0, x^\star(t)]$ belongs to the stopping set $\mathcal{D}$.

**Proof.** According to the definition of $x^\star$, Remark 5.2.24 and the fact that $t \mapsto G(t, x)$ is decreasing,

$$G(t, x^\star(t)) \geq \max_{x \in (0, \infty)} G(t + \tau, x) \geq G(t + \tau, x),$$

and by domination, taking expectation under measure $P_{t, x^\star(t)}$,

$$G(t, x^\star(t)) \geq E_{t, x^\star(t)}\left(e^{-r\tau}G(t + \tau, X_{t+\tau})\right),$$

so that $G(t, x^\star(t)) \geq V(t, x^\star(t))$, but $V \geq G$ forces the equation $G(t, x^\star(t)) = V(t, x^\star(t))$, and thus showing that all $(t, x^\star(t)) \in \mathcal{D}$. Furthermore, due to Lemma 5.2.23 and Remark 5.2.24, an analogous argument as Lemma 5.2.13 does the rest for us.
Figure 5.4: This figure displays the maps $t \mapsto x^*(t)$ with chosen parameters $r = 0.05$, $\sigma = 0.4$, $L = 6.5$, $K = 7$, $T = 5$, $t_* \approx 4.98$.

Figure 5.5: Two sketches to illuminate Lemmas 5.2.25 for $t_* > 0$. The areas coloured in green is $S \subset D$ and the blue dashed lines are plotting function $x^* : [0, t_*] \mapsto (0, L]$. The right figure is to illuminate the possible case stated in Remark 5.2.36.

Remark 5.2.26. It has been high-lighted in multiple literatures (if not all of them, such as [52, Page 113]) that if either the maps $t \mapsto H(t, x)$ or $x \mapsto H(t, x)$ are monotone (the later monotonicity is used to establish the up-connectedness of the sets), we will not be able to compare two integrals with different signs and proves most of the properties possessed by the optimal stopping boundary, which is probably the only downfall of such analytical method but opens the door to many interesting research possibilities; this is, unfortunately, the story of $K \leq L$. By narrowing down the scale of the
5.2 Finite-time Horizon

continuation set (the idea behind this is that, in the American put option pricing problem, the lower bound for the continuation set when \( T < \infty \) is the optimal stopping level obtained from pricing the perpetual American put option), we can make an additional assumption to further develop some insights towards the shape of sets \( C \) and \( D \).

**Assumption 5.2.27.** Parameters \( r \) and \( \sigma \) can be chosen such that the map of \( t \rightarrow H(t, x) \) is decreasing on the sets \([0, t_s] \times [x^*(t), \infty) \cup [t_s, T] \times [B(t), \infty) \cup \{t_s\} \times [x^*(t_s), L]\).

**Remark 5.2.28.** Note that the above assumption holds true in the state space \([0, t_s] \times [x^*(t), L]\) generally, see Figure 5.6. In addition, for \( x > L \), \( H = -rK \).

![Figure 5.6](image)

**Figure 5.6:** This figure displays the maps \( x \mapsto H_t(t, x) \) with chosen parameters \( r = 0.05, \sigma = 0.4, L = 6.5, K = 7, T = 5 \). Different-coloured lines indicate different \( t \) (and thus different \( x^*(t) \)).

For future reference, set
\[
S = [0, t_s] \times (0, x^*(t)] \cup \{t_s\} \times [x^*(t_s), L] \cup [t_s, T] \times (0, B(t)),
\]
so that \( S \subseteq D \).

**Proposition 5.2.29.** The map \( t \mapsto V(t, x) - G(t, x) \) is decreasing on \([0, T]\).

**Proof.** Let \( 0 \leq t_1 < t_2 \leq T \) and consider the stopping times:
\[
\tau_S = \inf\{0 \leq u \leq T - t_2 : X_u^T \in S\}
\]
and \( \tau_2 \) be the optimal stopping time for \( V(t_2, x) \), by the fact that \( S \subseteq D \), a direct result is \( \tau_2 \leq \tau_S \). Then, we have
\[
V(t_1, x) - V(t_2, x) \geq E(e^{-r\tau_2}G(t_1 + \tau_2, X_{t_2}^T)) - E(e^{-r\tau_2}G(t_2 + \tau_2, X_{t_2}^T))
\]
\[ G(t_1, x) - G(t_2, x) + E \left( \int_0^{t_2} e^{-ru} \left( -rKZ - \sigma^2 X_u^2 Z_x \right) (t_1 + u, X_u) \mathbb{1}\{X_u^x < L\} du \right) - G(t_2, x) - G(t_2, x) + E \left( -\frac{1}{2} \int_0^{t_2} e^{-ru} (K - L) Z_x(t_1, L-) dt_u^L(X^x) \right) - G(t_2, x)
\]

where the second inequality holds true via the map \( t \mapsto H(t, x) \) is decreasing and \( t \mapsto Z_x(t, L-) \) is increasing (see A.1.2). In particular, we have arrived at the inequality:

\[ V(t_1, x) - G(t_1, x) \geq V(t_2, x) - G(t_2, x), \]

which is precisely the claim.

The usefulness of Proposition 5.2.29 is well demonstrated by the following results:

**Proposition 5.2.30.** (i) The stopping set \( D \) is right-connected, i.e. increasing w.r.t time \( t \);

(ii) The continuation set \( C \) is left-connected, i.e. decreasing w.r.t time \( t \);

(iii) The stopping set \( D \) is down-connected;

(iv) The continuation set \( C \) is up-connected.

**Proof.** Assertions (i) and (ii) are immediate from Proposition 5.2.29, from which, (iii) and (iv) follow.

**Remark 5.2.31.** Once again, because the gain function is not smooth in the stopping set \( D \) at level \( L \), the presence of the local-time term in the pricing formula should not come to us as a surprise; above all, the twist comes when the optimal stopping boundary is crossing or remaining for some time at \( L \) (consider the situation for \( B(t) < L < K \) and \( x^*(t) = L \) for all \( t \in [0, t^*] \), then \{ \( (t, x) \in [0, T] \times [0, L] \) \( \subset D \) \), the smooth-fit condition fails to hold. See [21, Page 19] for similar treatment of such problem.

**Remark 5.2.32.** The minimal conditions under which the smooth-fit condition can hold in greater generality are the regularity of the diffusion process \( X \) and the differentiability of the gain function \( G \). For further contribution and examples, we refer to [64, Page 155].

To further construct the continuity of the value function, we need a definition first:
Definition 5.2.33. Let times $t_b$ and $t^b$ be defined as follows:
\[
\begin{cases}
  b(t) < L, & t \in [0, t_b), \\
  b(t) = L, & t \in [t_b, t^b], \\
  b(t) > L, & t \in (t^b, T],
\end{cases}
\]
and if $b(t) \geq L$ for all $t \in [0, T]$, we set $t_b = 0$. Note also that if $t_b = t^b$, then $t^b < T$.

As the reader has, hopefully noticed in the proof of Lemma 5.2.17 and Proposition 5.2.18, the difficulty in the proof of continuity of $V$ and smooth-fit condition lies in the singular point $(L, T)$ of $Z_x(t, x)$ (and thereby in the applicability of the dominated convergence theorem). To overcome this, the key idea is to keep the state space away from such a point by taking the continuity and the monotonicity of $B$ in $[t_u, T]$ into account (and these properties come for free!).

Lemma 5.2.34. The value function is continuous on $(0, T) \times (0, \infty)$.

Proof. The proof once again involves two statements, but with statement (i) of Lemma 5.2.17 remained unchanged, we confine ourselves by proving (ii) the map $x \to V(t, x)$ is continuous on $(0, \infty)$ for all $t \in [0, T]$. As before, the up-down connectedness of $\mathcal{C}$ and $\mathcal{D}$ suggests that for any $x_1 < x_2$ in $(0, \infty)$,
\[
G(t, x_2) - G(t, x_1) \leq V(t, x_2) - V(t, x_1),
\]
and let $\tau_2$ be optimal for $V(t, x_2)$ so that
\[
G(t, x_2) - G(t, x_1) \leq V(t, x_2) - V(t, x_1) \leq E \left( e^{-r\tau_2} G \left( t + \tau_2, X^{x_2}_{\tau_2} \right) \right) - E \left( e^{-r\tau_2} G \left( t + \tau_2, X^{x_1}_{\tau_2} \right) \right),
\]
in which, by letting $x_1 \to x_2$, Theorem 2.1.9 yields $V(t, x_1) - V(t, x_2) \to 0$ and thereby showing the left-continuity of $V$ w.r.t. $x$.

If $x_2 \leq \max\{x^*(t), B(t)\}$ for $t \in [0, T]$, Lemma 5.2.25 (see Figure 5.5) suggests that $\tau_2 = 0$, which, together with the continuity of $G$, shows that $V(t, x_2) - V(t, x_1) \to 0$ as $x_2 \to x_1$. If, on the other hand, $x_2 > \max\{x^*(t), B(t)\}$ for $t \in [0, T]$, then
\[
V(t, x_2) - V(t, x_1) \leq E \left( e^{-r\tau_2} G \left( t + \tau_2, X^{x_2}_{\tau_2} \right) - G \left( t + \tau_2, X^{x_1}_{\tau_2} \right) \right) \\
= E \left( e^{-r\tau_2} \left( (K - X^{x_2}_{\tau_2})^+ - (K - X^{x_1}_{\tau_2})^+ \right) Z \left( t + \tau_2, X^{x_2}_{\tau_2} \right) \right) \\
+ E \left( e^{-r\tau_2} \left( K - X^{x_1}_{\tau_2} \right)^+ \left( Z \left( t + \tau_2, X^{x_2}_{\tau_2} \right) - Z \left( t + \tau_2, X^{x_1}_{\tau_2} \right) \right) \right) \\
\leq KE \left( e^{-r\tau_2} \left( Z \left( t + \tau_2, X^{x_2}_{\tau_2} \right) - Z \left( t + \tau_2, X^{x_1}_{\tau_2} \right) \right) \right) \\
\leq KE \left( e^{-r\tau_2} I \{ X^{x_2}_{\tau_2} > L + \delta \} \left( Z \left( t + \tau_2, X^{x_2}_{\tau_2} \right) - Z \left( t + \tau_2, X^{x_1}_{\tau_2} \right) \right) \right) \\
+ KE \left( e^{-r\tau_2} I \{ X^{x_2}_{\tau_2} \leq L + \delta \} \left( Z \left( t + \tau_2, X^{x_2}_{\tau_2} \right) - Z \left( t + \tau_2, X^{x_1}_{\tau_2} \right) \right) \right),
\]

\footnote{Note that despite monotone convergence theorem allow the limit to be infinite, this does not suit for payoff function designed in this context.}
where the second inequality is because of inequality (5.2.21) and $X_{\tau_2}^{x_2} > X_{\tau_2}^{x_1}$.

Before we proceed, recall that $L \geq B(t) \geq K$ for $t \in [t_s, T]$ (and $L < K$). By employing the continuity and monotonicity of $B$, intermediate value theorem tells us that there exists $t_\delta \in (t_s, T)$ so that $B(t_\delta) = L + \delta$ and because $\tau_2 < \tau_S$, where $\tau_S = \inf\{u \in [0, T - t] \in X_u^{x_2} \in \mathcal{S}\}$, we know that $X_{\tau_2}^{x_2} \leq L + \delta$ implies $t + \tau_2 < t_\delta$ a.s. (see Figure 5.3), equation (5.2.32) therefore equals

$$V(t, x_2) - V(t, x_1) \leq KE\left(e^{-rt_2}I\{X_{\tau_2}^{x_2} > L + \delta\} (Z (t + \tau_2, X_{\tau_2}^{x_2}) - Z (t + \tau_2, X_{\tau_2}^{x_1}))\right)$$

$$+ KE\left(e^{-rt_2}I\{t + \tau_2 < t_\delta\} (Z (t + \tau_2, X_{\tau_2}^{x_2}) - Z (t + \tau_2, X_{\tau_2}^{x_1}))\right)$$

$$= KE\left(e^{-rt_2}I\{X_{\tau_2}^{x_2} > L + \delta\} (Z (t + \tau_2, X_{\tau_2}^{x_2}) - Z (t + \tau_2, X_{\tau_2}^{x_1}))\right)$$

$$+ K(x_2 - x_1)E\left(e^{-\frac{\sigma^2}{2}t_2 + \sigma W_2 I\{t + \tau_2 < t_\delta\} Z_x (t + \tau_2, X_{\tau_2}^{x_2})}\right)$$

$$= KE\left(e^{-rt_2}I\{X_{\tau_2}^{x_2} > L + \delta\} (Z (t + \tau_2, X_{\tau_2}^{x_2}) - Z (t + \tau_2, X_{\tau_2}^{x_1}))\right) + KC_2(x_2 - x_1)$$

(5.2.33)

where the last two steps are due to the mean value theorem for $x_3 \in [x_1, x_2]$, the strong solution of GBM and the fact that $Z_x(x, t) \leq C_2$ in $[0, t_\delta) \times (0, L + \delta]$ (see A.1.2).

Then, we can choose $\delta > 0$ so that $X_{\tau_2}^{x_1} > L$ almost surely whenever $X_{\tau_2}^{x_2} > L + \delta$ and $x_2 - x_1 < \frac{z_\delta}{\sigma^2} \leq (\text{the justification is the same as that in the proof of Lemma 5.2.17, we omit further details})$, by setting such $\delta$ in (5.2.33), joining with the fact that $Z(t, x) = 1$ for $x \geq L$,

$$G(t, x_2) - G(t, x_1) \leq V(t, x_2) - V(t, x_1) \leq KC_2(x_2 - x_1),$$

in which, let $x_2 \to x_1$, we have $V(t, x_2) - V(t, x_1) \to 0$, that is, $V$ is right continuous w.r.t $x$. The conclusion follows as $V$ is right and left continuous for any $x \in (0, \infty)$ for each $t \in [0, T)$ given and fixed.

