Optimization problems in financial mathematics : explicit solutions for diffusion models

Boguslavskaya, E.

Optimization Problems in Financial Mathematics: Explicit Solutions for Diffusion Models

ACADEMISCH PROEFSCHRIFT

ter verkrijging van de graad van doctor
aan de Universiteit van Amsterdam
op gezag van de Rector Magnificus prof. mr. P. F. van der Heijden
ten overstaan van een door het College voor Promoties ingestelde commissie,
in het openbaar te verdedigen in de Aula der Universiteit
op donderdag 4 mei 2006, te 14.00 uur

door

Elena Vladimirovna (Efimova) Boguslavskaya

geboren te Moskou, Rusland

Promotiecommissie:

Promotores:	prof. dr. A. N. Shiryaev prof. dr. C. A. J. Klaassen
Copromotor:	dr. A. A. Balkema
Overige leden:	prof. dr. A. Bagchi prof. dr. H. P. Boswijk prof. dr. M. Keane prof. dr. T. H. Koornwinder prof. dr. J. H. van Schuppen dr. P. J. C. Spreij dr. M. Zervos

Korteweg–de Vries Instituut voor Wiskunde
Faculteit der Natuurwetenschappen, Wiskunde en Informatica

ISBN .....

Cover: Collage of various vintage bonds and shares, arranged by Natalia Rogachevskaya.

To the memory of my grandmother
Nina Viktorovna Efimova,
(May, 4 1922 - January, 14 2005),
who always encouraged me to achieve.

Contents

Introduction

0.1 Summary

In this thesis, we consider several optimal control problems for diffusion processes. The processes considered in Chapter 1 arise from *optimization of dividend flows* for a company. In Chapter 2, the processes correspond to *optimization of investments in real options*. The solutions to both problems are of *singular* type, i.e. they have a “bang-bang” character. The processes considered in Chapter 3 model financial asset behaviour. We give an explicit solution for the optimal trading strategy. Unlike the problems considered in Chapter 1 and 2, this problem is of regular type, i.e. the optimal control is “smooth”. The aim of this thesis is to outline a method for finding *explicit* solutions for optimal control problems in one-dimensional diffusions. General stochastic control theory for diffusions is treated in [Krylov], [Oks], [FlemSoner].

The results of Chapter 1 have been submitted to Stochastics [Bog3]. The results of Chapter 2 have been published in Theory of Probability and Its Applications [Bog2] and Russian Mathematical Surveys [Bog1]. The results of Chapter 3 appeared in Risk Magazine [Bog4].

0.2 Overview

Here we provide an overview of the thesis.

In Chapter 1, we consider a model for a firm whose reserve $X = (X_t)_{t\geq 0}$ evolves according to the stochastic differential equation

$$dX_t = \mu dt + \sigma dW_t - dZ_t, \tag{1}$$

where $W = (W_t)_{t\geq 0}$ is a standard Wiener process, and μ and σ are known positive constants. The control process $Z = (Z_t)_{t\geq 0}$ represents the cumulative amount of dividend paid out up to time t. The major requirements on the control process $Z = (Z_t)_{t\geq 0}$ are that it is nonnegative, nondecreasing, and adapted to the filtration. The bankruptcy time τ is defined as $\tau = \inf\{t \geq 0 : X_t \leq 0\}$. It is assumed that the initial reserve x_0 is positive and that the liquidation value S, i.e. the salvage value of the firm's assets at the time of bankruptcy, is nonnegative. With λ the constant discount rate, the expected total pay-off to shareholders together with the discounted liquidation value equals

$$V(x, Z) = \mathbf{E}_x\{\int_0^\tau e^{-\lambda t} dZ_t + Se^{-\lambda\tau}\}. \tag{2}$$

We present explicit formulae for the optimal admissible process $\tilde{Z}$, i.e. the control process that maximizes $V(x, Z)$, for the case of bounded dividend rates, for the case of discrete dividends with transaction costs, and for the general case when the dividend process is allowed to be any nonnegative, nondecreasing, right-continuous process.

In Chapter 2, we consider a model for long-term irreversible investments. The predicted cost remaining at time t to complete the project is given by

$$dX_t = -I_t dt + \beta\sqrt{X_t I_t} dW_t + \gamma X_t d\tilde{W}_t, \tag{3}$$

where W_t and $\tilde{W}_t$ are uncorrelated standard Wiener processes, β and γ are nonnegative constants, and $I(X_t)$ represents the investment rate. The time τ at which the project is completed, is defined by $\tau = \inf\{t \geq 0 : X_t \leq 0\}$. It is assumed that the initial cost of completing the project, $X_0 = x$, is positive, and that the value of the project after completion is given by a positive constant V. With r the constant interest rate, the expected total profit on the project equals

$$F(x, I) = \mathbf{E}\{\int_0^\tau (-I(X_t)e^{-rt})dt + Ve^{-r\tau}\}. \tag{4}$$

We present explicit solutions for the optimal investment rate I^*, i.e. the investment rate that maximizes $F(x, I)$ for three separate cases: technical uncertainty ($\beta = 0, \gamma \neq 0$), input cost uncertainty ($\beta \neq 0, \gamma = 0$), and the case where both uncertainties are present ($\beta \neq 0, \gamma \neq 0$).

In Chapter 3 we consider the position management problem for an agent trading a mean-reverting asset. This problem arises in many statistical and fundamental arbitrage trading situations when the short-term returns on an asset are predictable, but limited risk-bearing capacity does not allow the agent to fully exploit this predictability. We use an Ornstein-Uhlenbeck process to model the price process $X = (X_t)_{t\geq 0}$

$$dX_t = -kX_t dt + \sigma dB_t,$$

where k and σ are positive constants and B_t is a standard Wiener process. The model reproduces some realistic patterns of trader behaviour. The control α_t represents the trader's position at time t, i.e. the number of units of the asset held. Assuming zero interest rates and no market friction, the wealth dynamics for a given control α_t is given by

$$dW_t = \alpha_t dX_t = -k\alpha_t X_t dt + \alpha_t \sigma dB_t.$$

We assume that there are no restrictions on α, so short selling is allowed, and there are no marginal requirements on the wealth W. Introduce the power utility function $\Psi(w) = (w^\gamma - 1)/\gamma$, $w \geq 0$, for some $\gamma \in (-\infty, 1)$. The expected utility at time T conditionally on the information available at time t is

$$J(W_t, X_t, t) = \mathbf{E}_t \Psi(W_T).$$

We present an explicit solution for the optimal position α^*, i.e. the position that maximizes $J(W_t, X_t, t)$.

0.3 Acknowledgments

I would like to express my deep gratitude to my advisors, prof. Albert Shiryaev and prof. Chris Klaassen for their attention, help and patience while I was writing this thesis. I thank prof. Shiryaev for introducing me into the field and for sharing his experience and excitement about the most interesting world of stochastic analysis. I thank prof. Klaassen for giving me the opportunity to work in a highly creative atmosphere of the Korteweg-de Vries Institute for Mathematics at the University of Amsterdam, for his support, encouragement and his belief in me. Furthermore, I would like to thank my co-promotor dr. Guus Balkema for helpful discussions and careful reading of the manuscript.

I am grateful for the financial support I received from 'IBIS UvA BV' (see http://www.ibisuva.nl/uk/index.htm) and the Korteweg-de Vries Institute for Mathematics.

I thank prof. Tom Koornwinder, prof. Arjen Doelman, prof. Jan Wiegerinck, and Oscar Lemmers for valuable discussions.

Cheers to Paul Beneker, Erdal Emsiz, and Nienke Valkhoff.

I thank Eveline Wallet who arranged a parking pass for me, which saved me at least one working hour a day.

Finally, I would like to thank my husband and co-author Michael and my sons Stepan and Matvey for their love, patience, and support.

Chapter 1

Dividend optimization models

In this chapter, we consider several models for dividend optimization. We consider separately three cases: the case of bounded dividend rates, the case of discrete dividends with transaction costs, and the case when the dividend process is any nonnegative, nondecreasing, right-continuous process. In the Introduction, Section 1.1, we formulate the problem and give an overview. In Section 1.2, we present the main results. The Sections 1.3, 1.4 and 1.5 present the proofs of optimality for the cases of bounded dividend rate, discrete and cadlag dividends respectively. Appendix 1.6 contains the notation reference, and Appendix 1.7 presents the technical material we used for proving optimality of the solution.

1.1 Introduction

1.1.1 Choice of reserve process and structure of dividend process

Radner and Shepp (1996) [Radner] have proposed a model for a firm whose reserve X_t in the absence of dividends evolves as an arithmetic Brownian motion[1]

$$\begin{cases} dX_t &= \mu dt + \sigma dW_t, \\ X_0 &= x_0, \end{cases} \tag{1.1}$$

where μ and σ are known positive constants, $W = (W_t)_{t\geq 0}$ is a standard Wiener process, and x_0 is the initial reserve. All processes are assumed to be adapted to the standard filtration $(\mathcal{F}_t)_{t\geq 0}$ of the Brownian motion W. In particular $\mathcal{F}_0$ is the σ-field generated by the null sets.

The firm's board of directors influences the stochastic fluctuations of the company's reserve by choosing the timing and the size of dividend payments. The dynamics of the reserve become

$$\begin{cases} dX_t &= \mu dt + \sigma dW_t - dZ_t, \\ X_0 &= x_0 - z_0, \end{cases} \tag{1.2}$$

[1] They also show why geometric Brownian motion is an inappropriate model for the reserve.

where the control process $Z = (Z_t)_{t\geq 0}$ represents the cumulative amount of dividends paid out up to time t with the initial dividend payment $Z_0 = z_0$. The initial payment z_0 does not exceed the initial reserve, i.e. $0 \leq z_0 \leq x_0$. The control process Z is assumed to be nonnegative, nondecreasing, and right-continuous.

We will distinguish three classes of admissible control processes. The first class consists of control processes with bounded dividend rates. In other words, the rate of the dividend payments is bounded by a positive constant K. More precisely, an admissible control process Z in this class is given by

$$Z_t = z_0 + \int_0^t U_s ds, \ t \geq 0, \tag{1.3}$$

where the dividend rate process $U = (U_t)_{t\geq 0}$ is assumed to be optional and to be bounded by a finite constant K, say; $0 \leq U_t \leq K < \infty$. As a second class we consider the case when the admissible control process Z with $Z_0 = z_0$ is given via a bivariate point process $(T_i, \xi_i)_{i=0}^{\infty}$ by

$$Z_t = \xi_0 + \sum_{i\geq 1} \xi_i I(T_i < t), \ t > 0, \tag{1.4}$$

where $\xi_0 = z_0$ is the initial gross dividend payment, and where ξ_i denotes the gross dividend payment at time T_i. There is a transaction cost $\gamma > 0$ associated with every dividend payment, including the initial payment at time $t = 0$ provided it is positive. A gross dividend payment, i.e. nett dividend payment summed with the transaction cost, does not exceed the current reserve but if nonzero should be bigger than transaction cost γ, i.e. $z_0 = 0$ or $\gamma \leq z_0 \leq x_0$, and $\gamma < \xi_i \leq X_{T_{i-}}$ for $i = 1, 2, \ldots$. Consequently, the nett dividend payments are $z_0 - \gamma$, if $z_0 \neq 0$, and $(\xi_i - \gamma)$. It is assumed here that $0 = T_0 < T_1 < T_2 < \ldots$ holds a.s. The assumptions imply that the admissible control process Z in (1.4) is nonnegative, nondecreasing, and right continuous.

The third class of admissible controls consists of all nonnegative, non-decreasing, right-continuous processes Z that start at z_0, $z_0 \leq x_0$, and satisfy $Z_t \leq X_{t-}$ for all $t > 0$.

The firm exists from time zero until the first moment τ, at which the cash reserve falls down to zero; $\tau = \inf\{t : X_t \leq 0\}$. The moment τ is called *the bankruptcy time.*

1.1.2 The value function

The aim of the board of directors is to maximize the expected total discounted nett dividend paid out during the existence of the firm together with the discounted salvage value at the time of bankruptcy. The maximum will be called the *initial value* of the firm, and denoted by v_0. It depends only on the initial reserve x_0. We introduce the *value function* of our optimization problem $\tilde{V} = \tilde{V}(\cdot)$ by

$$\tilde{V}(x) = \sup \mathbf{E}_x \left\{ \int_0^\tau e^{-\lambda t} dZ_t + S e^{-\lambda \tau} \right\}, \tag{1.5}$$

with some adaptation in the case of discrete payments to take care of the transaction cost. Here the expectation is taken conditionally on $X_0 = x$ and

$$\int_0^\tau e^{-\lambda t} dZ_t = \int_{(0,\tau)} e^{-\lambda t} dZ_t.$$

The constant S is the liquidation value, i.e. the salvage value of the firm's assets at the time of bankruptcy, the discount rate λ is a positive constant, and $Z_0 = z_0$ is the initial dividend payment. The supremum is taken over all admissible control processes $Z = (Z_t)_{t\geq 0}$.

In our analysis we will encounter the value function $\tilde{V} = \tilde{V}(\cdot)$. But the board of directors is interested to know the initial value of the firm v_0. If there is no initial payment z_0 then $X_0 = x_0$ and $v_0 = \tilde{V}(x_0)$. Else $X_0 = x_0 - z_0$ by right continuity of the process X, and $v_0 = z_0 - \gamma + \tilde{V}(x_0 - z_0)$ where z_0 is the gross payments at time $t = 0$, and γ is the transaction cost.

1.1.3 Objective

Our aim is to find, for each initial reserve x_0, the initial value of the firm v_0, the value function $\tilde{V}(\cdot)$, and the optimal dividend policy $\tilde{Z} = (\tilde{Z}_t)_{t\geq 0}$.

1.1.4 Normalization

Without loss of generality, we may and will assume that $\sigma^2 = 1$ by changing the time scale if necessary. This is achieved on replacing (μ, λ, K) by $(\mu/\sigma^2, \lambda/\sigma^2, K/\sigma^2)$.

1.1.5 Overview

A survey on the subject may be found in Taksar (1999) [Taksar]. The case of zero liquidation value $S = 0$ was solved by Jeanblanc and Shiryaev (1995) [JeanShir]. Here we present the general solution for any constant S. The salvage value S even may be negative.

1.2 The main result

Here we present the main results. A reader who is interested only in the results, but not in the proofs can stop after this section.

1.2.1 The case of bounded dividend rates

Suppose we are given the initial reserve $x_0 \geq 0$, the liquidation value S, the discount rate $\lambda > 0$, the upper bound K on the dividend the rate U in the case of bounded dividend rates, $K \geq U \geq 0$. Let us set

$$r_1 = -\mu + \sqrt{\mu^2 + 2\lambda} > 0, \tag{1.6}$$
$$r_2 = \mu + \sqrt{\mu^2 + 2\lambda} > 0, \tag{1.7}$$
$$\rho = -(K-\mu) + \sqrt{(K-\mu)^2 + 2\lambda} > 0, \tag{1.8}$$
$$Q = \frac{K}{\lambda} - \frac{1}{\rho}. \tag{1.9}$$

Define the functions $u_1(\cdot)$, $u_2(\cdot)$, $w(\cdot)$, $f(\cdot)$ and $A(\cdot)$ on $[0, \infty)$ by

$$u_1(x) = e^{-r_2 x}, \tag{1.10}$$
$$u_2(x) = e^{r_1 x} - e^{-r_2 x}, \tag{1.11}$$
$$v(x) = e^{-\rho x}, \tag{1.12}$$
$$w(x) = u_1(x)u_2'(x) - u_1'(x)u_2(x), \tag{1.13}$$
$$f(x) = S\frac{w(x)}{u_2'(x)} + \frac{u_2(x)}{u_2'(x)}, \tag{1.14}$$
$$A(x) = \frac{1 - Su_1'(x)}{u_2'(x)}, \; x \geq 0. \tag{1.15}$$

The optimal control process in the case of a bounded dividend rate is of the "bang-bang" type and may be described now as follows.

Theorem 1.2.1 *The initial value of the firm, the value function $\tilde{V} = \tilde{V}(\cdot)$ and the optimal dividend strategy in the model for bounded dividend rates are the following.*

1) If $S \geq Q$ holds, then the optimal strategy is to liquidate the company immediately and to distribute the initial reserve x_0 as dividend. In this case $\tau = 0$, $z_0 = x_0$, and the initial value of the firm is $v_0 = x_0 + S$.

2) If $S < Q$ holds, then the equation $f(x) = Q$ has a unique positive solution $\tilde{x}$, and the optimal initial dividend payment is

$$z_0 = \begin{cases} 0, & x_0 < \tilde{x}, \\ x_0 - \tilde{x}, & x_0 \geq \tilde{x}; \end{cases}$$

the optimal dividend rate at time t is $\tilde{U}_t = \tilde{u}(X_t)$ with

$$\tilde{u}(x) = \begin{cases} 0, & x < \tilde{x}, \\ K, & x \geq \tilde{x}; \end{cases}$$

the value function $\tilde{V}(\cdot)$ is

$$\tilde{V}(x) = \begin{cases} Su_1(x) + A(\tilde{x})u_2(x), & x < \tilde{x}, \\ K/\lambda + v(x)/v'(\tilde{x}), & x \geq \tilde{x}; \end{cases}$$

and the initial value of the firm is

$$v_0 = \begin{cases} \tilde{V}(x_0), & x_0 < \tilde{x}, \\ x_0 - \tilde{x} + Q, & x_0 \geq \tilde{x}; \end{cases}$$

This theorem states:**as soon as the reserve X_t hits the optimal value $\tilde{x}$, one should start distributing dividends at the maximum possible rate K until the reserve falls back below $\tilde{x}$.** This is illustrated in Fig. 1.1.

Note 1.2.1 *In the special case $S = 0$ the statements $S < Q$ and $S \geq Q$ are equivalent to $K > \lambda/(2\mu)$ and $K \leq \lambda/(2\mu)$ respectively (see Lemma 1.7.1). Returning to the original parameters these inequalities become $K > \sigma^2\lambda/(2\mu)$ and $K \leq \sigma^2\lambda/(2\mu)$. These inequalities might have nice economic explanations.*

1.2.2 The case of discrete dividends

Suppose we are given the initial reserve $x_0 \geq 0$, liquidation value S, discount rate $\lambda > 0$ and transaction cost $\gamma > 0$. Let r_1, r_2, $u_1(\cdot)$, $u_2(\cdot)$, $f(\cdot)$, $A(\cdot)$ be defined as above by (1.6), (1.7), (1.10), (1.11), (1.14), (1.15), respectively.

Suppose $S < \mu/\lambda$. The function $A = A(\cdot)$ is increasing-decreasing on $[0, \infty)$ with a unique maximum at $\bar{x}$, see Corollary 1.7.4. There exists a maximal interval (a_{min}, b_{max}) so that $A(a_{min}) = A(b_{max})$, and $a_{min} = 0$ or $b_{max} = \infty$. For $b \in (\bar{x}, b_{max})$ let $a = a(b)$ be the unique point in $(a_{min}, \bar{x})$ where $A(a) = A(b)$, and define

$$\begin{aligned} \Lambda(b) &= \int_{a(b)}^{b} u_2'(x)(A(x) - A(b))dy, \\ \gamma_{max} &= \Lambda(b_{max}). \end{aligned}$$

For details see Corollary 1.7.7 and Section 1.7.3. The following theorem gives the optimal strategy for the case of discrete dividends.

Theorem 1.2.2 *The initial value of the firm, the value function, and the optimal dividend strategy in the model for discrete dividends with transaction cost are the following.*

1) *If $S \geq \mu/\lambda$ and $\gamma < x_0$ hold, then the optimal strategy is to liquidate the company immediately and to distribute the initial reserve x_0 as dividend. In this case $\tau = 0$, $z_0 = x_0$, and the initial value of the firm is $v_0 = x_0 - \gamma + S$,*

2) *If $-1/r_2 < S < \mu/\lambda$ and $\gamma \geq \gamma_{max}$ hold or if $S \geq \mu/\lambda$ and $\gamma \geq x_0$ hold, then the equation $f(x) = x - \gamma + S$ has a unique positive solution x^*, and the initial dividend payment is*

$$z_0 = \begin{cases} 0, & x_0 < x^*, \\ x_0, & x_0 \geq x^*; \end{cases}$$

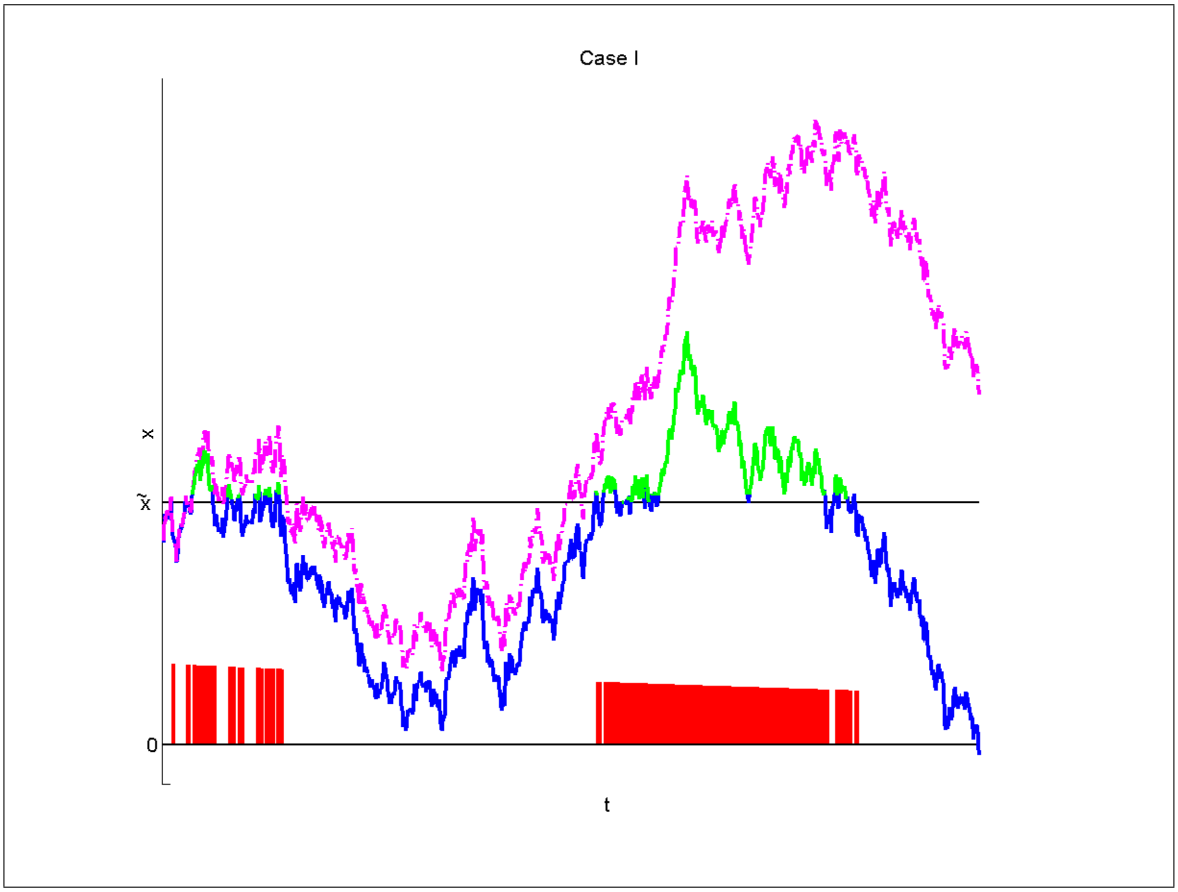

Figure 1.1: An example of the reserve and the dividend payments for the case of bounded dividend rate; *dashed*: an uncontrolled trajectory; *black:* reserve when the control is off; *light grey:* reserve when the control is on, *grey bars:* discounted dividends paid.

the optimal dividend strategy $(\tilde{\xi}_i, \tilde{T}_i)$ is to wait until the reserve reaches the critical level x^* and then immediately distribute the whole reserve as dividend $\tilde{\xi}_1 = x^*$, $T_1 = \inf\{t > 0 : X_t = x^*\}$, and liquidate the company.

The value function $\tilde{V} = \tilde{V}(\cdot)$ is given by

$$\tilde{V}(x) \;=\; \begin{cases} Su_1(x) + A(x^*)u_2(x), & x < x^*, \\ x - \gamma + S, & x \geq x^*; \end{cases}$$

and the initial value of the firm is $v_0 = \tilde{V}(x_0)$.

3) *If $-1/r_2 < S < \mu/\lambda$ and $\gamma < \gamma_{max}$ or if $S \leq -1/r_2$ then the the equation $\Lambda(b) = \gamma$ has a unique solution $\tilde{b}$, $\tilde{b} \geq \bar{x}$ (see Section 1.7.3). With $\tilde{a} = a(\tilde{b})$ the optimal initial dividend payment is*

$$z_0 \;=\; \begin{cases} 0, & x_0 < \tilde{b}, \\ x_0 - \tilde{a}, & x_0 \geq \tilde{b}; \end{cases}$$

and the optimal dividend strategy $(\tilde{\xi}_i, \tilde{T}_i)$ is to pay the amount $(\tilde{b} - \tilde{a})$ as groos dividend as soon as the reserve reaches the critical level $\tilde{b}$, i.e. $\tilde{\xi}_i = \tilde{b} - \tilde{a}$, $T_i = \inf\{t > T_{i-1} : X_{T_i} = \tilde{b}\}$.

The value function is

$$\tilde{V}(x) \;=\; \begin{cases} Su_1(x) + A(\tilde{b})u_2(x), & x < \tilde{b}, \\ \tilde{V}(\tilde{a}) + x - \tilde{a} - \gamma, & x \geq \tilde{b}; \end{cases}$$

and the initial value of the firm is $v_0 = \tilde{V}(x_0)$.

This theorem states: **whenever the reserve X_t hits a certain level $\tilde{b}$, one should pay $\tilde{b} - \tilde{a}$ in dividend pushing the value of X_t instantaneously to the value $\tilde{a}$ if only transaction costs are reasonable and the liquidation value is not very big.** This is illustrated in Fig. 1.2.

1.2.3 The case when the dividend process is any nonnegative, nondecreasing, right-continuous process

Suppose as before we are given an initial reserve $x_0 \geq 0$ and a liquidation value S, but assume now that the dividend process is any nonnegative, nondecreasing, and right-continuous process. Let r_1, r_2, $u_1(\cdot)$, $u_2(\cdot)$, $f(\cdot)$, $A(\cdot)$ be defined as above by (1.6), (1.7), (1.10), (1.11), (1.14), (1.15) respectively.

