# Universitext

Bernt Øksendal · Agnès Sulem

# Applied Stochastic Control of Jump Diffusions

With 24 Figures

Springer

Bernt Øksendal
University of Oslo
Center of Mathematics for Applications (CMA)
Department of Mathematics
0316 Oslo
Norway
*e-mail:* oksendal@math.uio.no

Agnès Sulem
INRIA Rocquencourt
Domaine de Voluceau
78153 Le Chesnay Cedex
France
*e-mail:* agnes.sulem@inria.fr

Cover figure is taken from page 91.

To my family

Eva, Elise, Anders and Karina

B. Ø.

A tous ceux qui m'accompagnent

A. S.

# Preface

Jump diffusions are solutions of stochastic differential equations driven by
Lévy processes. Since a Lévy process $\eta(t)$ can be written as a linear combina-
tion of $t$, a Brownian motion $B(t)$ and a pure jump process, jump diffusions
represent a natural and useful generalization of Itô diffusions. They have re-
ceived a lot of attention in the last years because of their many applications,
particularly in economics.

There exist today several excellent monographs on Lévy processes. How-
ever, very few of them - if any - discuss the optimal control, optimal stopping
and impulse control of the corresponding jump diffusions, which is the subject
of this book. Moreover, our presentation differs from these books in that it
emphazises the applied aspect of the theory. Therefore we focus mostly on
useful verification theorems and we illustrate the use of the theory by giving
examples and exercises throughout the text. Detailed solutions of some of the
exercises are given in the end of the book. The exercices to which a solution
is provided, are marked with an asterix $*$. It is our hope that this book will
fill a gap in the literature and that it will be a useful text for students, re-
searchers and practitioners in stochastic analysis and its many applications.
Although most of our results are motivated by examples in economics and
finance, the results are general and can be applied in a wide variety of sit-
uations. To emphasize this, we have also included examples in biology and
physics/engineering.

This book is partially based on courses given at the Norwegian School of
Economics and Business Administration (NHH) in Bergen, Norway, during the
Spring semesters 2000 and 2002, at INSEA in Rabat, Morocco in September
2000, at Odense University in August 2001 and at ENSAE in Paris in February
2002.

**Acknowledgments**   We are grateful to many people who in various ways have contributed to these lecture notes. In particular we thank Knut Aase, Fred Espen Benth, Jean-Philippe Chancelier, Rama Cont, Hans Marius Eikseth, Nils Christian Framstad, Jørgen Haug, Monique Jeanblanc, Kenneth Karlsen, Thilo Meyer-Brandis, Cloud Makasu, Sure Mataramvura, Peter Tankov and Jan Ubøe for their valuable help. We also thank Francesca Biagini for useful comments and suggestions to the text and her detailed solutions of some of the exercises. We are grateful to Dina Haraldsson and Martine Verneuille for proficient typing and Eivind Brodal for his kind assistance. We acknowledge with gratitude the support by the French-Norwegian cooperation project Stochastic Control and Applications, Aur 99–050.

Oslo and Paris,  August 2004                     Bernt Øksendal  and  Agnès Sulem

# Contents

# 1
## Stochastic Calculus with Jump diffusions

### 1.1 Basic definitions and results on Lévy Processes

In this chapter we present the basic concepts and results needed for the applied calculus of jump diffusions. Since there are several excellent books which give a detailed account of this basic theory, we will just briefly review it here and refer the reader to these books for more information.

**Definition 1.1.** *Let $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t \geq 0}, P)$ be a filtered probability space. An $\mathcal{F}_t$-adapted process $\{\eta(t)\}_{t \geq 0} = \{\eta_t\}_{t \geq 0} \subset \mathbb{R}$ with $\eta_0 = 0$ a.s. is called a Lévy process if $\eta_t$ is continuous in probability and has stationary, independent increments.*

**Theorem 1.2.** *Let $\{\eta_t\}$ be a Lévy process. Then $\eta_t$ has a cadlag version (right continuous with left limits) which is also a Lévy process.*

*Proof.* See e.g. [P], [S]. □

In view of this result we will from now on assume that the Lévy processes we work with are cadlag.

The *jump* of $\eta_t$ at $t \geq 0$ is defined by

$$\Delta \eta_t = \eta_t - \eta_{t-} \ . \tag{1.1.1}$$

Let $\mathbf{B}_0$ be the family of Borel sets $U \subset \mathbb{R}$ whose closure $\bar{U}$ does not contain 0. For $U \in \mathbf{B}_0$ we define

$$N(t, U) = N(t, U, \omega) = \sum_{s : 0 < s \leq t} \mathcal{X}_U(\Delta \eta_s) \ . \tag{1.1.2}$$

In other words, $N(t, U)$ is the number of jumps of size $\Delta \eta_s \in U$ which occur before or at time $t$. $N(t, U)$ is called the *Poisson random measure (or jump measure)* of $\eta(\cdot)$. The differential form of this measure is written $N(dt, dz)$.

**Remark 1.3.** Note that $N(t, U)$ is *finite* for all $U \in \mathbf{B}_0$. To see this we proceed as follows: Define

$$T_1(\omega) = \inf\{t > 0; \eta_t \in U\}$$

We claim that $T_1(\omega) > 0$ a.s. To prove this note that by right continuity of paths we have

$$\lim_{t \to 0^+} \eta(t) = \eta(0) = 0 \quad \text{a.s.}$$

Therefore, for all $\varepsilon > 0$ there exists $t(\varepsilon) > 0$ such that $|\eta(t)| < \varepsilon$ for all $t < t(\varepsilon)$. This implies that $\eta(t) \notin U$ for all $t < t(\varepsilon)$, if $\varepsilon < \text{dist}(0, U)$.