To this point, we have been working really hard on the discussion of the stopping and continuation sets and now we can finally return to our main purpose, i.e. the formulation of free-boundary problem,

**Proposition 5.2.35 (The Free-boundary Problem).** The stopping set and the continuation set are of the forms:

$$\mathcal{C} = \{(t, x) \in [0, T) \times (0, \infty): x > b(t)\},$$

$$\mathcal{D} = \{(t, x) \in [0, T) \times (0, \infty): x \leq b(t)\},$$

where the map $t \mapsto b(t)$ is increasing. The free-boundary problem in section 2.2 is therefore reshuffled accordingly as

$$V_t + \mathbb{L}_x V - rV = 0 \quad \text{for} \ (t, x) \in \mathcal{C},$$

$$V(t, x) = G(t, x) \quad \text{for} \ x = b(t) \text{ (instantaneous stopping)},$$

$$V_x(t, x) = G_x(t, x) \quad \text{for} \ x = b(t) < L \text{ and } x = b(t) > L \text{ (smooth-fit)},$$

(5.2.34) (5.2.35) (5.2.36)

---

10 Set $\mathcal{S}$ is the area coloured in green in Figure 5.5.
5.2 Finite-time Horizon

\[ V_x(t, L \pm) = G_x(t, L \pm) \quad \text{for } x = b(t) = L, \text{ (smooth-fit breaking down)}, \quad (5.2.37) \]
\[ V(t, x) > G(t, x) \quad \text{for } (t, x) \in \mathcal{C}, \quad (5.2.38) \]
\[ V(t, x) = G(t, x) \quad \text{for } (t, x) \in \mathcal{D}. \quad (5.2.39) \]

**Proof of the Smooth-fit Condition (5.2.36).** The proof for \( b(t) > L \) is essentially the same as that of Proposition 5.2.18. It thus remains to show (5.2.36) holds true for \( b(t) < L \). Let a point \((t, x) \in [0, T) \times (0, \infty)\) lying on the boundary \( b \) be fixed, i.e. \( x = b(t) < L \). Then, let \( \epsilon > 0 \) such that \( x + \epsilon < L \), and

\[
\frac{V(t, x + \epsilon) - V(t, x)}{\epsilon} \geq \frac{G(t, x + \epsilon) - G(t, x)}{\epsilon},
\]

where the inequality is due to \((t, x + \epsilon) \in \mathcal{C}\) and \((t, x) \in \mathcal{D}\); of which, we take the limit as \( \epsilon \to 0 \) and obtain

\[
\frac{\partial^+}{\partial x} V(t, x) \geq -Z(t, x) + (K - x) \frac{\partial}{\partial x} Z(t, x). \quad (5.2.40)
\]

In order to prove the converse inequality, we first note that (the exact same argument is presented in the proof of Lemma 5.2.34, so we omit some details here) there exists \( \delta > 0 \) s.t. \( B(t_\delta) = L + \delta \) for \( t_\delta \in (t_\epsilon, T) \). Then, let \( \tau_\epsilon \) be optimal for \( V(t, x + \epsilon) \), where we can choose \( \epsilon \in \left(0, \frac{\delta}{2}\right) \) so that \( X_{\tau_\epsilon}^x > L \) almost surely whenever \( X_{\tau_\epsilon}^x + \epsilon > L + \delta \) (the justification follows the same pattern as that in the proof of Lemma 5.2.17),

\[
\frac{V(t, x + \epsilon) - V(t, x)}{\epsilon} \leq \frac{1}{\epsilon} E \left( e^{-r t_\epsilon} \left( G(t + \tau_\epsilon, X_{\tau_\epsilon}^{x+\epsilon}) - G(t + \tau_\epsilon, X_{\tau_\epsilon}^x) \right) \right)
\]

\[
\leq -E \left( e^{-\frac{\sigma^2}{2} t_\epsilon t_\epsilon + \sigma W t_\epsilon} I \left\{ X_{\tau_\epsilon}^{x+\epsilon} < K \right\} Z(t + \tau_\epsilon, X_{\tau_\epsilon}^{x+\epsilon}) \right)
\]

\[
+ \frac{1}{\epsilon} E \left( e^{-r t_\epsilon} \left( K - X_{\tau_\epsilon}^x \right) I \left\{ X_{\tau_\epsilon}^{x+\epsilon} < \right\} \left( Z(t + \tau_\epsilon, X_{\tau_\epsilon}^{x+\epsilon}) - Z(t + \tau_\epsilon, X_{\tau_\epsilon}^x) \right) \right)
\]

\[
+ \frac{1}{\epsilon} E \left( e^{-r t_\epsilon} \left( K - X_{\tau_\epsilon}^x \right) I \left\{ X_{\tau_\epsilon}^{x+\epsilon} > \right\} \left( Z(t + \tau_\epsilon, X_{\tau_\epsilon}^{x+\epsilon}) - Z(t + \tau_\epsilon, X_{\tau_\epsilon}^x) \right) \right)
\]

\[
= -E \left( e^{-\frac{\sigma^2}{2} t_\epsilon t_\epsilon + \sigma W t_\epsilon} I \left\{ X_{\tau_\epsilon}^{x+\epsilon} < K \right\} Z(t + \tau_\epsilon, X_{\tau_\epsilon}^{x+\epsilon}) \right)
\]

\[
+ \frac{1}{\epsilon} E \left( e^{-r t_\epsilon} \left( K - X_{\tau_\epsilon}^x \right) I \left\{ X_{\tau_\epsilon}^{x+\epsilon} < \right\} \left( Z(t + \tau_\epsilon, X_{\tau_\epsilon}^{x+\epsilon}) - Z(t + \tau_\epsilon, X_{\tau_\epsilon}^x) \right) \right)
\]

\[
= -E \left( e^{-\frac{\sigma^2}{2} t_\epsilon t_\epsilon + \sigma W t_\epsilon} I \left\{ X_{\tau_\epsilon}^{x+\epsilon} < K \right\} Z(t + \tau_\epsilon, X_{\tau_\epsilon}^{x+\epsilon}) \right)
\]

\[
+ \frac{1}{\epsilon} E \left( e^{-r t_\epsilon} \left( K - X_{\tau_\epsilon}^x \right) I \left\{ X_{\tau_\epsilon}^{x+\epsilon} < \right\} \left( Z(t + \tau_\epsilon, X_{\tau_\epsilon}^{x+\epsilon}) - Z(t + \tau_\epsilon, X_{\tau_\epsilon}^x) \right) \right)
\]

\[
= -E \left( e^{-\frac{\sigma^2}{2} t_\epsilon t_\epsilon + \sigma W t_\epsilon} I \left\{ X_{\tau_\epsilon}^{x+\epsilon} < K \right\} Z(t + \tau_\epsilon, X_{\tau_\epsilon}^{x+\epsilon}) \right)
\]

where the second inequality holds true because of (5.2.28) and the first equality follows from the facts that for \( \epsilon \) of our choice, \( I \{ X_{\tau_\epsilon}^{x+\epsilon} > L + \delta \} = 1 \) entails \( I \{ X_{\tau_\epsilon}^x > L \} = 1 \), the second equality
is derived from the mean value theorem for $y \in [x, x + \epsilon]$. Recalling that the random variable $X$ is dominated by a positive integrable random variable, letting $\epsilon \to 0$, $\tau_\epsilon \to 0$ a.s. and exploiting Theorem 2.1.9 (the boundedness of $Z_x(t, x)$ for $(t, x) \in [0, t_\delta] \times (0, L + \delta]$ is verified in A.1.2) yield

$$\frac{\partial^+}{\partial x} V(t, x) \leq -Z(t, x) + \frac{\partial}{\partial x} Z(t, x),$$

which, together with (5.2.40) and the fact that $(t, x), (t, x - \epsilon) \in D$ so that

$$\frac{V(t, x) - V(t, x - \epsilon)}{\epsilon} = G(t, x) - G(t, x - \epsilon),$$

and this allows us to obtain (5.2.36).

Proof of (5.2.37). (i) Let $x = b(t) = L$ such that there exists $\epsilon > 0$, $x - \epsilon < L$ and $(t, x - \epsilon) \in D$. The first part of (5.2.37) is a consequence of

$$\frac{V(t, x) - V(t, x - \epsilon)}{\epsilon} = \frac{G(t, x) - G(t, x - \epsilon)}{\epsilon}$$

and via letting $\epsilon \to 0$, we see $\lim_{\epsilon \to 0} \frac{V(t, x) - V(t, x - \epsilon)}{\epsilon} = -1 + (K - L)Z_x(t, L-)$$

(ii) Again, let $x = b(t) = L$ so that there exists $0 < \epsilon < \frac{\epsilon}{t_\delta}$ (where $\delta > 0$ is defined as so that $B(t_\delta) = L + \delta$ for $t_\delta \in (t_\delta, T)$), $L < x + \epsilon < K$ and $(t, x + \epsilon) \in C$ and that

$$\frac{V(t, x + \epsilon) - V(t, x)}{\epsilon} \geq \frac{G(t, x + \epsilon) - G(t, x)}{\epsilon} = -1.$$

Furthermore, let $\tau_\epsilon$ be the optimal stopping time for $V(t, x + \epsilon)$ so that $X_{\tau_\epsilon}^{x+\epsilon} \geq L$ and that

$$\frac{V(t, x + \epsilon) - V(t, x)}{\epsilon} \leq \frac{1}{\epsilon} E\left(e^{-\frac{1}{2} t_\delta - \sigma W_\tau \epsilon} \left(G\left(t + \tau_\epsilon, X^{x+\epsilon}_{\tau_\epsilon}\right) - G\left(t + \tau_\epsilon, X^x_{\tau_\epsilon}\right)\right)\right)$$

$$\leq -E\left(e^{-\frac{1}{2} t_\delta - \sigma W_\tau \epsilon} Z\left(t + \tau_\epsilon, X^x_{\tau_\epsilon}\right) I\{X_{\tau_\epsilon}^{x+\epsilon} < K\}\right)$$

$$+ \frac{1}{\epsilon} E\left(e^{-\tau_\epsilon} I\{X^x_{\tau_\epsilon} = L\} \left(K - X^{x+\epsilon}_{\tau_\epsilon}\right)^+ \left(Z\left(t + \tau_\epsilon, X^{x+\epsilon}_{\tau_\epsilon}\right) - Z\left(t + \tau_\epsilon, X^x_{\tau_\epsilon}\right)\right)\right)$$

$$+ \frac{1}{\epsilon} E\left(e^{-\tau_\epsilon} I\{X^x_{\tau_\epsilon} > L + \delta\} \left(K - X^{x+\epsilon}_{\tau_\epsilon}\right)^+ \left(Z\left(t + \tau_\epsilon, X^{x+\epsilon}_{\tau_\epsilon}\right) - Z\left(t + \tau_\epsilon, X^x_{\tau_\epsilon}\right)\right)\right)$$

$$+ \frac{1}{\epsilon} E\left(e^{-\tau_\epsilon} I\{L < X^{x+\epsilon}_{\tau_\epsilon} < L + \delta\} \left(K - X^{x+\epsilon}_{\tau_\epsilon}\right)^+ \left(Z\left(t + \tau_\epsilon, X^{x+\epsilon}_{\tau_\epsilon}\right) - Z\left(t + \tau_\epsilon, X^x_{\tau_\epsilon}\right)\right)\right)$$

$$= -E\left(e^{-\frac{1}{2} t_\delta - \sigma W_\tau \epsilon} Z\left(t + \tau_\epsilon, X^x_{\tau_\epsilon}\right) I\{X_{\tau_\epsilon}^{x+\epsilon} < K\}\right)$$

$$+ \frac{1}{\epsilon} E\left(e^{-\tau_\epsilon} I\{t_\delta < t + \tau_\epsilon \leq t_\delta\} \left(K - X^{x+\epsilon}_{\tau_\epsilon}\right)^+ \left(Z\left(t + \tau_\epsilon, X^{x+\epsilon}_{\tau_\epsilon}\right) - Z\left(t + \tau_\epsilon, X^x_{\tau_\epsilon}\right)\right)\right)$$

$$+ \frac{1}{\epsilon} E\left(e^{-\tau_\epsilon} I\{X^x_{\tau_\epsilon} > L + \delta\} \left(K - X^{x+\epsilon}_{\tau_\epsilon}\right)^+ \left(Z\left(t + \tau_\epsilon, X^{x+\epsilon}_{\tau_\epsilon}\right) - Z\left(t + \tau_\epsilon, X^x_{\tau_\epsilon}\right)\right)\right)$$
where the second inequality is due to (5.2.28) and in the first equality, we use the fact that

\[
\text{Remark 5.2.38. The essential tool to prove Lemma 5.2.20 is the smooth-fit condition, which, may, at times, either be difficult to verify or hold true. Following is a rather convenient technique to prove the continuity based on the regularity of functions } V \text{ and } G \text{ and the monotonicity of } b.
\]

\text{Lemma 5.2.37. The optimal stopping boundary } b \text{ is continuous on } [0, T] \text{ and } b(T^-) = K.

\text{Proof. To avoid using smooth-fit condition, we use the argument provided in [16, Page 173-175].}

(i) The boundary } b \text{ is right-continuous on } [0, t_b]. \text{ The conclusion is immediate from the fact that the stopping set is closed and } b \text{ is increasing on } [0, t_b].

(ii) The boundary } b \text{ is left-continuous on } [0, t_b]. \text{ We assume that, for contradiction, there exists } t_0 \in (0, t_b) \text{ so that a discontinuity of } b \text{ occurs; that is, at } t_0, \text{ one has } b(t_0-) < b(t_0), \text{ where } b(t_0-) \text{ denotes the left limit of the boundary at } t_0, \text{ which always exists as } b(t) \text{ is monotonically increasing in } (0, t_b).

Let } x_1, x_2 \text{ be fixed such that } b(t_0-) < x_1 < x_2 < b(t_0). \text{ Then, for fixed } t' \in (0, t_0), \text{ we define an open bounded domain } \mathcal{R} \subset \mathcal{C} \text{ with } \mathcal{R} = \{(t', t_0) \times (x_1, x_2)\}. \text{ Hence, by the structure of } \mathcal{C}, \text{ we know the value function } V, \text{ which is } C^{1,2} \text{ in } \mathcal{R}, \text{ solves the Cauchy problem}

\[
V_t + L_x V - \tau V = 0. \quad (5.2.42)
\]

Then, we denote } C^\infty_c([x_1, x_2]) \text{ as the set of function with infinitely many continuous derivatives and compact support in } [x_1, x_2], \text{ in which, we define } \psi \geq 0 \text{ such that } \int_{x_1}^{x_2} \psi(y)dy = 1. \text{ By further
multiplying (5.2.42) by function $\psi$, and integrating the product over $(t, t_0) \times (x_1, x_2)$ for some 
$t \in (t', t_0)$,
\[
\int_t^{t_0} \int_{x_1}^{x_2} (V_t + \mathbb{L}_X V)(s, y) \psi(y)dyds + \left( - \int_t^{t_0} \int_{x_1}^{x_2} rV(s, y)\psi(y)dyds \right) = 0,
\tag{5.2.43}
\]
where the second double integral on the left-hand side is strictly negative, as $V > G \geq 0$ on $\mathcal{C}$. Therefore, we want to investigate the sign of first double integral by estimating its upper bound, which if is negative, the proof will be completed by reaching a contradiction. To do so, let us first observe that
\[
\int_t^{t_0} \int_{x_1}^{x_2} V_t(s, y)\psi(y)dyds = \int_{x_1}^{x_2} (V(t_0, y) - V(t, y))\psi(y)dy
\leq \int_{x_1}^{x_2} (G(t_0, y) - G(t, y))\psi(y)dy = \int_t^{t_0} \int_{x_1}^{x_2} G_t(s, y)\psi(y)dyds,
\tag{5.2.44}
\]
where the inequality is the consequence of $(t_0, y) \in \mathcal{D}$ and $(t, y) \in \mathcal{C}$.