Before we formulate the theorem let us first introduce reflecting Brownian motion with drift for the barrier $\bar{x}$. Let $x \leq \bar{x}$ hold. Consider the solution $(\bar{X}, L) = (\bar{X}_t, L_t)_{t \geq 0}$ to

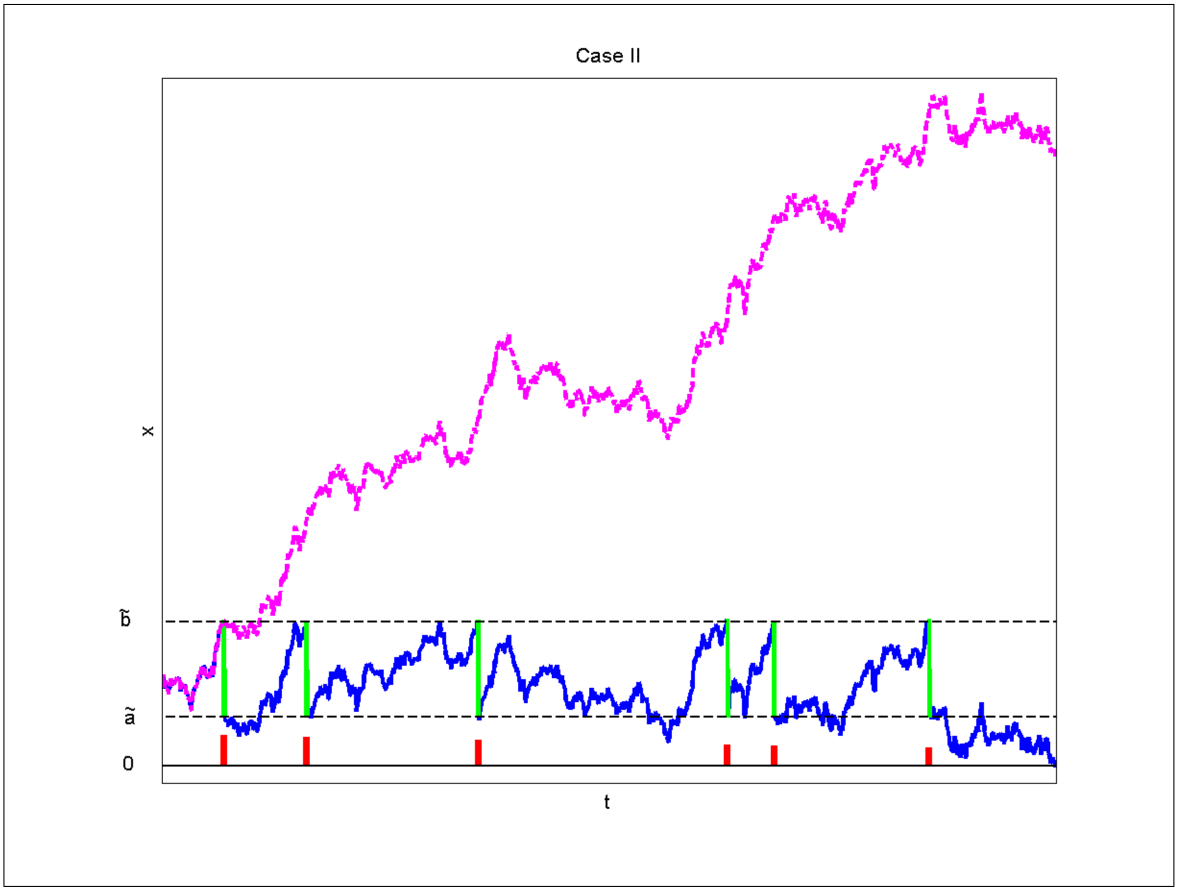

Figure 1.2: An example of the reserve and the dividend payments for the case of discrete dividend with transaction costs; *dashed*: an uncontrolled trajectory; *black:* reserve when the control is off; *light grey:* reserve when the control is on, *grey bars:* discounted dividends paid.

the stochastic differential equation with reflection (see [RevuzYor] Chapter IX, §2, Exercise 2.14)

$$\bar{X}_t = x + \mu t + W_t - L_t, \tag{1.16}$$

where $L = (L_t)_{t\geq 0}$ is a continuous nondecreasing $\mathcal{F}$-adapted process with $L_0 = 0$ such that

$$L_t = \int_0^t I(\bar{X}_s = \bar{x}) dL_s \tag{1.17}$$

and $X_t \leq \bar{x}$.

It is well known that a solution $(\bar{x}, L)$ exists, that the barrier $\bar{x}$ is a reflecting boundary for the process $\bar{X}$, and that L is the local time spent by the process $\bar{X}$ at the boundary $\bar{x}$.

Now let us formulate again the theorem that describes the optimal strategy.

Theorem 1.2.3 *The initial value of the firm, the value function, and the optimal dividend strategy for the case when the dividend process is any nonnegative, nondecreasing, right-continuous process are the following.*

1) *If $S \geq \mu/\lambda$ then the optimal strategy will be to distribute the initial reserve x_0 as dividend immediately, so $z_0 = x_0$, and hence to go bankrupt at the start, so $\tau = 0$ and $X_t = 0$ for all $t > 0$. This strategy yields the initial value of the firm, $v_0 = x_0 + S$.*

2) *If $S < \mu/\lambda$ holds then the equation $f(x) = \mu/\lambda$ has a unique positive solution $\bar{x}$, and the initial dividend payment is*

$$z_0 = \begin{cases} 0, & x_0 < \bar{x}, \\ x_0 - \bar{x}, & x_0 \geq \bar{x}. \end{cases}$$

The optimal dividend strategy is to wait until the reserve hits the critical level $\bar{x}$, then the process of dividend payments $L = (L_t)_{t\geq 0}$ and the reserve process $\bar{X} = (\bar{X}_t)_{t\geq 0}$ are the solutions of a SDE with reflection $\bar{X}_t = x + \mu t + W_t - L_t$, where $L = (L_t)_{t\geq 0}$ is a continuous nondecreasing and adapted process with $L_0 = 0$. The value function is

$$\tilde{V}(x) = \begin{cases} Su_1(x) + A(\bar{x})u_2(x), & x < \bar{x}, \\ x - \bar{x} + \mu/\lambda, & x \geq \bar{x}; \end{cases}$$

and the initial value of the firm is $v_0 = \tilde{V}(x_0)$.

This means that whenever the reserve X_t hits a certain value $\bar{x}$ from below one should start distributing the dividend so fast that the reserve reflects at $\bar{x}$. The dividend comes from the local time spent by the reserve at $\bar{x}$.

1.3 The case of bounded dividend rate

In the case of bounded dividend rates the stochastic optimal control problem defined by (1.2), (1.3) and (1.5) may be reformulated as follows. The reserve of the company is given by

$$\begin{cases} dX_t &= (\mu - U_t)dt + dW_t, \\ X_0 &= x_0 - z_0, \end{cases} \tag{1.18}$$

where x_0 is the initial reserve, z_0 is the initial dividend payment, and $U = (U_t)_{t \geq 0}$ is the bounded dividend rate process, $0 \leq U \leq K$. Here we define the value function $\tilde{V}(\cdot)$ by

$$\tilde{V}(x) = \sup_{0 \leq U \leq K} E_x \left\{ \int_{(0,\tau)} e^{-\lambda t} U_t dt + S e^{-\lambda \tau} \right\}, x \geq 0, \tag{1.19}$$

where the supremum is taken over all dividend rate processes U bounded by K, and where λ is the discount rate, S is the salvage value, and τ the bankruptcy time.

The initial value of the firm v_0 equals

$$v_0 = \max_z \left\{ z + \tilde{V}(x_0 - z) \right\},$$

where z is some initial dividend payment, $0 \leq z \leq x_0$. Note that the value function corresponds to the situation where no initial dividend payment is allocated, i.e. $z_0 \equiv 0$.

1.3.1 The candidate value function and the candidate optimal control

Suppose we found a candidate value function, say $V^* = V^*(\cdot)$. For now we leave the optimality check of the candidate value function, but discuss its properties.

Definition of the candidate value function and the candidate optimal control

Let us define the candidate value function $V^*(\cdot)$ on $[0, \infty)$ by

$$V^*(x) = \begin{cases} S u_1(x) + A(\tilde{x}) u_2(x), & x < \tilde{x}, \\ K/\lambda - (1/\rho) e^{-\rho(x - \tilde{x})}, & x \geq \tilde{x}, \end{cases} \tag{1.20}$$

where $\tilde{x}$ is the nonnegative solution of the equation $f(x) = S \vee Q$.

The boundary $\tilde{x}$ solves a free boundary problem, a so-called *Stefan problem*. Roughly speaking, the Stefan problem considers two regions and a free boundary between them. The boundary moves to maximize/minimize the associated functional (to maximize the utility function in our case). To make it more visual one can imagine two regions: one with ice and another with water and a moving boundary between them. The moving boundary eventually reaches the optimal state. In our case the Stefan problem is quite trivial, as it is only one-dimensional and perpetual in time. The two regions are $[0,\tilde{x}]$ and $[\tilde{x},\infty)$, and the moving boundary $\tilde{x}$ is a number that does not depend on time.

The two regions $[0,\tilde{x})$ and $[\tilde{x},\infty)$ have the crucial property: the candidate optimal control u^* is 0 in $[0,\tilde{x})$ and K in $[\tilde{x},\infty)$, i.e.

$$u^*(x) \quad = \quad \begin{cases} 0, & x<\tilde{x}, \\ K, & x\geq\tilde{x}. \end{cases} \tag{1.21}$$

It is easy to see that the candidate optimal control is of "bang-bang" type and assumes as its values the extreme values of the set of admissible controls $[0,K]$. In the region $[0,\tilde{x})$ we only "observe" and don't pay any dividends, but as soon as the process reaches the region $[\tilde{x},\infty)$ we start paying out dividends at the maximum possible speed.

Properties of the candidate value function

As we will see later the candidate value function V^* and the candidate control u^* introduced above were chosen to satisfy the Bellman equation (1.29) and the Bellman inequality (1.28). To show this we need properties of the candidate value function V^* in the regions $[0,\tilde{x}]$ and $[\tilde{x},\infty)$.

Denote by L the differential operator

$$L=\mu\frac{d}{dx}+\frac{1}{2}\frac{d^2}{dx^2}-\lambda.$$

Then the following proposition holds.

Proposition 1.3.1 *The function $V^*(\cdot)$ defined by (1.20) satisfies*

$$\begin{cases} LV^*(x)=0, & x<\tilde{x}, \\ LV^*(x)=K(V^{*\prime}(x)-1), & x\geq\tilde{x}. \end{cases}$$

Proof. The first equality follows from the fact that the functions $u_1(\cdot)$ and $u_2(\cdot)$ from definition (1.20) of V^* are linearly independent solutions of the differential equation $Lu=0$. The second equality we obtain from the knowledge that $(K/\lambda-Ce^{-\rho x})$, where C is any real constant, is a solution to $u_{xx}/2+(\mu-K)u_x-\lambda u+K=0$. □

The following lemma describes the behaviour of $V^{*\prime}(x)-1$ in the regions $[0,\tilde{x}]$ and $[\tilde{x},\infty)$.

Lemma 1.3.1 *With $\tilde{x}$ defined as the root of equation $f(x)=S\vee Q$, the function $V^*(\cdot)$ given by (1.20) is C^2 on $[0,\infty)$ and satisfies*

$$V^{*\prime}(x)-1>0,\ \ x<\tilde{x}, \tag{1.22}$$

$$V^{*\prime}(x)-1=0,\ \ x=\tilde{x}, \tag{1.23}$$

$$V^{*\prime}(x)-1<0,\ \ x>\tilde{x}. \tag{1.24}$$

PROOF. We discern two cases.

- Suppose $S \geq Q$. We have $\tilde{x} = 0$ by Lemma 1.7.2. Clearly, the function V^* is C^2 on $[0, \infty)$ then and we have

$$V^{*\prime}(x) - 1 = -\rho \left(S - \frac{K}{\lambda} \right) e^{-\rho x} - 1 \leq e^{-\rho x} - 1 < 0, \ x > 0,$$

since the statement $S \geq Q$ is equivalent to $\rho \left(S - \frac{K}{\lambda} \right) \geq -1$ by definition of Q.

- Suppose $S < Q$. It is obvious that the function V^* is C^2 on $[0, \tilde{x})$ and on $(\tilde{x}, \infty)$. Furthermore, $V^*(\tilde{x}-) = Su_1(\tilde{x}) + A(\tilde{x})u_2(\tilde{x}) = f(\tilde{x}) = Q$ and on the other hand $V^*(\tilde{x}+) = K/\lambda - 1/\rho = Q$ hold. For the first derivative we have

$$\begin{aligned} V^{*\prime}(\tilde{x}-) &= Su_1{}'(\tilde{x}) + A(\tilde{x})u_2{}'(\tilde{x}) = Su_1{}'(\tilde{x}) + \frac{1 - Su_1{}'(\tilde{x})}{u_2{}'(\tilde{x})} u_2{}'(\tilde{x}) = 1 \\ V^{*\prime}(\tilde{x}+) &= e^{-\rho(\tilde{x} - \tilde{x})} = 1. \end{aligned}$$

We have checked that V^* is C^1 on $[0, \infty)$ (note that we also proved (1.23)). Hence LV^* is continuous by Proposition 1.3.1. By definition of L this implies that V^* is C^2 on $[0, \infty)$. Now let us check (1.24) and (1.22). Again by Lemma 1.7.2 we have $\tilde{x} > 0$. For $x > \tilde{x}$ we have

$$V^{*\prime}(x) - 1 = e^{-\rho(x - \tilde{x})} - 1 < 0,$$

and hence (1.24). Let $\bar{x}$ be a solution to the equation $f(x) = \mu/\lambda$ (see Section 1.7.2). Note that $\tilde{x} < \bar{x}$ holds, as they are the solutions of $f(x) = Q$ and $f(x) = \mu/\lambda$ respectively, and $Q < \mu/\lambda$ holds by Lemma 1.7.1. Also note, that $f(x)$ is an increasing function for $x < \bar{x}$. Then by results from Corollary 1.7.5 we have for $0 \leq x < \tilde{x} < \bar{x}$

$$V^{*\prime}(x) - 1 = Su_1'(x) + A(\tilde{x})u_2'(x) - 1 = u_2'(x) \left(A(\tilde{x}) - A(x) \right) > 0.$$

□

1.3.2 Optimality of the initial payment

Now let us find the candidate initial value of the firm.

The initial value of the firm was defined as $v_0 = \max_z \left\{ z + \tilde{V}(x_0 - z) \right\}$, where z is the initial dividend payment. We are looking for a candidate initial value of the firm

$$v_0^* = \max_z \left\{ z + V^*(x_0 - z) \right\}, \tag{1.25}$$

and for a candidate initial dividend payment z_0^* such that $v_0^* = z_0^* + V^*(x_0 - z_0^*)$.

Write $x = x_0 - z$ and $\psi(x) = x_0 - x + V^*(x)$. In order to determine v_0^* in (1.25) we have to maximize $\psi(x)$ over $0 \le x \le x_0$. Since $\psi'(x) = V^{*\prime}(x) - 1 \le 0$ unless $S < Q$ and $x < \tilde{x}$ (by Lemma 1.3.1), the maximum of ψ is attained at $\tilde{x}$. Recall that $\tilde{x}$ was defined as the solution of $f(x) = Q$ if $S < Q$, and as 0 if $S \ge Q$. Hence returning to the previous notation we obtain

1. For $S < Q$ and $x_0 < \tilde{x}$ we have $z_0^* = 0$ and $v_0^* = V^*(x_0)$.
2. For $S < Q$ and $x_0 \ge \tilde{x}$ we have $z_0^* = x_0 - \tilde{x}$ and $v_0^* = x_0 - \tilde{x} + V^*(\tilde{x}) = x_0 - \tilde{x} + \mathrm{Q}$.
3. For $S \ge Q$ we have $z_0^* = x_0$ and $v_0^* = x_0 + V^*(0) = x_0 + S$.

Collecting all results, we see that we have found the initial value of the firm $v_0 = v_0^*$ and the initial dividend payment $z_0 = z_0^*$ provided we prove that the candidate value function is the best performing candidate and corresponds to the best control process. We will do this in the next section.

1.3.3 Value function optimality

Stochastic control verification properties

Suppose the board of directors uses the dividend strategy $U = (U_t)_{t\ge 0}$. Let $V(x, U)$ be the expected total discounted dividend corresponding to the board's strategy added together with the discounted liquidation value received upon bankruptcy with reserve x, $x = x_0 - z_0$

$$V(x, U) = \mathbf{E}_x \left\{ \int_{(0,\tau)} e^{-\lambda t} U_t dt + S e^{-\lambda \tau} \right\}.$$

Assume we have a candidate U^* for the optimal strategy and a candidate $V^*(\cdot)$ for the value function. To prove that these candidates are optimal indeed, it is enough to check the standard stochastic control *verification properties*:

(A) For any admissible control U

$$V(x, U) \le V^*(x), \text{ for all } x \ge 0. \tag{1.26}$$

(B) The control U^* satisfies

$$V(x, U^*) = V^*(x), \text{ for all } x \ge 0. \tag{1.27}$$

Indeed, if these verification properties are satisfied, they just show that no admissible control can beat U^*, i.e. $\tilde{V}(x) = V^*(x) = V^(x, U^*)$ for all $x \ge 0$.

The variational inequalities

Here we will show that the verification properties are satisfied once the variational inequalities described below are. This means that for proving optimality of U^* and V^*, it is enough to check these variational inequalities, which are named after Bellman.

Assume there exist a bounded function $V^*(\cdot)$ in $C^2[0,\infty)$ and an adapted control U^* such that the following variational inequalities hold:

- *Bellman inequality.* For any admissible control U

$$LV^*(x) + U_t\left(1 - V^{*\prime}(x)\right) \leq 0, \text{ for any } x \geq 0. \tag{1.28}$$

- *Bellman equation.*

$$LV^*(x) + U_t^*\left(1 - V^{*\prime}(x)\right) = 0, \text{ for any } x \geq 0. \tag{1.29}$$

Let us show that the verification properties (1.26) and (1.27) are satisfied if V^* and U^* satisfy (1.28) and (1.29).

Apply Itô's formula to $(e^{-\lambda t}V^*(X_t))_{t\geq 0}$, yielding

$$e^{-\lambda(t\wedge\tau)}V^*(X_{t\wedge\tau}) = V^*(x) + \int_0^{t\wedge\tau} e^{-\lambda s}\left[LV^*(X_s) - U_s V^{*\prime}(X_s)\right] ds + \int_0^{t\wedge\tau} e^{-\lambda s} V^{*\prime}(X_s) dW_s.$$

Taking the expectation $\mathbf{E}_x$ and using Bellman's inequality (1.28) we obtain

$$\begin{aligned} V^*(x) &= \mathbf{E}_x e^{-\lambda(t\wedge\tau)}V^*(X_{t\wedge\tau}) - \mathbf{E}_x \int_0^{t\wedge\tau} e^{-\lambda s}\left[LV^*(X_s) - U_s V^{*\prime}(X_s)\right] ds \\ &\qquad - \mathbf{E}_x \int_0^{t\wedge\tau} e^{-\lambda s} V^{*\prime}(X_s) dW_s \\ &\geq \mathbf{E}_x e^{-\lambda(t\wedge\tau)}V^*(X_{t\wedge\tau}) + \mathbf{E}_x \int_0^{t\wedge\tau} e^{-\lambda s} U_s ds - \mathbf{E}_x \int_0^{t\wedge\tau} e^{-\lambda s} V^{*\prime}(X_s) dW_s. \end{aligned} \tag{1.30}$$

The last stochastic integral in (1.30) is a martingale. Therefore its expectation is equal to zero. Letting $t \to \infty$ we find $e^{-\lambda(t\wedge\tau)}V^*(X_{t\wedge\tau}) \to e^{-\lambda\tau}S$. Since V^* and U are bounded, we may apply Lebesgue's theorem on dominated convergence and obtain

$$V^*(x) \geq \mathbf{E}_x e^{-\lambda\tau}S + \mathbf{E}_x \int_0^{\tau} e^{-\lambda s} U_s ds = V(x, U).$$

It is obvious that property (A) is satisfied. Property (B) can be checked in the same way but instead of the Bellman inequality we use the Bellman equation (1.29).

Verifying the Bellman inequalities

In the previous section we have proved that if for some function $V^* = V^*(\cdot)$ from $C^2[0,\infty)$ to $[0,\infty)$ and an adapted control U^* variational inequalities (1.28) and (1.29) are satisfied, then $V^* = V^*(\cdot)$ and $U^* = u^*(\cdot)$ give the optimal solution to our optimal control problem.

Recall that the boundary $\tilde{x}$ was defined as $\tilde{x} = 0$ if $S \geq Q$, otherwise $\tilde{x}$ is the unique solution of equation $f(x) = Q$. Then the following lemma completes the argument.

Lemma 1.3.2 *Functions $V^*(\cdot)$ and $u^*(\cdot)$ given by (1.20) and (1.21) correspondingly satisfy variational inequalities (1.28) and (1.29), i.e. for any admissible U_t*

$$\begin{aligned} LV^*(x) + U_t\left(1 - V^{*\prime}(x)\right) &\leq 0, \text{ for any } x \geq 0, \\ LV^*(x) + u^*(x)\left(1 - V^{*\prime}(x)\right) &= 0, \text{ for any } x \geq 0. \end{aligned}$$

PROOF. We discern two cases.

- For $x < \tilde{x}$ we have

$$LV^*(x) + U_t\left(1 - V^{*\prime}(x)\right) = U_t\left(1 - V^{*\prime}(x)\right) \leq 0, \tag{1.31}$$

as $LV^*(x) = 0$ by Proposition 1.3.1, $0 \leq U_t \leq K$ by definition, and $(1 - V^{*\prime}(x)) < 0$ by Lemma 1.3.1. Moreover, we see that the Bellman equation is satisfied if $u^*(x) = 0$ for all $0 \leq x < \tilde{x}$.

- For $x \geq \tilde{x}$ we have

$$\begin{aligned} LV^*(x) + U_t\left(1 - V^{*\prime}(x)\right) &= K(V^{*\prime}(x) - 1) + U_t\left(1 - V^{*\prime}(x)\right) \\ &= (K - U_t)(V^{*\prime}(x) - 1) \geq 0, \end{aligned}$$

by Proposition 1.3.1, by (1.23) and (1.24), and since $0 \leq U_t \leq K$ holds'. by definition. The equality is satisfied if $U_t = u^*(x) = K$. □

Thus we see that Lemma 1.3.2 completes the proof of Theorem 1.2.1. Existence of the optimal solution, i.e. the existence of $\tilde{x}$, follows from Corollary 1.7.2, see Appendix B.

1.4 The case of discrete dividend with transaction costs

Suppose the board of directors uses a dividend strategy with discrete payments $\pi = (T_i, \xi_i)_{i\geq 1}$. Here dividend payments are represented by ex-dividend times $0 = T_0 < T_1 < T_2 < \ldots$ and by the amount of the gross dividends $\xi_0 = z_0, \xi_1, \ldots$, which are not allowed to exceed the present reserve, $\xi_0 = z_0 \leq x_0$, $\xi_i \leq X_{T_i-}$, $i = 1, \ldots$. The dividend process is given by

$$Z_t = z_0 + \sum_{i\geq 1} \xi_i I(T_i \leq t \wedge \tau), \tag{1.32}$$

where τ is the bankruptcy time $\tau = \inf\{s : X_s \leq 0\}$.

The reserve of the company follows from

$$\begin{cases} dX_t &= \mu dt + dW_t - dZ_t, \\ X_0 &= x_0 - z_0, \end{cases} \tag{1.33}$$

where x_0 is the initial reserve, ξ_i is the gross dividend payment made at time T_i, and z_0 is the initial dividend payment. A transaction cost γ is associated with each dividend payment. Consequently, the share holder gets nett dividend payments $\xi_i - \gamma$, while the company pays gross dividends ξ_i. By common sense the nett dividend payments $\xi_i - \gamma$ must be positive. For notation convenience the initial dividend payment $\xi_0 = z_0$ is always present but could be zero, $z_0 = 0$, in which case no transaction cost is charged.

We measure the performance of a dividend policy by a functional of the dividend process $\pi = (\xi_i, T_i)_{i\geq 1}$ and of the transaction cost γ via

$$V(x,\pi) = E_x\left(\sum_{i\geq 1} e^{-\lambda T_i}(\xi_i - \gamma) I(T_i \leq \tau) + Se^{-\lambda\tau}\right), \tag{1.34}$$

where S is the liquidation value and the expectation is taken conditionally on $x = X_0 = x_0 - z_0$.

In a similar way as in the preceding Section 1.3 we define the value function $\tilde{V}(\cdot)$ by

$$\tilde{V}(x) = \sup_{\pi} V(x,\pi),\ x \geq 0,$$

and the initial value of the firm is

$$v_0 = \max_{0\leq z\leq x_0}\left\{((z-\gamma)\vee 0) + \tilde{V}(x_0 - z)\right\},$$

with x_0 the initial reserve. Observe that $\tilde{V}(0) = S$ as in the case of a bounded dividend rate, since $x = 0$ implies $\tau = 0$.

1.4.1 The candidate value function and the candidate optimal control

Suppose we found a candidate value function, say $V^* = V^*(\cdot)$. For now we leave the optimality check of the candidate value function for the following section, but discuss its properties.

N	CASE	$\tilde{a}$	$\tilde{b}$
1	$-1/r_2 < S < \mu/\lambda$ $\gamma \geq \gamma_{max}$	0	solution to $f(x) = x - \gamma + S$
2	$S \geq \mu/\lambda$ $\gamma \geq x_0$	0	solution to $f(x) = x - \gamma + S$
3	$-1/r_2 < S < \mu/\lambda$ $\gamma < \gamma_{max}$	$a(\tilde{b})$	solution to $\Lambda(x) = \gamma$
4	$S \leq -1/r_2$	$a(\tilde{b})$	solution to $\Lambda(x) = \gamma$
5	$S \geq \mu/\lambda, \gamma < x_0$	0	0

Table 1.1: Boundaries $\tilde{a}$ and $\tilde{b}$ in the case of discrete dividends, $\gamma > 0$.

Definition of the candidate value function and the candidate optimal control

Let $\tilde{a}$ and $\tilde{b}$ be given by Table 1.1. To recall the notation see Section 1.6. Define the candidate value function

$$V^*(x) \quad = \quad \begin{cases} Su_1(x) + A(\tilde{b})u_2(x), & x < \tilde{b}, \\ V^*(\tilde{a}) + x - \tilde{a} - \gamma, & x \geq \tilde{b}, \end{cases} \tag{1.35}$$

with the exception of case 5 of Table 1.1 with $S \geq \mu/\lambda, \gamma < x_0$, when we set $V^*(x) = S$ for $x \geq 0$.

Again, as in the case of bounded dividends the candidate optimal control is zero on $[0, \tilde{b})$. And again, the boundary $\tilde{b}$ defined by Table 1.1 solves a free boundary problem. But unlike the case of bounded dividends there are in fact two boundaries $\tilde{a}$ and $\tilde{b}$ tied to each other by some relation with the transaction cost γ. As we will see later the boundaries were chosen to satisfy the relation $\int_{\tilde{a}}^{\tilde{b}}(1 - V^{*\prime}(y))dy = \gamma$.

It is also useful to point out that unlike the case of bounded dividends, in the case of discrete dividends there are no restrictions on the speed of dividend payments. Thus, if we pay dividends with "the maximum possible speed" and start from below the boundary $\tilde{b}$

then the corresponding reserve process does not leave $[0, \tilde{b}]$. In other words, the candidate optimal control is designed to keep the reserve process within the strip $[0, \tilde{b}]$. As soon as the reserve reaches the boundary $\tilde{b}$ the amount $\tilde{b} - \tilde{a}$ is paid as a dividend, and the reserve process starts again from $\tilde{a}$. More precisely, the candidate optimal control π^* is defined as a multivariate point process $\pi^* = (T_i^*, \xi_i^*)_{i \geq 1}$ consisting of candidate optimal payment times $T_i^* = \inf(t > T_{i-1}^* : X_{t-} = \tilde{b})$, and of candidate optimal dividend payments $\tilde{\xi}_i = \tilde{b} - \tilde{a}$, $i = 1, 2, \ldots$.

Thus, it seems reasonable to consider three regions instead of two, namely $[0, \tilde{a})$, $[\tilde{a}, \tilde{b})$, and $[\tilde{b}, \infty)$.

The properties of the candidate value function

The following lemma describes the behaviour of $V^{*\prime}(x) - 1$ in the regions $[0, \tilde{a})$, $[\tilde{a}, \tilde{b})$, and $[\tilde{b}, \infty)$.