Next define inductively

$$T_{n+1}(\omega) = \inf\{t > T_n(\omega); \Delta\eta_t \in U\}.$$

Then by the above argument $T_{n+1} > T_n$ a.s. We claim that

$$T_n \to \infty \quad \text{as } n \to \infty, \text{ a.s.}$$

Assume not. Then $T_n \to T < \infty$. But then

$$\lim_{s \to T^-} \eta(s) \quad \text{cannot exist.}$$

contradicting the existence of left limits of the paths.

It is well-known that Brownian motion $\{B(t)\}_{t \geq 0}$ has stationary and independent increments. Thus $B(t)$ is a Lévy process. Another important example is the following:

**Example 1.4 (The Poisson process).** The Poisson process $\pi(t)$ of intensity $\lambda > 0$ is a Lévy process taking values in $\mathbb{N} \cup \{0\}$ and such that

$$P[\pi(t) = n] = \frac{(\lambda t)^n}{n!} e^{-\lambda t} \; ; \qquad n = 0, 1, 2, \ldots$$

**Theorem 1.5.** *[P, Theorem 1.35].*
**(i)** *The set function $U \to N(t, U, \omega)$ defines a $\sigma$-finite measure on $\mathbf{B}_0$ for each fixed $t, \omega$.*

**(ii)** *The set function*

$$\nu(U) = E[N(1, U)] \tag{1.1.3}$$

*where $E = E_P$ denotes expectation with respect to $P$, also defines a $\sigma$-finite measure on $\mathbf{B}_0$, called the Lévy measure of $\{\eta_t\}$.*

**(iii)** *Fix $U \in \mathbf{B}_0$. Then the process*

$$\pi_U(t) := \pi_U(t, \omega) := N(t, U, \omega)$$

*is a Poisson process of intensity $\lambda = \nu(U)$.*

**Example 1.6 (The compound Poisson process ).** Let $X(n); n \in \mathbb{N}$ be a sequence of i.i.d. random variables taking values in $\mathbb{R}$ with common distribution $\mu_{X(1)} = \mu_X$ and let $\pi(t)$ be a Poisson process of intensity $\lambda$, independent of all the $X(n)$'s.

The *compound Poisson process* $Y(t)$ is defined by

$$Y(t) = X(1) + \cdots + X(\pi(t)) ; \qquad t \geq 0 . \tag{1.1.4}$$

An increment of this process is given by

$$Y(s) - Y(t) = \sum_{k=\pi(t+1)}^{\pi(s)} X(k) ; \qquad s > t .$$

This is independent of $X(1), \ldots, X(\pi(t))$, and depends only on the difference $(s - t)$. Thus $Y(t)$ is a Lévy process.

To find the Lévy measure $\nu$ of $Y(t)$ note that if $U \in \mathbf{B}_0$ then

$$\nu(U) = E[N(1, U)] = E\Big[ \sum_{s; 0 \leq s \leq 1} \mathcal{X}_U(\Delta Y(s)) \Big]$$

$$= E[(\text{number of jumps}) \cdot \mathcal{X}_U(\text{jump})] = E[\pi(1)\mathcal{X}_U(X)] = \lambda\mu_X(U) ,$$

by independence. We conclude that

$$\nu = \lambda\mu_X . \tag{1.1.5}$$

This shows that a Lévy process can be represented by a compound Poisson process if and only if its Lévy measure is finite. Note, however, that there are many interesting Lévy processes with infinite Lévy measure. See e.g. [B].

**Theorem 1.7 (Lévy decomposition [JS]).** *Let $\{\eta_t\}$ be a Lévy process. Then $\eta_t$ has the decomposition*

$$\eta_t = \alpha t + \beta B(t) + \int_{|z|<R} z\widetilde{N}(t, dz) + \int_{|z|\geq R} zN(t, dz) , \tag{1.1.6}$$

*for some constants $\alpha \in \mathbb{R}$, $\beta \in \mathbb{R}$, $R \in [0, \infty]$. Here*

$$\widetilde{N}(dt, dz) = N(dt, dz) - \nu(dz)dt \tag{1.1.7}$$

*is the* compensated Poisson random measure *of $\eta(\cdot)$ and $B(t)$ is an independent Brownian motion. For each $A \in \mathbf{B}_0$ the process*

$$M_t := \widetilde{N}(t, A) \quad \textit{is a martingale.} \tag{1.1.8}$$

*If $\alpha = 0$ and $R = \infty$, we call $\eta_t$ a Lévy martingale .*

**Theorem 1.8.** *We can always choose $R = 1$. If*

$$E[\eta_t] < \infty \qquad \text{for all } t \geq 0$$

*then we may choose $R = \infty$ and hence write*

$$\eta_t = \alpha t + \beta B(t) + \int_{\mathbb{R}} z \widetilde{N}(t, dz).$$

*(See [S, Theorem 25.3]).*

**Theorem 1.9.** *[P]. A Lévy process is a strong Markov process.*

**Theorem 1.10 (The Lévy-Khintchine formula [P]).** *Let $\{\eta_t\}$ be a Lévy process with Lévy measure $\nu$. Then $\int_{\mathbb{R}} \min(1, z^2)\nu(dz) < \infty$ and*

$$E[e^{iu\eta_t}] = e^{t\psi(u)} , \qquad u \in \mathbb{R} \tag{1.1.9}$$

*where*

$$\psi(u) = -\tfrac{1}{2}\sigma^2 u^2 + i\alpha u + \int_{|z|<R} \{e^{iuz} - 1 - iuz\}\nu(dz) + \int_{|z|\geq R} (e^{iuz} - 1)\nu(dz) .$$
$$\tag{1.1.10}$$

*Conversely, given constants $\alpha, \sigma^2$ and a measure $\nu$ on $\mathbb{R}$ s.t.*

$$\int_{\mathbb{R}} \min(1, z^2)\nu(dz) < \infty ,$$

*there exists a Lévy process $\eta(t)$ (unique in law) such that (1.1.9–1.1.10) hold.*

**Note:** It is possible that $\int_{|z|\leq R} |z|\nu(dz) = \infty$.

**Theorem 1.11.** *[P, Corollary p. 48]. A Lévy process is a* semimartingale.

**Definition 1.12.** *[P]. Let $\mathbf{D}_{\text{ucp}}$ denote the space of cadlag adapted processes, equipped with the topology of uniform convergence on compacts in probability* (ucp) : $H_n \to H$ *ucp if for all $t > 0$* $\sup_{0 \leq s \leq t} |H_n(s) - H(s)| \to 0$ *in probability* $(A_n \to A$ *in probability if for all $\varepsilon > 0$ there exists $n_\varepsilon \in \mathbb{N}$ such that $n \geq n_\varepsilon \Rightarrow$* Prob.$(|A_n - A| > \varepsilon) < \varepsilon)$.