Next in line is partial integrations (three times),
\[
\int_t^{t_0} \int_{x_1}^{x_2} \mathbb{L}_X V(s, y)\psi(y)dyds = \int_t^{t_0} \int_{x_1}^{x_2} V(s, y)\mathbb{L}_X^*\psi(y)dyds,
\tag{5.2.45}
\]
where, if we also impose the condition $\psi(x_1) = \psi(x_2) = \psi'(x_1) = \psi'(x_2) = 0$ for the boundary terms disappear, $\mathbb{L}_X^*$ is the formal adjoint of $\mathbb{L}_X$ given as
\[
\mathbb{L}_X^* \psi(x) = \frac{1}{2} \frac{\partial^2}{\partial x^2} (\sigma^2 x^2 \psi(x)) - \frac{\partial}{\partial x} (r(x) \psi(x)).
\]

Since $G \in C^{1,2}(\mathcal{R})$,
\[
|V(s, y) - G(s, y)| = |V(s, y) - V(t_0, y) + G(t_0, y) - G(s, y)|
\leq |V(s, y) - V(t_0, y)| + |G(t_0, y) - G(s, y)|
\leq C_1(t_0 - s)^\beta + C_2(t_0 - s),
\tag{5.2.46}
\]
where $C_1, C_2$ are constants and $\beta > 0$, the first equality follows via again $(t_0, y) \in \mathcal{D}$ and the second inequality is due to the fulfilment of the conditions proposed on [16, Page 172, (C.3)].

In addition, we consider $|t_0 - t| < 1$ and set $C_3 = C_1 + C_2$, upon using $V \geq G$,
\[
0 \leq G(s, y) \leq V(s, y) \leq G(s, y) + C_3(t_0 - s)^\beta,
\]
of which, as a result, for any $s \in (t, t_0)$, in the view of (5.2.45),
\[
\int_t^{t_0} \int_{x_1}^{x_2} V(s, y)\mathbb{L}_X^*\psi(y)dyds = \int_t^{t_0} \int_{x_1}^{x_2} I\{\mathbb{L}_X^* \psi(y) \geq 0\} V(s, y)\mathbb{L}_X^* \psi(y)dyds
+ \int_t^{t_0} \int_{x_1}^{x_2} I\{\mathbb{L}_X^* \psi(y) < 0\} V(s, y)\mathbb{L}_X^* \psi(y)dyds
\leq \int_t^{t_0} \int_{x_1}^{x_2} I\{\mathbb{L}_X^* \psi(y) \geq 0\} \left( G(s, y) + C_3(t_0 - s)^\beta \right) \mathbb{L}_X^* \psi(y)dyds
\]

Therefore, we can conclude that the right-hand side of (5.2.43) is negative, as required.
where the last equality follows by partial integrations again.

Joining (5.2.44) with (5.2.47), from the view of equation (5.2.43),

\[ 0 \leq \int_t^{t_0} \int_{x_1}^{x_2} (G_t + \mathbb{L}_x G - rG)(s, y)\psi(y)dyds \]
\[ + C_3 \int_t^{t_0} (t_0 - s)^{\frac{\alpha}{2}} \int_{x_1}^{x_2} I\{\mathbb{L}_x^* \psi(y) \geq 0\} \mathbb{L}_x^* \psi(y)dyds \]
\[ = \int_t^{t_0} \int_{x_1}^{x_2} (G_t + \mathbb{L}_x G - rG)(s, y)\psi(y)dyds \]
\[ + C_3 \int_t^{t_0} (t_0 - s)^{\frac{\alpha}{2}} \int_{x_1}^{x_2} I\{\mathbb{L}_x^* \psi(y) > 0\} \mathbb{L}_x^* \psi(y)dyds \]
\[ = \int_t^{t_0} \int_{x_1}^{x_2} (-l_\epsilon)\psi(y)dyds + \frac{2\gamma}{2 + \beta} (t_0 - t)^{1 + \frac{\beta}{2}} \]
\[ = -l_\epsilon (t_0 - t) + \frac{2\gamma}{2 + \beta} (t_0 - t)^{1 + \frac{\beta}{2}}, \] (5.2.49)

where the second inequality follows in the sense that in such interval \([0, t^*] \times [x_1, x_2]\), there exists some constant \(l_\epsilon\) so that

\[ (G_t + \mathbb{L}_x G - rG)(t, x) = -(rKZ + \sigma^2 x^2 Z_x)(t, x) \leq -l_\epsilon < 0 \]

and that \(\int_{x_1}^{x_2} \psi(y)dy = 1\); since the second term in (5.2.48) is strictly positive, we can set

\[ \gamma = \gamma(\psi; x_1, x_2) := C_3 \int_{x_1}^{x_2} I\{\mathbb{L}_x^* \psi(y) > 0\} \mathbb{L}_x^* \psi(y)dy > 0, \]

from which, the second equality holds.

Taking the limit \(t_0 \to t\) thus leads to a contradiction in the sense that the second term on the last equality of (5.2.49) is vanishing more rapidly than its first term (that is, \(0 \leq -l_\epsilon (t_0 - t) < 0\)), hence the jump may not occur. \(\square\)
5.2.4 The Optimal Stopping Rule

In the Case of $t^* = 0$

In order to prepare for the main result, we shall first verify the conditions for Theorem 2.2.8 to be applicable.

**Lemma 5.2.39.** Let the function $F(t, x) = e^{-rt}V(t, x)$ defined on $[0, T] \times (0, \infty)$, $D_1 = \{(t, x) \in [0, T] \times (L, b(t))\}$ and $D_2 = \{(t, x) \in [0, T] \times (0, L)\}$. Then, the local time-space formula is applicable to $F$ if the following conditions are met:

(a) $F(t, x)$ is $C^{1, 2}$ on $\mathcal{C} \cup \mathcal{D}_1 \cup \mathcal{D}_2$;

(b) $F_t + \mathbb{L}_X F$ is locally bounded on $\mathcal{C} \cup \mathcal{D}_1 \cup \mathcal{D}_2$;

(c) $t \mapsto F_x(t, b(t) \pm)$ is continuous;

(d) $t \mapsto F_x(t, L \pm)$ is continuous;

(e) $F_{xx} = F_1 + F_2$ on $\mathcal{C} \cup \mathcal{D}_1 \cup \mathcal{D}_2$, where $F_1$ is non-negative and $F_2$ is continuous.

**Proof.** By the definition of $F$,

$$F(t, x) = \begin{cases} e^{-rt}V(t, x), & (t, x) \in \mathcal{C}, \\ e^{-rt}G(t, x), & (t, x) \in \mathcal{D}_1 \cup \mathcal{D}_2, \end{cases}$$

statement (a) is immediate.

To prove statement (b) is to show that $F_t + \mathbb{L}_X F$ is locally bounded on $(\mathcal{C} \cup \mathcal{D}_1 \cup \mathcal{D}_2) \cap K$ for each compact set $K$. Since

$$F_t + \mathbb{L}_X F = \begin{cases} 0, & (t, x) \in \mathcal{C}, \\ -rK, & (t, x) \in \mathcal{D}_1, \\ (-rKZ - \sigma^2 x^2 Z_x)(t, x), & (t, x) \in \mathcal{D}_2, \end{cases}$$

where the last case is locally bounded as it is a continuous function on $\mathcal{D}_2 \cap K$.

Statements (c) and (d) follows from (5.2.25) and the direct computation of

$$e^{-rt}G_x(t, L+) = -e^{-rt},$$

$$e^{-rt}G_x(t, L-) = e^{-rt} \left[ -1 + (K - L)Z_x(t, L-) \right],$$

(see A.1.2). Finally, for (e),

$$F_{xx}(t, x) = \begin{cases} \frac{2e^{-rt\sigma}}{\sigma^2} (rV_x - rxV_x + V_t)(t, x), & (t, x) \in \mathcal{C}, \\ 0, & (t, x) \in \mathcal{D}_1, \\ 2e^{-rt\sigma} (-G_t + rxG_x - \sigma^2 x^2 Z_x - rxZ)(t, x), & (t, x) \in \mathcal{D}_2, \end{cases}$$

where the first case indicate that $F$ is convex on $\mathcal{C}$, which (both convexity and concavity, see [66, Page 526]) is a rather stronger condition than (e), the second case is immediate and in the third case, $G_t(t, x) \leq 0$ and the rest terms are continuous. \qed
Following is one of the main results in the current section.

**Theorem 5.2.40.** The optimal stopping boundary in problem (5.2.5) can be characterised as the unique solution of the free-boundary equation:

\[
G(t, b(t)) = E_{t,b(t)} \left( e^{-r(T-t)} G(T, X_T) + rK \int_0^{T-t} e^{-ru} I\{L < X_{t+u} \leq b(t+u)\} du \right.
+ \int_0^{T-t} e^{-ru} \left( rKZ + \sigma^2 X_{t+u}^2 Z_x \right) (t + u, X_{t+u}) I\{X_{t+u} < L\} du
+ \frac{1}{2} \int_0^{T-t} e^{-ru}(K - L)Z_x(t + u, L-)dl^L_u(X),
\]

(5.2.50)
in the class of continuous increasing function \( c : [0, T] \to \mathbb{R}_+ \) satisfying \( 0 < c(t) < K \) for all \( 0 < t < T \).

The value of the current contract admits the following "early exercise premium" representation:

\[
V(t, x) = E_{t,x} \left( e^{-r(T-t)} G(T, X_T) + rK \int_0^{T-t} e^{-ru} I\{L < X_{t+u} \leq b(t+u)\} du \right.
+ \int_0^{T-t} e^{-ru} \left( rKZ + \sigma^2 X_{t+u}^2 Z_x \right) (t + u, X_{t+u}) I\{X_{t+u} < L\} du
+ \frac{1}{2} \int_0^{T-t} e^{-ru}(K - L)Z_x(t + u, L-)dl^L_u(X) \bigg) \bigg|_{t = T-x}
\]

(5.2.51)
for all \((t, x) \in [0, T] \times (0, \infty)\).

The proof follows essentially from [64, Page 385-392].

**Proof of The Formulas (5.2.50) and (5.2.51).** Since the function \( F \) fulfils the conditions on Lemma 5.2.39, an application of Theorem 2.2.7 yields for \( t \in [0, T] \) and \( s \in [0, T-t] \),

\[
e^{-rs} V(t + s, X_{t+s}^x) = V(t, x) + M_s
+ \int_0^s e^{-ru} \left( -rV + V_t + \mathbb{L}_x V \right) (t + u, X_{t+u}) I\{X_{t+u} \neq b(t+u), X_{t+u} \neq L\} du
+ \frac{1}{2} \int_0^s e^{-ru}(V_x(t + u, L-) - V_x(t + u, L+))dl^L_u(X_x)
+ \frac{1}{2} \int_0^s e^{-ru}(V_x(t + u, b(t+u)+) - V_x(t + u, b(t+u)-))dl^b_u(X_x)
= V(t, x) + M_s + \int_0^s e^{-ru} \left( -rG + G_t + \mathbb{L}_x G \right) (t + u, X_{t+u}) I\{X_{t+u} < b(t+u)\} du
+ \frac{1}{2} \int_0^s e^{-ru}(G_x(t + u, L+) - G_x(t + u, L-))dl^L_u(X_x),
\]

(5.2.52)
where \( M_s = \int_0^s e^{-ru} \sigma X_{t+u} V_x(t + u, X_{t+u}) I\{X_{t+u} \neq b(t+u), X_{t+u} \neq L\} dW_u \) is a martingale under measure \( P_{t,x} \), and from the smooth-fit condition and \( V = G \) on \( D \), the rest equalities are
rather immediate. Then, upon taking the expectation under measure \( P_{t,x} \) of (5.2.52) and invoking the optional sampling theorem, we obtain

\[
E_{t,x} (e^{-rs} V(t + s, X_{t+s})) = V(t,x) + E_{t,x} \left( \int_0^s e^{-ru} H(t + u, X_{t+u}) I\{X_{t+u} < b(t + u)\} du \right)
- E_{t,x} \left( \frac{1}{2} \int_0^s e^{-ru} (K - L) Z_x(t + u, L-) dl^L_u(X^x) \right),
\]  

(5.2.53)

where

\[
H(t, x) = (-rG + G_t + \mathbb{I}_x G)(t, x) = \begin{cases} 
-rK, & L < x < b(t), \\
(-rKZ - \sigma^2 x^2 Z_x)(t, x), & x < L.
\end{cases}
\]

Next, let \( s = T - t \) so that \( V(T, X_T) = G(T, X_T) \), that is,

\[
E_{t,x} (e^{-r(T-t)} V(T, X_T)) = E_{t,x} \left( e^{-r(T-t)} (K - X_T) I\{L < X_T < K\} \right),
\]

which, together with (5.2.53), shows that

\[
V(t, x) = E_{t,x} \left( e^{-r(T-t)} G(T, X_T) \right) - E_{t,x} \left( \int_0^{T-t} e^{-ru} H(t + u, X_{t+u}) I\{X_{t+u} < b(t + u)\} du \right)
+ E_{t,x} \left( \frac{1}{2} \int_0^{T-t} e^{-ru} (K - L) Z_x(t + u, L-) dl^L_u(X^x) \right),
\]

in which, we set \( x = b(t) \) and because \( V(t, b(t)) = G(t, b(t)) \), equation (5.2.50) follows.

Proof of uniqueness of the boundary. The key idea is to show that if there exists a continuous increasing function \( c : [0, T] \mapsto (0, \infty) \) satisfying \( L < c(t) < K \) and solving (5.2.50), then such \( c \) must coincide with the optimal stopping boundary \( b \). We proceed via steps:

(i) From the view of (5.2.51), we introduce the following functions:

\[
U(t, x) = E_{t,x} \left( e^{-r(T-t)} G(T, X_T) \right) - E_{t,x} \left( \int_0^{T-t} e^{-ru} H(t + u, X_{t+u}) I\{X_{t+u} < c(t + u)\} du \right)
+ E_{t,x} \left( \frac{1}{2} \int_0^{T-t} e^{-ru} (K - L) Z_x(t + u, L-) dl^L_u(X^x) \right)
\]

for all \((t, x) \in [0, T) \times (0, \infty)\), and

\[
V^c(t, x) = \begin{cases} 
U(t, x), & x > c(t), \\
G(t, x), & x \leq c(t),
\end{cases}
\]

which is continuous on \([0, T) \times (0, \infty)\) in the sense that \( c \) solves (5.2.50), implying \( U(t, c(t)) = G(t, c(t)) \). In addition, we reconstruct the continuation and stopping sets by the means of \( c \):

\[
\mathcal{C}_c = \{(t, x) \in [0, T) \times (0, \infty) : x > c(t)\},
\]

\[
\mathcal{D}_c = \{(t, x) \in [0, T) \times (0, \infty) : x \leq c(t)\}.
\]
From the definition, $V^c$ has the following properties, which we state without proof:\(^{11}\):

(a) $V^c = U$ is $C^{1,2}$ and satisfies the equation $V^c_t + \mathbb{L}_X V^c = r V^c$ on $C_c$;

(b) $V^c = G$ is $C^{1,2}$ on $D_c \setminus \{(t, x) \in [0, T] \times \{L, c(t)\}\}$.