Lemma 1.4.1 *With $\tilde{a}$ and $\tilde{b}$ defined by Table 1.1, the candidate value function $V^*(\cdot)$ given by (1.35) is C^1 on $[0, \infty)$ and satisfies with the exception of case 5 of Table 1.1,*

$$V^{*\prime}(x) - 1 > 0, \qquad 0 \leq x < \tilde{a}, \tag{1.36}$$

$$V^{*\prime}(x) - 1 = 0, \qquad x = \tilde{a}, \ \textit{for } \tilde{a} > 0, \tag{1.37}$$

$$V^{*\prime}(x) - 1 < 0, \qquad \tilde{a} < x < \tilde{b}, \tag{1.38}$$

$$V^{*\prime}(x) - 1 = 0, \qquad x \geq \tilde{b}. \tag{1.39}$$

Proof. It is obvious that the function V^* is C^1 on $[0, \tilde{b})$ and on $(\tilde{b}, \infty)$. Let us check this property at $\tilde{b}$.

- Suppose $(-1/r_2 < S < \mu/\lambda$ and $\gamma \geq \gamma_{max})$ or $(S \geq \mu/\lambda$ and $\gamma \geq x_0)$, see Table 1.1 cases 1 and 2. Then $V^*(\tilde{b}-) = Su_1(\tilde{b}) + A(\tilde{b})u_2(\tilde{b}) = f(\tilde{b})$. On the other hand $V(\tilde{b}+) = S + \tilde{b} - \gamma$. Moreover $f(\tilde{b}) = S + \tilde{b} - \gamma$ by definition of $\tilde{b}$. Hence $V^*(\tilde{b}-) = V^*(\tilde{b}+)$.

- Suppose $(-1/r_2 < S < \mu/\lambda$ and $\gamma < \gamma_{max})$ or $S \geq -1/r_2$, see Table 1.1 cases 3 and 4. By definition of $\tilde{a}$ and $\tilde{b}$ we have

$$\begin{aligned} \gamma &= \int_{\tilde{a}}^{\tilde{b}} u_2{}'(y) \left(A(y) - A(\tilde{b}) \right) dy = \int_{\tilde{a}}^{\tilde{b}} \left(1 - V^{*\prime}(y) \right) dy \\ &= \tilde{b} - \tilde{a} - (V^*(\tilde{b}-) - V^*(\tilde{a})). \end{aligned}$$

  Hence $V^*(\tilde{b}-) = V^*(\tilde{a}) + \tilde{b} - \tilde{a} - \gamma = V^*(\tilde{b}+)$.

For the left derivative we have

$$V^{*\prime}(\tilde{b}-) = Su_1{}'(\tilde{b}) + A(\tilde{b})u_2{}'(\tilde{b}) = Su_1{}'(\tilde{b}) + \frac{1 - Su_1{}'(\tilde{b})}{u_2{}'(\tilde{b})}u_2{}'(\tilde{b}) = 1. \tag{1.40}$$

Note, that (1.40) can be repeated for $\tilde{a} > 0$, since $A(\tilde{b}) = A(\tilde{a})$, and, therefore, we also have $V^{*\prime}(\tilde{a}) = 1$. Moreover, for the right derivative as well as for all $x \geq \tilde{b}$ we obtain $V^{*\prime}(x) - 1 = 0$, since $V^*(x) = V^*(\tilde{a}) + x - \tilde{a} - \gamma$. Consequently, $V^{*\prime}(\tilde{b}-) = 1 = V^{*\prime}(\tilde{b}+)$. In this way we have seen that V^* is C^1 on $[0, \infty)$. Now let us check the remaining properties (1.36) and (1.38). Indeed, for $x < \tilde{b}$, $\tilde{b} > 0$ we have by definition of $A(\cdot)$ (see 1.74)

$$\begin{aligned} V^{*\prime}(x) - 1 &= Su_1'(x) + A(\tilde{b})u_2'(x) - 1 \\ &= u_2'(x)\left(A(\tilde{b}) - A(x)\right) \\ &= -\phi(\tilde{b}, x). \end{aligned}$$

Thus, by Lemma 1.7.6 we obtain the required results. □

1.4.2 Firm's initial value optimality

Here we find explicitly the candidate initial value of the firm. The candidate initial value of the firm is given by

$$v_0 = \max_{0 \leq z \leq x_0} \left\{((z - \gamma) \vee 0) + V^*(x_0 - z)\right\}, \tag{1.41}$$

where $V^*(\cdot)$ is given by (1.35) for the cases 1 — 4 of Table 1.1, and $V^*(x) = S$, $x \geq 0$ for the case 5.

For the case 5, when V^* is a constant, we get $z_0 = x_0$ and $v_0 = x_0 - \gamma + S$.

Now let us study the remaining cases 1 — 4. Denote $y = x_0 - z$ and $y_0 = x_0 - z_0^*$. Define the function $v = v(\cdot)$ as $v(y) = ((x_0 - \gamma - y) \vee 0) + V^*(y)$. We reformulate the problem (1.41) as

$$v_0 = \max_{0 \leq y \leq x_0} v(y) = \max_{0 \leq y \leq x_0} \left\{((x_0 - \gamma - y) \vee 0) + V^*(y).\right\}.$$

Consider separately the following cases

- Suppose $x_0 - \gamma \leq 0$. Taking in mind that V^* is an increasing function we get $y_0 = x_0$ and $v_0 = V^*(x_0)$.

- Suppose $x_0 - \gamma > 0$. The derivative of v is

$$v'(y) = \left\{ \begin{array}{cc} -1 + V^{*\prime}(y), & y < x_0 - \gamma, \\ V^{*\prime}(y), & y > x_0 - \gamma \end{array} \right. . \tag{1.42}$$

  V^* is an increasing function. Moreover $V^{*\prime}(y) - 1 > 0$ for $y < \tilde{a}$, and $V^{*\prime}(y) - 1 \leq 0$ for $y \geq \tilde{a}$ by the results of Lemma 1.4.1. Thus, from (1.42) we get:

- Suppose $\tilde{a} > x_0 - \gamma$. Then $v = v(\cdot)$ is an increasing function. Consequently, $y_0 = x_0$, i.e. $z_0^* = 0$, and $v_0 = V^*(x_0)$.
- Now suppose $\tilde{a} < x_0 - \gamma$. Here the function $v = v(\cdot)$ has the following dynamics

$$v(y) = \begin{cases} \nearrow, & 0 < y < \tilde{a}, \\ \searrow, & \tilde{a} < y < x_0 - \gamma, \\ \nearrow, & x_0 - \gamma < y < x_0. \end{cases}$$

Thus, in order to find the maximum, we have to compare $v(\tilde{a})$ and $v(x_0)$, i.e. to compare $x_0 - \tilde{a} - \gamma + V^*(\tilde{a})$ and $V^*(x_0)$.

If $x_0 \geq \tilde{b}$, then by definition of V^* we have $V^*(x_0) = x_0 - \tilde{a} - \gamma + V^*(\tilde{a})$. We choose the candidate initial dividend payment z_0^* to be equal to the maximum possible optimal value $x_0 - \tilde{a}$.

Now suppose $x_0 < \tilde{b}$. Let us consider the following cases.

  - The cases 1 and 2 of Table 1.1, $x_0 < \tilde{b}$. Here $\tilde{a} = 0$. Consequently, by Lemma 1.4.1 and the definition of $\tilde{b}$, we get

$$V^*(x_0) - x_0 \geq V^*(\tilde{b}) - \tilde{b} = S - \gamma,$$

  or rearranging the terms

$$V^*(x_0) \geq x_0 - \gamma + S = x_0 - \tilde{a} - \gamma + V^*(\tilde{a}).$$

  - Consider the cases 3 and 4 of Table 1.1, $x_0 < \tilde{b}$. Recall that $\tilde{a}$ and $\tilde{b}$ were defined to satisfy the following equality (see Section 1.7.3)

$$\gamma = \int_{\tilde{a}}^{\tilde{b}} \big(1 - V^{*\prime}(y)\big)\, dy.$$

  Moreover, by Lemma 1.4.1 we have $1 - V^{*\prime}(y) > 0$ for $\tilde{a} < y < \tilde{b}$. It follows

$$\gamma = \int_{\tilde{a}}^{\tilde{b}} \big(1 - V^{*\prime}(y)\big)\, dy \geq \int_{\tilde{a}}^{x_0} \big(1 - V^{*\prime}(y)\big)\, dy, \text{ for } x_0 < \tilde{b}.$$

  Rewriting the latter integral we get $\gamma \geq x_0 - V^*(x_0) - \tilde{a} + V^*(\tilde{a})$, or rearranging the terms $V^*(x_0) \geq x_0 - \tilde{a} - \gamma + V^*(\tilde{a})$.

Thus, $v_0^* = V^*(x_0)$ for $x_0 < \tilde{b}$. It also means that the candidate initial dividend payment z_0^* in case $x_0 < \tilde{b}$ is zero.

Collecting all the results, we see that, if given the best performing candidate value function, or, more precisely, if given $V^*(x)$ such that $V^*(x) \geq V(x, \pi)$ for any $x \geq 0$ and any admissible strategy π, then we immediately have the initial value of the firm, $v_0 = v_0^* = V^*(x_0)$, and the initial dividend payment $z_0 = z_0^*$, where z_0^* is presented by the formulae

$$z_0^* = \begin{cases} 0, & x_0 < \tilde{b}, \\ x_0 - \tilde{a}, & x_0 \geq \tilde{b}, \end{cases}$$

with $\tilde{a}$ and $\tilde{b}$ given by Table 1.1.

We are going to check the optimality of the candidate value function in the following section.

1.4.3 Value function optimality

Let the process $X = (X_t)_{t \geq 0}$ correspond to a chosen dividend strategy $\pi = (T_i, \xi_i)_{i \geq 1}$, i.e.

$$\begin{cases} dX_t &= \mu dt + dW_t - dZ_t, \\ X_0 &= x_0 - z_0, \end{cases} \tag{1.43}$$

where

$$Z_t = \sum_{i \geq 1} \xi_i I(T_i \leq t \wedge \tau).$$

Note that if the initial dividend payment coincides with the candidate initial dividend payment $z_0 = z_0^*$, where

$$z_0^* = \begin{cases} 0, & x_0 < \tilde{b}, \\ x_0 - \tilde{a}, & x_0 \geq \tilde{b}, \end{cases}$$

with $\tilde{a}$ and $\tilde{b}$ given by Table 1.1, then it follows $X_0 = x_0 - z_0^* \leq \tilde{b}$.

Stochastic control verification properties

Let $M^\pi \subset \mathbb{R}_+$ be the range of the process $X = (X_t)_{t \geq 0}$ corresponding to the dividend strategy $\pi = (T_i, \xi_i)_{i \geq 1}$. Assume we have a candidate $\pi^* = (T_i^*, \xi_i^*)$ for the optimal strategy and a candidate $V^*(\cdot)$ for the value function. To prove that the strategy π^* and the function $V^*(\cdot)$ are optimal, it is enough to check the standard stochastic control *verification properties*:

(A_d) For any admissible control $\pi = (T_i, \xi_i)_{i \geq 1}$

$$V(x, \pi) \leq V^*(x), \text{ for all } x \geq 0. \tag{1.44}$$

(B_d) The control π^* satisfies

$$V(x,\pi^*) = V^*(x), \text{ for all } x \in M^{\pi^*}. \tag{1.45}$$

The property (A_d) says that no other function can be better than the candidate V^* on the range of all possible processes X controlled by π, i.e. $[0,\infty)$. But the property (B_d) points out a certain control π^* on which the corresponding utility function $V(x,\pi^*)$ coincides with the best performing candidate V^*. Under this control π^* the range of the corresponding controlled process X could be smaller compared to the original range $[0,\infty)$. To prove optimality it is enough to consider the range M^{π^*}. As we will see later, in our case M^{π^*} is the strip $[0,\tilde{b}]$, where $\tilde{b}$ is given by Table 1.1.

An analogue of the variational inequalities

Recall the notation

$$L = \mu\frac{d}{dx} + \frac{1}{2}\frac{d^2}{dx^2} - \lambda. \tag{1.46}$$

Assume there exist a C^2 function $V^*(\cdot)$ of $[0,\infty]$ to $[0,\infty)$ and an adapted control $\pi^* = (T_i^*, \xi_i^*)_{i\geq 0}$ such that the following analogues of the Bellman inequality (1.28) and the Bellman equation (1.29) hold:

I For all admissible controls $\pi = (T_i, \xi_i)_{i\geq 1}$ and the candidate value function $V^*(\cdot)$ we have

a) $LV^*(x) \leq 0$, for all $x \geq 0$,

b) $\int_{X_{T_i}}^{X_{T_i}-}(1 - V^{*\prime}(y))dy \leq \gamma$, where $\xi_i = X_{T_i-} - X_{T_i}$, $i = 1, 2, \ldots$.

II The candidate optimal control $\pi^* = (T_i^*, \xi_i^*)_{i\geq 1}$ and the candidate value function $V^*(\cdot)$ satisfy

a) $LV^*(x) = 0$, for all $x \in M^{\pi^*}$,

b) $\int_{X_{T_i^*}}^{X_{T_i^*}-}(1 - V^{*\prime}(y))dy = \gamma$, where $\xi_i^* = X_{T_i^*-} - X_{T_i^*}$, $i = 1, \ldots$.

Let us show now that the verification properties (1.44) and (1.45) are satisfied for such $V^*(\cdot)$ and $\pi^* = (T_i^*, \xi_i^*)_{i\geq 1}$.

Apply Itô's formula to the semimartingale $(e^{-\lambda t}V^*(X_t))_{t\geq 0}$:

$$
\begin{aligned}
&e^{-\lambda(t\wedge\tau)}V^*(X_{t\wedge\tau}) = V^*(X_0) + \int_0^{t\wedge\tau}(-\lambda e^{-\lambda s}V^*(X_s))ds + \int_0^{t\wedge\tau} e^{-\lambda s}V^{*\prime}(X_{s-})dX_s \\
&+\frac{1}{2}\int_0^{t\wedge\tau} e^{-\lambda s}V^{*\prime\prime}(X_s)ds + \sum_{0<s\le t\wedge\tau} e^{-\lambda s}\left\{V^*(X_s) - V^*(X_{s-}) - V^{*\prime}(X_{s-})\Delta X_s\right\} \\
&= V^*(X_0) - \int_0^{t\wedge\tau} e^{-\lambda s}V^{*\prime}(X_{s-})dZ_s + \int_0^{t\wedge\tau} e^{-\lambda s}LV^*(X_s)ds \\
&+ \int_0^{t\wedge\tau} e^{-\lambda s}V^{*\prime}(X_s)dW_s + \sum_{0<s\le t\wedge\tau} e^{-\lambda s}\left\{V^*(X_s) - V^*(X_{s-}) - V^{*\prime}(X_{s-})\Delta X_s\right\} \\
&\overset{(\Delta X_{T_i} = -\xi_i)}{=} V^*(X_0) + \int_0^{t\wedge\tau} e^{-\lambda s}LV^*(X_s)ds \qquad (1.47) \\
&+ \int_0^{t\wedge\tau} e^{-\lambda s}V^{*\prime}(X_s)dW_s + \sum_{0<s\le t\wedge\tau} e^{-\lambda s}\left\{V^*(X_s) - V^*(X_{s-})\right\}.
\end{aligned}
$$

Note that the last term in (1.47) can be represented as

$$
\begin{aligned}
&\sum_{0<s\le t\wedge\tau} e^{-\lambda s}\left\{V^*(X_s) - V^*(X_{s-})\right\} = -\sum_{i\ge 1, T_i\le t\wedge\tau} e^{-\lambda T_i}(\xi_i-\gamma)I(\xi_i>\gamma) \\
&+ \sum_{i\ge 1, T_i\le t\wedge\tau} e^{-\lambda T_i}[V^*(X_{T_i}) - V^*(X_{T_i -}) - (\Delta X_{T_i}+\gamma)]I(\xi_i>\gamma) \\
&= -\sum_{i\ge 1, T_i\le t\wedge\tau} e^{-\lambda T_i}(\xi_i-\gamma)I(\xi_i>\gamma) \\
&+ \sum_{i\ge 1, T_i\le t\wedge\tau} e^{-\lambda T_i}\left[\int_{X_{T_i}}^{X_{T_i -}}[-V^{*\prime}(u)+1]du - \gamma\right] I(\xi_i>\gamma).
\end{aligned}
$$

Thus, taking the expectation of (1.47) conditionally on $X_0 = x$ we obtain

$$
\begin{aligned}
V^*(x) \;\; = \;\; & \mathbf{E}_x e^{-\lambda(t\wedge\tau)}V^*(X_{t\wedge\tau}) - \mathbf{E}_x\int_0^{t\wedge\tau} e^{-\lambda s}LV^*(X_s)ds \\
& -\mathbf{E}_x\int_0^{t\wedge\tau} e^{-\lambda s}V^{*\prime}(X_s)dW_s + \mathbf{E}_x\left\{\sum_{i\ge 1, T_i\le t\wedge\tau} e^{-\lambda T_i}(\xi_i-\gamma)I(\xi_i>0)\right\} \qquad (1.48) \\
& - \mathbf{E}_x\left\{\sum_{i\ge 1, T_i\le\tau} e^{-\lambda T_i}\left[\int_{X_{T_i}}^{X_{T_i -}}[-V^{*\prime}(u)+1]du - \gamma\right] I(\xi_i>\gamma)\right\}.
\end{aligned}
$$

The stochastic integral in (1.48) is a martingale, as $V^{*\prime}(x)$ is bounded on $[0,\tilde{b}]$ as a continuous function, and $V^{*\prime}(x) = 1$ on $[\tilde{b},\infty)$. Therefore its expectation vanishes. Also, letting $t \to \infty$ and using the fact $V^*(0) = S$, we obtain

$$\begin{aligned} V^*(x) &= \mathbf{E}_x e^{-\lambda\tau} S + \mathbf{E}_x \left\{ \sum_{i\geq 1, T_i \leq \tau} e^{-\lambda T_i} (\xi_i - \gamma) \right\} - \mathbf{E}_x \int_0^\tau e^{-\lambda s} LV^*(X_s) ds \\ &\quad - \mathbf{E}_x \left\{ \sum_{i\geq 1, T_i \leq \tau} e^{-\lambda T_i} \left[\int_{X_{T_i}}^{X_{T_i-}} [-V^{*\prime}(u) + 1] du - \gamma \right] \right\}. \end{aligned}$$

By **I a)** and **I b)** this yields $V^*(x) \geq V(x, \pi)$, i.e. (A_d). Property (B_d) is obtained in the same way using **II a)** and **II b)**.

Verifying the analogue of the Bellman inequalities

In the previous section we have proved that if for some C^2 function $V^*(\cdot)$ from $[0, \infty)$ to $[0, \infty)$ and an adapted control $\pi^* = (T_i^*, \xi_i^*)_{i\geq 1}$ the analogues of the variational inequalities **Ia,b** and **IIa,b** from this Section 1.4.3 are satisfied, then $V^* = V^*(\cdot)$ and $\pi^* = (T_i^*, \xi_i^*)_{i\geq 1}$ are the optimal solution to our optimal control problem. Let us check those variational inequalities for the function $V^*(\cdot)$ and the process π^*.

Lemma 1.4.2 *Let $\tilde{a}$, $\tilde{b}$ be given by Table 1.1; let $V^*(\cdot)$ be the candidate value function given by (1.35); and let π^* be the candidate optimal dividend process, which is a multivariate point process $\pi^* = (T_i^*, \xi_i^*)_{i\geq 1}$ consisting of candidate optimal times $T_i^* = \inf(t > T_{i-1}^* : X_{t-} = \tilde{b})$, and of candidate optimal dividend payments $\tilde{\xi}_i = \tilde{b} - \tilde{a}$, $i = 1, 2, \ldots$.*

Then the candidate value function $V^* = V^*(\cdot)$ and the candidate optimal dividend process $\pi^* = (T_i^*, \xi_i^*)_{i\geq 1}$ satisfy the analogue of the variational inequalities **I a)**, **I b)**, **II a)** *and* **II b)** *from this Section 1.4.3.*

Proof.

I a)

- Suppose $x < \tilde{b}$. Then by definition of $V^*(\cdot)$ we have $V^*(x) = Su_1(x) + A(\tilde{b})u_2(x)$, see (1.35). But $u_1(\cdot)$ and $u_2(\cdot)$ linearly independent solutions to the differential equation $Lu = 0$, see Note 1.6.1. It follows that $LV^*(x) = 0$ for $0 \leq x < \tilde{b}$.
- Suppose $x \geq \tilde{b}$. Then by definition of $V^*(\cdot)$ we have $V^*(x) = V^*(\tilde{a}) + x - \tilde{a} - \gamma$, see (1.35). Thus,

$$\begin{aligned} LV^*(x) &= \mu - \lambda V^*(x) = \lambda \left(\frac{\mu}{\lambda} - V^*(x) \right) \\ &\leq \lambda \left(\frac{\mu}{\lambda} - V^*(\tilde{b}) \right). \end{aligned}$$

Therefore, if we show that $V^*(\tilde{b}) \geq \mu/\lambda$, where $\tilde{b}$ is given by Table 1.1, then we will have proved the desired property **I a)** for $x \geq \tilde{b}$.

Indeed, if $S \geq \mu/\lambda$ then

$$\frac{\mu}{\lambda} - V^*(\tilde{b}) \leq \frac{\mu}{\lambda} - V^*(0) = \frac{\mu}{\lambda} - S \leq 0.$$

Now suppose $S < \mu/\lambda$. Recall that $\mu/\lambda = f(\bar{x})$ by definition of $\bar{x}$, see Table 1.2, and $V^*(\tilde{b}) = f(\tilde{b})$ by definition of the functions V^* and f. Also note that $\tilde{b} > \bar{x}$, as it follows from Lemma 1.7.4 for the case $\gamma \geq \gamma_{max}$, and from construction of $\tilde{b}$ for the case $\gamma < \gamma_{max}$. Thus, by Corollary 1.7.1,

$$\frac{\mu}{\lambda} - V^*(\tilde{b}) = f(\bar{x}) - f(\tilde{b}) \leq 0.$$

I b) The function $\phi = \phi(b, \cdot)$ is defined by (1.76) as

$$\phi(b, x) = u_2'(x)(A(x) - A(b)).$$

It follows $1 - V^{*\prime}(u) = \phi(\tilde{b}, u)$, where $\tilde{b}$ is given by Table 1.1. Then by Lemma 1.7.6 $\phi(\tilde{b}, x)$ is negative for any $x \geq 0$ except for x with $\tilde{a} \leq x \leq \tilde{b}$, where $\tilde{a}$ and $\tilde{b}$ are given by Table 1.1. Thus, by Lemma 1.7.6, we have for any $0 \leq x < y$

$$\begin{aligned}\int_x^y (1 - V^{*\prime}(u))du &= \int_x^y \phi(\tilde{b}, u)du \leq \int_{\tilde{a}}^{\tilde{b}} \phi(\tilde{b}, u)du \\ &= \int_{\tilde{a}}^{\tilde{b}} (1 - V^{*\prime}(u))du = \gamma.\end{aligned}$$

II a) We are going to prove $LV^*(x) = 0$ for all $x \in M^{\pi^*}$. We start by describing the range M^{π^*}. By proving the property **I** we showed that no utility function $V(x, \pi)$ can beat V^*. Thus, the initial dividend payment z_0 coincides with the candidate initial dividend payment, or more precisely it equals 0, if $x_0 < \tilde{b}$ and it equals $x_0 - \tilde{a}$ if $x_0 \geq \tilde{b}$. It means that $X_0 = x_0 - z_0 \leq \tilde{b}$, or in other words the underlying best controlled process X starts from below $\tilde{b}$. Moreover, from the definition of π^* we see that once started from below $\tilde{b}$ and being controlled by π^* the process X will never leave the strip $[0, \tilde{b}]$. In other words the range M^{π^*} is contained in $[0, \tilde{b}]$. Now it is easy to see that $LV^*(x) = 0$ for all $x \in [0, \tilde{b}]$ by definition of V^*.

II b) With $X_{T_i^*} = \tilde{a}$ and $X_{T_i-)^*} = \tilde{b}$ we have by construction of $\tilde{a}$ and $\tilde{b}$

$$\int_{T_i^*}^{T_{i^*}-} (1 - V^{*\prime}(u))du = \int_{\tilde{a}}^{\tilde{b}} (1 - V^{*\prime}(u))du = \gamma.$$

This completes the proof. □

Thus we see that the analogues of the Bellman (in)equalities are satisfied, and, consequently, Theorem 1.2.2 is proved. The matter of existence of the optimal solution (the existence of $\tilde{a}$ and $\tilde{b}$, given by Table 1.1) is discussed in Lemma 1.7.4 and in Lemma 1.7.5.

1.5 The case when the dividend process is any nonnegative, nondecreasing, and right continuous process

Let us consider the case when the control process Z is allowed to be any nonnegative, nondecreasing, and right continuous process, adapted to the Brownian filtration. The reserve of the company is given by

$$\begin{cases} dX_t &= \mu dt + dW_t - dZ_t, \\ X_0 &= x_0 - z_0, \end{cases} \tag{1.49}$$

where x_0 is the initial reserve, z_0 is the initial dividend payment, and $Z = (Z_t)_{t\geq 0}$ is the dividend process.

Let S be the liquidation or salvage value and λ the discount rate. We define the value function $\tilde{V}(\cdot)$ by

$$\tilde{V}(x) = \sup_Z \mathbf{E}_x \left\{ \int_{(0,\tau)} e^{-\lambda t} dZ_t + S e^{-\lambda \tau} \right\}, x \geq 0, \tag{1.50}$$

where the supremum is taken over all admissible processes Z, and the expectation is taken with respect to $X_0 = x_0 - z_0 = x$. The initial value of the firm v_0 equals

$$v_0 = z_0 + \tilde{V}(x_0 - z_0) = \sup_{0\leq z\leq x_0} \left\{ z + \tilde{V}(x_0 - z) \right\},$$

with x_0 the initial reserve.

1.5.1 The candidate value function and the candidate optimal control

Suppose we found a candidate value function, say $V^* = V^*(\cdot)$. As in the case of discrete dividend, we leave the optimality check of the candidate value function for the next section, but discuss its properties here.

Definition of the candidate value function and the candidate optimal control

Let the candidate value function $V^* = V^*(\cdot)$ be defined by

$$V^*(x) = \begin{cases} Su_1(x) + A(\bar{x})u_2(x), & x < \bar{x}. \\ x - \bar{x} + \mu/\lambda, & x \geq \bar{x}, \end{cases} \tag{1.51}$$

where $\bar{x}$ is the root of the equation $f(x) = (S \vee \mu/\lambda)$.

The candidate optimal control is a local time process at $\bar{x}$. Let $x \leq \bar{x}$ hold. Consider the solution $(\bar{X}, L) = (\bar{X}_t, L_t)_{t\geq 0}$ to the stochastic differential equation with reflection (see [RevuzYor] Chapter IX, §2, Exercise 2.14)

$$\bar{X}_t = x + \mu t + W_t - L_t, \tag{1.52}$$

where $L = (L_t)_{t\geq 0}$ is a continuous nondecreasing $\mathcal{F}$-adapted process with $L_0 = 0$ such that $L_t = \int_0^t I(\bar{X}_s = \bar{x}) dL_s$, and $\bar{X}_t \leq \bar{x}$.

As is well known a solution $(\bar{x}, L)$ exists, the barrier $\bar{x}$ is a reflecting boundary for the process $\bar{X}$, and the process L is the local time spent by process $\bar{X}$ at the boundary $\bar{x}$.