*Let $\mathbf{L}_{\text{ucp}}$ denote the space of adapted* caglad *processes (left continuous with right limits), equipped with the* ucp *topology. If $H(t)$ is a step function of the form*

$$H(t) = H_0 \mathcal{X}_{\{0\}}(t) + \sum_i H_i \mathcal{X}_{(T_i, T_{i+1}]}(t) ,$$

where $H_i \in \mathcal{F}_{T_i}$ and $0 = T_0 \le T_1 \le \cdots \le T_{n+1} < \infty$ are $\mathcal{F}_t$-stopping times and $X$ is cadlag, we define

$$J_X H(t) := \int_0^t H_s dX_s := H_0 X_0 + \sum_i H_i (X_{T_{i+1} \wedge t} - X_{T_i \wedge t}) \,; \qquad t \ge 0.$$

**Theorem 1.13.** *[P, p. 51]. Let $X$ be a semimartingale. Then the mapping $J_X$ can be extended to a continuous linear map*

$$J_X : \mathbf{L}_{\mathrm{ucp}} \to \mathbf{D}_{\mathrm{ucp}} \,.$$

This construction allows us to define stochastic integrals of the form

$$\int_0^t H(s) d\eta_s$$

for all $H \in \mathbf{L}_{\mathrm{ucp}}$. (See also Remark 1.18). In view of the decomposition (1.1.6) this integral can be split into integrals with respect to $ds$, $dB(s)$, $\widetilde{N}(ds, dz)$ and $N(ds, dz)$. This makes it natural to consider the more general stochastic integrals of the form

$$X(t) = X(0) + \int_0^t \alpha(s, \omega) ds + \int_0^t \beta(s, \omega) dB(s) + \int_0^t \int_{\mathbb{R}} \gamma(s, z, \omega) \bar{N}(ds, dz)$$

$$(1.1.11)$$

where the integrands are satisfying the appropriate conditions for the integrals to exist and we for simplicity have put

$$\bar{N}(ds, dz) = \begin{cases} N(ds, dz) - \nu(dz) ds & \text{if } |z| < R \\ N(ds, dz) & \text{if } |z| \ge R \,, \end{cases}$$

with $R$ as in Theorem 1.7. As is customary we will use the following short hand differential notation for processes $X(t)$ satisfying (1.1.11):

$$dX(t) = \alpha(t) dt + \beta(t) dB(t) + \int_{\mathbb{R}} \gamma(t, z) \bar{N}(dt, dz) \,. \qquad (1.1.12)$$

We call such processes *Itô-Lévy processes* .

## 1.2 The Itô formula and related results

We now come to the important Itô formula for Itô-Lévy processes:

If $X(t)$ is given by (1.1.12) and $f : \mathbb{R}^2 \to \mathbb{R}$ is a $C^2$ function, is the process $Y(t) := f(t, X(t))$ again an Itô-Lévy process and if so, how do we represent it in the form (1.1.12)?

If we argue heuristically and use our knowledge of the classical Itô formula it is easy to guess what the answer is:

Let $X^{(c)}(t)$ be the continuous part of $X(t)$, i.e. $X^{(c)}(t)$ is obtained by removing the jumps from $X(t)$. Then an increment in $Y(t)$ stems from an increment in $X^{(c)}(t)$ plus the jumps (coming from $N(\cdot, \cdot)$). Hence in view of the classical Itô formula we would guess that

$$
dY(t) = \frac{\partial f}{\partial t}(t, X(t))dt + \frac{\partial f}{\partial x}(t, X(t))dX^{(c)}(t) + \frac{1}{2}\frac{\partial^2 f}{\partial x^2}(t, X(t)) \cdot \beta^2(t)dt
$$
$$
+ \int_{\mathbb{R}} \{f(t, X(t^-) + \gamma(t, z)) - f(t, X(t^-))\}N(dt, dz) .
$$

It can be proved that our guess is correct. Since

$$
dX^{(c)}(t) = \left(\alpha(t) - \int_{|z| < R} \gamma(t, z)\nu(dz)\right)dt + \beta(t)dB(t) ,
$$

this gives the following result:

**Theorem 1.14 (The 1-dimensional Itô formula [BL], [A], [P]).** *Suppose $X(t) \in \mathbb{R}$ is an Itô-Lévy process of the form*

$$
dX(t) = \alpha(t, \omega)dt + \beta(t, \omega)dB(t) + \int_{\mathbb{R}} \gamma(t, z, \omega)\bar{N}(dt, dz) , \qquad (1.2.1)
$$

*where*

$$
\bar{N}(dt, dz) = \begin{cases} N(dt, dz) - \nu(dz)dt & if \quad |z| < R \\ N(dt, dz) & if \quad |z| \geq R \end{cases} \qquad (1.2.2)
$$

*for some $R \in [0, \infty]$.*

*Let $f \in C^2(\mathbb{R}^2)$ and define $Y(t) = f(t, X(t))$. Then $Y(t)$ is again an Itô-Lévy process and*

$$
dY(t) = \frac{\partial f}{\partial t}(t, X(t))dt + \frac{\partial f}{\partial x}(t, X(t))[\alpha(t, \omega)dt + \beta(t, \omega)dB(t)]
$$
$$
+ \frac{1}{2}\beta^2(t, \omega)\frac{\partial^2 f}{\partial x^2}(t, X(t))dt
$$
$$
+ \int_{|z| < R} \Big\{f(t, X(t^-) + \gamma(t, z)) - f(t, X(t^-))
$$
$$
- \frac{\partial f}{\partial x}(t, X(t^-))\gamma(t, z)\Big\}\nu(dz)dt
$$
$$
+ \int_{\mathbb{R}} \{f(t, X(t^-) + \gamma(t, z)) - f(t, X(t^-))\}\bar{N}(dt, dz) . \qquad (1.2.3)
$$