(ii) Following exactly the same pattern on Lemma 5.2.39, we can verify that Theorem 2.2.8 is applicable to $e^{-rt}V^c(t, X_t)$ so that

\[
e^{-rs}V^c(t+s, X^x_{t+s}) = V^c(t, x) + M^c_x
\]

\[
+ \int_0^s e^{-ru} (-rV^c + V^c_t + \mathbb{L}_X V^c) (t + u, X^x_{t+u}) I\{X^x_{t+u} = c(t+u), X^x_{t+u} \neq L\} du
\]

\[
+ \frac{1}{2} \int_0^s e^{-ru} (V^c_x (t + u, L+) - V^c_x (t + u, L-)) dL^u_x(X^x)
\]

\[
+ \frac{1}{2} \int_0^s e^{-ru} (V^c_x (t + u, c(t+u)+) - V^c_x (t + u, c(t+u)-)) dL^u_x(X^x)
\]

\[
= V^c(t, x) + M^c_x
\]

\[
+ \int_0^s e^{-ru} (-rG + G_t + \mathbb{L}_X G) (t + u, X^x_{t+u}) I\{X^x_{t+u} = c(t+u), X^x_{t+u} \neq L\} du
\]

\[
+ \frac{1}{2} \int_0^s e^{-ru} (G_x (t + u, L+) - G_x (t + u, L-)) dL^u_x(X^x)
\]

\[
+ \frac{1}{2} \int_0^s e^{-ru} (V^c_x (t + u, c(t+u)+) - V^c_x (t + u, c(t+u)-)) dL^u_x(X^x),
\]

where $M^c_x = \int_0^s e^{-ru} \sigma X^x_{t+u} V^c(t+u, X^x_{t+u}) I\{X^x_{t+u} = c(t+u), X^x_{t+u} \neq L\} du$ and it is a martingale under measure $P_{t,x}$ so that $E_{t,x}(M^c_x) = 0$ for each $0 \leq s \leq T-t$. Let us take the expectation under measure $P_{t,x}$ upon recalling the definition of $V^c$,

\[
E_{t,x}(e^{-rs}V^c(t+s, X^x_{t+s})) = V^c(t, x)
\]

\[
+ E_{t,x}\left(\int_0^s e^{-ru} (-rG + G_t + \mathbb{L}_X G) (t + u, X^x_{t+u}) I\{X^x_{t+u} = c(t+u)\} du\right)
\]

\[
+ E_{t,x}\left(\frac{1}{2} \int_0^s e^{-ru} (G_x (t + u, L+) - G_x (t + u, L-)) dL^u_x(X^x)\right)
\]

\[
+ E_{t,x}\left(\frac{1}{2} \int_0^s e^{-ru} (V^c_x (t + u, c(t+u)+) - V^c_x (t + u, c(t+u)-)) dL^u_x(X^x)\right). (5.2.54)
\]

(iii) The smooth-fit condition (5.2.25) then makes it reasonable to guess that $x \mapsto V^c(t, x)$ is $C^1$ at $c(t)$ for $0 \leq t < T$. This will not only be a guess if we could somehow show that $U(t, x) = G(t, x)$.

The crucial observation is that, from the strong Markov property and time-homogeneous property of $X$ and smoothing lemma (Lemma 2.3.12),

\[
E_{t,x}\left(e^{-rs}U(t+s, X^x_{t+s}) - \int_0^s e^{-ru} H(t+u, X^x_{t+u}) I\{X^x_{t+u} = c(t+u)\} du\right)
\]

\[
+ \frac{1}{2} \int_0^s e^{-ru} (K - L) Z_x(t+u, L-) dL^u_x(X^x)\right)
\]

\[
= E\left(e^{-rs}E\left(e^{-r(T-t-s)} G(T, X_T) - \int_{t+s}^T e^{-r(u-t-s)} H(u, X_u) I\{X_u = c(u)\} du\right)\right)
\]

\(^{11}\)By expressing the expectation using Law of the unconscious statistician, this shall be clear.
+ \frac{1}{2} \int_{t+s}^{T} e^{-r(u-t-s)} (K - L) Z_x(u, L-) \, du + \frac{K - L}{2} \int_{t+s}^{T} e^{-r(u-t)} Z_x(u, L-) \, du \right) |_{F_t} \\
+ E_{t,x} \left( - \int_{t}^{t+s} e^{-r(u-t)} H(u, X_u) I\{X_u < c(u)\} \, du + \frac{K - L}{2} \int_{t}^{t+s} e^{-r(u-t)} Z_x(u, L-) \, du \right)

= E_{t,x} \left( e^{-r(T-t)} G(T, X_T) - \int_{t+s}^{T} e^{-r(u-t)} H(t + u, X_{t+u}) I\{X_{t+u} < c(t+u)\} \, du \\
+ \frac{1}{2} \int_{t+s}^{T} e^{-r(u-t)} (K - L) Z_x(u, L-) \, du \right) = U(t, x),

after which, the following equation emerges

\[ E_{t,x} \left( e^{-r(s)} U(t + s, X_{t+s}) \right) - U(t, x) = E_{t,x} \left( \int_{0}^{s} e^{-ru} H(t + u, X_{t+u}) I\{X_{t+u} < c(t+u)\} \, du \\
- \frac{1}{2} \int_{0}^{s} e^{-ru} (K - L) Z_x(t + u, L-) \, du \right). \tag{5.2.55} \]

With the above observation in mind, we now ready to show that \( U(t, x) = G(t, x) \) for all \( 0 < x \leq c(t) \). We first consider the stopping time,

\[ \sigma_c = \inf \{0 \leq s \leq T - t : X_{t+s} \geq c(t+s)\}, \]

and note that, since \( c \) solves (5.2.50), \( U(t, c(t)) = G(t, c(t)) \) for all \( 0 \leq t < T \) and because of the definition of \( U, U(T, x) = G(T, x) \). From (5.2.55), we see that for \( 0 < x \leq c(t) \),

\[ U(t, x) = E_{t,x} \left( e^{-\sigma_c} U(t + \sigma_c, X_{t+\sigma_c}) - \int_{0}^{\sigma_c} e^{-ru} H(t + u, X_{t+u}) I\{X_{t+u} < c(t+u)\} \, du \\
+ \frac{1}{2} \int_{0}^{\sigma_c} e^{-ru} (K - L) Z_x(t + u, L-) \, du \right) \\
= E_{t,x} \left( e^{-\sigma_c} G(t + \sigma_c, X_{t+\sigma_c}) - \int_{0}^{\sigma_c} e^{-ru} H(t + u, X_{t+u}) I\{X_{t+u} < c(t+u)\} \, du \\
+ \frac{1}{2} \int_{0}^{\sigma_c} e^{-ru} (K - L) Z_x(t + u, L-) \, du \right) \\
= G(t, x) + E_{t,x} \left( \int_{0}^{\sigma_c} e^{-ru} H(t + u, X_{t+u}) \, du - \frac{K - L}{2} \int_{0}^{\sigma_c} e^{-ru} Z_x(t + u, L-) \, du \right) \\
- E_{t,x} \left( \int_{0}^{\sigma_c} e^{-ru} H(t + u, X_{t+u}) \, du - \frac{K - L}{2} \int_{0}^{\sigma_c} e^{-ru} Z_x(t + u, L-) \, du \right) = G(t, x), \]
where the third equality is because of (5.2.19) and the definition of \( \sigma_c \), that is, for \( u \in [0, \sigma_c] \), 
\[ I\{X_{t+u} < c(t + u)\} = 1 \]; after which, the desired assertion follows, so that (5.2.54) equals

\[
E_{t,x} \left( e^{-r\tau} V^c (t + s, X_{t+s}) \right) = V^c (t, x) \\
+ E_{t,x} \left( \int_0^s e^{-ru}H (t + u, X_{t+u}) I\{X_{t+u} < c(t + u)\} du \\
+ \frac{1}{2} \int_0^s e^{-ru} (G_x (t + u, L+) - G_x (t + u, L-)) dl^L_u (X) \right). \tag{5.2.56}
\]

(iv) Before moving on, we pause for another direct consequence of (5.2.56): \( V^c (t, x) \leq V (t, x) \).
To see this relation, we begin by considering the stopping time

\[ \tau_c = \inf\{0 \leq s \leq T - t : X_{t+s} \leq c(t + s)\} \]

With \( s \) replaced by \( \tau_c \) on (5.2.56), we find that

\[
E_{t,x} \left( e^{-r\tau_c} V^c (t + \tau_c, X_{t+\tau_c}) \right) = E_{t,x} \left( e^{-r\tau_c} G(t + \tau_c, X_{t+\tau_c}) \right) = V^c (t, x),
\]

in the sense that \( L < c < K \), joining which, with the definition of \( V \) in (5.2.5), proves \( V^c \leq V \).

(v) With the aid of the above discussions, now is time for the justification of the uniqueness of \( b \). Here comes the first relation \( c \geq b \) on \([0, T]\). Let us fix \((t, x) \in (0, T) \times (0, \infty)\) so that \( x < b(t) \wedge c(t) \) and consider the stopping time

\[ \sigma_b = \inf\{0 \leq s \leq T - t : X_{t+s} \geq b(t + s)\} \]

Via inserting \( \sigma_b \) in place of \( s \) on equations (5.2.53) and (5.2.56), we obtain

\[
E_{t,x} \left( e^{-r\sigma_b} V (t + \sigma_b, X_{t+\sigma_b}) \right) = V (t, x) + E_{t,x} \left( \int_0^{\sigma_b} e^{-ru}H (t + u, X_{t+u}) du \\
- \frac{1}{2} \int_0^{\sigma_b} e^{-ru} (K - L) Z_x (t + u, L-) dl^L_u (X) \right),
\]

\[
E_{t,x} \left( e^{-r\sigma_b} V^c (t + \sigma_b, X_{t+\sigma_b}) \right) = V^c (t, x) + E_{t,x} \left( \int_0^{\sigma_b} e^{-ru}H (t + u, X_{t+u}) I\{X_{t+u} < c(t + u)\} du \\
- \frac{1}{2} \int_0^{\sigma_b} e^{-ru} (K - L) Z_x (t + u, L-) dl^L_u (X) \right),
\]

which, from the view of the relation \( V^c \leq V \), implies that

\[
E_{t,x} \left( \int_0^{\sigma_b} e^{-ru}H (t + u, X_{t+u}) du \right) \geq E_{t,x} \left( \int_0^{\sigma_b} e^{-ru}H (t + u, X_{t+u}) I\{X_{t+u} \leq c(t + u)\} du \right),
\]

that is,

\[
E_{t,x} \left( \int_0^{\sigma_b} e^{-ru}H (t + u, X_{t+u}) I\{X_{t+u} \geq c(t + u)\} du \right) \geq 0,
\]

but \( H < 0 \), forcing \( c \geq b \) on \([0, T]\).
(vi) In order to fill the last slot in the proof, we assume that, there exists \( t \in (0, T) \) such that \( c(t) > b(t) \) and choose \( x \in (b(t), c(t)) \). Now consider the stopping time

\[
\tau_b = \inf\{0 \leq u \leq T - t : X_{t+u} \leq b(t+u)\},
\]

and once again, insert \( \tau_b \) in the place of \( s \) on equations (5.2.53) and (5.2.56),

\[
E_{t,x} (e^{-r\tau_b} G(t + \tau_b, X_{t+\tau_b})) = V(t, x) + E_{t,x} \left( \frac{1}{2} \int_0^{\tau_b} e^{-ru}(K - L)Z_x(t + u, L-)dL_u^X(X) \right)
\]

\[
E_{t,x} (e^{-r\tau_b} G(t + \tau_b, X_{t+\tau_b})) = V^c(t, x) + E_{t,x} \left( \int_0^{\tau_b} e^{-ru}H(t + u, X_{t+u})I\{X_{t+u} \leq c(t + u)\}du \right.
\]
\[
\left. - \frac{1}{2} \int_0^{\tau_b} e^{-ru}(K - L)Z_x(t + u, L-)dL_u^X(X) \right).
\]

which, from the view of the relation \( V^c \leq V \), entails that

\[
E_{t,x} \left( \int_0^{\tau_b} e^{-ru}H(t + u, X_{t+u})I\{X_{t+u} \leq c(t + u)\}du \right) \geq 0, \quad (5.2.57)
\]

but \( H < 0 \), leading to a contradiction so that such point cannot exist. This completes the proof. \( \square \)

\[\text{Figure 5.7: This figure displays the maps } t \mapsto b(t) \text{ and } t \mapsto B(t) \text{ on } [0, T] \text{ with the same chosen parameters } r = 0.05, \sigma = 0.4, L = 2, K = 7, T = 10, \text{ where functions } b \text{ and } B \text{ are defined as the optimal stopping boundaries for problem (5.2.1) in the case of } L < B(0) \text{ and the standard American put option (in page 82) respectively.}\]
5.2 Finite-time Horizon

In analogy with the arguments that preceded Lemma 5.2.39, technical conditions for Theorem 2.2.8 to be applied have to be verified before we can really take off.

**Lemma 5.2.41.** Let the function $F(t, x) = e^{-rt}V(t, x)$ defined on $[0, T) \times (0, \infty)$,

- $D_1 = \{(t, x) \in [0, t_b] \times (0, b(t))\}$,
- $D_2 = \{(t, x) \in [t_b, t^g] \times (0, L)\}$,
- $D_3 = \{(t, x) \in [t^g, T] \times (0, L) \cup (L, b(t))\}$.

Then, Theorem 2.2.7 is applicable to $F$ if the following conditions are met:

(a) $F(t, x)$ is $C^{1,2}$ on $C \cup D_1 \cup D_2 \cup D_3$;
(b) $F_t + L_b F$ is locally bounded;
(c) $t \mapsto F_x(t, b(t)\pm)$ is continuous;
(d) $t \mapsto F_x(t, L\pm)$ is continuous;
(e) $F_{xx} = F_1 + F_2$ on $C \cup D_1 \cup D_2 \cup D_3$, where $F_1$ is non-negative and $F_2$ is continuous.