It can also be noted that the candidate optimal dividend process can be obtained as limit of the case of a bounded dividend rate as $K \to \infty$, and of the case of discrete dividends as $\gamma \to 0$. Indeed, the definitions of $\tilde{x} = \tilde{x}(K)$ and $\bar{x}$ just below formulas (1.20) and (1.51), respectively, and formula (1.79) of Lemma 1.7.1 show that $\tilde{x}(K) \to \bar{x}$ as $K \to \infty$, by (relevant formulas) the values $\tilde{a}$ and $\tilde{b}$ draw together, converging to $\bar{x}$, as γ goes to 0. Thus, it is not surprising that as we will see later the optimal initial dividend payment is $z_0 = (x_0 - \bar{x}) \vee 0$, and that the optimal dividend process $\tilde{Z} = (\tilde{Z}_t)_{t\geq 0}$ corresponds to the process X with a reflecting barrier at $\bar{x}$. The dividends come from the local time the process X spends at the reflecting boundary $\bar{x}$.

For some other optimal reflection problems for linear diffusions see [Alvarez].

Properties of the candidate value function

The following lemma shows the behaviour of $V^{*\prime}(x) - 1$ in the regions $[0, \bar{x})$ and $[\bar{x}, \infty)$.

Lemma 1.5.1 *With $\bar{x}$ being the root of equation $f(x) = S \vee (\mu/\lambda)$, the candidate value function $V^*(\cdot)$ defined by (1.51) is C^1 on $[0, \infty)$, and satisfies the following relations*

$$V^{*\prime}(x) - 1 > 0, \ \ x < \bar{x}, \tag{1.53}$$
$$V^{*\prime}(x) - 1 = 0, \ \ x = \bar{x}, \tag{1.54}$$
$$V^{*\prime}(x) - 1 = 0, \ \ x > \bar{x}. \tag{1.55}$$

Proof. It is obvious that the function V^* is C^1 on $[0, \bar{x})$ and on $(\bar{x}, \infty)$. Let us check continuity and continuous differentiability of V^* at $\bar{x}$. Indeed, $V^*(\bar{x}-) = Su_1(\bar{x}) + A(\bar{x})u_2(\bar{x}) = f(\bar{x}) = \mu/\lambda$. On the other hand $V^*(\bar{x}+) = \bar{x} - \bar{x} + \mu/\lambda = \mu/\lambda$. For the left derivative we have

$$V^{*\prime}(\bar{x}-) = Su_1{}'(\bar{x}) + A(\bar{x})u_2{}'(\bar{x}) = Su_1{}'(\bar{x}) + \frac{1 - Su_1{}'(\bar{x})}{u_2{}'(\bar{x})} u_2{}'(\bar{x}) = 1.$$

For the right derivative, and for all $x \geq \bar{x}$, we get $V^{*\prime}(\bar{x}) = 1$, since $V^*(x) = x - \bar{x} + \mu/\lambda$. Therefore V^* is C^1 on $[0, \infty)$, and in passing we also proved (1.54) and (1.55).

Now let us show (1.53). Since $\bar{x} = 0$ holds in case $S \geq \mu/\lambda$, the interval $[0, \bar{x})$ is nonempty only if $S < \mu/\lambda$. Then, by Corollary 1.7.5 we obtain

$$\begin{aligned} V^{*\prime}(x) - 1 &= S{u_1}'(x) + A(\bar{x}){u_2}'(x) - 1 \\ &= u_2'(x)\left(A(\bar{x}) - A(x)\right) > 0. \end{aligned}$$

□

1.5.2 Firm's initial value optimality

Now we are ready to find explicitly the candidate initial value of the firm. The candidate initial value of the firm is defined as

$$v_0^* = \max_z \left\{ z + V^*(x_0 - z) \right\}, \tag{1.56}$$

where z is the initial dividend payment, $0 \leq z \leq x_0$.

Write $x = x_0 - z$ and $\psi(x) = x_0 - x + V^*(x)$. In order to determine v_0^* in (1.56) we have to maximize $\psi(x)$ over $0 \leq x \leq x_0$. Note that among all possible values at which the maximum is attained we take the minimal possible value. Since $\psi'(x) = V^{*\prime}(x) - 1 = 0$ unless $S < \mu/\lambda$ and $x < \bar{x}$ (by Lemma 1.5.1), the maximum of ψ is attained at (the minimal possible) $\bar{x}$ if $S < \mu/\lambda$, and at 0 if $S \geq \mu/\lambda$. Hence returning to the previous notation we obtain

1. For $S < \mu/\lambda$ and $x_0 < \bar{x}$ we have $z_0^* = 0$ and $v_0^* = V^*(x_0)$.
2. For $S < \mu/\lambda$ and $x_0 \geq \bar{x}$ we have $z_0^* = x_0 - \bar{x}$ and $v_0^* = x_0 - \bar{x} + V^*(\bar{x}) = x_0 - \bar{x} + \mu/\lambda$.
3. For $S \geq \mu/\lambda$ we have $z_0^* = x_0$ and $v_0^* = x_0 + V^*(0) = x_0 + S$.

Thus we have found the initial value of the firm $v_0 = v_0^*$ and the initial dividend payment $z_0 = z_0^*$ once we will have proved that the candidate value function is the best performing candidate, i.e. once we will have proved $V^*(x) \geq V(x, Z)$ for any $x \geq 0$ and any admissible strategy Z; cf. (1.44).

We are going to check the optimality of the candidate value function in this way in the following section.

1.5.3 Value function optimality

Let the process $X = (X_t)_{t \geq 0}$ correspond to a chosen dividend strategy $Z = (Z_t)_{t \geq 0}$, i.e.

$$\left\{ \begin{aligned} dX_t &= \mu dt + dW_t - dZ_t, \\ X_0 &= x_0 - z_0, \end{aligned} \right. \tag{1.57}$$

where $Z = (Z_t)_{t\geq 0}$ is any nonnegative, nondecreasing, and right continuous process with z_0 the initial dividend payment. Note, that if the initial dividend payment coincides with the candidate initial dividend payment, then $z_0 = z_0^*$, where z_0^* is defined as

$$z_0^* = \begin{cases} 0, & x_0 < \bar{x}, \\ x_0 - \bar{x}, & x_0 \geq \bar{x}, \end{cases}$$

with $\bar{x}$ the nonnegative root of equation $f(x) = S \vee (\mu/\lambda)$. Then it follows that $X_0 = x_0 - z_0 \leq \bar{x}$.

Stochastic control verification properties

As in the case with discrete dividend let $M^Z \subset \mathbb{R}_+$ be the range of the process $X = (X_t)_{t\geq 0}$ corresponding to the chosen dividend strategy $Z = (Z_t)_{t\geq 0}$. Suppose the board of directors uses the dividend strategy $Z = (Z_t)_{t\geq 0}$. Let $V(x, Z)$ be the expected total discounted dividend corresponding to the board's strategy added together with the discounted liquidation value received upon bankruptcy with reserve x, $x = x_0 - z_0$

$$V(x, Z) = \mathbf{E}_x \left\{ \int_{(0,\tau)} e^{-\lambda t} dZ_t + S e^{-\lambda \tau} \right\}.$$

Assume we have a candidate $Z^* = (Z_t^*)_{t\geq 0}$ for the optimal strategy and a candidate $V^*(\cdot)$ for the value function. To prove that the strategy Z^* and the function $V^*(\cdot)$ are optimal, it is enough to check the standard stochastic control *verification properties*:

(A_g) For any admissible control $Z = (Z_t)_{t\geq 0}$, i.e. any nonnegative, nondecreasing, cadlag process Z,
$$V(x, Z) \leq V^*(x), \text{ for all } x \in M^Z. \tag{1.58}$$

(B_g) The control Z^* satisfies

$$V(x, Z^*) = V^*(x) \text{ for all } x \in M^{Z^*}. \tag{1.59}$$

Note that the range M^Z coincides with $[0, \infty)$. The property (A_g) says that no other function can be better than the candidate V^* on the range of all possible controlled by Z processes X, i.e. $[0, \infty)$. But the property (B_g) points out a certain control Z^* on which the corresponding utility function $V(x, Z^*)$ coincides with the best performing candidate V^*. Under this control Z^* the range of the corresponding controlled process X could be smaller than the original range $[0, \infty)$. To prove optimality it is enough to consider the range M^{Z^*}. As we will see later, in our case M^{Z^*} will be the strip $[0, \bar{x}]$, where $\bar{x}$ is the solution to $f(x) = S \vee \mu/\lambda$.

Analogue of variational inequalities

Recall the notation

$$L = \mu \frac{d}{dx} + \frac{1}{2}\frac{d^2}{dx^2} - \lambda. \tag{1.60}$$

Assume there exist a C^2 candidate value function $V^*(\cdot)$ from $[0,\infty]$ to $[0,\infty)$ and an adapted candidate control $Z^* = (Z_t^*)_{t\geq 0}$ such that the following analogues of the Bellman inequality (1.28) and the Bellman equation (1.29) hold:

I For all admissible controls $Z = (Z_t)_{t\geq 0}$ and the candidate value function $V^*(\cdot)$ we have

a) $LV^*(x) \leq 0$, for all $x \geq 0$,

b) $\int_{X_{T_i}}^{X_{T_i}-} (1 - V^{*\prime}(y))dy \leq 0$, where $\xi_i = X_{T_i-} - X_{T_i}$, $i = 1, 2, \ldots$.

II The *continuous* candidate optimal control $Z^* = (Z_t^*)_{t\geq 0}$ and the candidate value function $V^* = V^*(\cdot)$ satisfy

a) $LV^*(x) = 0$, for all $x \in M^{Z^*}$,

b) $\int_0^{t\wedge\tau} e^{-\lambda s}\,(1 - V^{*\prime}(X_{s-}))\,dZ_s^* = 0$.

Let us show now that the verification properties are satisfied for such $V^*(\cdot)$ and $Z^* = (Z_t^*)_{t\geq 0}$. Apply Itô's formula to the semimartingale $(e^{-\lambda t}V^*(X_t))_{t\geq 0}$

$$\begin{aligned} e^{-\lambda(t\wedge\tau)}V^*(X_{t\wedge\tau}) &= V^*(X_0) + \int_0^{t\wedge\tau}(-\lambda e^{-\lambda s}V^*(X_s))ds + \int_0^{t\wedge\tau} e^{-\lambda s}V^{*\prime}(X_{s-})dX_s \\ &+\frac{1}{2}\int_0^{t\wedge\tau} e^{-\lambda s}V^{*\prime\prime}(X_s)ds + \sum_{0<s\leq t\wedge\tau} e^{-\lambda s}\left\{V^*(X_s) - V^*(X_{s-}) - V^{*\prime}(X_{s-})\Delta X_s\right\}. \end{aligned}$$

Together with $X_t = x + \mu t + W_t - Z_t$, $x = X_0 = x_0 - z_0$, we get

$$\begin{aligned} V^*(x) &= e^{-\lambda(t\wedge\tau)}V^*(X_{t\wedge\tau}) + \int_0^{t\wedge\tau} e^{-\lambda s}V^{*\prime}(X_{s-})dZ_s - \int_0^{t\wedge\tau} e^{-\lambda s}LV^*(X_s)ds \\ &- \int_0^{t\wedge\tau} e^{-\lambda s}V^{*\prime}(X_s)dW_s - \sum_{0<s\leq t\wedge\tau} e^{-\lambda s}\left\{V^*(X_s) - V^*(X_{s-}) + V^{*\prime}(X_{s-})\Delta X_s\right\}. \end{aligned} \tag{1.61}$$

Taking the expectation of (1.61) conditionally on $X_0 = x$ we obtain

$$\begin{aligned} V^*(x) &= \mathbf{E}_x e^{-\lambda(t\wedge\tau)} V^*(X_{t\wedge\tau}) + \mathbf{E}_x \int_0^{t\wedge\tau} e^{-\lambda s} dZ_s - \mathbf{E}_x \int_0^{t\wedge\tau} e^{-\lambda s}(1 - V^{*\prime}(X_{s-}))dZ_s \\ &- \mathbf{E}_x \int_0^{t\wedge\tau} e^{-\lambda s} LV^*(X_s) ds - \mathbf{E}_x \int_0^{t\wedge\tau} e^{-\lambda s} V^{*\prime}(X_s) dW_s \qquad (1.62) \\ &- \mathbf{E}_x \sum_{0<s\le t\wedge\tau} e^{-\lambda s} \left\{ V^*(X_s) - V^*(X_{s-}) - V^{*\prime}(X_{s-})\Delta X_s \right\}. \end{aligned}$$

The stochastic integral in (1.62) is a martingale, as $V^{*\prime}(x)$ is bounded on $[0, \bar{x}]$ and $V^{*\prime}(x) = 1$ on $[\bar{x}, \infty)$. Therefore its expectation is equal to zero. Also, letting $t \to \infty$ and using the fact $V^*(0) = S$, we get

$$\begin{aligned} V^*(x) &= \mathbf{E}_x e^{-\lambda\tau} S + \mathbf{E}_x \int_0^{\tau} e^{-\lambda s} dZ_s \\ &- \mathbf{E}_x \int_0^{\tau} e^{-\lambda s}(1 - V^{*\prime}(X_{s-}))dZ_s - \mathbf{E}_x \int_0^{t\wedge\tau} e^{-\lambda s} LV^*(X_s) ds \\ &- \mathbf{E}_x \sum_{0<s\le t\wedge\tau} e^{-\lambda s} \left\{ V^*(X_s) - V^*(X_{s-}) - V^{*\prime}(X_{s-})\Delta X_s \right\}. \qquad (1.63) \end{aligned}$$

Note that Lemma 1.5.1 implies

$$V^{*\prime}(x) \ge 1 \qquad (1.64)$$

and that consequently **I b)** yields the following inequality

$$\begin{aligned} V^*(X_s) - V^*(X_{s-}) - V^{*\prime}(X_{s-})\Delta X_s &\le V^*(X_s) - V^*(X_{s-}) - \Delta X_s \\ &= \int_{X_s}^{X_{s-}} \left(1 - V^{*\prime}(u)\right) du \le 0. \qquad (1.65) \end{aligned}$$

Thus, applying (1.65), **I a)**, and (1.64) to 1.63) we get the property (A_g), i.e. $V^*(x) \ge V(x, Z)$. To get the property (B_g) note that under the candidate control Z^* the underlying process X is continuous, and consequently $V^*(X_s) - V^*(X_{s-}) - V^{*\prime}(X_{s-})\Delta X_s = 0$. Thus, by using **II a)** and **II b)** we obtain property (B_g) from (1.63). This completes the proof.

Verifying the analogue of the Bellman inequalities

In the previous section we have proved that if for some C^2 function $V^*(\cdot)$ on $[0, \infty)$ and an adapted control $Z^* = (Z_t^*)_{t\ge 0}$ the analogues of variational inequalities **Ia),b)** and **IIa),b)** from this Section 1.5.3 are satisfied, then $V^*(\cdot)$ and $Z^* = (Z_t^*)_{t\ge 0}$ are the optimal solution to our optimal control problem. Let us check those variational inequalities for the function $V^*(\cdot)$ and the process Z^*.

Lemma 1.5.2 *Let $\bar{x}$ be the solution to $f(x) = S \vee (\mu/\lambda)$ and $V^*(\cdot)$ the candidate value function given by (1.51). Let Z^* be the candidate optimal dividend process, $Z_t^* = Z_0 + L_t$, where $L = (L_t)_{t\geq 0}$ is a continuous nondecreasing $\mathcal{F}$-adapted process with $L_0 = 0$ such that $L_t = \int_0^t I(\bar{X}_s = \bar{x})dL_s$, i.e. L is the local time spent by process $\bar{X}$ at the boundary $\bar{x}$. Then the candidate value function $V^* = V^*(\cdot)$ and the candidate optimal dividend process $Z^* = (Z_t^*)_{t\geq 0}$ satisfy the analogues of the variational inequalities* **I a)***,* **I b)***,* **II a)** *and* **II b)** *from this Section 1.5.3.*

PROOF.

I a)
- Suppose $x < \bar{x}$. It follows from the definition of $V^* = V^*(\cdot)$ and the fact that $u_1(\cdot)$ and $u_2(\cdot)$ are linearly independent solutions to the differential equation $Lu = 0$ (see Section 1.6), that $LV^* = 0$ holds.
- Suppose $x \geq \bar{x}$. Then we have

$$\begin{aligned} LV^*(x) &= \mu - \lambda V^*(x) = \lambda \left(\frac{\mu}{\lambda} - V^*(x)\right) \\ &\leq \lambda \left(\frac{\mu}{\lambda} - V^*(\bar{x})\right) = 0. \end{aligned}$$

I b) This property follows from Lemma 1.5.1.

II a) We are going to prove $LV^*(x) = 0$ for all $x \in M^{Z^*}$. Let us start by describing the range M^{Z^*}. By proving the property **I** we showed that no utility function $V(x, Z)$ can beat V^*, i.e. V^* is the best performing candidate. Thus, as we discussed in Section 1.5.2, we can figure out the initial dividend payment. The initial dividend payment z_0 is 0 if $x_0 < \bar{x}$, and $x_0 - \bar{x}$ if $x_0 \geq \bar{x}$. It means that $X_0 = x_0 - z_0 \leq \bar{x}$, or in other words the underlying best controlled process X starts from below $\bar{x}$. Moreover, from the definition of Z^* we see that once started from below $\bar{x}$ and being controlled by Z^* the process X will never leave the strip $[0, \bar{x}]$, but will reflect at the boundary $\bar{x}$. In other words the range M^{Z^*} coincides with $[0, \bar{x}]$. Now it is easy to see that $LV^*(x) = 0$ for all $x \in [0, \bar{x}]$ by definition of V^*.

II b) The candidate optimal control $Z^* = (Z_t^*)_{t\geq 0}$, defined as $Z_t^* = z_0 + L_t$ with $L_0 = 0$ and $L_t = \int_0^t I(\bar{X}_s = \bar{x})dL_s$, is continuous by construction. Moreover, $V^{*\prime}(y) = 1$ at $y = \bar{x}$ by direct calculation. Thus it follows that $\int_0^{t\wedge\tau} e^{-\lambda s}\left(1 - V^{*\prime}(X_{s-})\right) dZ_s^* = 0$. □

Thus, we see that the analogues of the Bellman (in)equalities are satisfied, and, consequently, Theorem 1.2.3 is proved. The issue of existence of the optimal solution, i.e. the existence of $\bar{x}$, is discussed in Corollary 1.7.3.

1.6 Appendix A. Notation reference

In this section for reference purposes we present the notation we used in this chapter. First recall that we were given parameters $\lambda > 0, K > 0,\ \mu \geq 0, \gamma \geq 0,\ S \in \mathbb{R}$. Now let us define the following constants and functions

$$
\begin{aligned}
r_1 &= -\mu + \sqrt{\mu^2 + 2\lambda} > 0, && (1.66)\\
r_2 &= \mu + \sqrt{\mu^2 + 2\lambda} > 0, && (1.67)\\
\rho &= -(K - \mu) + \sqrt{(K - \mu)^2 + 2\lambda}, && (1.68)\\
Q &= \frac{K}{\lambda} - \frac{1}{\rho}, && (1.69)\\
u_1(x) &= e^{-r_2 x}, && (1.70)\\
u_2(x) &= e^{r_1 x} - e^{-r_2 x}, && (1.71)\\
v(x) &= e^{-\rho x} && (1.72)\\
w(x) &= u_1(x)u_2'(x) - u_1'(x)u_2(x), && (1.73)\\
A(x) &= \frac{1 - Su_1'(x)}{u_2'(x)}, && (1.74)\\
f(x) &= S\frac{w(x)}{u_2'(x)} + \frac{u_2(x)}{u_2'(x)}, && (1.75)\\
\phi(b, x) &= u_2'(x)(A(x) - A(b)). && (1.76)
\end{aligned}
$$

A straightforward computation gives

$$
\begin{aligned}
f(x) &= Su_1(x) + A(x)u_2(x),\\
\frac{r_2 - r_1}{r_1 r_2} &= \frac{1}{r_1} - \frac{1}{r_2} = \frac{\mu}{\lambda}.
\end{aligned}
$$

Note 1.6.1 *The functions $u_1(\cdot)$ and $u_2(\cdot)$ defined by (1.70) and by (1.71) are linearly independent solutions of the differential equation $\mu u' + \frac{1}{2}u'' - \lambda u = 0$.*

The rest of the notation is created under assumption $S < \mu/\lambda$, and presented in Table 1.2

1.7 Appendix B. Technical lemmas

In this section we are going to study the properties of the functions we used for the representations of the results.

$\bar{x}$	constant	the positive root of equation $f(x) = \mu/\lambda$
$a = a(\cdot)$	function	$a = a(b)$, $b \in [\bar{x}, \infty)$ such that $A(b) = A(a(b))$
a_{min}	constant	$a_{min} = \begin{cases} 0, & S \geq -\frac{1}{r_2} \\ \frac{ln(-Sr_2)}{r_2}, & S < -\frac{1}{r_2} \end{cases}$
b_{max}	constant	$b_{max} \geq \bar{x}$, $a(b_{max}) = a_{min}$
$\Lambda = \Lambda(\cdot)$	function	$\Lambda(b) = \int_{a(b)}^{b} u_2'(y)\,(A(y) - A(b))\,dy$, $b \geq \bar{x}$,
$\tilde{b}$	constant	the root of $\Lambda(b) = \gamma$, $b \geq \bar{x}$
$\tilde{a}$	constant	$\tilde{a} = a(\tilde{b})$
γ_{max}	constant	$\gamma_{max} = \Lambda(b_{max})$

Table 1.2: The notation created under assumption $S < \mu/\lambda$.

1.7.1 $Q = Q(K)$ as a function of K

Recall that $Q = K/\lambda - 1/\rho$ where $\rho = -(K-\lambda) + \sqrt{(K-\lambda)^2 + 2\lambda}$, see (1.68) and (1.69). See Fig. 1.3 for a plot of Q as a function of K.

Lemma 1.7.1 *Viewed as a function of K, the expression Q defined by (1.9) is denoted by $Q(K)$ and is strictly increasing. Moreover, we have*

$$Q(0) = -\frac{1}{r_2} < 0, \tag{1.77}$$

$$Q(\frac{\lambda}{2\mu}) = 0, \tag{1.78}$$

$$Q(\infty) = \frac{\mu}{\lambda} > 0. \tag{1.79}$$

Proof. Straightforward computation shows (1.77) and (1.78). Rewrite

$$Q(K) = (2\lambda)^{-1}\left(K + \mu - \sqrt{(K-\mu)^2 + 2\lambda}\right),$$

and note that the derivative

$$Q'(K) = (2\lambda)^{-1}\left(1 - \frac{K-\mu}{\sqrt{(K-\mu)^2 + 2\lambda}}\right)$$

is positive. This proves that $K \mapsto Q(K)$ is strictly increasing. Finally note

$$\lim_{K\to\infty} Q(K) = (2\lambda)^{-1} \lim_{K\to\infty}\left(K + \mu - (K-\mu) + O(\frac{1}{K})\right) = \frac{\mu}{\lambda}.\triangle$$

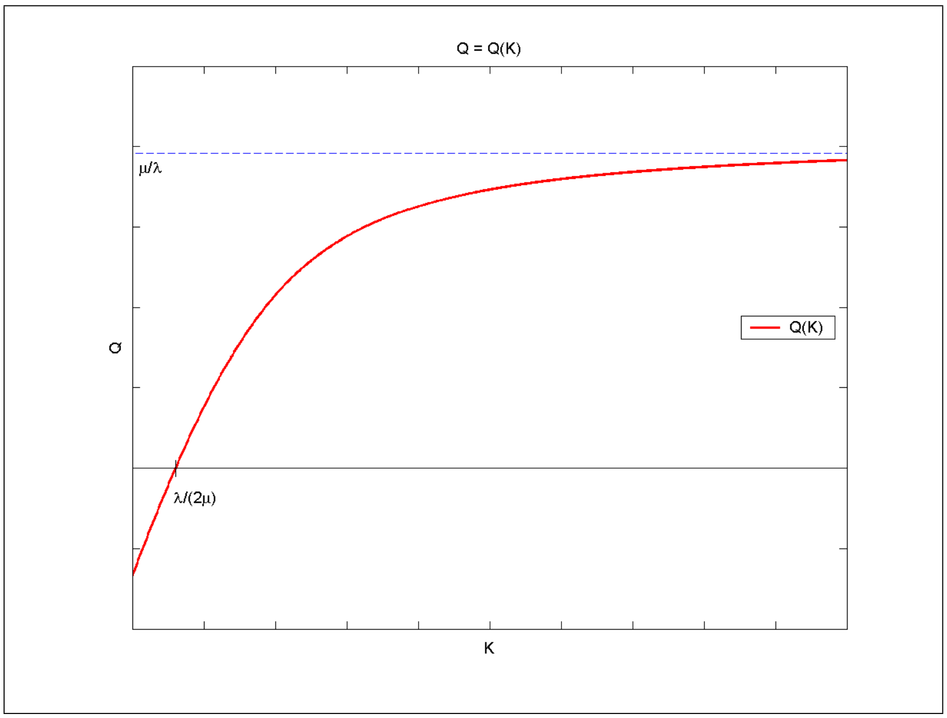

Figure 1.3: The graph of $Q(K)$.

1.7.2 The functions $f = f(\cdot)$ and $A = A(\cdot)$

Here we study the function $f = f(x)$. See Fig. 1.4 for the case $0 < S < \mu/\lambda$.

Lemma 1.7.2 *The following properties hold for $f = f(\cdot)$*

1 $f(0) = S$,

2 $f(x)$ *is continuous,*

3 $f(\infty) = 1/r_1$,

4 *If $S > 0$, then $f(x)$ is a strictly increasing-decreasing function on $[0, \infty)$ and the maximum is attained at the unique positive root of $Su_2(x) = (r_1 + r_2)/(r_1 r_2)$. If $S \le 0$, then $f(x)$ is strictly increasing on $[0, \infty)$.*

Proof. Recall

$$\begin{aligned} f(x) &= S\frac{w(x)}{u_2'(x)} + \frac{u_2(x)}{u_2'(x)} \\ &= S\frac{(r_1+r_2)e^{(r_1-r_2)x}}{r_1 e^{r_1 x} + r_2 e^{-r_2 x}} + \frac{e^{r_1 x} - e^{-r_2 x}}{r_1 e^{r_1 x} + r_2 e^{-r_2 x}}. \end{aligned}$$

Properties **1-3** can be easily checked by direct calculation. To prove property **4**, let us write down a closed form expression for the derivative

$$f'(x) = r_1 r_2 \frac{(r_1+r_2)e^{(r_1-r_2)x}}{(r_1 e^{r_1 x} + r_2 e^{-r_2 x})^2}\left(\frac{r_1+r_2}{r_1 r_2} - Su_2(x)\right).$$

Thus, since $u_2(x)$ is strictly increasing function from $[0, \infty)$ onto $[0, \infty)$, we obtain the desired property. □

Corollary 1.7.1 *Suppose $S < \mu/\lambda$. Then $f(y) > f(\bar{x})$ for $y > \bar{x}$.*

Proof. This is straightforward in the case $S \le 0$, as $f(\cdot)$ is a strictly increasing function then. Suppose now $0 < S < \mu/\lambda$.Note that $f(\bar{x}) = \mu/\lambda < 1/r_1 = f(\infty)$ holds by Table 1.2, some computation, and Lemma 1.7.2. Since $f(\cdot)$ is a continuous increasing-decreasing function this implies $f(y) > f(\bar{x})$ for $y > \bar{x}$. □

Corollary 1.7.2 *The equation $f(x) = Q$ has a positive solution if and only if $S < Q$ holds.*

Proof.