**Note:**  If $R = 0$ then $\bar{N} = N$ everywhere
If $R = \infty$ then $\bar{N} = \widetilde{N}$ everywhere.

**Example 1.15 (The geometric Lévy process ).** Consider the stochastic differential equation

$$dX(t) = X(t^-)\Big[\alpha dt + \beta dB(t) + \int_{\mathbb{R}} \gamma(t, z)\bar{N}(dt, dz)\Big], \qquad (1.2.4)$$

where $\alpha, \beta$ are constants and $\gamma(t, z) \geq -1$. To find the solution $X(t)$ of this equation we rewrite it as follows:

$$\frac{dX(t)}{X(t^-)} = \alpha dt + \beta dB(t) + \int_{\mathbb{R}} \gamma(t, z)\bar{N}(dt, dz) \ .$$

Now define

$$Y(t) = \ln X(t) \ .$$

Then by Itô's formula,

$$dY(t) = \frac{X(t)}{X(t)}[\alpha dt + \beta dB(t)] - \tfrac{1}{2}\beta^2 X^{-2}(t)X^2(t)dt$$

$$+ \int_{|z|<R} \{\ln(X(t^-) + \gamma(t, z)X(t^-)) - \ln(X(t^-))$$

$$- X^{-1}(t^-)\gamma(t, z)X(t^-)\}\nu(dz)dt$$

$$+ \int_{\mathbb{R}} \{\ln(X(t^-) + \gamma(t, z)X(t^-)) - \ln(X(t^-))\}\bar{N}(dt, dz)$$

$$= \big(\alpha - \tfrac{1}{2}\beta^2\big)dt + \beta dB(t) + \int_{|z|<R} \{\ln(1+\gamma(t, z)) - \gamma(t, z)\}\nu(dz)dt$$

$$+ \int_{\mathbb{R}} \ln(1+\gamma(t, z))\bar{N}(dt, dz).$$

Hence

$$Y(t) = Y(0) + \big(\alpha - \tfrac{1}{2}\beta^2\big)t + \beta B(t) + \int_0^t \int_{|z|<R} \{\ln(1 + \gamma(s, z)) - \gamma(s, z)\}\nu(dz)ds$$

$$+ \int_0^t \int_{\mathbb{R}} \ln(1+\gamma(s, z))\bar{N}(ds, dz)$$

and this gives the solution

$$X(t) = X(0) \exp \left\{ \left( \alpha - \tfrac{1}{2}\beta^2 \right)t + \beta B(t) \right.$$

$$+ \int_0^t \int_{|z|<R} \{\ln(1 + \gamma(s,z)) - \gamma(s,z)\}\nu(dz)ds$$

$$\left. + \int_0^t \int_{\mathbb{R}} \ln(1 + \gamma(s,z))\bar{N}(ds,dz) \right\}. \tag{1.2.5}$$

In analogy with the diffusion case ($N = 0$) we call this process $X(t)$ a *geometric Lévy process*. It is often used as a model for stock prices. See e.g. [B].

Next we formulate the corresponding multi-dimensional version of Theorem 1.14:

**Theorem 1.16 (The multi-dimensional Itô formula).** *Let $X(t) \in \mathbb{R}^n$ be an Itô-Lévy process of the form*

$$dX(t) = \alpha(t,\omega)dt + \sigma(t,\omega)dB(t) + \int_{\mathbb{R}^n} \gamma(t,z,\omega)\bar{N}(dt,dz) , \tag{1.2.6}$$

*where $\alpha : [0,T] \times \Omega \to \mathbb{R}^n$, $\sigma : [0,T] \times \Omega \to \mathbb{R}^{n \times m}$ and $\gamma : [0,T] \times \mathbb{R}^n \times \Omega \to \mathbb{R}^{n \times \ell}$ are adapted processes such that the integrals exist. Here $B(t)$ is an $m$-dimensional Brownian motion and*

$$\bar{N}(dt,dz)^T = (\bar{N}_1(dt,dz_1), \dots, \bar{N}_\ell(dt,dz_\ell))$$
$$= (N_1(dt,dz_1) - \mathcal{X}_{|z_1|<R_1}\nu_1(dz_1)dt, \dots, N_\ell(dt,dz_\ell) - \mathcal{X}_{|z_\ell|<R_\ell}\nu_\ell(dz_\ell)dt) ,$$

*where $\{N_j\}$ are independent Poisson random measures with Lévy measures $\nu_j$ coming from $\ell$ independent (1-dimensional) Lévy processes $\eta_1, \dots, \eta_\ell$.*

*Note that each column $\gamma^{(k)}$ of the $n \times \ell$ matrix $\gamma = [\gamma_{ij}]$ depends on $z$ only through the $k^{\text{th}}$ coordinate $z_k$, i.e.*

$$\gamma^{(k)}(t,z,\omega) = \gamma^{(k)}(t,z_k,\omega) ; \ z = (z_1, \cdots, z_\ell) \in \mathbb{R}^\ell.$$

*Thus the integral on the right of (1.2.6) is just a shorthand matrix notation. When written out in detail component number $i$ of $X(t)$ in (1.2.6), $X_i(t)$, gets the form*

$$dX_i(t) = \alpha_i(t,\omega)dt + \sum_{j=1}^m \sigma_{ij}(t,\omega)dB_j(t)$$
$$+ \sum_{j=1}^\ell \int_{\mathbb{R}} \gamma_{ij}(t,z_j,\omega)\bar{N}_j(dt,dz_j) ; \ 1 \le i \le n. \tag{1.2.7}$$