**Proof.** Statements (a), (b) and (c) follow by repeating same proof as that of statements (a) and (b) in Lemma 5.2.39.

The maps $t \mapsto F_x(t, b(t)\pm)$ for $t \in [0, T)$ and $t \mapsto F_x(t, L\pm)$ for $t \in [0, t_b]$ are continuous as the result of $V$ is $C^{1,2}$ on the continuation set $C$, see [21, Page 8] and Lemma 3.3.2. Moreover, the map $t \mapsto F_x(t, b(t)\pm)$ is continuous as $F_x(t, b(t)\pm) = e^{-rt}G_x(t, b(t)\pm)$ and the map $t \mapsto b(t)$ is continuous, which, joining with $F_x(t, L\pm) = e^{-rt}G(t, L\pm)$ for $t \in [t^g, T]$ (see A.1.4), justifies statements (c) and (d). \qed

**Theorem 5.2.42.** The optimal stopping boundary $b : [0, T] \mapsto \mathbb{R}$ in problem (5.2.5) is given as $b(t) = b_1(t) \vee b_2(t)$ where $b_1(t) = x^*(t)$ for $t \in [0, t_*]$ with $x^* : [0, t_*] \times (0, L)$ being continuous non-decreasing function defined in Lemma 5.2.23, $b_2 : [0, T] \mapsto \mathbb{R}$ is in the class of continuous non-decreasing functions $c : [0, T] \mapsto \mathbb{R}$ satisfying $0 < c(t) < K$ for all $t \in [0, T)$ and $b_2$ can be characterised as the unique solution of the following nonlinear integral equation\footnote{The computation of expectation of the local time term is on A.1.2 or see [23, Page 1523], [21, Page 9].}:

\[
G(t, b_2(t)) = E_{t, b_2(t)} \left( e^{-r(T-t)}G(T, X_T) - \int_0^s e^{-ru} H(t + u, X_{t+u}) I\{X_{t+u} < b_2(t+u)\} du \right) \\
- \frac{1}{2} \int_0^s e^{-ru} I\{b_2(t+u) > L\} (G_x(t + u, L+) - G_x(t + u, L-)) dE_{t,x} \left( \psi^L_u(X) \right) \\
- \frac{1}{2} \int_0^s e^{-ru} I\{b_2(t+u) = L\} (G_x(t + u, L+) - G_x(t + u, L-)) dE_{t,x} \left( \psi^L_u(X) \right),
\]

satisfying $b(T) = K$. The value of the current contract admits the following “early exercise premium” representation:

\[
V(t, x) = E_{t,x} \left( e^{-r(T-t)}G(T, X_T) - \int_0^{T-t} e^{-ru} H(t + u, X_{t+u}) I\{X_{t+u} < b(t+u)\} du \right)
\]
- \frac{1}{2} \int_0^{T-t} e^{-ru} I\{b(t + u) > L\} (G_x(t + u, L^+) - G_x(t + u, L^-)) dE_{t,x} \left( l_u^L(X) \right)
- \frac{1}{2} \int_0^{T-t} e^{-ru} I\{b(t + u) = L\} (G_x(t + u, L^+) - G_x(t + u, L^-)) dE_{t,x} \left( l_u^L(X) \right) \quad (5.2.59)

for all \((t, x) \in [0, T] \times (0, \infty)\).

**Proof of The Formulas (5.2.58) and (5.2.59).**

Since the function \(F\) fulfils the conditions on Lemma 5.2.41, an application of Theorem 2.2.7 yields

\[
e^{-rs} V(t + s, X_{t+s}^s) = V(t, x) + M_s
\]

\[
+ \int_0^s e^{-ru} \left(- ru + V_t + \mathbb{L}_X V\right) (t + u, X_{t+u}^s) I\{X_{t+u}^s < b(t + u), X_{t+u}^s \neq L\} du
\]

\[
+ \frac{1}{2} \int_0^s e^{-ru} (V_x(t + u, L^+) - V_x(t + u, L^-)) dl_u^L(X^x)
\]

\[
+ \frac{1}{2} \int_0^s e^{-ru} \{b(t + u) > L\} (G_x(t + u, L^+) - G_x(t + u, L^-)) dl_u^L(X^x)
\]

\[
+ \frac{1}{2} \int_0^s e^{-ru} \{b(t + u) = L\} (V_x(t + u, L^+) - V_x(t + u, L^-)) dl_u^L(X^x), \quad (5.2.60)
\]

where \(M_s = \int_0^s e^{-ru} \sigma X_{t+u} V_x(t + u, X_{t+u}) I\{X_{t+u}^s \neq b(t + u), X_{t+u}^s \neq L\} dW_u\) is a martingale under measure \(P_{t,x}\) for each \(s \in [0, T - t]\), and the second equality is due to the fact that the smooth-fit condition fails as \(b(t) = L\) and the gain function is not smooth as \(b(t) > L\) for \(t \in [0, T]\).

Then, upon taking the expectation under measure \(P_{t,x}\) of (5.2.60) and invoking the optional sampling theorem, we obtain

\[
E_{t,x} \left(e^{-rs} V(t + s, X_{t+s}^s)\right) = V(t, x) + E_{t,x} \left( \int_0^s e^{-ru} H(t + u, X_{t+u}) I\{X_{t+u} < b(t + u)\} du \right)
\]

\[
+ \frac{1}{2} \int_0^s e^{-ru} I\{b(t + u) > L\} (G_x(t + u, L^+) - G_x(t + u, L^-)) dE_{t,x} \left( l_u^L(X) \right)
\]

\[
+ \frac{1}{2} \int_0^s e^{-ru} I\{b(t + u) = L\} (V_x(t + u, L^+) - V_x(t + u, L^-)) dE_{t,x} \left( l_u^L(X) \right), \quad (5.2.61)
\]

where

\[
H(t, x) = (-ru + G_t + \mathbb{L}_X G)(t, x) = \begin{cases} -RK, & L < x < b(t), \\ (-KZ - \sigma^2x^2Z_x)(t, x), & x < L. \end{cases}
\]

Next, let \(s = T - t\) so that \(E_{t,x} \left( e^{-r(T-t)} V(T, X_T) \right) = E_{t,x} \left( e^{-r(T-t)} G(T, X_T) \right)\), which, together with (5.2.61) shows that

\[
V(t, x) = E_{t,x} \left( e^{-r(T-t)} G(T, X_T) - \int_0^{T-t} e^{-ru} H(t + u, X_{t+u}) I\{X_{t+u} < b(t + u)\} du \right)
\]
\[
- \frac{1}{2} \int_0^{T-t} e^{-ru} I\{b(t+u) > L\} (G_x (t+u, L^+) - G_x (t+u, L^-)) dE_{t,x} \left( l^{L}_u (X) \right) \\
- \frac{1}{2} \int_0^{T-t} e^{-ru} I\{b(t+u) = L\} (G_x (t+u, L^+) - G_x (t+u, L^-)) dE_{t,x} \left( l^{L}_u (X) \right)
\]

where \( G(T, x) = (K - x) I\{L < x < K\} \); after which, let \( x = b(t) \) and since \( V(t, b(t)) = G(t, b(t)) \), (5.2.58) follows.

**Remark 5.2.43.** To avoid the danger of confusion in notations, in what follows, we simply assume, without loss of generality, that \( b_2(t) \geq b_1(t) \) for all \( t \in [0, T] \) (see Figure 5.9) such that the optimal stopping boundary is just the unique solution for (5.2.58), that is \( b = b_2 \).

**Proof of uniqueness of the solution in (5.2.58).**

The proof is almost the same as that of Theorem 5.2.40 and the difference is coming from the local-time term. Henceforth, using that recipe, we set

\[
U(t, x) = E_{t,x} \left( e^{-r(T-t)} G(T, X_T) - \int_0^{T-t} e^{-ru} H(t+u, X_{t+u}) I\{X_{t+u} < c(u)\} du \right)
\]

\[
- \frac{1}{2} \int_0^{T-t} e^{-ru} I\{c(t+u) > L\} (G_x (t+u, L^+) - G_x (t+u, L^-)) dE_{t,x} \left( l^{L}_u (X) \right)
\]

\[
- \frac{1}{2} \int_0^{T-t} e^{-ru} I\{c(t+u) = L\} (U_x (t+u, L^+) - U_x (t+u, L^-)) dE_{t,x} \left( l^{L}_u (X) \right), \quad (5.2.62)
\]

where

\[
\begin{align*}
U_x(t, L^+) &= -1, & c(t) &= L, \\
U_x(t, L^-) &= -1 + (K - L) Z_x(t, L^-), & c(t) &\geq L,
\end{align*}
\]

and since \( c \) solves (5.2.58), \( U(t, c(t)) = G(t, c(t)) \), after which, steps (i) and (ii) in the proof of Theorem 5.2.40 can be carried over without changes.

(iii) Another interpretation of step (iii) in Theorem 5.2.40 is that \( V^c(t, x) = U(t, x) \) for \( (t, x) \in [0, T] \times (0, \infty) \), which will be well illustrated here. As before, we observe that, via equation (5.2.62), strong Markov property and smoothing lemma (Lemma 2.3.12),

\[
E_{t,x} \left( e^{-rs} U(t+s, X_{t+s}) - \int_0^s e^{-ru} H(t+u, X_{t+u}) I\{X_{t+u} < c(u)\} du \right)
\]

\[
- \frac{1}{2} \int_0^s e^{-ru} I\{c(t+u) \geq L\} (G_x (t+u, L^+) - G_x (t+u, L^-)) dE^{L}_u (X)
\]

\[
= E \left( e^{-r(T-t-s)} G(T, X_T) - \int_{t+s}^T e^{-r(u-t-s)} H(u, X_u) I\{X_u < c(u)\} du \right)
\]

\[
- \frac{1}{2} \int_{t+s}^T e^{-r(u-t-s)} I\{c(u) \geq L\} (G_x (u, L^+) - G_x (u, L^-)) dE^{L}_u (X) \mid F_{t+s}
\]

\[
- \int_t^{t+s} e^{-r(u-t)} H(u, X_u) I\{X_u < c(u)\} du \\
- \frac{1}{2} \int_t^{t+s} e^{-r(u-t)} I\{c(u) \geq L\} (G_x (u, L^+) - G_x (u, L^-)) dE^{L}_u (X) \mid F_t
\]
exhibiting its martingale property and the relation:

\[ U(t,x) = \mathbb{E}_t \left( e^{-r(T-t)} G(T, X_T) - \int_{t+s}^T e^{-r(u-t)} H(u, X_u) I\{X_u < c(u)\} du \right) \]

\[ - \frac{1}{2} \int_{t+s}^T e^{-r(u-t)} I\{c(u) \geq L\} (G_x(u, L+) - G_x(u, L-)) dl^L_u(X) \]

\[ - \int_{t}^{t+s} e^{-r(u-t)} H(u, X_u) I\{X_u < c(u)\} du \]

\[ - \frac{1}{2} \int_{t}^{t+s} e^{-r(u-t)} I\{c(u) > L\} (G_x(u, L+) - G_x(u, L-)) dl^L_u(X) \]

\[ = \mathbb{E}_t \left( e^{-r(T-t)} G(T, X_T) - \int_{0}^{T-t} e^{-ru} H(t+u, X_{t+u}) I\{X_{t+u} < c(t+u)\} du \right) \]

\[ + \frac{K - L}{2} \int_{0}^{T-t} e^{-ru} I\{c(t+u) \geq L\} Z_x(t+u, L-) dl^L_u(X) \]

\[ = U(t,x), \]

exhibiting its martingale property and the relation:

\[ E_{t,x} \left( e^{-rs} U(t+s, X_{t+s}) \right) - U(t,x) \]

\[ = E_{t,x} \left( \int_{0}^{s} e^{-ru} H(t+u, X_{t+u}) I\{X_{t+u} < c(t+u)\} du \right) \]

\[ - \frac{K - L}{2} \int_{0}^{s} e^{-ru} I\{c(t+u) \geq L\} Z_x(t+u, L-) dl^L_u(X) \].

(5.2.63)

We are now have made the preparation needed to show that \( U(t,x) = G(t,x) \) for \((t,x) \in \mathcal{D}_c\). Let us consider the stopping time,

\[ \sigma_c = \inf \{0 \leq s \leq T - t : X_{t+s} \geq c(t+s)\}, \]

so that, from the fact that \(c(t)\) solves (5.2.58), \( U(t + \sigma_c, X_{t+\sigma_c}) = G(t + \sigma_c, X_{t+\sigma_c}) \).

Then, let \(s = \sigma_c\) in equation (5.2.63) so that for \(x \in (0, c(t)]\),

\[ U(t,x) = \mathbb{E}_{t,x} \left( e^{-r\sigma_c} U(t + \sigma_c, X_{t+\sigma_c}) - \int_{0}^{\sigma_c} e^{-ru} H(t+u, X_{t+u}) I\{X_{t+u} < c(t+u)\} du \right) \]

\[ - \frac{1}{2} \int_{0}^{\sigma_c} e^{-ru} I\{c(t+u) \geq L\} (G_x(t+u, L+) - G_x(t+u, L-)) dl^L_u(X) \]

\[ = \mathbb{E}_{t,x} \left( e^{-r\sigma_c} G(t + \sigma_c, X_{t+\sigma_c}) - \int_{0}^{\sigma_c} e^{-ru} H(t+u, X_{t+u}) du \right) \]

\[ - \frac{1}{2} \int_{0}^{\sigma_c} e^{-ru} I\{c(t+u) > L\} (G_x(t+u, L+) - G_x(t+u, L-)) dl^L_u(X) \]

\[ = G(t,x) + \mathbb{E}_{t,x} \left( \int_{0}^{\sigma_c} e^{-ru} H(t+u, X_{t+u}) du \right) \]

\[ - \mathbb{E}_{t,x} \left( \frac{K - L}{2} \int_{0}^{\sigma_c} e^{-ru} Z_x(t+u, L-) dl^L_u(X) \right) \]

\[ - \mathbb{E}_{t,x} \left( \int_{0}^{\sigma_c} e^{-ru} H(t+u, X_{t+u}) du \right) \]

\[ + \mathbb{E}_{t,x} \left( \frac{K - L}{2} \int_{0}^{\sigma_c} e^{-ru} I\{c(t+u) > L\} Z_x(t+u, L-) dl^L_u(X) \right) \]
\[ G(t, x) - E_{t,x} \left( \frac{K - L}{2} \int_0^{\sigma_c} e^{-ru} I\{c(t + u) \leq L\} Z_x(t + u, L-)dL_u(X) \right) = G(t, x), \]

where the first and last equality is due to the definition of \( \sigma_c \), in particular, for \( c \leq L, X \) will either never reach \( L \) before the process stops or reach \( L \) but stop simultaneously; and the third equality holds true via (5.2.19).

This proves that \( U(t, x) = G(t, x) \) for \( (t, x) \in \mathcal{D}^c \) as desired; correspondingly, from the construction of \( V^c(t, x) \), the conclusion that \( V^c(t, x) = U(t, x) \) on \([0, T] \times (0, \infty)\) can be drawn so that

\[
E_{t,x} \left( e^{-rs} V^c(t + s, X_{t+s}) \right) - V^c(t, x) \\
= E_{t,x} \left( \int_0^s e^{-ru} H(t + u, X_{t+u}) I\{X_{t+u} < c(t + u)\} du \\
- \frac{K - L}{2} \int_0^s e^{-ru} I\{c(t + u) \geq L\} Z_x(t + u, L-)dL_u(X) \right). \quad (5.2.64)
\]

(iv) Re-rerunning step (iv) in the proof of Theorem 5.2.40 with minor modifications of considering the stopping time \( \tau_c = \inf\{0 \leq s \leq T - t : X_{t+s} \leq c(t + s)\} \) and replacing \( s \) with \( \tau_c \) in (5.2.64), we once again arrives at the relation \( V^c \leq V \). Namely,

\[
V^c(t, x) = E_{t,x} \left( e^{-r\tau_c} V^c(t + \tau_c, X_{t+\tau_c}) - \int_0^{\tau_c} e^{-ru} H(t + u, X_{t+u}) I\{X_{t+u} < c(t + u)\} du \\
+ \frac{K - L}{2} \int_0^{\tau_c} e^{-ru} I\{c(t + u) \geq L\} Z_x(t + u, L-)dL_u(X) \right) \\
= E_{t,x} \left( e^{-r\tau_c} V^c(t + \tau_c, X_{t+\tau_c}) \right) \\
= E_{t,x} \left( e^{-r\tau_c} G(t + \tau_c, X_{t+\tau_c}) \right) \leq V(t, x),
\]

where the second equality is based on the definition of \( \tau_c \): if \( c(t) > L \), then \( X \) will never hit \( L \) before stopping; on the other hand, if \( c(t) = L \), it will spend zero time on level \( L \).