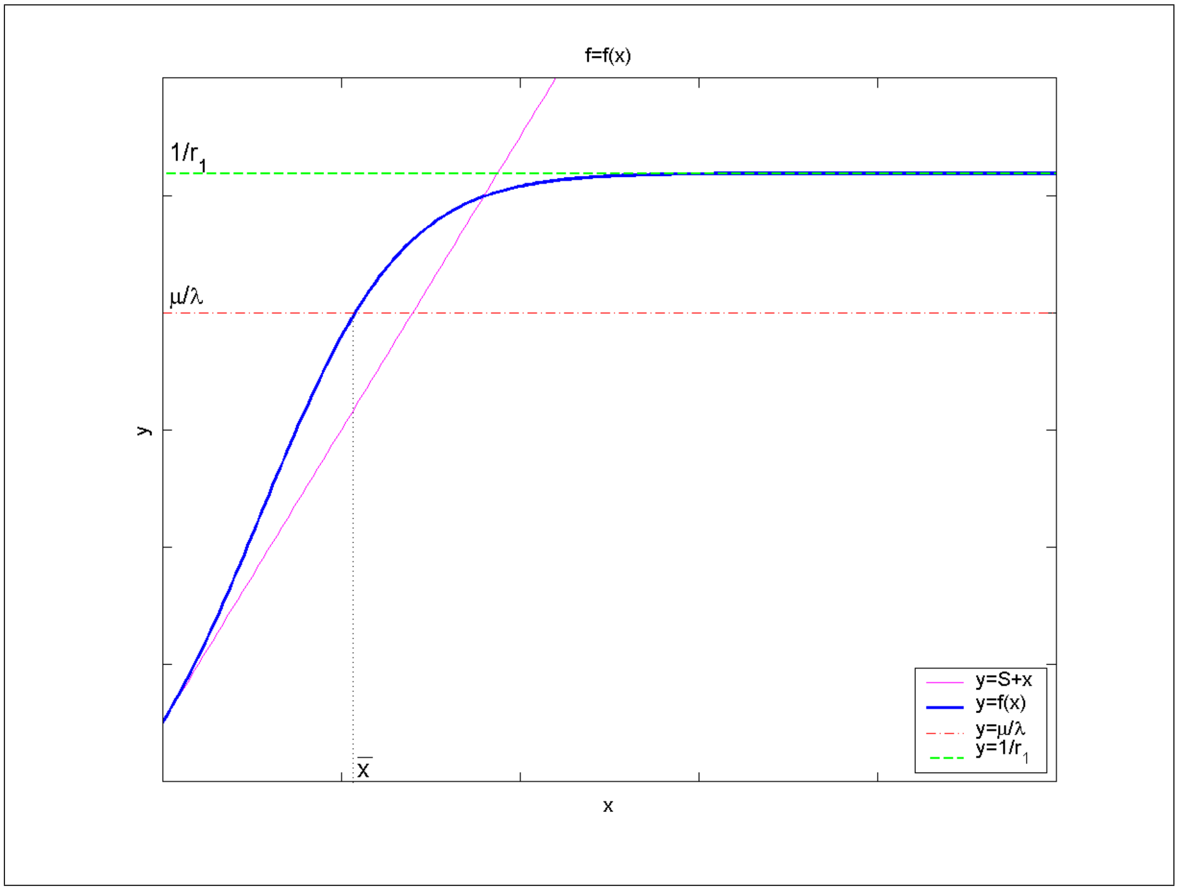

Figure 1.4: Function $f = f(x)$ in the case of reasonable liquidation value $0 < S < \mu/\lambda$.

- Suppose $S < Q$. By Lemma 1.7.2 and Lemma 1.7.1 we have

$$f(0) = S < Q < \frac{\mu}{\lambda} = \frac{r_2 - r_1}{r_1 r_2} < \frac{1}{r_1} = f(\infty).$$

  Since f is increasing-decreasing or strictly increasing on $[0, \infty)$, by Lemma 1.7.2, this implies that $f(x) = Q$ has a unique positive solution.

- Suppose $S \geq Q$. Since f is strictly increasing-decreasing on $[0, \infty)$ (see Lemma 1.7.2) with $f(0) = S \geq Q$ and $f(\infty) = 1/r_1 > \mu/\lambda = Q(\infty) \geq Q$ (see Lemma 1.7.1), it cannot take the value Q at a point $x > 0$. □

Corollary 1.7.3 *The equation $f(x) = \mu/\lambda$ has a positive solution if and only if $S < \mu/\lambda$ holds.*

Proof. The proof is analogous to the proof of Corollary 1.7.2. □

Recall now that by $\bar{x}$ we denote the positive root of equation $f(x) = S \vee (\mu/\lambda)$, which exists and differ from zero if and only if $S < \mu/\lambda$.

Lemma 1.7.3 *Suppose $S < \mu/\lambda$. The following statements are equivalent*

1 *$A(x)$ has a stationary point $\bar{x}$.*

2 *$\bar{x}$ satisfies $f(x) = \mu/\lambda$.*

3 *$\bar{x}$ satisfies $f'(x) = 1$.*

Proof. First observe the following equalities, which are easy to check by direct calculation

$$u_1'(x)u_2''(x) - u_1''(x)u_2'(x) = -r_1 r_2 w(x), \tag{1.80}$$

$$u_2'' = r_1 r_2 u_2 - (r_2 - r_1)u_2'(x). \tag{1.81}$$

Now let us check the equivalence of the statements:

$1 \Leftrightarrow 2$ To prove this we take the derivative of $A(x)$

$$\begin{aligned}
A'(x) &= \frac{-Su_1''(x)u_2'(x) - (1 - Su_1'(x))u_2''(x)}{(u_2'(x))^2} \\
&= \frac{S(u_1'(x)u_2''(x) - u_1''(x)u_2'(x)) - u_2''(x)}{(u_2'(x))^2} \\
&\overset{(1.80),(1.81)}{=} \frac{-r_1 r_2 S w(x) - r_1 r_2 u_2(x) + (r_2 - r_1)u_2'(x)}{(u_2'(x))^2} \\
&= \frac{r_1 r_2}{u_2'(x)} \left(\frac{r_2 - r_1}{r_1 r_2} - \frac{Sw(x) + u_2(x)}{u_2'(x)} \right) \\
&= \frac{r_1 r_2}{u_2'(x)} \left(\frac{\mu}{\lambda} - f(x) \right).
\end{aligned}$$

$3 \Leftrightarrow 1$

$$\begin{aligned} f'(x) &= (Su_1(x) + A(x)u_2(x))' \\ &= Su_1'(x) + A'(x)u_2(x) + A(x)u_2'(x) \\ &= Su_1'(x) + A'(x)u_2(x) + 1 - Su_1'(x) \\ &= 1 + A'(x)u_2(x). \end{aligned}$$

□

Together, Lemma 1.7.2 and 1.7.3 have the following consequences.

Corollary 1.7.4 *If $S < \mu/\lambda$ then $A'(x) > 0$ for $x < \bar{x}$, and $A'(x) < 0$ for $x > \bar{x}$. If $S \geq \mu/\lambda$ then $A'(x) < 0$ for $x > 0$.*

Corollary 1.7.5 *If $S < \mu/\lambda$, $A(\cdot)$ is strictly increasing on $[0, \bar{x}]$, and strictly decreasing on $[\bar{x}, \infty)$. If $S \geq \mu/\lambda$ the function $A(\cdot)$ is strictly decreasing on $[0, \infty)$.*

Corollary 1.7.6 *If $S < \mu/\lambda$ then $f'(x) > 1$ for $0 < x < \bar{x}$, and $f'(x) < 1$ for $x > \bar{x}$, $f'(\bar{x}) = 1$. If $S \geq \mu/\lambda$ then $f'(x) < 1$ for $x > 0$.*

Corollary 1.7.7 *For each b, $\bar{x} \leq b < b_{max}$, there exists a unique a, $a_{min} < a \leq \bar{x}$, such that $A(a) = A(b)$ if and only if $S < \mu/\lambda$.*

Lemma 1.7.4 *Let x_s be a positive solution to $f(x) = S + x - \gamma$. Then this solution is unique, and moreover if $S < \mu/\lambda$ then $x_s > \bar{x}$, where $\bar{x}$ solves $f(x) = \mu/\lambda$.*

Proof. In view of $f(0) = S$, Corollary 1.7.6 shows that the function $\psi(x) = f(x) - S - x$, $x \geq 0$, is increasing on $[0, \bar{x}]$ and decreasing on $[\bar{x}, \infty)$ if $S < \mu/\lambda$, and is decreasing on $[0, \infty)$ if $S \geq \mu/\lambda$. Moreover, it satisfies $\psi(0) = 0$, $\psi(\infty) = -\infty$. Consequently, $\psi(x) = -\gamma$ has a unique solution x_s, and if $S < \mu/\lambda$ then $x_s > \bar{x}$ for $\gamma > 0$. □

The behaviour of $A = A(x)$ as a function of x when $-1/r_2 < S < \mu/\lambda$, is shown in Fig.1.5.

1.7.3 The functions $a = a(\cdot)$ and $\Lambda = \Lambda(\cdot)$

Suppose $S < \mu/\lambda$. In this case the function $A = A(\cdot)$ is increasing-decreasing with a positive maximum in $\bar{x} > 0$, and there exists a maximal interval $(a_{min}, b_{max}) \subset [0, \infty)$ so that $A(a_{min}) = A(b_{max})$. Indeed $A(0) = (1 + r_2 S)/(r_1 + r_2)$ and $A(\infty) = 0$ gives

$$a_{min} = \begin{cases} 0, & -1/r_2 < S < \mu/\lambda \\ ln(-Sr_2)/r_2, & S \leq -1/r_2. \end{cases}$$

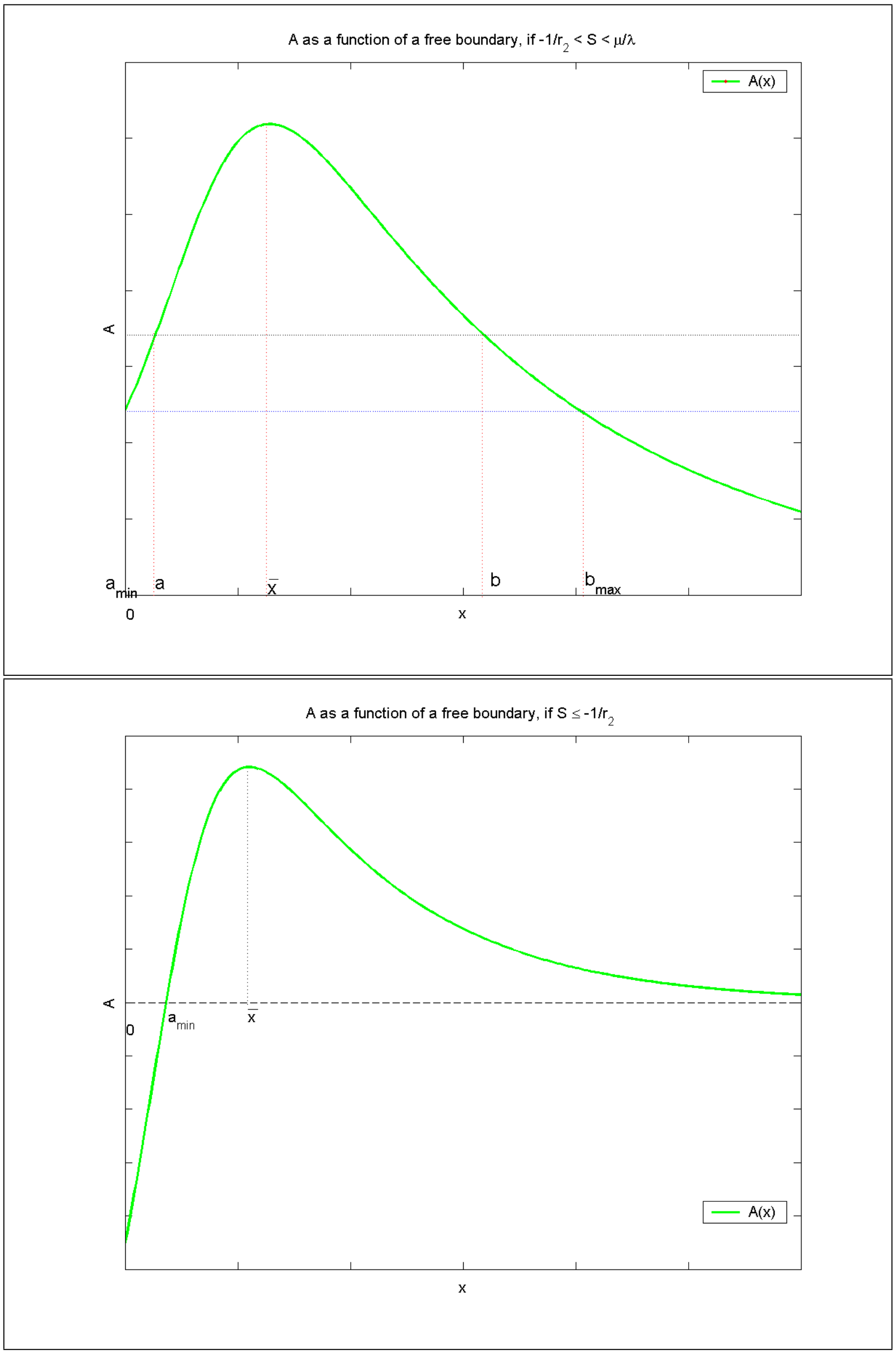

Figure 1.5: The graphs of $A = A(x)$ if $-1/r_2 < S < \mu/\lambda$ and if $S \leq -1/r_2$.

The constant b_{max}, $b_{max} \geq \bar{x}$, can be obtained from the expression $A(a_{min}) = A(b_{max})$. Note that in the case $S \leq -1/r_2$ the value of b_{max} equals infinity.

For $S < \mu/\lambda$ there is a continuous decreasing function $b \mapsto a(b)$ mapping $b \in [\bar{x}, b_{max})$ into $a = a(b) \in (a_{min}, \bar{x}]$ so that

$$A(a) = A(b),\ a = a(b).$$

For $S < \mu/\lambda$ by $\Lambda = \Lambda(\cdot)$ we denote a map from $[\bar{x}, b_{max})$ to $[0, \infty)$ given by the formula

$$\Lambda(b) = \int_{a(b)}^{b} u_2'(y)\,(A(y) - A(b))\,dy.$$

Proposition 1.7.1 *Suppose $S < \mu/\lambda$. The function $b \mapsto \Lambda(b)$ from $[\bar{x}, b_{max})$ to $[0, \infty)$ is continuous and strictly increasing, and $\Lambda(\bar{x}) = 0$. If $b_{max} = \infty$ then $\Lambda = \Lambda(\cdot)$ is unbounded.*

PROOF. It is easy to see from the definition that $\Lambda = \Lambda(\cdot)$ is a continuous function, and $\Lambda(\bar{x}) = 0$ as $a(\bar{x}) = \bar{x}$.

Moreover $u_2'(y)\,(A(y) - A(b)) \nearrow$ as $b \nearrow$ from $\bar{x}$ to b_{max}, because $A(b) \searrow$ as $b \nearrow$ from $\bar{x}$ to b_{max}. Also $(b - a(b)) \nearrow$ as $b \nearrow$ from $\bar{x}$ to b_{max}. Therefore, we conclude that $\Lambda = \Lambda(\cdot)$ is a strictly increasing function.

Let us show that in the case $S \leq -1/r_2$ the function $\Lambda(\cdot)$ is unbounded on $[\bar{x}, b_{max})$. Indeed, here $b_{max} = \infty$, and consequently $A(b_{max}) = 0$. Thus we have for $\Lambda(b_{max})$

$$\begin{aligned} \Lambda(b_{max}) &= \int_{a_{min}}^{b_{max}} u_2'(y)\,(A(y) - A(b_{max}))\,dy = \int_{a_{min}}^{\infty} u_2'(y)A(y)dy \\ &= \int_{a_{min}}^{\infty} (1 - Su_1'(x))\,dx = \int_{a_{min}}^{\infty} \left(1 + Sr_2 e^{-r_2 x}\right)dx = \infty. \end{aligned}$$

$\square$

Lemma 1.7.5 *Suppose $S < \mu/\lambda$. The equation $\Lambda(b) = \gamma$ has a unique solution $\tilde{b}$, $\tilde{b} \geq \bar{x}$, if and only if $\gamma < \gamma_{max}$ or $S \leq -1/r_2$.*

PROOF. By the results of Proposition 1.7.1 it follows that for any $\gamma > 0$ starting from $b = \bar{x}$ and continuously increasing b from $\bar{x}$ to ∞, we find the unique values $\tilde{b}$ and $a(\tilde{b})$ such that $\Lambda(\tilde{b}) = \gamma$. In case $-1/r_2 < S < \mu/\lambda$ while we are continuously increasing b from $\bar{x}$ to b_{max} (in this case b_{max} is finite) we might reach "the maximal possible γ":

$$\gamma_{max} = \Lambda(b_{max}) = b_{max} + S + \frac{(1 - Sr_1)e^{-r_2 b_{max}} - (1 + Sr_2)e^{r_1 b_{max}}}{(r_1 + r_2)}.$$

Thus the condition $\gamma < \gamma_{max}$ must be satisfied for the optimal boundaries pair $(\tilde{a}, \tilde{b})$, $0 < \tilde{a} < \bar{x} < \tilde{b}$, to exist. This concludes the proof. $\square$

1.7.4 The behaviour of $\phi(b,\cdot)$.

Recall $\phi(b,x)$ was defined by (1.76) as

$$\phi(b,x) = u_2'(x)(A(x) - A(b)).$$

We have the following results.

Lemma 1.7.6 *The function $\phi(b,x)$, defined by (1.76), has the following properties*

- *Suppose $S < \mu/\lambda$ and $\bar{x} < b < b_{max}$. Then*

$$\begin{cases} \phi(b,x) < 0, & 0 < x < a(b), \\ \phi(b,x) > 0, & a(b) < x < b, \\ \phi(b,x) < 0, & b < x. \end{cases}$$

- *Suppose $S < \mu/\lambda$ and $b \geq b_{max}$. Then*

$$\begin{cases} \phi(b,x) > 0, & x < b, \\ \phi(b,x) < 0, & x > b. \end{cases}$$

- *Suppose $S \geq \mu/\lambda$. Then for any $y > 0$ we have*

$$\begin{cases} \phi(y,x) > 0, & 0 < x < y, \\ \phi(y,x) < 0, & x > y. \end{cases}$$

Proof. Observe that u_2' is strictly positive. Hence $\phi(b,x)$ has the sign of $A(x) - A(b)$.

- Suppose $S < \mu/\lambda$ and $\bar{x} < b < b_{max}$. Then the results follow from $A(x) < A(a) < A(y) > A(b) > A(z)$ for $0 < x < a < y < b < z$.
- Suppose $S < \mu/\lambda$ and $b \geq b_{max}$. The desired follows from $A(x) > A(b)$ for $x < b$ and $A(x) < A(b)$ for $x > b$.
- Suppose $S \geq \mu/\lambda$. The function $A(x)$ is strictly decreasing in this case. The inequalities follow. $\square$

Chapter 2

Investment optimization model

In this chapter we consider a model for investment optimization. We treat three cases: the case of technical uncertainty only, the case of input cost uncertainty only, and the case where both uncertainties are present.

In introduction Section 2.1, we formulate the problem. In Section 2.2 we show how we have found the solution. In Section 2.3 we prove optimality by checking the verification properties. In Sections 2.4.1, 2.4.2, 2.4.3, and 2.4 we include some well-known properties of the special functions which we use.

2.1 Introduction, formulation of the problem

We consider the following mathematical model which was proposed by Pindyck in [Pind1]. Let the nonnegative random variable X_t be the predicted cost remaining at time t to complete a project (for example to build a nuclear power plant). We split the uncertainty about the predicted remaining investment into two components: *a technical uncertainty* that depends only on the firm's strategy, and *an input cost uncertainty* that depends on external circumstances.

Let us assume that the stochastic process $X = (X_t)_{t\geq 0}$ satisfies the following stochastic differential equation,

$$dX_t = -I_t dt + \beta\sqrt{X_t I_t} dW_t + \gamma X_t d\tilde{W}_t, \tag{2.1}$$

where I_t is a nonanticipating nonnegative function (investment rate), β and γ are known nonnegative constants, and $W = (W_t)_{t\geq 0}$ and $\tilde{W} = (\tilde{W}_t)_{t\geq 0}$ are uncorrelated Wiener processes.

Equation (2.1) shows that the amount needed to complete the project decreases as investment proceeds. Simultaneously this amount is affected by two different types of uncertainty.

A. The case $\beta = 0, \gamma > 0$ corresponds to *input cost uncertainty* governed by the Wiener process $\tilde{W}$. The fluctuations are proportional to the remaining predicted cost X_t. Examples are fluctuations in cost of labor and material, and in government regulations. These can influence X irrespectively of what the firm does. Here X can fluctuate even when there is no investment. The conditional expectation of the quadratic variation $\mathsf{E}\left((dX)^2 \mid X\right) = \gamma^2 X^2 dt$ does not depend on the investment rate I.

B. The case $\beta > 0, \gamma = 0$ corresponds to *the technical uncertainty* governed by the Wiener process W. Here X can fluctuate only if investments are taking place. The conditional expectation of the quadratic variation $\mathsf{E}\left((dX)^2 \mid X\right) = \beta^2 IX dt$ is linear in I.

C. The case $\beta > 0, \gamma > 0$ allows for *both types of uncertainties.*

In all three cases the total amount of realized investments is known only after completion of the project. The parameters β and γ are chosen in accordance with the given applied problem (see [Pind1]).

The rate of investment is the control in our problem. Our purpose is to find an optimal investment strategy for the project. We assume that the rate of investment I_t is a nonanticipating, nonnegative, bounded random function, $0 \leq I_t \leq I_{max}$, where I_{max} is a constant. A control that satisfies these conditions is called *an admissible control.* Let us denote the class of admissible controls by $\mathbf{U}$. Nonnegativity of the investment rate means that we cannot take back any part of our money once we have invested it. The maximal value of the investment rate I_{max} is specified by two factors. The first one is that an investor has limited liquidity and can invest his money only at a bounded rate. For example, a company cannot invest in a project more than 1 million $ a year. Besides that there can be external limitations on the investment rate. For example, construction works cannot be completed within certain time limits, or research on a new drug cannot be speeded up by additional investments.

Let $F(x, I)$ denote *a utility functional,*

$$F(x, I) = \mathsf{E}_x\{-\int_0^\tau I e^{-rt} dt + V e^{-r\tau}\}, \tag{2.2}$$

where E_x is the conditional expectation given the initial value $X_0 = x$, and where $\tau = \inf\{t : X_t = 0\}$ denotes the point in time of completion of the project. Furthermore, the positive constant V is *the value of the project upon completion*, which is known in advance, and the nonnegative constant r is the investment rate.

The utility functional measures the discounted "expected gain from the project", i.e. the expectation of the value minus the investments. The integral in (2.2) denotes the discounted investment in the project, and the second summand is the discounted gain from the project.

Let us introduce *a value function*

$$F(x) = \sup_{I \in \mathbf{U}} F(x, I), \ \ x \geq 0. \tag{2.3}$$

We consider $F(x)$ as a criterion for investment optimality. Note, that if there is no restriction on the rate of investment (i.e. $I_{max} = \infty$), then the optimal strategy is to invest immediately all the capital needed to complete the project. ("If one invests then at the maximal possible rate.")

Furthermore, note the scaling property $F(x, I, V) = I_{max} F\left(x/I_{max}, I/I_{max}, V/I_{max}\right)$, in an obvious notation. Thus we can put $I_{max} = 1$ without loss of generality.

In this chapter we want to find an optimal investment strategy $\tilde{I} = \tilde{I}(x)$ and the value function $F = F(x)$. Generally, problems of this type (see [JeanShir], [KramMor]) are solved in the following way. Using heuristic arguments one finds a strategy that is suspected to be optimal. Subsequently one computes the corresponding value of the utility functional. Finally, one proves that the strategy is optimal and that the function constructed is indeed the value function. We will prove that the following strategy is optimal: *if the capital needed to complete the project is less than a certain value x^*, then we should invest at the maximal rate; if at any moment the capital needed to complete the project exceeds x^*, we should stop investing in the project.*

Pindyck [Pind1] has found the solution for the case of technical uncertainty when the interest rate is equal to zero ($r = 0$). It is interesting to note that in the case of input cost uncertainty and in the presence of both input cost and technical uncertainty the optimal strategy will be *to not invest at all*, but to wait while the project is completed "by itself"! Indeed, here time is "free" ($r = 0$) and the probability that the process (2.1) with $I_t \equiv 0$ hits zero in finite time (the probability of completing the project in finite time) is equal to 1. We obtain an explicit solution for the general problem when $r > 0$. For more information on the economic background of this model we refer to Pindyck [Pind1].

2.2 Finding the solution

2.2.1 Formulating the Stefan problem with a free boundary

The value function $F = F(x)$ is the criterion of optimality in our model:

$$F(x) = \sup_{I \in \mathbf{U}} \mathsf{E}_x \{ \int_0^\tau (-I) e^{-rt} dt + V e^{-r\tau} \}. \tag{2.4}$$

Recall that $\mathbf{U} = \{I : 0 \leq I_t \leq 1\}$ is the class of admissible controls, and $\tau = \inf\{t : X_t = 0\}$ is the time at which the project is finished. One can easily see from (2.4) that $F(0) = V$ and $F(x) \leq V$ for all $x \geq 0$.

Let us introduce operators L_1 and L_2 acting on functions $G = G(x), x \geq 0$, from $C^2((0,\infty)\backslash\{x^*\}) \bigcup C^1(0,\infty)$, (where $x^* \in (0,\infty)$) according to the formula

$$\begin{aligned} L_1 G &= -\frac{dG}{dx} + \frac{1}{2}\beta^2 x \frac{d^2G}{dx^2}, \\ L_2 G &= \frac{1}{2}\gamma^2 x^2 \frac{d^2 G}{dx^2} - rG. \end{aligned} \tag{2.5}$$

Let us assume that there exists a function $\tilde{F} = \tilde{F}(x)$ and a control $\tilde{I} = \tilde{I}(x)$ such that the Bellman equation is satisfied.

Let us write the Bellman equation (see [FlemSoner] for the reference how to obtain the Bellman equation)

$$\sup_{0 \le I \le 1} \{I(L_1\tilde{F}(x) - 1) + L_2\tilde{F}(x)\} = 0. \tag{2.6}$$

As one can see, equation (2.6) is linear in the control I. Therefore, we may assume that the optimal control is the following:

$$\tilde{I}(x) \quad = \quad \begin{cases} 1, & L_1\tilde{F}(x) - 1 \ge 0 \\ 0, & L_1\tilde{F}(x) - 1 < 0 \end{cases} . \tag{2.7}$$

In this way we obtain that for those x, where $\tilde{I}(x) = 1$, the following equation is satisfied

$$L_1\tilde{F}(x) + L_2\tilde{F}(x) - 1 = 0, \tag{2.8}$$

and for those x, where $\tilde{I}(x) = 0$, the following equation holds

$$L_2\tilde{F}(x) = 0. \tag{2.9}$$

Taking into account intuitive considerations on the structure of the optimal control, we assume that there exists a constant x^*, such that $\tilde{I}(x) = 1$ for $x < x^*$, and $\tilde{I}(x) = 0$ for $x \ge x^*$ (i.e. we must invest at maximal rate if the cost of the project is "reasonable", and abandon the project when it becomes too "expensive"). In other words we have to solve the following free boundary Stefan problem: find a value x^* and a smooth bounded function $\tilde{F}(\cdot)$ such that

$$\begin{aligned} \tilde{F}(0) &= V, && (2.10)\\ L_1\tilde{F}(x) + L_2\tilde{F}(x) - 1 &= 0,\ 0 \le x \le x^*, && (2.11)\\ L_2\tilde{F}(x) &= 0,\ x \ge x^*. && (2.12) \end{aligned}$$

2.2.2 Solution to the Stefan problem

One can easily check a bounded solution for the Stefan problem, (2.10)–(2.12), satisfies

$$\tilde{F}(x) = \frac{x}{(b-1)(\frac{1}{2}\beta^2 b + 1)} \left(\frac{x}{x^*}\right)^{-b},\ x \ge x^*,\ \text{where } b = \frac{1}{2}\left(1 + \sqrt{1 + \frac{8r}{\gamma^2}}\right). \tag{2.13}$$

Indeed, the general solution of (2.12) is $c_1 x^{1-b} + c_2 x^b$. In order for the solution to be bounded we must put $c_2 = 0$. We can obtain c_1 from (2.11) with $x = x^*$. Note that the case $\gamma = 0$ may be considered as a limit case by letting $\gamma \downarrow 0$, resulting in $\tilde{F}(x) = 0$, $x \ge x^*$.