*Let $f \in C^{1,2}([0,T]) \times \mathbb{R}^n; \mathbb{R})$. Put $Y(t) = f(t, X(t))$. Then*

$$dY(t) = \frac{\partial f}{\partial t}dt + \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(\alpha_i dt + \sigma_i dB(t)) + \frac{1}{2} \sum_{i,j=1}^{n} (\sigma\sigma^T)_{ij} \frac{\partial^2 f}{\partial x_i \partial x_j}dt$$

$$+ \sum_{k=1}^{\ell} \int_{|z_k|<R_k} \left\{ f(t, X(t^-) + \gamma^{(k)}(t, z_k)) - f(t, X(t^-)) \right.$$

$$\left. - \sum_{i=1}^{n} \gamma_i^{(k)}(t, z_k) \frac{\partial f}{\partial x_i}(X(t^-)) \right\} \nu_k(dz_k)dt$$

$$+ \sum_{k=1}^{\ell} \int_{\mathbb{R}} \{ f(t, X(t^-) + \gamma^{(k)}(t, z_k)) - f(t, X(t^-)) \} \bar{N}_k(dt, dz_k)$$

$$(1.2.8)$$

*where $\gamma^{(k)} \in \mathbb{R}^n$ is column number $k$ of the $n \times \ell$ matrix $\gamma = [\gamma_{ik}]$ and $\gamma_i^{(k)} = \gamma_{ik}$ is the coordinate number $i$ of $\gamma^{(k)}$.*

**Theorem 1.17 (The Itô-Lévy isometry).** *Let $X(t) \in \mathbb{R}^n$ be as in (1.2.6) but with $X(0) = 0$ and $\alpha = 0$. Then*

$$E[X^2(T)] = E\left[ \int_0^T \left\{ \sum_{i=1}^{n}\sum_{j=1}^{m} \sigma_{ij}^2(t) + \sum_{i=1}^{n}\sum_{j=1}^{\ell} \int_{\mathbb{R}} \gamma_{ij}^2(t, z_j)\nu_j(dz_j) \right\} dt \right]$$

$$= \sum_{i=1}^{n} E\left[ \int_0^T \left\{ \sum_{j=1}^{m} \sigma_{ij}^2(t) + \sum_{j=1}^{\ell} \int_{\mathbb{R}} \gamma_{ij}^2(t, z_j)\nu_j(dz_j) \right\} dt \right] \quad (1.2.9)$$

*provided that the right hand side is finite.*

*Proof.* This follows from the Itô formula applied to $f(t, x) = x^2 = |x|^2$. We omit the details. $\square$

**Remark 1.18.** As a special case of Theorem 1.17 assume that

$$X(t) = \eta(t) = \int_{\mathbb{R}} z\widetilde{N}(dt, dz) \in \mathbb{R}$$

with $E[X^2(T)] = T \int_{\mathbb{R}} z^2\nu(dz) < \infty$. Then we get the isometry

$$E\left[ \left( \int_0^T H(t)d\eta(t) \right)^2 \right] = E\left[ \int_0^T H^2(t)dt \right] \cdot \int_{\mathbb{R}} z^2\nu(dz)$$

for all $H \in \mathbf{L}_{\text{ucp}}$ (see Definition 1.12) such that $H \in L^2([0,T] \times \Omega)$, i.e. such that

$$\big\|H\big\|^2_{L^2([0,T]\times\Omega)} := E\Big[\int\limits_0^T H^2(t)dt\Big] < \infty\ .$$

Using this we can in the usual way extend the definition of the integral

$$\int\limits_0^T Y(t)d\eta(t) \in L^2(\Omega)$$

to all processes $Y(t)$ which are limits in $L^2([0,T]\times\Omega)$ of processes $H_n(t) \in \mathbf{L}_{\text{ucp}} \cap L^2([0,T]\times\Omega)$. We will call such processes $Y(t)$ *predictable processes* .

## 1.3 Lévy stochastic differential equations

The geometric Lévy process is an example of a *Lévy diffusion* , i.e. the solution of a stochastic differential equation (SDE) driven by Lévy processes.

**Theorem 1.19 (Existence and uniqueness of solutions of Lévy SDEs).**
*Consider the following Lévy SDE in $\mathbb{R}^n$: $X(0) = x_0 \in \mathbb{R}^n$ and*

$$dX(t) = \alpha(t,X(t))dt + \sigma(t,X(t))dB(t) + \int\limits_{\mathbb{R}^n} \gamma(t,X(t^-),z)\widetilde{N}(dt,dz) \quad (1.3.1)$$

*where $\alpha : [0,T]\times\mathbb{R}^n \to \mathbb{R}^n$, $\sigma : [0,T]\times\mathbb{R}^n \to \mathbb{R}^{n\times m}$ and $\gamma : [0,T]\times\mathbb{R}^n\times\mathbb{R}^n \to \mathbb{R}^{n\times\ell}$ satisfy the following conditions*

*(At most linear growth) There exists a constant $C_1 < \infty$ such that*

$$\|\sigma(t,x)\|^2 + |\alpha(t,x)|^2 + \int\limits_{\mathbb{R}} \sum_{k=1}^{\ell} |\gamma_k(t,x,z)|^2 \nu_k(dz_k) \le C_1(1+|x|^2)$$

*for all $x \in \mathbb{R}^n$*

*(Lipschitz continuity) There exists a constant $C_2 < \infty$ such that*

$$\|\sigma(t,x) - \sigma(t,y)\|^2 + |\alpha(t,x) - \alpha(t,y)|^2$$

$$+ \sum_{k=1}^{\ell} \int\limits_{\mathbb{R}} |\gamma^{(k)}(t,x,z_k) - \gamma^{(k)}(t,y,z_k)|^2 \nu_k(dz_k) \le C_2|x-y|^2\ ;$$

*for all $x, y \in \mathbb{R}^n$ .*

*Then there exists a unique cadlag adapted solution $X(t)$ such that*

$$E[|X(t)|^2] < \infty \qquad for\ all\ t\ .$$

Solutions of Lévy SDEs in the *time homogeneous* case, i.e. when $\alpha(t,x) = \alpha(x)$, $\sigma(t,x) = \sigma(x)$ and $\gamma(t,x,z) = \gamma(x,z)$, are called *jump diffusions* (or *Lévy diffusions*).