(v) Analogously, our task in this step is to prove the relation \( c \geq b \) holds true on \([0, T] \).

Let us fix \( (t, x) \in (0, T) \times (0, \infty) \) such that \( x < b(t) \wedge c(t) \) and consider the stopping time

\[
\sigma_b = \inf\{0 \leq s \leq T - t : X_{t+s} \geq b(t + s)\}.
\]

Replacing \( s \) by \( \sigma_b \) in (5.2.64) and in (5.2.61) yields,

\[
E_{t,x} \left( e^{-r\sigma_b} V^c(t + \sigma_b, X_{t+\sigma_b}) \right) - V^c(t, x) \\
= E_{t,x} \left( \int_0^{\sigma_b} e^{-ru} H(t + u, X_{t+u}) I\{X_{t+u} < c(t + u)\} du \\
- \frac{K - L}{2} \int_0^{\sigma_b} e^{-ru} I\{c(t + u) \geq L\} Z_x(t + u, L-)dL_u(X) \right)
\]

and

\[
E_{t,x} \left( e^{-r\sigma_b} V(t + \sigma_b, X_{t+\sigma_b}) \right) - V(t, x)
\]
From early step (iv), we know that
\[ E_{t,x} \left( e^{-r_0} V^{c'}(t+\sigma_b, X_{t+\tau}) \right) = E_{t,x} \left( e^{-r_0} V(t+\sigma_b, X_{t+\tau}) \right), \]
and that \( V^c(t, x) = V(t, x) = G(t, x) \), it then follows that

\[
E_{t,x} \left( \int_0^{\sigma_b} e^{-r_0} H(t+u, X_{t+u}) I\{X_{t+u} > c(t+u)\} du \right) \]
\[
- \frac{K-L}{2} \int_0^{\sigma_b} e^{-r_0} I\{b(t+u) \geq L\} Z_x(t+u, L-) dl^L_u(X) \geq 0. \tag{5.2.65}
\]

The desired claim is that \( c \geq b \), so let us assume the contrary that \( c < b \). It then follows that \( I\{c(t+u) \geq L\} = 1 \) entails \( I\{b(t+u) \geq L\} = 1 \). In addition, note that \( H < 0 \), leading to a contradiction and thereby proving \( c \geq b \).

(vi) In this final step, we wish to show that \( c = b \) on \([0, T]\). To achieve this, suppose that \( c = b \) is violated such that \( c > b \) (from the previous step). Then, by choosing \( x \in (b(t), c(t)) \) and considering the stopping time
\[ \tau_b = \inf \{0 \leq s \leq T - t : X_{t+s} \leq b(t+s)\} \]
and with \( s \) in (5.2.64) and in (5.2.61) replaced by \( \tau_b \), noticing that \( (t + \tau_b, X_{t+\tau_b}) \in D \subset D_c \), we arrive at

\[
E_{t,x} \left( e^{-r_0} G(t + \tau_b, X_{t+\tau_b}) \right) = V(t, x),
\]
\[
E_{t,x} \left( e^{-r_0} G(t + \tau_b, X_{t+\tau_b}) \right) = V^c(t, x) + E_{t,x} \left( \int_0^{\sigma_b} e^{-r_0} H(t+u, X_{t+u}) I\{X_{t+u} < c(t+u)\} du \right) \]
\[
- E_{t,x} \left( \frac{K-L}{2} \int_0^{\sigma_b} e^{-r_0} I\{c(t+u) \geq L\} Z_x(t+u, L-) dl^L_u(X) \right). \]

The relation \( V^c \leq V \) therefore suggests that

\[
E_{t,x} \left( \int_0^{\sigma_b} e^{-r_0} H(t+u, X_{t+u}) I\{X_{t+u} < c(t+u)\} du \right) \]
\[
- \frac{K-L}{2} \int_0^{\sigma_b} e^{-r_0} I\{c(t+u) \geq L\} Z_x(t+u, L-) dl^L_u(X) \geq 0. \tag{5.2.66}
\]

which leads to a contradiction since (5.2.66) is strictly negative, in particular, \( H < 0 \) so are the local time terms. This establishes the uniqueness of the solution in (5.2.58).

The proof of Theorem 5.2.42 is finally complete.

\( \square \)

Remark 5.2.44. The algorithm, that solves equations (5.2.50) and (5.2.58) numerically in order to obtain the optimal stopping boundary \( b \), is provided in A.6.
5.2 Finite-time Horizon

Figure 5.8: This figure displays the maps $t \mapsto b(t)$ and $t \mapsto B(t)$ on $[0, T]$ with the same chosen parameters $r = 0.05, \sigma = 0.4, L = 6.5, K = 7, T = 10$, where functions $b$ and $B$ are defined as the optimal stopping boundaries for problem (5.2.1) in the case of $B(0) < L$ and the standard American put option (in page 82) respectively.

Figure 5.9: This figure displays the maps $t \mapsto b(t)$, $t \mapsto B(t)$ on $[0, T]$ and $t \mapsto x^*(t)$ on $[0, t_\ast]$ with the same chosen parameters $r = 0.05, \sigma = 0.4, L = 5, K = 7, T = 10$, where functions $b, B$ are defined as that in Figure 5.8 and the detailed definition of function $x^*$ is given in Lemma 5.2.23.
Chapter 6

Additional Remarks

In the early 20th century, the Argentinean writer Jorge Luis Borges envisioned a vast and fantastic, infinite-in-size library (the library of Babel, [9]) that would contain every possible book. Some of these books would contain great knowledge, such as the Grand Unified Theory of Physics, the first million digits of the decimal expansion of pi, the perfect game of chess but some of the books would also involve mistakes of one sort or another, such as a hundred false refutations of Einstein’s theory of relativity, a million false reference books for the library itself. The Borgesian library therefore simultaneously contains all the information in the world and none at all.

Our work is certainly nowhere near the fascination of the library of Babel but it also avoids the overall meaninglessness of taking a vase repository of information without any proper framework. As an inevitable tradeoff, not every methods and models are being investigated and presented carefully in details. This chapter presents some extra remarks hoping to make up for such limitation.

The Bolza-formulated Problem

The methodology developed in Chapter 5 can be potentially useful for the study of Bolza-formulated problem, which is of the form as follows:

\[
V(t, x) = \sup_{\tau \in [0, T-t]} E_{t,x} \left( G(t + \tau, X_{t+\tau}) + \int_0^{\tau} H(t + u, X_{t+u}) du \right),
\]

where the supremum is taken over stopping time \( \tau \) of \( X \) satisfying \( E_{0,x} \int_0^{\tau} H(t + u, X_{t+u}) du < \infty \) and as before, the gain function \( G \) is uniformly integrable.

**Remark 6.0.1.** The stopping set and continuation set of problem (6.0.1) are defined as

\[
D = \{(t, x) \in [0, T] \times \mathbb{R} : V(t, x) = G(t, x)\},
\]

\[
C = \{(t, x) \in [0, T] \times \mathbb{R} : V(t, x) > G(t, x)\}.
\]

For more, see [62, Page 215] and [64, Page 314].

Now consider the optimal stopping problem that we have discussed so far:

\[
\bar{V}(t, x) = \sup_{\tau \in [0, T-t]} E_{t,x} (G(t + \tau, X_{t+\tau})),
\]

(6.0.2)
and supposed that in both problems, the stopping and continuation sets are non-empty and the first entrance time of the stopping set is optimal. For notational convenience we denote the stopping set as \( \mathcal{D} \) and \( \overline{\mathcal{D}} \), continuation set as \( \mathcal{C} \) and \( \overline{\mathcal{C}} \) in problems (6.0.1) and (6.0.2) respectively.

**Lemma 6.0.2.** (i) If function \( H : [0, T] \times \mathbb{R} \to \mathbb{R} \) is non-negative, then \( \mathcal{D} \subseteq \overline{\mathcal{D}} \) and \( \overline{\mathcal{C}} \subseteq \mathcal{C} \); (ii) If function \( H : [0, T] \times \mathbb{R} \to \mathbb{R} \) is non-positive, then \( \overline{\mathcal{D}} \subseteq \mathcal{D} \) and \( \mathcal{C} \subseteq \overline{\mathcal{C}} \).

**Proof.** For \( H \geq 0 \), \( \overline{V}(t, x) - G(t, x) \leq V(t, x) - G(t, x) \), which, together with the definition of both sets, proves the first assertion. The second assertion follows from the similar argument that for \( H \leq 0 \), \( \overline{V}(t, x) - G(t, x) \geq V(t, x) - G(t, x) \). \( \square \)

**Example 6.0.3** (Anticipating the Future). Let us follow the exact same parameters’ setting as that in Chapter 5 and consider the optimal stopping problem:

\[
V = \sup_{\tau \in [0, T]} E_{0,x} \left( e^{-r(\tau \wedge \theta)} (K - X_{\tau \wedge \theta})^+ \right)
\]

and notice clearly that for \( \tau \geq \theta \), the rule to decide “when to exercise” gets seemingly mixed up with the future information because of \( \theta \). This example intends to investigate how this anticipation will affect the solution(s).

Exploiting the same trick as before for the first term while observing, from the fact that the predictable projection \( A_t \) is the \( \mathcal{F}_t \)-predictable compensator of \( I\{\theta \leq t\} \), i.e. the predictable increasing process (see [54, Page 12]) and the Doob-Meyer decomposition, that:

\[
E_{0,x} \left( e^{-r\theta} (K - X_\theta)^+ I\{\tau \geq \theta\} \mid \mathcal{F}_\tau \right) = E_{0,x} \left( \int_0^\tau e^{-ru} (K - X_u)^+ dA_u \mid \mathcal{F}_\tau \right) = \int_0^\tau e^{-ru}(K - X_u)^+ dA_u = \int_0^\tau \frac{K - L}{L} \left( \frac{1}{2} - \frac{r}{\sigma^2} \right) e^{-ru} dL_u \geq 0.
\]

In so doing, we arrive at the Bolza-formulated problem:

\[
V = \sup_{\tau \in [0, T]} E_{0,x} \left( e^{-r\tau} (K - X_\tau)^+ Z(\tau + \tau, X_\tau) + \int_0^\tau \frac{K - L}{L} \left( \frac{1}{2} - \frac{r}{\sigma^2} \right) e^{-ru} dL_u \right),
\]

and its “coupled” problem studied in Chapter 5 is

\[
\overline{V} = \sup_{\tau \in [0, T]} E_{0,x} \left( e^{-r\tau} (K - X_\tau)^+ Z(\tau, X_\tau) \right).
\]

The first statement in Lemma 6.0.2 then suggests that \( \mathcal{D} \subseteq \overline{\mathcal{D}} \) and \( \overline{\mathcal{C}} \subseteq \mathcal{C} \). However, further computation discloses that no optimal stopping strategy can be found when the optimal stopping boundary is below the fixed level \( L \), at the moment of writing, we have not come up with any proper financial explanation for such result yet.
Financial Interpretation

At this very end of our journey, let us try to make financial sense of Chapters 4 and 5 in the following two aspects:

1. Defaultable Contingent Claim with Zero Recovery

When a contingent claim is valued, it is conventional to assume that there is no risk that the counter-party will default, that is, the counter-party will always be able to make the promised payment. However, this assumption is far less defensible in the non-exchange-traded (or over-the-counter) market and the importance of considering default risk in over-the-counter markets is recognised by bank capital adequacy standards, see [42, Page 300] and [13] for a more detailed introduction of default risk.

According to [27, Pages 188, 190], under certain condition that the default market is complete, meaning that the defaultable contingent claim is hedgeable, the value functions (4.0.1) and (5.0.1) can be viewed as the value of defaultable contingent claims priced by uninformed agent who does not know the time when the default occurs, which is modelled by the last exit times; and if default does occur before the optimal stopping time, the payoff functions in our work suggest that the agent loses everything.

2. Temporal Risk Aversion

In [60], the temporal risk aversion is described as the unwillingness to accept an actuarially fair risk. Under this definition, agents are said to be (temporal) risk averse if they are willing to pay non-negative amount for the privilege of exchanging any risk for its actual value.

According to Theorem 6 in [60, Page 224], the optimal stopping rules chosen by agents actually reveals their temporal risk aversion. Let us ambiguously translate Theorem 6 into the “language" used in our context: assume there are two agents (A and B) considering the optimal stopping problem with only different reward $G_A$ and $G_B$ respectively and that in both problems, the optimal stopping times exist, denoted as $\tau_A$ and $\tau_B$, Agent A is said to be more (temporal) risk averse than Agent B, if and only if $\tau_A < \tau_B$.

Apparently, our results reveal that comparing to rational agents who face the optimal stopping problem with standard payoff of Russian and American put options, agents facing the optimal stopping problems due to their own belief preference in Chapter 4 and 5 are actually more (temporal) risk averse (lack of risk appetite), given that their stopping sets are set larger (for example, in Chapter 5, we have $D^A \subseteq D$), implying they tend to stop before rational agents do.

From the point of view in [47], the pricing problem is actually two-sided in the sense that “the buyer’s objective is that he starts out with the amount of $-x$ dollars (the fair price he is willing to pay for Ticket A) at time 0, and looks for a stopping time $\tau_s$ such that the payment he receives allows him to recover the debt he incurred at $t = 0$ to purchase Ticket A; the seller’s objective is to use $x$ dollars received from the buyer and find a portfolio that makes it possible for him to fulfil his obligation without risk (with probability 1) whenever the buyer choose to exercise Ticket A.” From this viewpoint, the aftermath of being irrational seemingly leaves some elbow room for arbitrage opportunities (for the seller) \(^1\), considering that irrational agent must have paid the fair price to own

\(^1\)This is why we only claim the price that irrational agent is willing to pay for, instead of claiming that we find the arbitrage-free price of the presuming contract in Chapter 4 and 5.
these American contingent claims at the very first place and decides later that he is going to exercise it before rational agent does.

**Further Words in Solving PDEs**

Despite no intention on going down this direction ourselves, we feel obliged to say a bit more about solving PDEs by using the Laplace-Carson transform (LCT) (For more details, readers are referred to [12], [51]). For the sake of argument, let us consider PDE (3.3.2) with the following boundary conditions:

\[ V(t, x) = G(t, x) \] at \( x = b(t) \), (continuous fit condition),
\[ V_x(t, x) = G_x(t, x) \] at \( x = b(t) \), (smooth fit condition),
\[ V(T, x) = G(T, x) \] (terminal condition).