Consider the following differential equation:

$$L_1 u(x) + L_2 u(x) = 0. \tag{2.14}$$

Suppose $u_1(x)$ and $u_2(x)$ are linearly independent solutions of (2.14), such that $u_1(0) = 1$, $u_2(0) = 0$. The following condition is necessary and sufficient for $\tilde{F}(x)$ with $0 \le x \le x^*$ to be a solution to the Stefan problem (2.10)-(2.12) on $[0, x^*]$

$$\tilde{F}(x) = \left(V + \frac{1}{r}\right)[u_1(x) - \Theta(x^*)u_2(x)] - \frac{1}{r},\ 0 \le x \le x^*, \tag{2.15}$$

where $\Theta(x)$ and x^* are obtained from the following conditions

$$\begin{aligned} \tilde{F}'(x)\,|_{x=x^*} &= -\frac{1}{\frac{1}{2}\beta^2 b + 1} \text{ (the "smooth pasting" condition)}, && (2.16)\\ L_2\tilde{F}(x)\,|_{x=x^*} &= 0. && (2.17) \end{aligned}$$

Substituting (2.15) in (2.16) and (2.17) we get

$$\Theta(x) = \frac{L_2u_1(x) - \left(\frac{1}{2}\beta^2 b + 1\right)u_1'}{L_2u_2(x) - \left(\frac{1}{2}\beta^2 b + 1\right)u_2'},\ 0 \le x \le x^* \tag{2.18}$$

Note that using (2.14) and (2.5), one can write

$$\begin{aligned} L_2u_i(x) - \left(\frac{1}{2}\beta^2 b + 1\right)u_i'(x) &= -L_1u_i(x) - \left(\frac{1}{2}\beta^2 b + 1\right)u_i'(x) \\ &= -\frac{1}{2}\beta^2\left(xu_i''(x) + bu_i'(x)\right),\ \text{for } i = 1, 2. \end{aligned}$$

Thus, we may rewrite (2.18) as

$$\Theta(x) = \frac{xu_1''(x) + bu_1'(x)}{xu_2''(x) + bu_2'(x)},\ 0 \le x \le x^*. \tag{2.19}$$

We shall obtain x^* as the minimal positive root of the equation $\Phi(x) = 0$, where

$$\Phi(x) = \left(V + \frac{1}{r}\right)[L_1u_1(x) - \Theta(x)L_1u_2(x)] - 1. \tag{2.20}$$

We shall show further that x^* is also the minimal positive root of the equation

$$L_1\tilde{F}(x) - 1 = 0 \text{ (see lemma 3.2)}.$$

To find an explicit solution in terms of special functions we have to consider the cases of technical uncertainty, input cost uncertainty, and the case of the presence of both uncertainties separately. As equation (2.14) is an equation of type 2.1.2.166 in [ZaiPol], we can write down its explicit solutions $u_1(x)$ and $u_2(x)$. To find an explicit expression for $\Theta(x)$ and $\Phi(x)$ from (2.18), (2.19), and (2.20), we shall use identities from Appendices A, B, and C in the cases of technical uncertainty, input cost uncertainty, and in the case of the presence of both uncertainties, respectively.

The case of technical uncertainty, $\gamma = 0$. Let $c = 1 + 2/\beta^2$. Replace x by $z = 2rx/\beta^2$ and write u_1 as a function $\tilde{u}_1$ of z, and likewise for other functions.

$$u_1(x) = \tilde{u}_1(z) = \frac{2}{\Gamma(c)} z^{c/2} K_c(2\sqrt{z}), \tag{2.21}$$

$$u_2(x) = \tilde{u}_2(z) = \frac{2}{\Gamma(c)} z^{c/2} I_c(2\sqrt{z}) \tag{2.22}$$

Note, that for $\gamma \searrow 0$ we have $b \nearrow \infty$. Consequently, we obtain from (2.19)

$$\Theta(x) = \tilde{\Theta}(z) = \frac{u_1'(z)}{u_2'(z)} \overset{(2.40),(2.41)}{=} -\frac{K_{c-1}(2\sqrt{z})}{I_{c-1}(2\sqrt{z})}, \tag{2.23}$$

where $I_\nu(x)$ and $K_\nu(x)$ are modified Bessel functions of the first and the second type respectively (see Section 2.4.1). Note that $L_1 u_i - r u_i = 0, i = 1, 2$, holds in view of (2.40), (2.14), and $\gamma = 0$. Thus we can write down the expression for $\Phi(x)$

$$\Phi(x) = \tilde{\Phi}(z) = (Vr + 1)\frac{2}{\Gamma(c)} z^{c/2} \left(K_c(2\sqrt{z}) + \frac{K_{c-1}(2\sqrt{z})}{I_{c-1}(2\sqrt{z})} I_c(2\sqrt{z}) \right) - 1. \tag{2.24}$$

The case of input cost uncertainty, $\beta = 0$. Let $z = rx/(b(b-1))$. Then

$$u_1(x) = \tilde{u}_1(z) = \frac{\Gamma(b+1)}{\Gamma(2b)} z^{1-b} e^{-1/z} M(b+1, 2b, z^{-1}), \tag{2.25}$$

$$u_2(x) = \tilde{u}_2(z) = \frac{\Gamma(b+1)}{\Gamma(2b)} z^{1-b} e^{-1/z} U(b+1, 2b, z^{-1}), \tag{2.26}$$

where $M(a, b, x)$ and $U(a, b, x)$ are confluent hypergeometric functions of the first and the second type (see Section 2.4.2). Use (2.51) and (2.52) to compute the first and the second derivatives of $\tilde{u}_1(z)$ and $\tilde{u}_2(z)$. Substitute the expressions for the derivatives in (2.19). Then (2.53) and (2.54) give

$$\Theta(x) = \tilde{\Theta}(z) = -\frac{b-1}{2} \frac{M(b, 2b+1, z^{-1})}{U(b, 2b+1, z^{-1})}. \tag{2.27}$$

Note that $L_1 \tilde{u}_i = -r/(b(b-1)) \frac{d\tilde{u}_i}{dz}, i = 1, 2$, as $\beta = 0$. Take derivatives using (2.51) and (2.52) to obtain

$$\Phi(x) = \tilde{\Phi}(z) = (Vr+1)\frac{\Gamma(b)}{\Gamma(2b)} z^{-b} e^{-1/z} \left(M(b, 2b, z^{-1}) - \frac{1}{2} \frac{M(b, 2b+1, z^{-1})}{U(b, 2b+1, z^{-1})} U(b, 2b, z^{-1}) \right) - 1. \tag{2.28}$$

The case of both uncertainties, $\beta > 0$, $\gamma > 0$. Let $z = rx(c-1)/(b(b-1))$ and, again, $c = 1 + 2\beta^{-2}$. We have

$$u_1(x) = \tilde{u}_1(z) = F(b-1, -b; 1-c; -z) \tag{2.29}$$
$$u_2(x) = \tilde{u}_2(z) = z^c F(b-1+c, -b+c; c+1; -z) \tag{2.30}$$

where $F(a, b; c, x)$ are hypergeometric functions (see Section 2.4.3). Using (2.59) and (2.60) we compute the first and the second derivatives of $u_1(z)$ and $u_2(z)$. Then we substitute the expressions for the derivatives in (2.19). Thus from (2.63) and (2.64) we obtain

$$\Theta(x) = \tilde{\Theta}(z) = -\frac{(b-1)b^2}{c(c-1)(b+c-1)} z^{1-c} \frac{F(1+b, 1-b; 2-c; -z)}{F(b+c, -b+c; c; -z)}. \tag{2.31}$$

Using (2.65) and the expression for the first and the second derivatives of $u_i(z), i = 1, 2$, it is not difficult to obtain expressions for $L_1 u_i(z), i = 1, 2$. Substituting the obtained expressions in (2.20) we get

$$\begin{aligned} \Phi(x) = \tilde{\Phi}(z) &= (Vr+1)[F(b, 1-b; 1-c; -z) + \\ &+ \frac{b(b-c)}{c(c-1)} z \frac{F(1+b, 1-b; 2-c; -z)}{F(b+c, -b+c; c; -z)} F(b+c, 1-b+c; c+1; -z)] - 1. \end{aligned} \tag{2.32}$$

Now we can summarize the results in the following theorem:

Theorem 2.2.1 *In the model described above the optimal control $\tilde{I} = \tilde{I}(x)$ and the value function $F = F(x)$ are the following:*

$$\tilde{I}_t = \begin{cases} 1, & x < x^* \\ 0, & x \geq x^* \end{cases}$$

$$F(x) = \begin{cases} (V + 1/r)\,[u_1(x) - \theta(x^*) u_2(x)] - 1/r, & x < x^* \\ x(b-1)^{-1}(\frac{1}{2}\beta^2 b + 1)^{-1}\,(x/x^*)^{-b}, & x \geq x^*, \end{cases}$$

where $b = \frac{1}{2}\left(1 + \sqrt{1 + (8r)/(\gamma^2)}\right)$, x^* is the minimal positive root of the equation $\Phi(x) = 0$, and $u_1(x), u_2(x), \Theta(x)$, and $\Phi(x)$ are defined by the equalities (2.21) —(2.32).

2.3 The proofs

In this section we prove theorem 2.2.1. We shall need two lemmas.

Lemma 2.3.1 *There exists at least one positive root of the equation $\tilde{\Phi}(z) = 0$.*[1]

PROOF OF THE LEMMA. Note, that for any $z \geq 0, b > 1, c > 1$ we have

$$\begin{aligned} I_{c-1}(2\sqrt{z}) &\quad > \quad 0, \\ U(b, 2b+1, z^{-1}) &\overset{(integral\ \ transform)}{=} \frac{1}{\Gamma(b)} \int_0^\infty e^{-t/z} t^{b-1} (1+t)^b dt > 0, \\ F(b+c, -b+c; c; -z) &\overset{(Euler\ \ transform)}{=} (1+z)^{-b-c} F(b+c, b; c; \frac{z}{z+1}) > 0. \end{aligned}$$

Together with the continuity of the Bessel, confluent hypergeometric and hypergeometric functions, these inequalities show that $\tilde{\Phi} = \tilde{\Phi}(z)$ is a continuous function. Taking the asymptotic representation of the corresponding special functions we get that $\tilde{\Phi}(0) = Vr > 0$ and $\tilde{\Phi}(+\infty) = -1 < 0$. By the intermediate value theorem there is a value z^* where $\tilde{\Phi}(z^*) = 0$. □

Lemma 2.3.2 *The following inequalities hold*

$$L_1 \tilde{F}(x) - 1 > 0, x < x^*, \tag{2.33}$$

$$L_1 \tilde{F}(x) - 1 \leq 0, x \geq x^*. \tag{2.34}$$

PROOF OF THE LEMMA. Let us prove (2.34). For $\gamma = 0$ we have $L_1 \tilde{F} - 1 = -1 < 0$. If $\gamma \neq 0$ we have

$$L_1 \tilde{F}(x) - 1 = \left(\frac{x^*}{x}\right)^b - 1 \leq 0, x \geq x^*.$$

Now let us prove (2.33) separately for each type of uncertainty.

Let $\tilde{\Psi}(z) = \Psi(x) \overset{def}{=} L_1 \tilde{F}(x) - 1$. Then in the case of **technical uncertainty** we have

$$\tilde{\Psi}(z) = (Vr+1) \frac{2}{\Gamma(c)} \left(z^{c/2} K_c(2\sqrt{z}) - \tilde{\Theta}(z^*) I_c(2\sqrt{z}) \right) - 1.$$

Take the derivative of $\tilde{\Psi}(z)$. Note that $I_{c-1}(2\sqrt{z}) > 0$, for $c > 1$. By corollary 2.4.1 from Appendix D we obtain the inequality

$$\begin{aligned} \tilde{\Psi}'_z(z) &= (Vr+1) \frac{2}{\Gamma(c)} \left[-z^{(c-1)/2} K_{c-1}(2\sqrt{z}) + \tilde{\Theta}(z^*) z^{(c-1)/2} I_{c-1}(2\sqrt{z}) \right] = \\ &= (Vr+1) \frac{2}{\Gamma(c)} z^{(c-1)/2} I_{c-1}(2\sqrt{z}) \left[\tilde{\Theta}(z) - \tilde{\Theta}(z^*) \right] \overset{(cor.2.4.1)}{<} 0. \end{aligned}$$

Thus, $\tilde{\Psi}(z)$ is a strictly decreasing continuous function for $z < z^*$. Moreover, by asymptotic properties of modified Bessel functions we have $\tilde{\Psi}(0) = Vr$. Also from the conditions of the lemma we have $\tilde{\Psi}(z^*) = 0$. Thus, $\tilde{\Psi}(z) > 0$ for $z < z^*$.

[1]Note that the changes of variables $z = z(x)$ depend on the cases of uncertainty.

In the case of **input cost uncertainty** we have:

$$\tilde{\Psi}(z) = (Vr+1)\frac{\Gamma(b)}{\Gamma(2b)}z^{-b}e^{-1/z}\left(M(b,2b,z^{-1}) + \frac{\tilde{\Theta}(z^*)}{b-1}U(b,2b+1,z^{-1})\right) - 1.$$

Let us consider the difference $\tilde{\Psi}(z) - \tilde{\Phi}(z)$. By corollary 2.4.2 from Appendix D we have

$$\tilde{\Psi}(z) - \tilde{\Phi}(z) = (Vr+1)\frac{\Gamma(b-1)}{\Gamma(2b)}z^{-b}e^{-1/z}U(b,2b;z^{-1})\left(\tilde{\Theta}(z^*) - \tilde{\Theta}(z)\right) > 0, z < z^*.$$

But for $z < z^*$ we have $\tilde{\Phi}(z) > 0$, as $\tilde{\Phi}(0) = Vr > 0$ and z^* is the minimal positive root of $\tilde{\Phi}(z) = 0$. Therefore, $\tilde{\Psi}(z) > \tilde{\Phi}(z) > 0$.

In the case of **two uncertainties** we have:

$$\tilde{\Psi}(z) = (Vr+1)\left(F(b,1-b;1-c;-z) - \right.$$
$$\left. - \frac{(b-c)(b+c+1)}{b(b-1)}\tilde{\Theta}(z^*)z^c F(b+c,1-b+c;c+1;-z)\right) - 1.$$

In order to prove that $\tilde{\Psi}(z)$ is positive for $z < z^*$ we consider separately the cases $b - c < 0$ and $b - c > 0$.

Suppose $b - c < 0$. Note, that

$$F(b+c,1-b+c;c+1;-z) \overset{Euler\ transform}{=} (1+z)^{-b-c}F(b+c,b;c+1;\frac{z}{1+z}) > 0$$

By corollary 2.4.3 from Appendix D we obtain:

$$\tilde{\Psi}(z) - \tilde{\Phi}(z) = (Vr+1)\frac{b(b-c+1)}{b(b-1)}z^c F(b+c,1-b+c;c+1;-z)\left[\tilde{\Theta}(z) - \tilde{\Theta}(z^*)\right] > 0.$$

Thus, $\tilde{\Psi}(z) > \tilde{\Phi}(z) > 0$ for $z < z^*$.

Now suppose $b - c > 0$. First we note, that

$$F(b+c,-b+c;c;-z) \overset{Euler\ transform}{=} (1+z)^{-b-c}F(b+c,b;c;\frac{z}{1+z}) > 0$$

Let us rewrite $\tilde{\Psi}(z)$ as

$$\tilde{\Psi}(z) = (Vr+1)(1+z)^{-b}N(z) - 1,$$

where

$$N(z) = (1+z)^b F(b,1-b;1-c;-z) - (1+z)^b \frac{(b-c)(b+c+1)}{b(b-1)}\tilde{\Theta}(z^*)z^c F(b+c,1-b+c;c+1;-z)$$

Let us take the derivative of $N(z)$ using (2.62),(2.61). By corollary 2.4.3 from Appendix D we obtain:

$$\begin{aligned} N_z'(z) &= \frac{b(b-c)}{1-c}(1+z)^{b-1}F(b+1,1-b;2-c;-z) + \\ &+ \tilde{\Theta}(z^*)\frac{(b-c)(b+c-1)c}{b(b-1)}z^c(1+z)^{b-1}F(b+c,-b+c;c;-z) = \\ &= \frac{(b-c)(b+c-1)c}{b(b-1)}z^{1-c}(1+z)^{b-1}F(b+c,-b+c;c;-z)\left[\tilde{\Theta}(z)-\tilde{\Theta}(z^*)\right] < 0 \end{aligned}$$

Thus, we have proved that $N(z)$ and $(1+z)^{-b}$ are strictly decreasing functions for $z < z^*$. Therefore $\tilde{\Psi}(z)$ is a strictly decreasing function for $z < z^*$. Besides that, $\tilde{\Psi}(0) = Vr > 0$ and $\tilde{\Psi}(z^*) = 0$.

Thus, $\tilde{\Psi}(z) > 0$ for $z < z^*$. □

Proof of the theorem. To prove the theorem we need to check if *the verification properties* hold.

According to the standard technique of stochastic optimal control *the verification properties* are the following:

(A) There exists a function $\tilde{F} = \tilde{F}(x)$ such that for any admissible control $I = I(x)$

$$F(x,I) \le \tilde{F}(x)$$

(B) There exists a control $\tilde{I} = \tilde{I}(x)$ such that

$$F(x,\tilde{I}) = \tilde{F}(x)$$

Let us show that property (A) holds.

Applying the Îto formula to $\left(e^{-rt}\tilde{F}(X_t)\right)_{t\ge 0}$ we obtain

$$\begin{aligned} e^{-r(t\wedge\tau)}\tilde{F}(X_{(t\wedge\tau)}) = \tilde{F}(X_0) + \int_0^{t\wedge\tau} e^{-rs}L(I)\tilde{F}(X_s)ds + \int_0^{t\wedge\tau} e^{-rs}\gamma X_s\tilde{F}'(X_s)dW_s \\ + \int_0^{t\wedge\tau} e^{-rs}\beta\sqrt{I_sX_s}\tilde{F}'(X_s)d\tilde{W}_s \end{aligned}$$

Let us notice that by lemma 2.3.2 we have for any admissible control I

$$IL_1\tilde{F}(x) + L_2\tilde{F}(x) - I = I(L_1\tilde{F}(x)-1) + L_2\tilde{F}(x) \le \tilde{I}\left(L_1\tilde{F}(x)-1\right) + L_2\tilde{F}(x) = 0. \quad (2.35)$$

Taking the mathematical expectation E_x of $e^{-rt}\tilde{F}(X_t)$, by (2.35) we obtain

$$\begin{aligned}
&\tilde{F}(x) = \mathsf{E}_x e^{-r(t\wedge\tau)}\tilde{F}(X_{t\wedge\tau}) - \mathsf{E}_x\int_0^{t\wedge\tau} e^{-rs}\left(IL_1+L_2\right)\tilde{F}(X_s)ds - \\
&\mathsf{E}_x\int_0^{t\wedge\tau} e^{-rs}\gamma X_s\tilde{F}'(X_s)dW_s - \mathsf{E}_x\int_0^{t\wedge\tau} e^{-rs}\beta\sqrt{I_sX_s}\tilde{F}'(X_s)d\tilde{W}_s \geq \\
&\geq \ \mathsf{E}_x e^{-r(t\wedge\tau)}\tilde{F}(X_{t\wedge\tau}) + \mathsf{E}_x\int_0^{t\wedge\tau}(-I)e^{-rs}ds - \mathsf{E}_x\int_0^{t\wedge\tau} e^{-rs}\gamma X_s\tilde{F}'(X_s)dW_s \\
&- \ \mathsf{E}_x\int_0^{t\wedge\tau} e^{-rs}\beta\sqrt{I_sX_s}\tilde{F}'(X_s)d\tilde{W}_s. \qquad (2.36)
\end{aligned}$$

Note, that the stochastic integrals in (2.36) are martingales. Therefore the mathematical expectation of those integrals is equal to zero. Letting t go to infinity in (2.36) we obtain $\tilde{F}(x) \geq F(x,I)$, as $e^{-r(t\wedge\tau)}\tilde{F}(X_{t\wedge\tau}) \to Ve^{-r\tau}$. Indeed, if $\tau < \infty$ then $\tilde{F}(X_\tau) = V$. And if $\tau = \infty$ then $\tilde{F}(X_t)$ is bounded, $e^{-r(t\wedge\tau)} \to 0$ and $Ve^{-r(t\wedge\tau)} \to$, $t \to \infty$. Thus it follows from (2.36) that $\tilde{F}(x) \geq F(x,I)$, i.e. the property (A) holds.

Let us check property (B) similarly. Applying Îto formula to $e^{-rt}\tilde{F}(X_t)$ and taking mathematical expectation we have:

$$\begin{aligned}
&\tilde{F}(x) = \mathsf{E}_x e^{-r(t\wedge\tau)}\tilde{F}(X_{t\wedge\tau}) - \mathsf{E}_x\int_0^{t\wedge\tau} e^{-rs}\left(\tilde{I}L_1+L_2\right)\tilde{F}(X_s)ds \\
&-\mathsf{E}_x\int_0^{t\wedge\tau} e^{-rs}\gamma X_s\tilde{F}'(X_s)dW_s - \mathsf{E}_x\int_0^{t\wedge\tau} e^{-rs}\beta\sqrt{I_sX_s}\tilde{F}'(X_s)d\tilde{W}_s = \\
&= \ \mathsf{E}_x e^{-r(t\wedge\tau)}\tilde{F}(X_{t\wedge\tau}) + \mathsf{E}_x\int_0^{t\wedge\tau}(-\tilde{I})e^{-rs}ds - \mathsf{E}_x\int_0^{t\wedge\tau} e^{-rs}\gamma X_s\tilde{F}'(X_s)dW_s - \\
&- \ \mathsf{E}_x\int_0^{t\wedge\tau} e^{-rs}\beta\sqrt{I_sX_s}\tilde{F}'(X_s)d\tilde{W}_s \qquad (2.37)
\end{aligned}$$

The mathematical expectation of two last terms in (2.37) is equal to zero. Thus, as $t \to \infty$ we obtain:

$$\tilde{F}(x) = \mathsf{E}_x\left(\int_0^\tau(-\tilde{I})e^{-rs}ds + Ve^{-r\tau}\right) = F(x,\tilde{I}.)$$

□

2.4 Appendix

We have used a number of special functions in this chapter. For the reader's convenience we include some definitions and properties of special functions in the next subsections.

2.4.1 Bessel functions

A function $I_\nu(z)$ is called a modified Bessel function of the first kind (see [AS], [Erd]) if

$$I_\nu(z) = (z/2)^\nu \sum_{n=0}^{\infty} \frac{(z^2/4)^n}{n!\Gamma(\nu+n+1)}. \tag{2.38}$$

A function $K_\nu(z)$ is called a modified Bessel function of the third kind if

$$K_\nu(z) = \pi/2 \frac{I_{-\nu}(z) - I_\nu(z)}{\sin(\nu\pi)}. \tag{2.39}$$

We use the following properties of Bessel functions:

$$\frac{d}{dz}\left(z^{\frac{\nu}{2}} I_\nu(2\sqrt{z})\right) = z^{\frac{\nu-1}{2}} I_{\nu-1}(2\sqrt{z}) \tag{2.40}$$

$$\frac{d}{dz}\left(z^{\frac{\nu}{2}} K_\nu(2\sqrt{z})\right) = -z^{\frac{\nu-1}{2}} K_{\nu-1}(2\sqrt{z}) \tag{2.41}$$

The asymptotic behaviour is given by

$$I_\nu(2\sqrt{z}) \sim z^{\frac{\nu}{2}}/\Gamma(\nu+1),\ (\nu \neq -1, -2, ...) \text{ as } z \to 0, \tag{2.42}$$

$$K_\nu(2\sqrt{z}) \sim \frac{1}{2}\Gamma(\nu) z^{-\frac{\nu}{2}}, (\nu > 0) \text{ as } z \to 0, \tag{2.43}$$

$$I_\nu(z) \sim \frac{e^z}{\sqrt{2\pi z}} \text{ as } z \to \infty, \tag{2.44}$$

$$K_\nu(z) \sim \sqrt{\frac{\pi}{2z}} e^{-z} \text{ as } z \to \infty. \tag{2.45}$$

Some computation shows

$$I_{\nu-2}(2\sqrt{z}) - (\nu-1) z^{-1/2} I_{\nu-1}(2\sqrt{z}) - I_\nu(2\sqrt{z}) = 0. \tag{2.46}$$

Together with (2.40) this implies that $G(z) = z^{\nu/2} I_\nu(2\sqrt{z})$ satisfies

$$zG''(z) - (\nu-1)G'(z) - G(z) = 0. \tag{2.47}$$

2.4.2 Kummer functions

In this section we introduce Kummer functions, also known as confluent hypergeometric functions.

A function $M(a,b,z)$ is called a confluent hypergeometric function of the first kind (see [AS], [Erd]) if

$$M(a,b,z) = \sum_{n=0}^{\infty} \frac{(a)_n}{(b)_n} \frac{z^n}{n!}, \tag{2.48}$$

$(a)_n = a(a+1)..(a+n-1) = \Gamma(a+n)/\Gamma(a)$, A function $U(a,b,z)$ is called a confluent hypergeometric function of the second kind if

$$U(a,b,z) = \frac{\pi}{\sin \pi b} \left(\frac{M(a,b,z)}{\Gamma(1+a-b)\Gamma(b)} - z^{1-b} \frac{M(1+a-b,2-b,z)}{\Gamma(a)\Gamma(2-b)} \right). \tag{2.49}$$

We use the following properties of confluent hypergeometric functions:

Integral representation

$$\Gamma(a)U(a,b,z) = \int_0^{\infty} e^{-zt} t^{a-1} (1+t)^{b-a-1} dt, a>0, z>0. \tag{2.50}$$

Differential relations

$$\frac{d}{dz}\left[e^{-z} z^{b-a} M(a,b,z)\right] = (b-a) e^{-z} z^{b-a-1} M(a-1,b,z) \tag{2.51}$$

$$\frac{d}{dz}\left[e^{-z} z^{b-a} U(a,b,z)\right] = -e^{-z} z^{b-a-1} U(a-1,b,z). \tag{2.52}$$

Some identities

$$M(a,b,z) - M(a-1,b,z) = \frac{z}{b} M(a,b+1,z) \tag{2.53}$$

$$(b-a)U(a,b,z) + U(a-1,b,z) = zU(a,b+1,z). \tag{2.54}$$

Asymptotic behaviour as $z \to 0$

$$M(a,b,z) \to 1 \text{ for } b \notin \{0,-1,-2,...\} \tag{2.55}$$

$$U(a,b,z) = \frac{\Gamma(b-1)}{\Gamma(a)} z^{1-b} + O(|z|^{b-2}), (\text{for } b>2). \tag{2.56}$$

Asymptotic behaviour as $z \to \infty$

$$M(a,b,z) = \frac{\Gamma(b)}{\Gamma(a)} e^z z^{a-b} [1 + O(|z|^{-1})], (\text{for } z>0), \tag{2.57}$$

$$U(a,b,z) = z^{-a}\left[1 + O(|z|^{-1})\right]. \tag{2.58}$$

2.4.3 Hypergeometric functions

A function $F(a,b;c;z)$ is called a hypergeometric function if :

$$F(a,b;c;z) = \sum_{n=0}^{\infty} \frac{(a)_n (b)_n}{(c)_n} \frac{z^n}{n!},$$

where $(a)_n = a(a+1)..(a+n-1)$.