**Theorem 1.20.** *A jump diffusion is a strong Markov process.*

*Proof.* See [P, Theorem V.32]. $\qquad\square$

**Definition 1.21.** *Let $X(t) \in \mathbb{R}^n$ be a jump diffusion. Then the generator $A$ of $X$ is defined on functions $f : \mathbb{R}^n \to \mathbb{R}$ by*

$$Af(x) = \lim_{t \to 0^+} \frac{1}{t} \big\{ E^x[f(X(t))] - f(x) \big\} \qquad \text{(if the limit exists),}$$

*where $E^x[f(X(t))] = E[f(X^{(x)}(t))]$, $X^{(x)}(0) = x$.*

**Theorem 1.22.** *Suppose $f \in C_0^2(\mathbb{R}^n)$. Then $Af(x)$ exists and is given by*

$$Af(x) = \sum_{i=1}^{n} \alpha_i(x) \frac{\partial f}{\partial x_i}(x) + \tfrac{1}{2} \sum_{i,j=1}^{n} (\sigma\sigma^T)_{ij}(x) \frac{\partial^2 f}{\partial x_i \partial x_j}(x)$$

$$+ \int_{\mathbb{R}} \sum_{k=1}^{\ell} \{ f(x + \gamma^{(k)}(x,z)) - f(x) - \nabla f(x) \cdot \gamma^{(k)}(x,z) \} \nu_k(dz_k) \,.$$

$$(1.3.2)$$

From now on we *define* $Af(x)$ by the expression (1.3.2) for all $f$ such that the partial derivatives of $f$ and the integrals in (1.3.2) exist at $x$.

**Theorem 1.23 (The Dynkin formula I).** *Let $X(t) \in \mathbb{R}^n$ be a jump diffusion and let $f \in C_0^2(\mathbb{R}^n)$. Let $\tau$ be a stopping time such that*

$$E^x[\tau] < \infty \,.$$

*Then*

$$E^x[f(X(\tau))] = f(x) + E^x\Big[ \int_0^{\tau} Af(X(s))ds \Big].$$

*Proof.* This follows by combining the Itô formula (1.2.8) with the formula (1.3.2) for $A$ and taking expectation. $\qquad\square$

This version is usually strong enough for applications in the case when there are no jumps ($N = 0$). However, for jump diffusions we need the following stronger, localized version:

**Theorem 1.24 (The Dynkin formula II).** *Let $X(t) \in \mathbb{R}^n$ be a jump diffusion, $G \subset \mathbb{R}^n$ be an open set and let $f \in C^2(G) \cap C(\bar{G})$. Let $\tau < \infty$ be a stopping time. Suppose that*

$$\tau \leq \tau_G := \inf\{t > 0; X(t) \notin G\} \tag{1.3.3}$$

$$X(\tau) \in \bar{G} \quad a.s. \tag{1.3.4}$$

$$E^x\Big[|f(X(\tau))| + \int_0^\tau \{|Af(X(t))| + |\sigma^T(X(t))\nabla f(X(t))|^2$$

$$+ \sum_{k=1}^\ell \int_{\mathbb{R}} |f(X(t) + \gamma^{(k)}(X(t), z_k)) - f(X(t))|^2 \nu_k(dz_k)\}dt\Big] < \infty \ . \tag{1.3.5}$$

*Then we have*

$$E^x[f(X(\tau))] = f(x) + E^x\Big[\int_0^\tau (Af)(X(t))dt\Big].$$

**Definition 1.25.** *In general, if $\{\psi_m\}_{m=1}^\infty$ and $g$ are functions defined on a set $G \subset \mathbb{R}^n$, we say that $\psi_m \to g$ pointwise dominatedly in $G$ if $\psi_m(x) \to g(x)$ for all $x \in G$ and there exists a constant $C < \infty$ such that*

$$|\psi_m(x)| \leq C|g(x)| \quad \text{for all } x \in G, \ m = 1, 2, \dots$$

*Proof of Theorem 1.24.*

Choose $f_m \in C_0^2(\mathbb{R}^n)$ such that $f_m \to f$ pointwise dominatedly in $\bar{G}$ and $\frac{\partial f_m}{\partial x_i} \to \frac{\partial f}{\partial x_i}$, $\frac{\partial^2 f_m}{\partial x_i \partial x_j} \to \frac{\partial^2 f}{\partial x_i \partial x_j}$ and $Af_m \to Af$ pointwise dominatedly in $G$ for all $i, j = 1, \dots, n$. Then apply Theorem 1.23 to each $f_m$ and $\tau \wedge k$, $k = 1, 2, \dots$ Let $m, k \to \infty$ and apply the dominated convergence theorem. $\qquad\square$

## 1.4 The Girsanov theorem and applications

The Girsanov theorem and the related concept of an equivalent local martingale measure (ELMM) are important in the applications of stochastic analysis to finance. In this chapter we first give a general semimartingale discussion and then we apply it to Itô-Lévy processes. We refer to [Ka] for more details.