For computational convenience, set \( s = T - t \) so that \( V(t, x) = V(T - s, x) = \tilde{V}(s, x) \) and \( b(t) = b(T - s) = \tilde{b}(s) \) and then we define LCT as follows for \( \lambda > 0 \):

\[ \hat{V}(\lambda, x) = \int_0^\infty \lambda e^{-\lambda s} \tilde{V}(s, x) ds. \]

Next in line is to take the LCT on both sides of (3.3.2)

\[ L_X \hat{V}(\lambda, x) - \int_0^\infty \lambda e^{-\lambda s} \tilde{V}_s(s, x) ds = 0, \]

where the second term on the left-hand side is from \( V_t(t, x) = -\tilde{V}_s(s, x) \); after which, integration by parts, joining with the terminal condition \( V(T, x) = \tilde{V}(0, x) = G(T, x) \) yields

\[
\int_0^\infty \lambda e^{-\lambda s} \tilde{V}_s(s, x) ds = \int_0^\infty \lambda e^{-\lambda s} d\tilde{V}(s, x) = -\lambda \tilde{V}(0, x) - \lambda \int_0^\infty \tilde{V}(s, x) d\lambda e^{-\lambda s}
\]

\[ = -\lambda V(T, x) + \lambda \int_0^\infty \lambda e^{-\lambda s} \tilde{V}(s, x) ds \]
\[ = -\lambda G(T, x) + \lambda \tilde{V}(\lambda, x). \]

In so doing, we arrived at the corresponding ODE with the boundary conditions:

\[ L_X \tilde{V}^{LCT}(\lambda, x) - \lambda \tilde{V}^{LCT}(\lambda, x) + \lambda G(T, x) = 0 \]
\[ \tilde{V}(\lambda, x) = \int_0^\infty \lambda e^{-\lambda s} \tilde{G}(s, x) ds \] for \( x = \tilde{b}(s) \), continuous fit condition,
\[ \tilde{V}_x(\lambda, x) = \int_0^\infty \lambda e^{-\lambda s} \tilde{G}_x(s, x) ds \] for \( x = \tilde{b}(s) \), smooth fit condition.

It may then be possible to obtain the explicit forms for \( \tilde{V}^{LCT}(\lambda, x) \) and \( \tilde{b}^{LCT}(\lambda) \) and then to inverse the transformation by the Gaver-Stehfest method (see [1, Page 52-54], [34], [80]).

Let us briefly apply this method to solve the problem in Example 3.3.5:
Example 6.0.4 (Rebound Example). The free-boundary problem of the Asian option is formulated as follows [64, Page 418]:

\[
V_t + (1 - rx)V_x + \frac{1}{2} \sigma^2 x^2 V_{xx} = 0, \quad x \geq b(t),
\]

\[
V(t, x) = \left(1 - \frac{x}{t}\right)^+, \quad x = b(t), t = T,
\]

\[
V_x(t, x) = -\frac{1}{t}, \quad x = b(t).
\]

After the Laplace-Carson Transformation, the following ODE can be obtained:

\[
\begin{cases}
\frac{1}{2} \sigma^2 x^2 \hat{V}_{xx} + (1 - rx)\hat{V}_x - \lambda \hat{V} + \lambda \left(1 - \frac{x}{t}\right) = 0, & b(T - s) \leq x < T - s, \\
\frac{1}{2} \sigma^2 x^2 \hat{V}_{xx} + (1 - rx)\hat{V}_x - \lambda \hat{V} = 0, & x \geq T - s,
\end{cases}
\]

Maple2020 then tells us that (as a matter of fact, it is always a good idea to ask Maple first when attempt to use such method),

\[
\hat{V}(\lambda, x) = \begin{cases}
C_1 M(\lambda, x) + C_2 U(\lambda, x) + u(\lambda, x), & b(T - s) \leq x < T - s \\
C_3 M(\lambda, x) + C_4 U(\lambda, x), & x \geq T - s
\end{cases}
\]

where \(M(\lambda, x)\) and \(U(\lambda, x)\) are the fundamental increasing and decreasing solutions with the explicit forms given in A1.9. Observe that for \(x \to \infty\), \(V(t, x) \to 0\), so immediately we have \(C_3 = 0\) and all that left (the most difficult part) is to determine the coefficients \(C_1, C_2, C_4\) by solving

\[
\begin{cases}
C_1 M(\lambda, x) + C_2 U(\lambda, x) + u(\lambda, x) = C_4 U(\lambda, x), & x = T - s, \\
C_1 M(\lambda, x) + C_2 U(\lambda, x) + u(\lambda, x) = \int_0^\infty \lambda e^{-\lambda s} \tilde{G}(s, x) ds, & x = b(T - s), \\
C_1 \frac{d}{dx} M(\lambda, x) + C_2 \frac{d}{dx} U(\lambda, x) + \frac{d}{dx} u(\lambda, x) = \frac{d}{dx} \int_0^\infty \lambda e^{-\lambda s} \tilde{G}(s, x) ds & x = b(T - s),
\end{cases}
\]

where the first equation is because of the continuity of the value function, we omit further details since this is not the main method applied in current work, but it certainly will be a healthy exercise to put it on paper.
Appendix A

Auxiliary Results

A.1 Computations Associated with Azéma Supermartingale

A.1.1 For Chapter 4

1. For notational convenience, we set:

\[ \alpha = \frac{2r}{\sigma^2} - 1, \quad b = r - \frac{\sigma^2}{2} \]

\[ m = -\frac{d_1^2}{2}, \quad n = -\frac{d_2^2}{2} \]

\[ d_1 = -\log y\sigma^{-1}(T-t)^{-\frac{1}{2}} + b\sigma^{-1}(T-t)^{\frac{1}{2}} \]

\[ d_2 = -\log y\sigma^{-1}(T-t)^{-\frac{1}{2}} - b\sigma^{-1}(T-t)^{\frac{1}{2}} \]

\[ \frac{\partial d_1}{\partial t} = -\frac{\log y}{2\sigma}(T-t)^{-\frac{3}{2}} - \frac{b}{2\sigma}(T-t)^{-\frac{1}{2}} \]

\[ \frac{\partial d_2}{\partial t} = -\frac{\log y}{2\sigma}(T-t)^{-\frac{3}{2}} + \frac{b}{2\sigma}(T-t)^{-\frac{1}{2}} \]

\[ \frac{\partial d_1}{\partial y} = \frac{\partial d_2}{\partial y} = -\sigma^{-1}y^{-1}(T-t)^{-\frac{1}{2}} \]

\[ \left( \frac{\partial d_1}{\partial y} \right)^2 = \left( \frac{\partial d_2}{\partial y} \right)^2 = \sigma^{-2}y^{-2}(T-t)^{-1} \]

\[ \frac{\partial^2 d_1}{\partial y^2} = \frac{\partial^2 d_2}{\partial y^2} = \sigma^{-1}y^{-2}(T-t)^{-\frac{1}{2}} \]

2. The derivative of \( Z \) w.r.t time \( t \):

\[ \frac{\partial Z}{\partial t}(t, y) = \frac{1}{\sqrt{2\pi}} \left( e^m \frac{\partial d_1}{\partial t} + y^n e^n \frac{\partial d_2}{\partial t} \right) \]
\[ = \frac{1}{\sqrt{8\pi}} \left( -e^m \sigma^{-1}(T-t)^{-\frac{3}{2}} \log y - e^n \sigma^{-1}(T-t)^{-\frac{3}{2}} y^\alpha \log y \right). \]

3. The derivative of \( Z \) w.r.t. space \( y \):

\[ \frac{\partial Z}{\partial y}(t, y) = \frac{1}{\sqrt{2\pi}} \left( e^m \frac{\partial d_1}{\partial y} + \alpha y^{\alpha-1} \Phi(d_2) + y^\alpha e^n \frac{\partial d_2}{\partial y} \right). \]

\[ = \frac{1}{\sqrt{2\pi}} \left( -\sigma^{-1} y^{-1}(T-t)^{-\frac{1}{2}} \left( -e^m e^n + e^m \right) + \alpha y^{\alpha-1} \Phi(d_2) \right). \]

4. The derivative is bounded for \( y > L > 1 \) and \( t \in [0, T] \):

\[ \left| \frac{\partial Z}{\partial y}(t, y) \right| = \sqrt{\frac{\pi}{\pi}} \sigma^{-1} y^{-1}(T-t)^{-\frac{1}{2}} e^{-\frac{d_1^2}{2}} - \alpha y^{\alpha-1} \Phi(d_2) \]

\[ \leq \sqrt{\frac{\pi}{\pi}} \sigma^{-1} (T-t)^{-\frac{1}{2}} e^{-\frac{d_1^2}{2}} - \alpha \]

\[ \leq \sqrt{\frac{\pi}{\pi}} \sigma^{-1} (T-t)^{-\frac{1}{2}} e^{-\left( -\log L \sigma^{-1}(T-t)^{-\frac{1}{2}} \right)^2} - \alpha = C_1 < \infty. \]

Note that for \( t \to T \), but observe that as \( (T-t)^{-\frac{1}{2}} \to \infty \), \( e^{-\frac{d_1^2}{2}} \to 0 \) which is decreasing much faster.

5. The second derivative of \( Z \) w.r.t. space \( y \):

\[ \frac{\sigma^2 y^2 \partial^2 Z}{2 \partial y^2}(t, y) = \frac{1}{\sqrt{8\pi}} \sigma^2 y^2 \left( -d_1 e^m \left( \frac{\partial d_1}{\partial y} \right)^2 + e^m \frac{\partial^2 d_1}{\partial y^2} + \alpha(\alpha - 1) y^{\alpha-2} \Phi(d_2) \right. \]

\[ + \alpha y^{\alpha-1} e^n \frac{\partial d_2}{\partial y} - d_2 y^\alpha e^n \left( \frac{\partial d_2}{\partial y} \right)^2 + y^\alpha e^n \frac{\partial^2 d_2}{\partial y^2} \]

\[ = \frac{1}{\sqrt{8\pi}} \left( -d_1 e^m (T-t)^{-1} + e^m \sigma(T-t)^{-\frac{1}{2}} + \alpha(\alpha - 1) \sigma^2 y^\alpha \Phi(d_2) \right. \]

\[ - 2\alpha \sigma y^\alpha \left( e^n (T-t)^{-\frac{1}{2}} - d_2 y^\alpha e^n (T-t)^{-1} + y^\alpha e^n \sigma(T-t)^{-\frac{1}{2}} \right) \].

A.1.2 For Chapter 5

1. As before, for notational convenience, we let:

\[ \alpha = \frac{2r}{\sigma^2} - 1 \]
A.1 Computations Associated with Azéma Supermartingale

\[ d_1 = \frac{-\log \frac{L}{x} + \left( r - \frac{\sigma^2}{2} \right) (T - t)}{\sigma \sqrt{T - t}} \]
\[ d_2 = \frac{-\log \frac{L}{x} - \left( r - \frac{\sigma^2}{2} \right) (T - t)}{\sigma \sqrt{T - t}} \]

2. The relation between \( d_1 \) and \( d_2 \):

\[ e^{-\frac{d_2^2}{2}} = \left( \frac{L}{x} \right)^\alpha e^{-\frac{d_1^2}{2}}. \]

3. The derivatives with respect to time:

The derivatives of \( d_1 \) and \( d_2 \) w.r.t time is given by:

\[ \frac{\partial}{\partial t} d_1 = -\frac{1}{2} \sigma^{-1} (T - t)^{-\frac{3}{2}} \log \frac{L}{x} - \frac{1}{2} \left( \frac{r}{\sigma} - \frac{\sigma}{2} \right) (T - t)^{-\frac{1}{2}}, \]
\[ \frac{\partial}{\partial t} d_2 = -\frac{1}{2} \sigma^{-1} (T - t)^{-\frac{3}{2}} \log \frac{L}{x} + \frac{1}{2} \left( \frac{r}{\sigma} - \frac{\sigma}{2} \right) (T - t)^{-\frac{1}{2}}, \]

The derivative of \( Z \) w.r.t time is given by:

\[ \frac{\partial}{\partial t} Z(t, x) = \begin{cases} 
-\sqrt{\frac{2}{\pi}} \log \left( \frac{L}{x} \right) e^{-\frac{d_1^2}{2}} \sigma^{-1} (T - t)^{-\frac{3}{2}}, & t < T, \\
0, & t = T.
\end{cases} \]

4. The derivatives with respect to space:

The derivatives of \( d_1 \) and \( d_2 \) w.r.t space is given by:

\[ \frac{\partial}{\partial x} d_1 = \frac{\partial}{\partial x} d_2 = x^{-1} \sigma^{-1} (T - t)^{-\frac{1}{2}}, \]

\[ \frac{\partial}{\partial x} Z(t, x) = \begin{cases} 
\sqrt{\frac{2}{\pi}} e^{-\frac{d_2^2}{2}} x^{-1} \sigma^{-1} (T - t)^{-\frac{3}{2}} - \alpha \left( \frac{L}{x} \right)^\alpha x^{-1} \Phi(d_2), & x < L \\
0, & x \geq L \text{ or } t = T,
\end{cases} \]

and thus, for \((t, x) \in [0, T) \times (0, L)\)

\[ \frac{\partial^2}{\partial x \partial t} Z(t, x) = \sqrt{\frac{2}{\pi}} x^{-1} \sigma^{-1} (T - t)^{-\frac{3}{2}} e^{-\frac{d_1^2}{2}} \left( 1 - \left( \log \frac{L}{x} \right)^2 (T - t)^{-1} \sigma^{-2} + \log \frac{L}{x} \sigma^{-1} \left( \frac{r}{\sigma} - \frac{\sigma}{2} \right) \right). \]
5. The bounded partial derivative w.r.t $x$ for $(t, x) \in [0, t_δ) \times (0, \infty)$:

$$0 \leq \frac{\partial}{\partial x} Z(t, x) = \sqrt{\frac{\sqrt{2}}{\pi}} e^{-\frac{x^2}{2}} x^{-1} \sigma^{-1} (T - t)^{-\frac{1}{2}} - \alpha \left( \frac{L}{x} \right)^\alpha x^{-1} \Phi(d_2)$$

$$< \sqrt{\frac{\sqrt{2}}{\pi}} \sigma^{-1} (T - t_δ)^{-\frac{1}{2}} x^{-1} e^{-\frac{x^2}{2}} \Phi(d_2) - \alpha L^\alpha x^{-\frac{2x}{\sigma}} \Phi(d_2) = C_2 < \infty.$$ 

Note that by L'Hôpital’s rule (and that $d_2$ is a function of $x$ and $\frac{2x}{\sigma} < 1$, $\lim_{x \to \infty} \frac{x}{e^x} \to 0$):

$$\lim_{x \to 0} \Phi(d_2) = \lim_{x \to 0} \sqrt{\frac{\sqrt{2}}{2\pi}} e^{-\frac{x^2}{2}} x^{-1} \sigma^{-1} (T - t)^{-\frac{1}{2}} = \sqrt{\frac{\sqrt{2}}{2\pi}} \frac{\sigma}{2r (T - t)^{-\frac{1}{2}}} \lim_{x \to 0} \frac{e^{-\frac{x^2}{2}}}{x^\sigma} \to 0$$

6. The derivative w.r.t time $t$ for $x < L, t < T$:

$$-rKZ_t - \sigma^2 x^2 Z_xt = \sqrt{\frac{1}{2\pi}} e^{-\frac{d_2^2}{2}} (T - t)^{-\frac{3}{2}} \left( \frac{r(K - x) \log \frac{x}{y}}{\sigma} - \sigma x + \sigma x \left( \frac{\log \frac{x}{y}}{\sigma \sqrt{T - t}} + \frac{\sigma x}{2} \log \frac{L}{x} \right) \right).$$

Note that the assumption $-rKZ_t - \sigma^2 x^2 Z_xt \leq 0$ on set $S$, together with $Z_t \leq 0$, suggests that $Z_xt \geq -\frac{rKZ_t}{\sigma x^2} \geq 0$.