We use the following properties of hypergeometric functions: *Differential relations*

$$\frac{d}{dz} F(a,b;c;z) = \frac{ab}{c} F(a+1,b+1;c+1;z) \tag{2.59}$$

$$\frac{d}{dz}\left[z^{c-1} F(a,b;c;z)\right] = (c-1) z^{c-2} F(a,b;c-1;z) \tag{2.60}$$

$$\frac{d}{dz}\left[(1-z)^{a-c+1} z^{c-1} F(a,b;c;z)\right] = (c-1) z^{c-2} (1-z)^{a-c} F(a,b-1;c-1;z) \tag{2.61}$$

$$\frac{d}{dz}\left[(1-z)^{a} F(a,b;c;z)\right] = -\frac{a(c-b)}{c} (1-z)^{a-1} F(a+1,b;c+1;z). \tag{2.62}$$

Some identities

$$z\frac{d}{dz} F(a,b;c;z) + aF(a,b;c;z) - aF(a+1,b;c;z) = 0 \tag{2.63}$$

$$(c-a-1)F(a,b;c;z) + aF(a+1,b;c;z) - (c-1)F(a,b;c-1;z) = 0 \tag{2.64}$$

$$z\frac{d}{dz} F(a,b;c;z) + (c-1)F(a,b;c;z) - (c-1)F(a,b;c-1;z) = 0. \tag{2.65}$$

Integral representation

$$F(a,b;c;z) = \frac{\Gamma(c)}{\Gamma(b)\Gamma(c-b)} \int_0^1 t^{b-1}(1-t)^{c-b-1}(1-tz)^{-a} dt, (c > b > 0). \tag{2.66}$$

Asymptotic behaviour as $z \to 0$

$$F(a,b;c;z) \to 1. \tag{2.67}$$

Asymptotic behaviour as $z \to \infty$

$$F(a,b;c;z) \sim \frac{\Gamma(c)\Gamma(a-b)}{\Gamma(a)\Gamma(c-b)} (-z)^{-b}, a > b. \tag{2.68}$$

Euler transform

$$F(a,b;c;z) = (1-z)^{-a} F\left(a, c-b; c; \frac{z}{z-1}\right).$$

2.4.4 Wronskian type property

Lemma 2.4.1 *Let $f \neq 0$ and $u = u(x)$ and $v = v(x)$ be two linearly independent solutions of the second order differential equation*

$$f(x)y'' + g(x)y' + h(x)y = 0. \tag{2.69}$$

Then $\Theta(x) = u(x)/v(x)$ is strictly monotone outside the zeros of v, and its derivative is

$$\Theta'(x) = \frac{const}{v^2(x)} \exp\left(-\int \frac{g(x)}{f(x)} dx\right). \tag{2.70}$$

Proof. Since u and v are two linearly independent solutions of (2.69), the following equalities hold:

$$f(x)u'' + g(x)u' + h(x)u = 0 \tag{2.71}$$
$$f(x)v'' + g(x)v' + h(x)v = 0. \tag{2.72}$$

Multiply (2.71) by v and (2.72) by u. By subtracting the second term from the first we obtain

$$f(x)\left(u''v - v''u\right) + g(x)\left(u'v - v'u\right) = 0. \tag{2.73}$$

Let $w = u'v - v'u$. Then $w' = u''v - v''u$. Therefore we can rewrite (2.73) as a first order differential equation in separated variables. The function $w(x) = const\ \exp\left(-\int g(x)/f(x)dx\right)$ is the solution of this equation. Thus we obtain

$$\Theta'(x) = \frac{u'v - v'u}{v^2} = \frac{const}{v^2}\left[\exp\left(-\int \frac{g(x)}{f(x)}\right)\right]. \tag{2.74}$$

$\square$

Corollary 2.4.1 *The function $\Theta(z) = -K_{c-1}(2\sqrt{z})/I_{c-1}(2\sqrt{z})$ is strictly increasing and $\Theta'_z(z) = \frac{1}{2z}\left[I_{c-1}(2\sqrt{z})\right]^{-2}$.*

Corollary 2.4.2 *For $b > 1$, the function $\Theta(z) = -(b-1)M(b, 2b+1, z^{-1})/(2U(b, 2b+1, z^{-1}))$ is strictly increasing and $\Theta'_z(z) = \frac{1}{2}e^{1/z}z^b\left[(b, 2b+1, z^{-1})\right]^{-2}$.*

Corollary 2.4.3 *For $c > 1$ and $b > 1$ the function $\Theta(z) = -\frac{(b-1)b^2}{c(c-1)(b+c-1)}z^{1-c}F(b+1, 1-b; 2-c; -z)/F(b+c, -b+c; c; -z)$ is strictly increasing and*

$$\Theta'_z(z) = \frac{(b-1)b^2}{c(b+c-1)}z^{-c}(1+z)^{-c-1}\left[F(b+c, -b+c; c; -z)\right]^{-2}.$$

Chapter 3

Optimal arbitrage trading

In this chapter we consider a model for arbitrage trading.

In section 3.1 we introduce the model. In section 3.2 we present the explicit solution for the optimal optimal trading strategy and the proof of optimality. In section 3.3 we analyze the behaviour of the solution from section 3.2 and discuss possible generalizations.

3.1 Introduction

3.1.1 Motivation

Many academic papers about optimal trading rules and portfolio selection assume that the assets follow geometric Brownian motions, or, more generally, random walks. These papers are typically concerned with portfolio selection problems faced by long-term investors. In this paper, we consider a problem where the asset price is driven by a mean-reverting process. With some exceptions (e.g. [Lo]), this kind of processes is not widely used to model stock or bond price dynamics. However similar portfolio selection problems arise naturally in many "relative value" strategies assuming some kind of mean reversion in a tradable asset.

Consider, for example, a limited capital speculator trading the spread (i.e. the difference) between two cointegrated assets or, more generally, an arbitrageur with a limited capital trading a mean-reverting asset. The trader knows the "correct" (long-term average) price of the asset, and he knows that the price will sooner or later revert to the correct level, but the risk is that the position losses may become unbearable for the trader before the reversion happens. The finite horizon assumption is quite realistic because the bonuses to traders and fees to hedge fund managers are usually paid yearly. Just to give an example, Fig. 3.1 shows the spread of the once famous BASF-Bayer stock pair.

Faced with a mean-reverting process, a trader would typically take a long (i.e. positive) position in the asset when the asset is below its long-term mean and a short (i.e. negative)

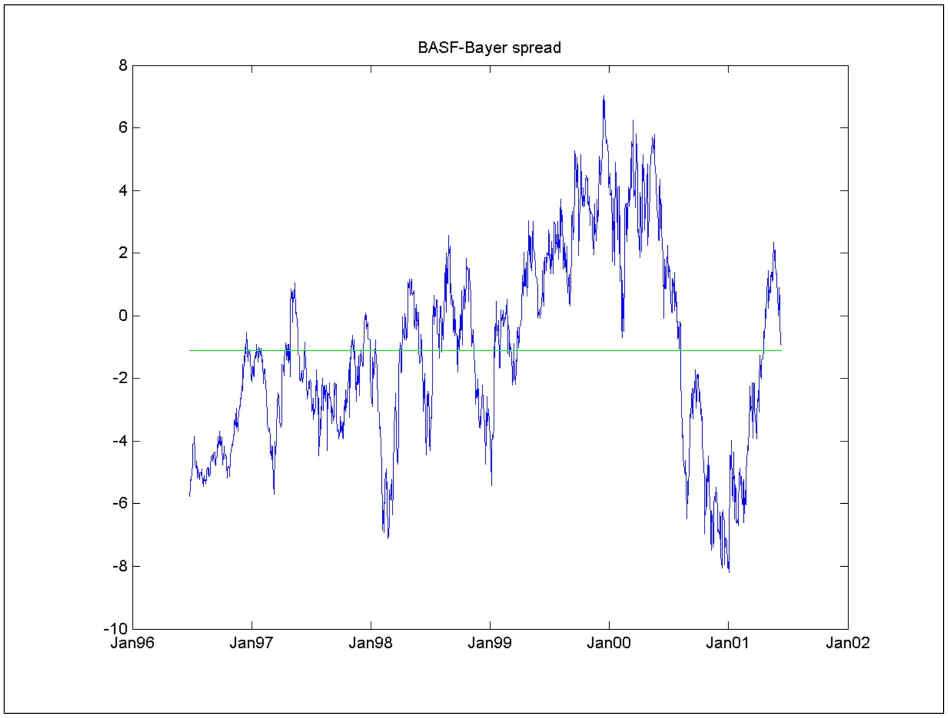

Figure 3.1: Difference of ordinary share close price for BASF AG and Bayer AG, 1997-2001.

position when the asset is below the long-term mean. He would then either liquidate the position when the price reverts closer to the mean and take the profit or he might have to close the position before the reversion happens and face the losses. The question is in the size of the position and how the position should be optimally managed as the price and the trader's wealth change and time passes by. An often used rule of thumb is that one opens the short position as soon as the spread is above one standard deviation from its mean and a long position as soon as the spread is below one standard deviation.

It is well known that capital and risk-bearing constraints may seriously limit arbitrage activities. Shleifer [Shl] built an equilibrium model for a market with limited capital arbitrageurs.

We solve the optimal problem assuming an Ornstein-Uhlenbeck process for the price and power utility over the final wealth for a finite horizon agent. This model was first formulated for the power utility case and solved for the log-utility case by Mendez-Vivez, Morton, and Naik, [Morton], [M-VMN].

Besides the quantitative result, there is a number of interesting qualitative questions to answer about the optimal strategy.

- When and how aggressively should one open the position?
- When should one cut a loosing position?
- Can a trader ever be happy when the spread widens against his position?
- What is the effect of process parameters on the optimal strategy?
- How does the trading strategy and the value function change as the time horizon approaches?
- What is the effect of risk aversion on position term dependence?
- How does the uncertainty in the process parameters affect the optimal strategy?

We answer these questions in section 3.3.

3.1.2 Choice of the price process

Without loss of generality we can assume that the long-term mean of the price process is zero. WE stick to the simplest example of a mean-reverting process, namely, the *Ornstein-Uhlenbeck process* given by

$$dX_t = -kX_t dt + \sigma dB_t, \tag{3.1}$$

where B_t is a Brownian motion, k and σ are positive constants. This process will revert to its long-term mean zero. More exactly, given X_t, the distribution of X_{t+s}, $s > 0$, is normal with parameters

$$E(X_{t+s}|X_t) = X_t e^{-ks}; \qquad Var(X_{t+s}|X_t) = \left(\frac{1-e^{-2ks}}{2k}\right)\sigma^2. \tag{3.2}$$

Informally, the constant k measures the speed of the mean-reversion and σ measures the strength of the noise component.

3.1.3 Choice of the utility function

For $-\infty < \gamma < 1$ we consider the power utility

$$U = U(W_T) = \frac{1}{\gamma} W_T^\gamma \tag{3.3}$$

over the terminal wealth W_T. This is a simple but rich enough family of utility functions. Utility functions are defined up to an additive constant. To include the log-utility as a special case, it is sometimes more convenient to consider the family of utility functions $U(W_T) =$

$\frac{1}{\gamma}(W_T^\gamma - 1)$. Taking the limit $\gamma \to 0$ we obtain the log-utility function $U(W_T) = \log(W_T)$. The log-utility version of our problem was solved by A. Morton [Morton].

The relative risk aversion is measured by $1 - \gamma$, so the bigger γ is, the less risk averse is the agent. In the limit $\gamma \to 1$ we have a linear utility function. In section 3.3 we study the effect of γ on the trading strategy.

3.1.4 The model

The problem can be treated in the general portfolio optimization framework of [Merton]. Suppose a traded asset follows an Ornstein-Uhlenbeck process (3.1). It is convenient to think about X_t as a "spread" between the price of an asset and its "fair value". Let α_t be a trader's position at time t, i.e. the number of units of the asset held. This parameter is the control in our optimization problem. Assuming zero interest rates and no market frictions, the wealth dynamics for a given control α_t is given by

$$dW_t = \alpha_t dX_t = -k\alpha_t X_t dt + \alpha_t \sigma dB_t. \tag{3.4}$$

We assume that there are no restrictions on α, so short selling is allowed and there are no marginal requirements on the wealth W.

We solve the expected terminal utility maximization problem for an agent with a prespecified time horizon T and initial wealth W_0. The utility function (3.3) is defined over the terminal wealth W_T. The value function $J(W_t, X_t, t)$ is the expectation of the terminal utility conditional on the information available at time t:

$$J(W_t, X_t, t) = \sup_{\alpha_t} \mathbf{E}_t \frac{1}{\gamma} W_T^\gamma. \tag{3.5}$$

3.1.5 Normalization

It is more convenient to work with dimensionless time and money. Let \$ be the dimension of X; we denote it by $[X] = \$$. By T we denote the dimension of time. From Eq. (3.1) it is clear that $[\sigma] = \$T^{-1/2}$ and $[k] = T^{-1}$. Renormalizing price X_t, position size α_t, and time t

$$\begin{aligned} X &\to \frac{X}{\sigma}\sqrt{k}, . \\ \alpha &\to \frac{\alpha}{\sqrt{k}}\sigma, \\ t &\to kt, \end{aligned} \tag{3.6}$$

we can assume that $k = 1$ and $\sigma = 1$. The wealth W does not change under this normalization. Note that this normalization is slightly different from the one used in [M-VMN].

3.1.6 Overview

In [M-VMN] it is proven that for $\gamma = 0$ (log-utility case) the optimal control is given simply by

$$\alpha_t = -W_t X_t.$$

The case $\gamma = 0$ is simpler than the general case because a log-utility agent does not hedge intertemporally (see [Merton]) and the equations are much simpler. The same paper also derives an approximate solution for the case $\gamma < 0$. The approximation does not behave particularly well.

In Section 3.2, we obtain an exact solution to the problem defined by Eqs. (3.1) – (3.5) for the general case $\gamma < 1$, $\gamma \neq 0$. The answer is given by Eqs. (3.15) and (3.16). In section 3.3, we analyze this solution. We are looking at how J and α change as the spread X changes and how the risk aversion affects the trader's strategy.

We will see that although our model is very simple, it reproduces some of the typical trader behavior patterns. For example, if a trader is more risk-averse than a log-utility one, then he will cut his position as the time horizon approaches. This behavior is similar to the anecdotical evidence on real position management practice.

Section 3.4 concludes with suggestions for possible generalizations.

3.2 Main result

3.2.1 The Hamilton-Jacobi-Bellman equation

We need to find the optimal control $\alpha^*(W_t, X_t, t)$ and the value function $J(W_t, X_t, t)$ as explicit functions of wealth W_t, price X_t, and time t.

The *Hamilton-Jacobi-Bellman equation*[1] is

$$\sup_{\alpha}\left(J_t - xJ_x - \alpha x J_w + \frac{1}{2}J_{xx} + \frac{1}{2}\alpha^2 J_{ww} + \alpha J_{xw}\right) = 0 \tag{3.7}$$

The first order optimality condition on control α^* is

$$\alpha^*(w, x, t) = x\frac{J_w}{J_{ww}} - \frac{J_{xw}}{J_{ww}}. \tag{3.8}$$

Substituting this condition into the Hamilton-Jacobi-Bellman equation for the value function, we obtain the non-linear PDE

$$J_t + \frac{1}{2}J_{xx} - xJ_x - \frac{1}{2}J_{ww}\left(\frac{J_{xw}}{J_{ww}} - x\frac{J_w}{J_{ww}}\right)^2 = 0. \tag{3.9}$$

[1] see e.g. [FlemSoner].

Note that the first summand at the right-hand side of Eq.(3.8) is the myopic demand term corresponding to a static optimization problem while the second term hedges from changes in the investment opportunity set. For a log-utility investor ($\gamma = 0$) the second term vanishes (see [Merton].)

3.2.2 Main theorem

Let

$$\tau = T - t \tag{3.10}$$

be the time left for trading and define the constant ν and time functions $C(\tau)$, $C'(\tau)$, and $D(\tau)$ by

$$\begin{aligned} \nu &= \frac{1}{\sqrt{1-\gamma}} && (3.11) \\ C(\tau) &= \cosh \nu\tau + \nu \sinh \nu\tau && (3.12) \\ C'(\tau) &= \frac{dC(\tau)}{d\tau} = \nu \sinh \nu\tau + \nu^2 \cosh \nu\tau && (3.13) \\ D(\tau) &= \frac{C'(\tau)}{C(\tau)}. && (3.14) \end{aligned}$$

As we shall see, the function $D(\tau)$ plays a crucial role in determining the optimal strategy.

Theorem 3.2.1 *Suppose that $\gamma < 0$ or $0 < \gamma < 1$. Then the optimal strategy for the problem (3.1) – (3.5) is given by*

$$\alpha_t^*(w, x, t) = -wxD(\tau). \tag{3.15}$$

The value function is given by

$$J(w, x, t) = \frac{1}{\gamma} w^\gamma \sqrt{e^\tau C(\tau)^{\gamma-1}} \exp\left(\frac{x^2}{2}\left(1 + (\gamma - 1)D(\tau)\right)\right), \tag{3.16}$$

where τ, $C(\tau)$, and $D(\tau)$ are defined by Eqs. (3.10) – (3.14) and $X_t = x$, $W_t = w$.

We will prove the theorem in Section 3.5.

Note that the optimal position is linear in both wealth W_t and spread X_t. The term in the last exponent in (3.16) measures the expected utility of the immediate trading opportunity. If $X_t = 0$ i.e. there are no immediate trading opportunities, the value function (3.16) simplifies to

$$J(w, 0, t) = \frac{1}{\gamma} w^\gamma \sqrt{e^\tau C(\tau)^{\gamma-1}}.$$

The $\frac{1}{\gamma} w^\gamma$ term is just the expected utility generated by the present wealth. The square root term can be thought of as the value of the time. We will analyze Eqs. (3.15) and (3.16) in more detail in section 3.3.

3.3 Analysis

In this section, we analyze the behavior of solution (3.15) - (3.16). Unless specified otherwise, the parameters used for illustrations are $k = 2$, $\sigma = 1$, and $\gamma = -2$. From Eq. (3.2), it follows that the long-term standard deviation of the price process value is 1/2, so, roughly, an absolute value of X greater than 0.5 presents a reasonable trading opportunity.

3.3.1 Position management

Let us look at how the value function and trading position change as X_t changes. Using Itø's lemma, we see from (3.15) that the diffusion term of $d\alpha_t$ is

$$-D(\tau)\left(W_t + \alpha_t X_t\right).$$

Thus, the covariance of $d\alpha$ and dX is

$$\mathrm{Cov}\left(d\alpha, dX\right) = -D(\tau)(W_t + \alpha_t X_t) = W_t D(\tau)\left(-1 + X_t^2 D(\tau)\right). \tag{3.17}$$

This is negative whenever

$$|X| \leq \sqrt{1/D(\tau)}.$$

Consequently, as X_t diverges from 0 either way, we start slowly building up the position α_t of the opposite sign than X_t. If X_t diverges further from 0, our position is making a loss, but we are still increasing the position until the squared spread X_t^2 reaches $1/D(\tau)$. If the spread widens beyond that value, we start cutting a loss-making position. Another interpretation of Eq. (3.17) is that we start cutting a loss-making position as soon as the *position spread* $-\alpha X_t$ exceeds total wealth W_t. Fig. 3.2 shows how $D(\tau)$ depends on the remaining time τ for different values of γ.

Not surprisingly, for the log-utility case $\gamma = 0$, the threshold $1/D(\tau)$ equals identically one and we are getting the same result as in [Morton] and [M-VMN].

3.3.2 Value function dynamics

Let us check now how the value function $J(W_t, X_t, t)$ evolves with X_t. In [M-VMN], it is shown that a log-utility agent's value function always decreases as the spread moves against his position. It might be the case that a more aggressive agent's value function sometimes increases as the spread X_t moves against his position because the investment opportunity set improves. Let us check whether this ever happens to a power utility agent.

Using Itô's lemma, we see from (3.16) that the diffusion term of dJ_t is

$$J_t X_t \left(1 - D(\tau)\right).$$

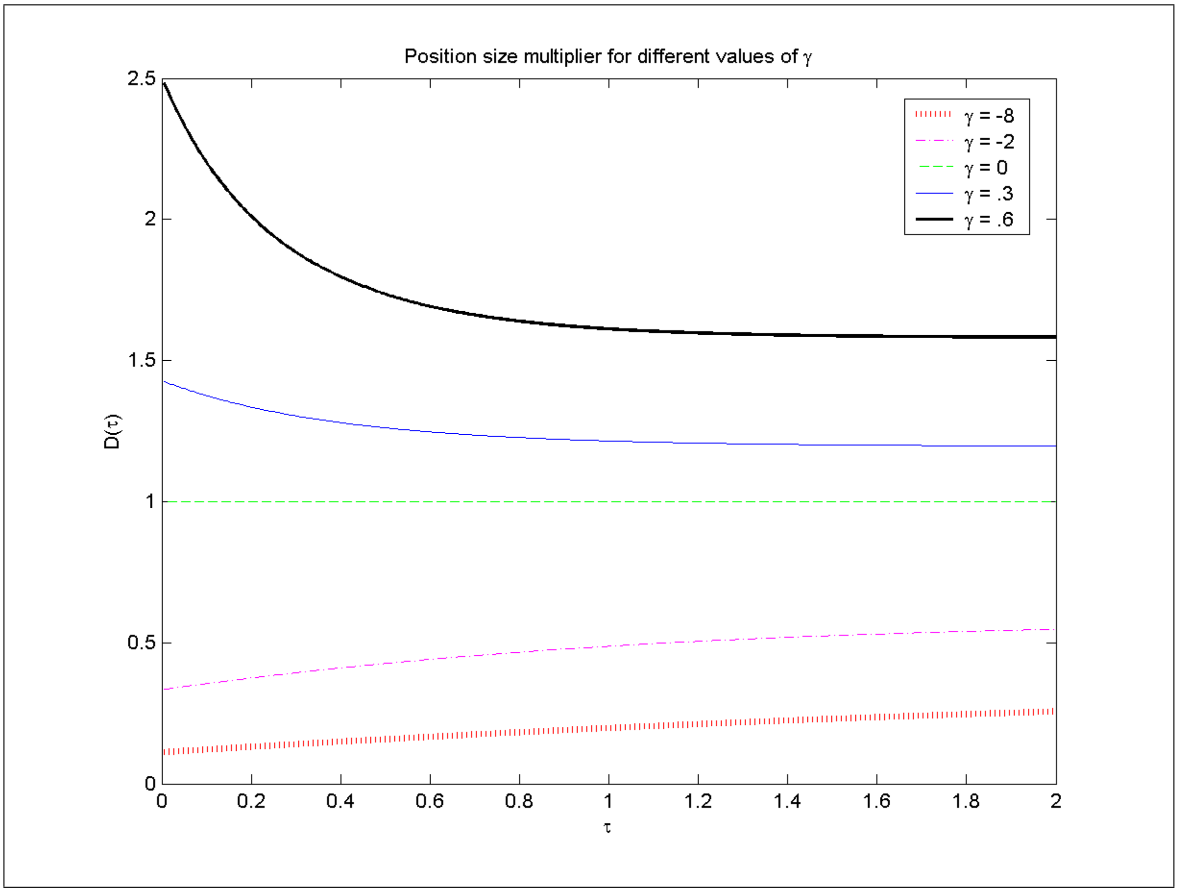

Figure 3.2: $D(\tau)$ as a function of the remaining time τ for five different values of γ.

Thus,

$$\text{Cov}\left(dJ, dX\right) = J_t X_t \left(1 - D(\tau)\right). \tag{3.18}$$

For $\gamma < 0$, the utility function is always negative, so the value function is also always negative. Similarly, for $\gamma > 0$ the value function J is always positive. It is easy to check that the sign of $1 - D(\tau)$ is opposite to the sign of γ for all τ. Thus, $\text{Cov}\left(dJ, dX\right)$ is positive for $X_t < 0$ and negative for $X_t > 0$. This means that any power utility agent suffers decrease in his value function J as the spread moves against his position. This is true even for an agent with an almost linear utility $\gamma \to 1$.

For $0 < \gamma < 1$ there is a non-zero bankruptcy probability.

3.3.3 Time value

Let us look once more at how the value function depends on the time left for trading. Recall that

$$J(w,x,t) = \underbrace{\frac{1}{\gamma}w^\gamma}_{A} \underbrace{\sqrt{e^\tau C(\tau)^{\gamma-1}}}_{B} \underbrace{\exp\left(\frac{x^2}{2}\left(1+(\gamma-1)D(\tau)\right)\right)}_{C}, \tag{3.19}$$

where τ, $C(\tau)$, and $D(\tau)$ are defined by Eqs. (3.10) – (3.14) and $X_t = x$, $W_t = w$. Thus, the value function J can be split into three multiplicative terms. Term A is the value derived from the present wealth, term B is the time value, and term C is the value of the *immediate* investment opportunity. Fig. 3.3 shows[2] the dependence of the value function J on the remaining time τ assuming that there is no immediate opportunity, i.e. $X = 0$.

Since the strategy of the log-utility agent does not depend on the time, his value function J grows linearly with time (the green line in Fig. 3.3.) Extension of the trading period beyond a certain minimal length does not significantly increase the value function of a sufficiently risk-averse agent (the pink and the red lines on Fig. 3.3.)

The time value grows roughly exponentially in X_t^2 if there is an immediate investment opportunity.

3.3.4 Effect of risk-aversion on time inhomogeneity

The ratio $D(\tau)$ defined by Eq. (3.14) plays a crucial role in most of our formulas: it determines the position size in Eq. (3.15), the threshold at which we start unwinding a loosing position (Eq. (3.17)), and it also enters equations (3.16) for the value function and (3.18) for the covariance of J and X. Fig. 3.2 shows the graphs of $D(\tau)$ for different values of γ. Recall that Eq. (3.15) implies that for given wealth W_t and spread X_t, position size is proportional to $D(\tau)$.

We see that for $\gamma = 0$ (log-utility) the optimal position does not depend on time. For $\gamma > 0$ the agent is less risk-averse than a log-utility agent. So, for given price X_t and wealth W_t, his position *increases* as the final time approaches. In practice, traders often tend to become *less aggressive* as the bonus time approaches. This is consistent with the optimal behavior of a power utility agent with $\gamma < 0$. For example, assume that $k = 8$ and $\gamma = -2$ and let us measure the time in years. Then for the same wealth W and spread X, the position just a week before the year-end is a third lower than it is at the beginning of the year.

[2]The figure shows the graphs of the function $J(w,x,\tau) - \frac{1}{\gamma}$ for $w = 1$, $x = 0$, $0 \leq \tau \leq 1$ and several different values of γ. Substraction of $\frac{1}{\gamma}$ from the value function makes the comparison easier with the log-utility case $\gamma = 0$.