Let $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t\geq 0}, P)$ be a filtered probability space. Let $Q$ be another probability measure on $\mathcal{F}_T$. We say that $Q$ is *equivalent* to $P \mid \mathcal{F}_T$ if $P \mid \mathcal{F}_T \ll Q$ and $Q \ll P \mid \mathcal{F}_T$, or, equivalently, if $P$ and $Q$ have the same zero sets in $\mathcal{F}_T$. By the Radon-Nikodym theorem this is the case if and only if we have

$$dQ(\omega) = Z(T)dP(\omega) \quad \text{and} \quad dP(\omega) = Z^{-1}(T)dQ(\omega) \quad \text{on } \mathcal{F}_T$$

for some $\mathcal{F}_T$-measurable random variable $Z(T) > 0$ a.s. $P$. In that case we also write

$$\frac{dQ}{dP} = Z(T) \quad \text{and} \quad \frac{dP}{dQ} = Z^{-1}(T) \quad \text{on } \mathcal{F}_T. \tag{1.4.1}$$

We first make a simple, but useful observation:

**Lemma 1.26.** *Suppose $Q \ll P$ with $\dfrac{dQ}{dP} = Z(T)$ on $\mathcal{F}_T$. Then*

$$Q \mid \mathcal{F}_t \ll P \mid \mathcal{F}_t \text{ for all } t \in [0, T] \text{ and}$$

$$Z(t) := \frac{d(Q \mid \mathcal{F}_t)}{d(P \mid \mathcal{F}_t)} = E_P[Z(T) \mid \mathcal{F}_t]; \qquad 0 \le t \le T. \tag{1.4.2}$$

*In particular, $Z(t)$ is a $P$-martingale.*

*Proof.* Since $P(G) = 0 \Rightarrow Q(G) = 0$ for all $G \in \mathcal{F}_T \supseteq \mathcal{F}_t$, it is clear that $Q \mid \mathcal{F}_t \ll P \mid \mathcal{F}_t$. Choose $F \in \mathcal{F}_t$. Then

$$E_P\big[F \cdot E[Z(T) \mid \mathcal{F}_t]\big] = E_P\big[E_P[FZ(T) \mid \mathcal{F}_t]\big]$$
$$= E_P[FZ(T)] = E_Q[F] = E_P[F \cdot Z(t)].$$

Since this holds for all $F \in \mathcal{F}_t$ we conclude that

$$E_P[Z(T) \mid \mathcal{F}_t] = Z(t), \quad \text{as claimed.} \qquad \square$$

**Definition 1.27.** *Let $X(t), Y(t) \in \mathbb{R}^n$ be two cadlag semimartingales. The quadratic covariation of $X(\cdot)$ and $Y(\cdot)$, denoted by $[X, Y](\cdot)$, is the unique semimartingale such that*

$$X(t) \cdot Y(t) = X(0) \cdot Y(0) + \int_0^t X(s^-) \cdot dY(s) + \int_0^t Y(s^-) \cdot dX(s) + [X, Y](t).$$

**Example 1.28.** Let

$$dX_i(t) = \alpha_i(t, \omega)dt + \sigma_i(t, \omega)dB(t) + \int_\mathbb{R} \gamma_i(t, z)\widetilde{N}(dt, dz); \qquad i = 1, 2$$

be two Itô-Lévy processes. Then by the Itô formula (Theorem 1.16) we have (see Exercise 1.7)

$$d(X_1(t)X_2(t)) = X_1(t^-)dX_2(t) + X_2(t^-)dX_1(t) + \sigma_1(t)\sigma_2(t)dt$$
$$+ \int_\mathbb{R} \gamma_1(t, z)\gamma_2(t, z)N(dt, dz).$$

Hence in this case the quadratic covariation is

$$[X_1, X_2](t) = \int\limits_0^t \sigma_1(s)\sigma_2(s)ds + \int\limits_0^t \int\limits_{\mathbb{R}} \gamma_1(s,z)\gamma_2(s,z)N(ds,dz)$$

$$= \int\limits_0^t \Big[ \sigma_1(s)\sigma_2(s) + \int\limits_{\mathbb{R}} \gamma_1(s,z)\gamma_2(s,z)\nu(dz) \Big] ds$$

$$+ \int\limits_0^t \int\limits_{\mathbb{R}} \gamma_1(s,z)\gamma_2(s,z)\widetilde{N}(ds,dz).$$

Recall that a semimartingale $M(t)$ is called a *local* martingale (with respect to $P$) if there exists an increasing sequence of $(\mathcal{F}_{t-})$ stopping times $\tau_n$ such that $\lim\limits_{n\to\infty} \tau_n = \infty$ a.s. and

$$M(t \wedge \tau_n) \quad \text{is a martingale with respect to } P \text{ for all } n.$$

**Theorem 1.29 (Girsanov theorem for semimartingales).** *Let $Q$ be a probability measure on $\mathcal{F}_T$ and assume that $Q$ is equivalent to $P$ on $\mathcal{F}_T$, with*

$$dQ(\omega) = Z(t)dP(\omega) \qquad on \ \mathcal{F}_t; \ t \in [0, T].$$

*(See Lemma 1.26). Assume that $Z(t)$ is continuous on $[0, T]$. Let $M(t)$ be a local $P$-martingale. Then the process $\widehat{M}(t)$ defined by*

$$\widehat{M}(t) := M(t) - \int\limits_0^t \frac{d[M, Z](s)}{Z(s)}$$

*is a local $Q$-martingale.*

SKETCH OF PROOF. By integration by parts we have

$$\widehat{M}(t)Z(t) - \widehat{M}(0)Z(0) = \int\limits_0^t \widehat{M}(s^-)dZ(s) + \int\limits_0^t Z(s)d\widehat{M}(s) + [M, Z](t)$$

$$= \int\limits_0^t \widehat{M}(s^-)dZ(s) + \int\limits_0^t Z(s)dM(s) - \int\limits_0^t Z(s)\frac{d[M, Z](s)}{Z(s)} + [M, Z](t)$$

$$= \int\limits_0^t \widehat{M}(s^-)dZ(s) + \int\limits_0^t Z(s)dM(s).$$

Therefore, since $Z(t)$ is a local $P$-martingale (Lemma 1.26) we conclude that $\widehat{M}(t)Z(t)$ is a local $P$-martingale. But then $\widehat{M}(t)$ is a local $Q$-martingale, since

$$E_Q[\widehat{M}(\tau) \mid \mathcal{F}_t] = \frac{E_P[\widehat{M}(\tau)Z(\tau) \mid \mathcal{F}_t]}{E_P[Z(\tau) \mid \mathcal{F}_t]} = \frac{\widehat{M}(t)Z(t)}{Z(t)} = \widehat{M}(t)$$

for stopping times $\tau \geq t$.