7. The expectation of the local time.

Observe that

$$\lim_{\epsilon \to 0} \frac{1}{\epsilon} P \left( \log \frac{L - \epsilon}{y} < \left( r - \frac{\sigma^2}{2} \right) u + \sigma W_u < \log \frac{L + \epsilon}{y} \right) = \frac{1}{\sigma \sqrt{u} L} \phi \left( \frac{\log \frac{L}{y} - \left( r - \frac{\sigma^2}{2} \right) u}{\sigma \sqrt{u} L} \right),$$

so that

$$dE_{t,x} \left[ \frac{L}{u} \right] = \sigma^2 L^2 \lim_{\epsilon \to 0} \frac{1}{\epsilon} P \left( \log \frac{L - \epsilon}{x} < \left( r - \frac{\sigma^2}{2} \right) u + \sigma W_u < \log \frac{L + \epsilon}{x} \right) du$$

$$= \frac{\sigma L}{\sqrt{u}} \phi \left( \frac{\log \frac{L}{x} - \left( r - \frac{\sigma^2}{2} \right) u}{\sigma \sqrt{u}} \right) du,$$

where $\phi(z) = \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}}$ is the probability density function of the standard normal law.
A.2 Useful Results for the Proof of Smooth-fit Condition

Lemma A.2.1 (For Chapter 4). As $\epsilon \to 0$, $\tau_\epsilon \to 0$ almost surely.

Proof. By construction of $\tau_\epsilon$ and the definition of $Y$,
\[
\tau_\epsilon = \inf \{ 0 \leq s \leq T - t : Y_{t+s}^{y-\epsilon} \geq b(t + s) \}
\]
\[= \inf \left\{ 0 \leq s \leq T - t : \frac{y - \epsilon}{X_s} \geq b(t + s) \right\}. \]

Note that \(\frac{y - \epsilon}{X_s} \geq b(t + s)\), so is \(g\) function of the Brownian motion, so is \(g\) for standard Brownian motion \(\tilde{W}\) under measure \(\tilde{P}\), implying that \(\tilde{P}(\tau_\epsilon = 0) = 1\), together with the fact that \(0 \leq \tau_\epsilon \leq \tilde{\tau}_\epsilon = 0\) a.s., we know \(\tau_\epsilon = 0\) a.s.

Lemma A.2.2 (For Chapter 5). As $\epsilon \to 0$, $\tau_\epsilon \to 0$ almost surely.

Proof. By the construction of $\tau_\epsilon$ and the strong solution of $X$,
\[
\tau_\epsilon = \inf \{ 0 \leq s \leq T - t : X_{t+s}^{x+\epsilon} \leq b(t + s) \}
\]
\[= \inf \left\{ 0 \leq s \leq T - t : W_s \leq \frac{\log \frac{b(t+s)}{b(t)} + \beta s}{\sigma} \right\}
\]
\[= \inf \{ 0 \leq s \leq T - t : -W_s \geq g(s) \}, \]
where \(\beta = r - \frac{\sigma^2}{2}\) and \(g(s) = -\frac{\log \frac{b(t+s)}{b(t)} + \beta s}{\sigma}\). As $\epsilon \to 0$ (recalling that $x = b(t)$),
\[g(s) \to -\frac{\log \frac{b(t+s)}{b(t)} + \beta s}{\sigma} \leq \frac{\beta s}{\sigma},\]
where the inequality follows from the map $t \mapsto b(t)$ being decreasing, and since $s \mapsto \frac{\beta s}{\sigma}$ is a lower function for standard Brownian motion $\tilde{W}$ under measure $\tilde{P}$, so is $g(s)$, implying that $\tilde{P}(\tau_\epsilon = 0) = 1$, and together with the fact that $0 \leq \tau_\epsilon \leq \tilde{\tau}_\epsilon = 0$ a.s., we know $\tau_\epsilon = 0$ a.s.
A.3 Property of the Value Function

Lemma A.3.1 (For Chapter 4). The map \( t \mapsto V(t, y) \) is decreasing.

\[ E_{t,y} \left( e^{-\lambda \tau} (Y_{t+\tau} - L)^+ Z(t + \tau, Y_{t+\tau}) \right) = E \left( e^{-\lambda \tau} (Y^y_{\tau} - L)^+ Z(t + \tau, Y^y_{\tau}) \right), \]

and since the map \( t \mapsto Z(t, y) \) is decreasing, we have \( t \mapsto E \left( e^{-\lambda \tau} (Y^y_{\tau} - L)^+ Z(t + \tau, Y^y_{\tau}) \right) \) is decreasing. Moreover, \( t \mapsto T - t \) is decreasing so that the supremum is taken over a smaller set as \( t \) increases and that by definition, \( t \mapsto V(t, y) \) is decreasing.

Lemma A.3.2 (For Chapter 5). The map \( t \mapsto V(t, x) \) is decreasing.

\[ E_{t,x} \left( e^{-r \tau} (K - X_{t+\tau})^+ Z(t + \tau, X_{t+\tau}) \right) = E \left( e^{-r \tau} (K - X^x_{\tau})^+ Z(t + \tau, X^x_{\tau}) \right), \]

and since the map \( t \mapsto Z(t, x) \) is decreasing, we have \( t \mapsto E \left( e^{-r \tau} (K - X^x_{\tau})^+ Z(t + \tau, X^x_{\tau}) \right) \) is decreasing. Moreover, \( t \mapsto T - t \) is decreasing so that the supremum is taken over a smaller set as \( t \) increases and that by definition, \( t \mapsto V(t, x) \) is decreasing.

A.4 The Uniform Integrability of \( Y \)

To apply Theorem 2.1.9, we observe that

\[ Y^y_t = \frac{y \vee \max_{0 \leq u \leq t} X_u}{X_t} \leq \frac{y \max_{0 \leq u \leq T} X_u}{X_t} = ye^{-(r + \frac{\sigma^2}{2})t - \sigma \tilde{W}_t} \max_{0 \leq u \leq T} X_u \]

\[ \leq ye^{\sigma \max_{0 \leq u \leq t} (-\tilde{W}_u)} e^{(r + \frac{\sigma^2}{2})T + \sigma \max_{0 \leq u \leq t} (\tilde{W}_u)} \]

\[ \leq \frac{1}{2} ye^{(r + \frac{\sigma^2}{2})T} \left( e^{2\sigma \max_{0 \leq u \leq T} (-\tilde{W}_u)} + e^{2\sigma \max_{0 \leq u \leq T} \tilde{W}_u} \right) = \frac{ye^{(r + \frac{\sigma^2}{2})T}}{e^{2\sigma \sqrt{T}X}}, \quad (A.4.1) \]

where the first inequality holds as \( y \geq 1, \max_{0 \leq u \leq T} X_u \geq \max_{0 \leq u \leq t} X_u \geq 1 \) and the last inequality follows from Young’s inequality and the random variable \( X \) has the probability density function given as follows

\[ f_X(x) = \sqrt{\frac{2}{\pi}} e^{-\frac{x^2}{2}}, \quad x > 0. \]

We further estimate the following

\[ E \left( e^{2\sigma \sqrt{T}X} \right) = \sqrt{\frac{2}{\pi}} \int_0^\infty e^{2\sigma \sqrt{T}x - \frac{x^2}{2}} dx \]
\[
\frac{\sqrt{2}}{\pi} \int_{0}^{\infty} e^{-\frac{1}{2}(x^2-4\sigma \sqrt{T}x+4\sigma^2T)+2\sigma^2T} dx \\
= \sqrt{\frac{1}{2\sigma^{2}}} e^{2\sigma^{2}T} \int_{-2\sigma\sqrt{T}}^{\infty} e^{-\frac{x^{2}}{2}} dy \quad \text{(Change of variable } y = x - 2\sigma \sqrt{T}) \\
= e^{2\sigma^{2}T} + \sqrt{\frac{1}{2\pi}} \int_{-2\sigma\sqrt{T}}^{0} e^{-\frac{y^{2}}{2}} dy < \infty.
\]

We therefore, in a somewhat lengthy way, have proven that \(Y_{t}^{y}\) is dominated by a positive integrable random variable.

### A.5 Other Simple Results

1. The martingale property of \(\left\{ e^{-\frac{\sigma^{2}}{2}t+\sigma W_{t}}, F_{t} \right\}, t \geq 0 \):

   Observe that for \(s \leq t\):

   \[
   E \left( e^{-\frac{\sigma^{2}}{2}t+\sigma W_{t}} | F_{s} \right) = E \left( e^{-\frac{\sigma^{2}}{2}s+\sigma W_{s}} e^{-\frac{\sigma^{2}}{2}(t-s)+\sigma(W_{t}-W_{s})} | F_{s} \right) \\
   = e^{-\frac{\sigma^{2}}{2}s+\sigma W_{s}} E \left( e^{-\frac{\sigma^{2}}{2}(t-s)+\sigma(W_{t}-W_{s})} | F_{s} \right) \\
   = e^{-\frac{\sigma^{2}}{2}s+\sigma W_{s}} E \left( e^{-\frac{\sigma^{2}}{2}(t-s)+\sigma(W_{t}-W_{s})} \right) \\
   = e^{-\frac{\sigma^{2}}{2}s+\sigma W_{s}} \sqrt{\frac{1}{2\pi}} \int_{-\infty}^{\infty} e^{-\frac{\sigma^{2}}{2}(t-s)+\sigma\sqrt{T}x} e^{-\frac{x^{2}}{2}} dx \\
   = e^{-\frac{\sigma^{2}}{2}s+\sigma W_{s}} \sqrt{\frac{1}{2\pi}} \int_{-\infty}^{\infty} e^{-\frac{1}{2}(x-\sigma \sqrt{T}x)^{2}} dx \\
   = e^{-\frac{\sigma^{2}}{2}s+\sigma W_{s}},
\]

where the second equality is from \(W_{s} \in F_{s}\) and the third equality is due to the fact that \(W_{t} - W_{s}\) is independent of \(F_{s}\), and the last equality will be more apparent if one sets \(y = x - \sigma \sqrt{T-s}\).

2. The fundamental linearly independent solutions in Example 6.0.4:

   Set \(a = \sqrt{2} \sqrt{2\sigma^{4}+(16\lambda+8r)\sigma^{2}+8\sigma^{2}-2\sigma^{2}-4r} \frac{4\sigma^{2}}{\sigma^{2}}\) and \(b = a + 1 + \frac{r}{\sigma^{2}}\)

   \[
   M(\lambda, x) = x^{-a} \text{KummerM} \left( a, b, \frac{2\sigma^{2}}{x} \right), \\
   U(\lambda, x) = x^{-a} \text{KummerU} \left( a, b, \frac{2\sigma^{2}}{x} \right),
   \]

   whereas \(u(\lambda, x) = \frac{(1+r)T-\lambda x-1}{(\lambda+r)^{T}}\).
A.6 Algorithmic Remarks

Given that the probability density function of geometric Brownian motion $X_u$ starting at $x$ equals

$$f(x_0, x, u) = \frac{1}{\sqrt{2\pi u \sigma x}} e^{-\frac{1}{2u} \left( \log \frac{x}{x_0} - \left( \frac{r}{\sigma^2} \right) u \right)^2}, \quad (A.6.1)$$

so that (5.2.50) can be re-formulated in terms of integrals rather than expectations via exploiting the Law of the Unconscious Statistician,

$$G(t, b(t)) = e^{-r(T-t)} \int_{K}^{L} (K - x) f(b(t), x, T-t)dx$$
$$- \int_{0}^{T-t} \int_{0}^{b(t+u)} e^{-ru} H(t + u, x) f(b(t), x, T-t)dxdu$$
$$+ \frac{(K - L)\sigma^2 L^2}{2} \int_{0}^{T-t} e^{-ru} Z_x(t + u, L-) f(b(t), L, u)du.$$

For simplicity, we further set

$$G(t, b(t)) = V(T - t, b(t)) + \int_{0}^{T-t} Y(t, b(t), t + u, b(t + u))du,$$

and that for $n = 500$, $h = \frac{T}{n}$, $i = 1, \ldots, n + 1$ we have $t_i = T - ih$ and $bb(1) = b(T), bb(i + 1) = b(t_i), \ldots, bb(n + 1) = b(0)$ so that

$$\int_{0}^{T-t_i} Y(t_i, b(t_i), t_i + u, b(t_i + u))du \approx \frac{h}{2} \left( Y(t_i, b(t_i), T, K) + Y(t_i, b(t_i), t_i, b(t_i)) \right)$$
$$+ h \sum_{j=1}^{i-1} Y(t_i, b(t_i), t_i + jh, b(t_i + jh))$$
$$= \frac{h}{2} \left( Y(T - ih, bb(i + 1), T, K) + Y(T - ih, bb(i + 1), T - ih, bb(i + 1)) \right)$$
$$+ h \sum_{j=1}^{i-1} Y(T - ih, bb(i + 1), T - (i - j)h, bb(i - j + 1)) .$$

As for computing the integral numerically, the build-in function `vapintegral` in Matlab will be up for the task. The algorithm is given on the next page.
**Algorithm 1** An algorithm for solving nonlinear integral equation in Mathlab

\[ K = 7 \]  \quad \triangleright \text{the strike price}

\[ n = 500 \]  \quad \triangleright \text{size of the vector}

\[ bb = (n + 1) \]  \quad \triangleright \text{vector for the optimal stopping boundary in each discrete time point}

\[ bb(1) = K \]  \quad \triangleright \text{bb(1)=b(T)=K}

\begin{verbatim}
for i = 2 : n + 1 do
    syms b
    bb(i) = func(b, i, bb);  \triangleright \text{go to function which solves equation}
end for
\end{verbatim}

**function** \( fb = \text{func}(b, c, d) \)

\[ T = 10, K = 7, \sigma = 0.4, r = 0.05, \ldots ; \]  \quad \triangleright \text{define all the parameters needed}

\[ i = c; \]

\[ h = T/n; \]

\[ Z(t, x), Z_{x}(t, x), H(t, x), f(z, u; x), \ldots ; \]  \quad \triangleright \text{define all the functions needed}

\[ \text{syms } b \]  \quad \triangleright \text{Create symbolic scalar variables: the optimal stopping boundary } bb(i + 1)

\[ \text{sum} = V(\text{ih}, b) + \frac{h}{2} [Y(\text{T} - \text{ih}, b, T, K) + Y(\text{T} - \text{ih}, b, T - \text{ih}, b)]; \]

\begin{verbatim}
for j = 1 : i - 1 do
    sum = sum + hY(\text{T} - \text{ih}, b, T - (i - j)h, bb(i - j + 1))
end for
\end{verbatim}

\[ \text{equation} = G(t_i, b) == \text{sum}; \]

\[ fb = \text{vpasolve(eq, b, K)}; \quad \triangleright \text{return function value, } K \text{ is the initial guess for the solution} \]

**end function**
Bibliography