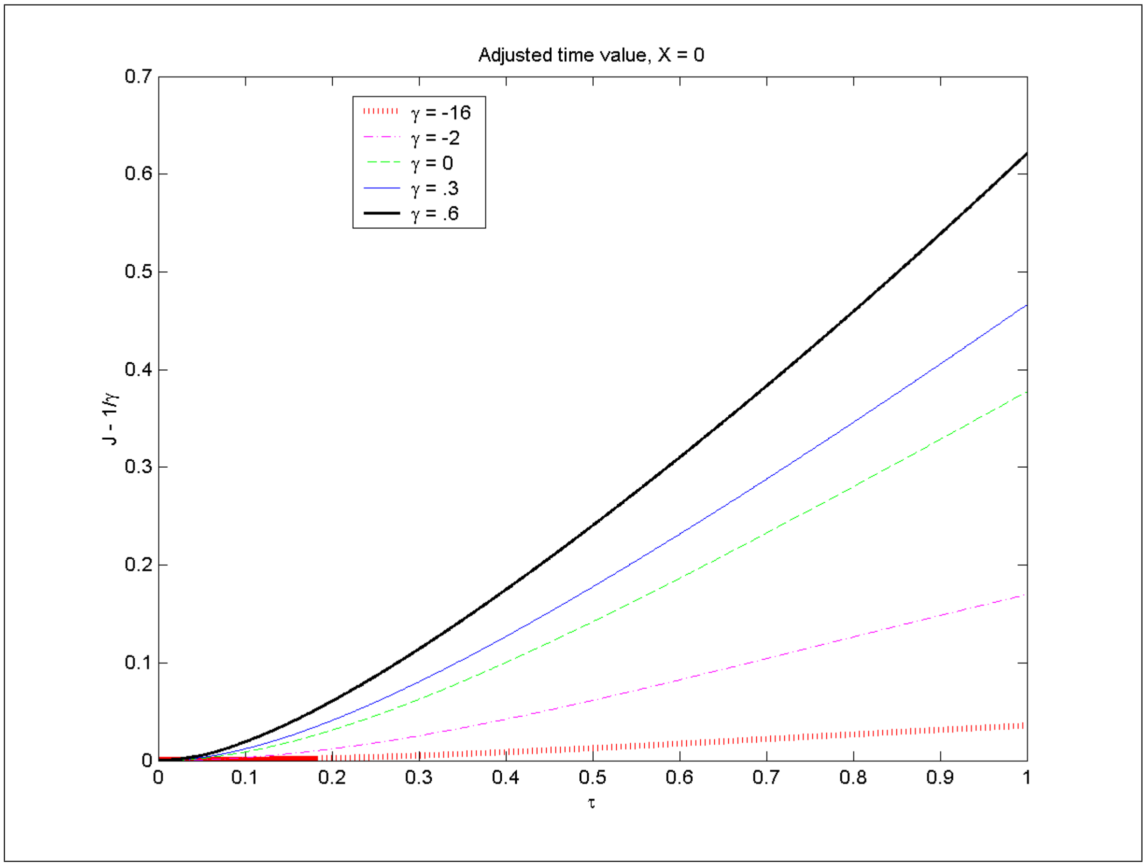

Figure 3.3: Adjusted time value, $X = 0$, $W = 1$.

3.3.5 Simulation results

To study the effect of parameter misspecification, we performed a Monte-Carlo simulation study. Fig. 3.4 shows a sample price trajectory with the corresponding optimal position and the wealth trajectories. A simulation without variance reduction also gives a good proxy to the discretisation and sampling errors, i.e. to the deviations of accumulated wealth from the predicted wealth due to the sampling error and non-continuous rehedging.

In reality, it is very hard to predict the mean-reversion parameter k. Even if we assume that the price series is stationary, k has to be estimated from the past data. Figure 3.5 shows the effect of trading with the wrong value of k.

In a Monte carlo simulation, we generated a set of Ornstein-Uhlenbeck process trajectories with $k = 2$, $\sigma = 1$ and then simulated trading with a wrong value of k. To look at the dependence of the optimal position α^* on the mean reversion coefficient K, it is convenient

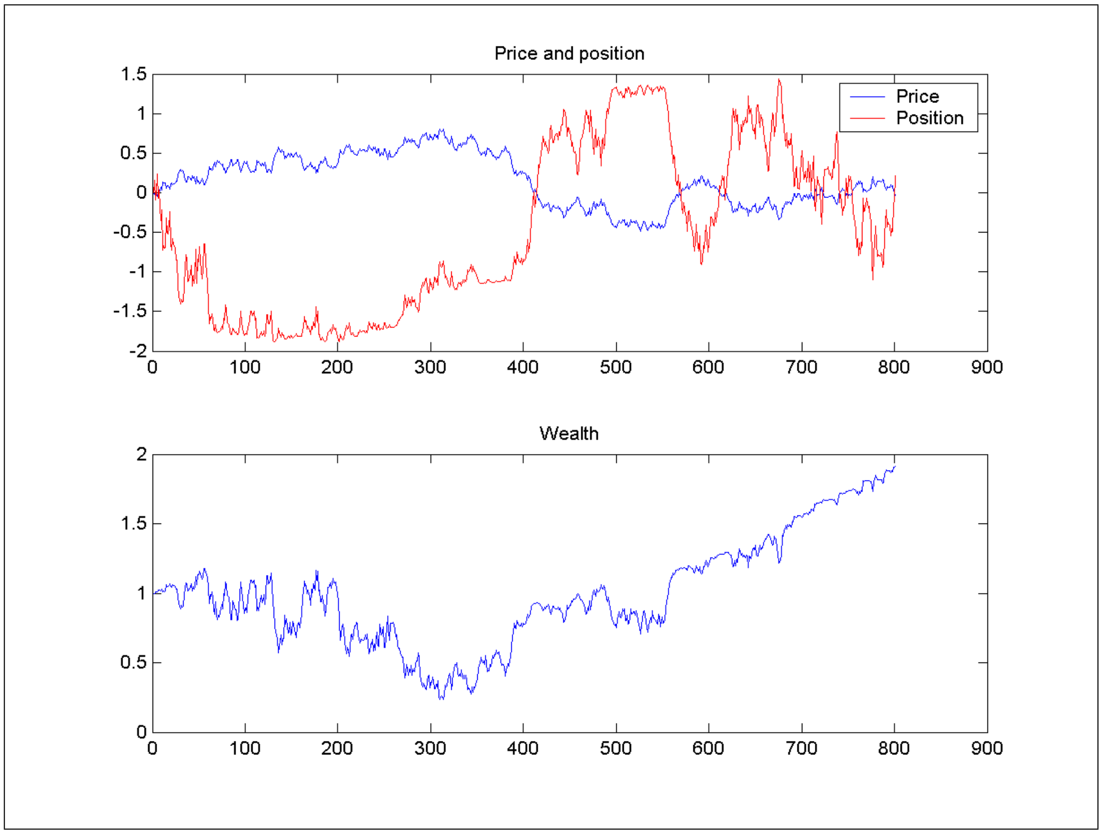

Figure 3.4: A simulated price sample together with position and wealth dynamics.

to invert the transforms (3.6) and to rewrite Eq. (3.15) as

$$\alpha = \frac{k}{\sigma^2} wxD(\tau/k). \tag{3.20}$$

Thus, we took K in the interval $(1, \ldots, 3.2)$ and simulated trading with position determined by (3.20), but with K substituted for k. On the horizontal axis of the graph we have $\log(K/k)$. Blue and red dashed lines show the two standard deviations confidence interval bounds for the mean terminal utility when trading with a given value of K. The black cross shows the value function from Eq. (3.16) for $K = k$.

We can see that the influence of misspecification of the mean reversion coefficient is asymmetric. Trading with a conservatively estimated k reduces greatly the utility uncertainty. Not surprisingly, overestimation of the mean reversion leads to excessively aggressive positions and big discretisation errors. It is much safer to underestimate k than to overestimate it.

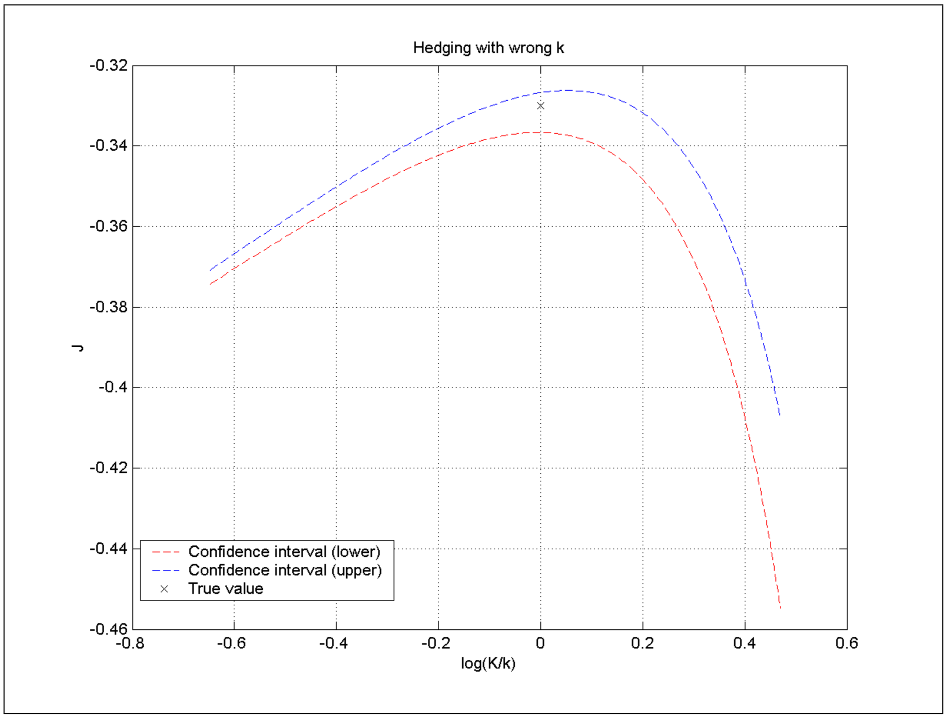

Figure 3.5: A simulated price sample together with position and wealth dynamics.

3.4 Conclusions and possible generalizations

We solved the optimal portfolio selection problem assuming that there is a single risky asset following an Ornstein-Uhlenbeck process with known parameters and there is a representative agent with given wealth, investment horizon, and power utility function. The other assumptions used were the absence of market frictions and perfect liquidity of the asset traded.

Most of these assumptions are similar to the ones made in the Black-Scholes model. Each of these assumptions is not quite realistic. Even when one manages to find a mean-reverting trading asset, one will need to estimate the parameters of the process. The prices usually seem to follow non-stationary processes, with periodic regime switches and jumps. Market frictions make continuous trading unviable, while the presence of other traders competing for the same trading opportunity and the feedback between trades and prices affect the optimal strategy. A trader usually does not commit all of his capital to trade a single asset, so the real-world problems involve multiple risky asset portfolio selection.

The model considered can be extended to include many of these more realistic features. The resulting PDE is not very likely to have an explicit solution, but singular perturbation theory may be used to obtain approximations by expansions around our solution. A similar problem in a discrete setting is considered in [Vigodner]. The discrete framework allows to introduce easily transaction costs but, in most cases, lacks explicit solutions. The attraction level of the mean reverting process X_t may be assumed not to be known *a priori* and to be inferred from observations of X_t. This problem can be treated in a Bayesian framework similarly to [Lakner].

On the other hand, our simple model can serve as a benchmark in practical situations. Quite often, practitioners prefer to introduce *ad hoc* corrections to a simple model rather than using a more involved model with a large number of parameters.

3.5 Appendix A.

3.5.1 A technical lemma

To prove the theorem we need the following lemma.

Lemma 3.5.1 *The functions $\alpha^* = \alpha^*(w,x,t)$ and $J = J(w,x,t)$ defined by Eqs. (3.15) and (3.16) have the following properties*

1. *$J = J(w,x,t)$ is a solution to Eq. (3.9);*
2. *boundary condition at T:*
$$J(w,x,T) = \frac{1}{\gamma}(w^\gamma - 1);$$
3. *concavity in current wealth:*
$$J_{ww} \leq 0, \text{ for all } w \geq 0,\ x \in \Re, 0 \leq t \leq T;$$
4. *α^* satisfies the first order optimality condition (3.8).*

Proof of the lemma. All properties can be checked by direct calculations. □

3.5.2 Proof of the theorem

Let $J(w,x,t,\alpha)$ be the expected terminal utility if the trader follows a particular strategy α. It is enough to show that $J(W_t,X_t,t)$ and $\alpha^*(W_t,X_t,t)$ given by Eqs. (3.15), (3.16) satisfy two standard stochastic optimal control conditions:

(A) For any control $\alpha = \alpha(w, x, t)$

$$J(w, x, t, \alpha) \le J(w, x, t) \text{ for all } x \in \Re, w \ge 0, 0 \le t \le T,$$

(B) The control $\alpha^* = \alpha^*(w, x, t)$ satisfies

$$J(w, x, t, \alpha^*) = J(w, x, t).$$

Condition (A). Applying Itô's formula to $J\left(W_s, X_s, s\right)_{t \le s \le T}$, we obtain

$$\begin{aligned} J(W_s, X_s, s) &= J(W_t, X_t, t) + \int_t^s \mathcal{L}(\alpha) J(W_u, X_u, u) du \\ &+ \int_t^s J_x(W_u, X_u, u) dB_u + \int_t^s \alpha_u J_w(W_u, X_u, u) dB_u, \end{aligned} \tag{3.21}$$

where

$$\mathcal{L}(\alpha) J = J_t - x J_x - \alpha x J_w + \frac{1}{2} J_{xx} + \frac{1}{2} \alpha^2 J_{ww} + \alpha J_{xw}. \tag{3.22}$$

Using the Lemma, we see that

$$\begin{aligned} &\mathcal{L}(\alpha) J = \\ &= \frac{1}{2} J_{ww} \left(\alpha - \left(x \frac{J_w}{J_{ww}} - \frac{J_{xw}}{J_{ww}} \right) \right)^2 + \left(J_t + \frac{1}{2} J_{xx} - x J_x - \frac{1}{2} J_{ww} \left(\frac{J_{xw}}{J_{ww}} - x \frac{J_w}{J_{ww}} \right)^2 \right) \\ &= \frac{1}{2} J_{ww} \left(\alpha - \left(x \frac{J_w}{J_{ww}} - \frac{J_{xw}}{J_{ww}} \right) \right)^2 \le 0. \end{aligned} \tag{3.23}$$

Taking the mathematical expectation $\mathbf{E}_t$ of $J(W_s, X_s, s)$ from (3.21) we obtain

$$\begin{aligned} \mathbf{E}_t J(W_s, X_s, s) &= J(W_t, X_t, t) + \mathbf{E}_t \int_t^s \mathcal{L}(\alpha) J(X_u, W_u, u) du \\ &+ \mathbf{E}_t \int_t^s J_x(W_u, X_u, u) dB_u + \mathbf{E}_t \int_t^s \alpha J_w(W_u, X_u, u) dB_u. \end{aligned} \tag{3.24}$$

The stochastic integrals in (3.24) are martingales, so the mathematical expectation of these integrals is zero. Thus, the last two summands in (3.24) vanish. Now let $t \to T$. Using (3.23), we can rewrite (3.24) as

$$\begin{aligned} J(W_t, X_t, t) &= \mathbf{E}_t J(W_T, X_T, T) - \mathbf{E}_t \int_t^T \mathcal{L}(\alpha) J(X_u, W_u, u) du \\ &\ge \mathbf{E}_t (\frac{1}{\gamma} W_T^\gamma) = J^\alpha(W_t, X_t, t, \alpha), \end{aligned}$$

i.e. condition (A) is satisfied.

Condition (B). It is clear from the Lemma that

$$\mathcal{L}(\alpha^*)J = 0.$$

So for $\alpha = \alpha^*$ we have

$$\begin{aligned} J(W_t, X_t, t) &= \mathbf{E}_t J(W_T, X_T, T) - \mathbf{E}_t \int_t^T \mathcal{L}(\alpha^*)J(X_u, W_u, u)du \\ &= \mathbf{E}_t(\frac{1}{\gamma} W_T^\gamma) = J^\alpha(W_t, X_t, t, \alpha^*), \end{aligned}$$

i.e. condition (B) is satisfied. This concludes the proof of the Theorem.

Bibliography

[Alvarez] Alvarez, L. H. R. (2001) Singular Stochastic Control, Linear Diffusions, and Optimal Stopping: A Class of Solvable Problems. *SIAM Journal on Control and Optimization Volume* **39** no. 6, 1697-1710.

[AS] Abramowitz, M.; Stegun, I. A. (eds.) (1972) *Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables.* (1972) National Bureau of Standards. 10th print. New York: Wiley-Inerscience, 1046 p.

[Bog1] Boguslavskaya, E. (1997) Explicit solution for one optimal control problem of a diffusion *Russ. Math.Surveys*, , (pages).

[Bog2] Boguslavskya, E. (2001) On optimization of long-term irreversible investments in a diffusion model. *Theory of Probability & Its Applications* **45** no. 4, 647-658.

[Bog3] Boguslavskaya, E. (2005) On optimization of dividend flow for a company in the presence of liquidation value. Submitted to *Math. Oper. Res.*

[Bog4] Boguslavsky, M.; Boguslavskaya, E. (2004) Optimal arbitrage trading,*Risk magazine*, no 6, 69-73; Reprinted in *Derivatives trading and option pricing*, ed. N. Dunbar, Risk books, 2005, pp 365-380.

[Lo] Campbell J.; Lo A.; MacKinley A. (1997) *The Econometrics of Financial Markets.* Princeton University Press, USA.

[Dixit] Dixit, A. (1992) Investment and hysteresis, *Journal of Economic Perspectives*, **6**, 107-132.

[DixPin] Dixit A. , Pindyck R. S. (1993) *Investment under uncertainty*, Princeton University Press, Princeton, NJ.

[DSS] Dubins, L.E.; Shepp, L. A.; Shiryaev, A. N. (1993) Optimal stopping rules and maximal inequalities for Bessel processes, *Theory Probab. Appl.* **38**, no.2, 226-261.

[Dynkin] Dynkin, E. (1960) *Theory of Markov processes*, Pergamon Press.

[Erd] Erdelyi, A.; Magnus, W.; Oberhettinger, F; Tricomi, F. G. (1981) *Higher transcendental functions. Vol. I. Based on notes left by Harry Bateman. Reprint of the 1955 original*, Robert E. Krieger Publishing Co. Inc., Melbourne, Fla., 302pp.

[FlemSoner] Fleming, W.; Soner, H.M. (1993) *Controlled Markov Processes and Viscosity Solutions*, Springer-Verlag.

[JacShir] Jacod, J.; Shiryaev, A. N. (1987) *Limit theorems for stochastic processes.* Berlin: Springer-Verlag.

[JeanShir] Jeanblanc, M.; Shiryaev, A.N. (1995) Optimization of the flow of dividends. *Russian Math. Surveys* **50**, no.2, 257-277.

[HojTaksar] Hojgaard, B.; Taksar, M. (2001) Optimal risk control for a large corporation in the presence of returns on investements, *Finance and Stochastics*, no. 5, 527-547.

[ItoMcKean] Ito, K.; McKean, H. (1965) *Diffusion processes and their sample paths*, Springer-verlag.

[KramMor] Kramkov, D.; Mordecki, E. (1994) An integral option, *Theory of Probability & Its Applications*, **39**, no. 1, 162 - 172.

[Krylov] Krylov, N.V. (1980) *Controlled diffusion processes*, Springer, New York.

[Lakner] Lakner, P. (1998) Optimal trading strategy for an investor: the case of partial information, *Stochastic Processes and their Applications*, **76**, 77–97.

[M-VMN] Mendez-Vivez, D.; Morton, A.; Naik, V. (2000) Trading Mean Reverting Process, unpublished manuscript.

[Merton] Merton, R.C. (1990) *Continuous-Time Finance*, Basil Blackwell Ltd.

[Morton] Morton, A. (2001) “When in trouble – double”, presentation at Global Derivatives 2001 conference, St.Juan - les Pins.

[Oks] Oksendal, B. (2003) *Stochastic differential equations: An introduction with applications.* Sixth edition, Springer.

[Pind1] Pindyck, R. S. (1993) Investments of uncertain cost, *Journal of Financial Economics*, **34**, 53-76, North Holland.

[Pind2] Pindyck, R. S. (1991) Irreversibility, uncertainty, and investment,*Journal of Economic Literature*, **29**, 1110-1152.

[Radner] Radner, R.; Shepp, L. (1996) Risk vs. profit potential: A model for corporate strategy. *J. Econ. Dynam. Control* **20**, no. 8, 1373-1393.

[RevuzYor] Revuz, D.; Yor, M. (1991) *Continuous Martingales and Brownian Motion* Springer-Verlag, Berlin.

[Shl] Schleifer, A. (2000) *Inefficient Markets: An Introduction to Behavioural Finance*, Oxford Univ. Press, Oxford.

[Shiryaev] Shiryaev, A. N. (1999) *Essentials of Stochastic Finance: Facts, Models, Theory*, World Scientific Publishing Company.

[Taksar] Taksar, M. (2000) Optimal Risk and dividend distribution control models for an insurance company. *Math. Methods of Operational Research* no. 1, 1-42.

[Vigodner] Vigodner, A. (2000) Dynamic Programming and Optimal Lookahead Strategies in High Frequency Trading with Transaction Costs, preprint.

[ZaiPol] Polyanin, A.D.; Zaitsev, V.F. (1995) *Handbook of exact solutions for ordinary differential equations.* CRC Press, Boca Raton, FL.

Curriculum Vitae

Elena Boguslavskaya was born in Moscow on March 14, 1973. She received her high school diploma from School 57 in Moscow in 1990. In the same year she began her studies at the department of Mechanics and Mathematics of the University of Moscow from which she graduated in June 1996. In November 1996, she became PhD student at the Steklov Institute in Moscow. From February 2000 till June 2002 she was working as a bursaal at Universiteit van Amsterdam. At present she is working as a researcher at City University, London.

Nederlandse Samenvatting

In dit proefschrift beschouwen we verscheidene controleproblemen voor diffusieprocessen, waarbij het de bedoeling is optimale controlestrategieën te vinden. De stochastische processen die we in Hoofdstuk 1 bestuderen, modelleren het uitkeren van dividend van een bedrijf. De processen in Hoofdstuk 2 corresponderen met het investeren in een project. De optimale oplossingen voor de problemen uit beide hoofdstukken zijn van het singuliere type, d.w.z. ze hebben een "alles of niets" karakter. De processen die we in Hoofdstuk 3 onder de loep nemen, zijn modellen voor het prijsgedrag van aandelen, waarbij dan een optimale handelsstrategie het doel is. In tegenstelling tot de problemen die in de Hoofdstukken 2 en 3 worden bestudeerd, is de optimale handelsstrategie in Hoofdstuk 3 van het reguliere type, d.w.z. de optimale controlestrategie is hier "glad". In dit proefschrift worden *expliciete* oplossingen voor de genoemde optimale controleproblemen in ééndimensionale diffusieprocessen gevonden.

In Hoofdstuk 1 beschouwen we een model voor een bedrijf waarvan de reserve $X = (X_t)_{t\geq 0}$ zich ontwikkelt volgens de stochastische differentiaalvergelijking

$$dX_t = \mu dt + \sigma dW_t - dZ_t, \tag{3.25}$$

waarbij $W = (W_t)_{t\geq 0}$ een standaard Wiener proces is en waarbij μ en σ bekende positieve constanten zijn. Het controleproces $Z = (Z_t)_{t\geq 0}$ stelt de cumulatieve hoeveelheid dividend voor die is uitbetaald tot en met tijdstip t. De belangrijkste eisen die aan het controle proces $Z = (Z_t)_{t\geq 0}$ worden gesteld, zijn dat het proces niet-negatief en niet-dalend is en dat het is aangepast aan de filtratie, d.w.z. dat het slechts gebaseerd is op het verleden. Het tijdstip τ waarop het bedrijf bankroet gaat, wordt gedefinieerd als $\tau = \inf\{t \geq 0 : X_t \leq 0\}$. Er wordt aangenomen dat het startkapitaal x_0 positief is en dat de liquidatiewaarde S, d.w.z. de waarde van de activa van het bedrijf op het moment van faillissement, niet-negatief is. Bij een constant rentepercentage λ is de totale verwachte uitbetaling aan de aandeelhouders tezamen met de geïndiceerde liquidatiewaarde gelijk aan

$$V(x, Z) = \mathbf{E}_x\{\int_0^\tau e^{-\lambda t} dZ_t + Se^{-\lambda\tau}\}. \tag{3.26}$$

We geven expliciete formules voor het optimale, toegestane controle proces $\tilde{Z}$, met andere woorden het controle proces dat de waarde $V(x, Z)$ maximaliseert, en wel voor drie gevallen:

1. het geval van een begrensde dividenduitkeringssnelheid,

2. het geval van discrete dividenduitkeringen met transactiekosten,

3. het algemene geval waarin het dividend proces elk niet-negatief, niet-dalend, rechtscontinu proces mag zijn.

In Hoofdstuk 2 bestuderen we een model voor onomkeerbare lange-termijn investeringen. Hierbij worden de op tijdstip t voorziene kosten om het project af te ronden gegeven door de stochastische differentiaalvergelijking

$$dX_t = -I_t dt + \beta\sqrt{X_t I_t} dW_t + \gamma X_t d\tilde{W}_t, \tag{3.27}$$

waarbij W_t en $\tilde{W}_t$ ongecorreleerde standaard Wiener processen zijn en β en γ niet-negatieve constanten en waarbij $I(X_t)$ de investeringssnelheid voorstelt. Het tijdstip τ waarop het project wordt voltooid, wordt gedefinieerd als $\tau = \inf\{t \geq 0 : X_t \leq 0\}$. Er wordt aangenomen dat de initieel voorziene kosten om het project af te ronden, $X_0 = x$, positief zijn en dat de waarde van het project na afronding wordt gegeven door een positieve constante V. Bij constante rente r is de totale verwachte met het project te behalen winst gelijk aan

$$F(x, I) = \mathbf{E}\{\int_0^\tau (-I(X_t)e^{-rt})dt + Ve^{-r\tau}\}. \tag{3.28}$$

We presenteren expliciete uitdrukkingen voor de optimale investeringssnelheid I^*, d.w.z. de investeringssnelheid die de verwachte winst $F(x, I)$ maximaliseert, en wel voor een drietal gevallen:

1. technische onzekerheid ($\beta = 0, \gamma \neq 0$),

2. onzekerheid over de investeringskosten ($\beta \neq 0, \gamma = 0$),

3. beide onzekerheden aanwezig ($\beta \neq 0, \gamma \neq 0$).

In Hoofdstuk 3 bekijken we het probleem van de positiebepaling voor een handelaar in een mean-reverting aandeel. Dit probleem komt in veel handelssituaties voor met statistische en fundamentele arbitrage wanneer de korte termijn inkomsten op een aandaal voorspelbaar zijn, maar wanneer een beperkt risicodragend vermogen de handelaar verhindert deze voorspelbaarheid volledig uit te buiten. We gebruiken een Ornstein-Uhlenbeck proces om het prijsproces $X = (X_t)_{t\geq 0}$ te modelleren:

$$dX_t = -kX_t dt + \sigma dB_t,$$

waarbij k en σ positieve constanten zijn en B_t een standaard Wiener proces is. Dit model reproduceert enkele realistische patronen in handelaargedrag. Het controleproces α_t representeert de positie van de handelaar op tijdstip t, d.w.z. het aantal aandelen dat wordt aangehouden. Wanneer we de rente op nul stellen en aannemen dat er geen wrijving is in de markt, wordt de dynamiek van het vermogen bij een zeker controleproces α_t gegeven door

$$dW_t = \alpha_t dX_t = -k\alpha_t X_t dt + \alpha_t \sigma dB_t.$$

We nemen aan dat er geen beperkingen zijn op α en dat dus in het bijzonder short gaan is toegestaan. We leggen ook geen beperkingen op aan het startkapitaal W. We gebruiken de nutsfunctie $\Psi(w) = (w^{\gamma} - 1)/\gamma$, $w \geq 0$, voor zekere $\gamma \in (-\infty, 1)$. Het verwachte nut op tijdstip T, voorwaardelijk gegeven de informatie beschikbaar op tijdstip t, is dan

$$J(W_t, X_t, t) = \mathbf{E}_t \Psi(W_T).$$

We presenteren een expliciete uitdrukking voor de optimale positie α^*, d.w.z. de positie die $J(W_t, X_t, t)$ maximaliseert.