We now apply this to two important special cases:

**Theorem 1.30 (Girsanov theorem I for Itô processes).** *Let $X(t)$ be an n-dimensional Itô process of the form*

$$dX(t) = \alpha(t, \omega)dt + \sigma(t, \omega)dB(t); \qquad 0 \leq t \leq T$$

*where $\alpha(t) = \alpha(t, \omega) \in \mathbb{R}^n$, $\sigma(t) = \sigma(t, \omega) \in \mathbb{R}^{n \times m}$ and $B(t) \in \mathbb{R}^m$. Assume that there exists a process $\theta(t) \in \mathbb{R}^m$ such that*

$$\sigma(t)\theta(t) = \alpha(t) \qquad \text{for a.a. } (t, \omega) \in [0, T] \times \Omega \qquad (1.4.3)$$

*and such that the process $Z(t)$ defined for $0 \leq t \leq T$ by*

$$Z(t) := \exp\left\{ -\int_0^t \theta(s)dB(s) - \frac{1}{2}\int_0^t \theta^2(s)ds \right\} \qquad (1.4.4)$$

*exists. Define a measure $Q$ on $\mathcal{F}_T$ by*

$$dQ(\omega) = Z(T)dP(\omega) \qquad \text{on } \mathcal{F}_T. \qquad (1.4.5)$$

*Assume that*

$$E_P[Z(T)] = 1. \qquad (1.4.6)$$

*Then $Q$ is a probability measure on $\mathcal{F}_T$, $Q$ is equivalent to $P$ and $X(t)$ is a local martingale with respect to $Q$.*

**Remark 1.31.** Such a measure $Q$ is called an *equivalent local martingale measure* for $X(t)$.

*Proof.* Put $M(t) = \int_0^t \sigma(s, \omega)dB(s)$; $0 \leq t \leq T$. Then by Theorem 1.29, the process

$$\widehat{M}(t) := \int_0^t \sigma(s)dB(s) - \frac{d[M, Z](t)}{Z(t)}$$

is a local $Q$-martingale. Since

$$dZ(t) = -\theta(t)Z(t)dB(t)$$

we have

$$\frac{d[M, Z](t)}{Z(t)} = \frac{-\sigma(t)\theta(t)Z(t)}{Z(t)}dt = -\sigma(t)\theta(t)dt = -\alpha(t)dt.$$

Hence

$$\widehat{M}(t) = \int\limits_0^t \sigma(s)dB(s) - \int\limits_0^t (-\alpha(s))ds = X(t)$$

is a local $Q$-martingale by Theorem 1.29. □

We also state without proof the following related version of the Girsanov theorem for Itô processes:

**Theorem 1.32 (Girsanov theorem II for Itô diffusions ).** *Let* $X(t), \theta(t)$ *and* $Q$ *be as defined in Theorem 1.30. Assume that the Novikov condition holds, i.e.*

$$E_P\left[\exp\left(\tfrac{1}{2}\int\limits_0^T \theta^2(s)ds\right)\right] < \infty . \tag{1.4.7}$$

*Then* $Q$ *is a probability measure on* $\mathcal{F}_T$, *the process*

$$\hat{B}(t) := \int\limits_0^t \theta(s)ds + B(t); \qquad 0 \leq t \leq T \tag{1.4.8}$$

*is a Brownian motion with respect to* $Q$ *and expressed in terms of* $\hat{B}(t)$ *we have*

$$dX(t) = \sigma(t,\omega)d\hat{B}(t); \qquad 0 \leq t \leq T. \tag{1.4.9}$$

Finally we turn to the Girsanov theorem for jump diffusions. First we need a result about how the probabilistic properties of the pure jump process $N(t,U)$ change under a change of probability law. The following result a special case of Theorem 3.24 and Theorem 5.19 in [JS] (see also [C, Theorem 3.2]). We refer the reader to these sources for a proof and more details.

**Lemma 1.33.** *Let* $\theta(s,z) \leq 1$ *be a process such that*

$$Z(t) := \exp\left\{ \int\limits_0^t \int\limits_{\mathbb{R}} \ln(1 - \theta(s,z))\widetilde{N}(ds, dz) \right.$$

$$\left. + \int\limits_0^t \int\limits_{\mathbb{R}} \{\ln(1 - \theta(s,z)) + \theta(s,z)\}\nu(dz)ds \right\}$$

*exists for* $0 \leq t \leq T$. *Define a measure* $Q$ *on* $\mathcal{F}_T$ *by*

$$dQ(\omega) = Z(T)dP(\omega).$$

*Assume that*

$$E_P[Z(T)] = 1.$$

Then $Q$ is a probability measure on $\mathcal{F}_T$ and if we define the random measure $\widetilde{N}^Q(\cdot, \cdot)$ by

$$\widetilde{N}^Q(dt, dz) := N(dt, dz) - (1 - \theta(t, z))\nu(dz)dt \qquad (1.4.10)$$

then

$$\int_0^t \int_{\mathbb{R}} \widetilde{N}^Q(ds, dz) = \int_0^t \int_{\mathbb{R}} N(ds, dz) - \int_0^t \int_{\mathbb{R}} (1 - \theta(s, z))\nu(dz)ds$$

is a $Q$-local martingale.

**Theorem 1.34 (Girsanov theorem I for jump processes).** *Let $X(t)$ be a 1-dimensional Itô-Lévy process of the form*

$$dX(t) = \alpha(t, \omega)dt + \int_{\mathbb{R}} \gamma(t, z)\widetilde{N}(dt, dz). \qquad (1.4.11)$$

*Assume that there exists a process $\theta(s, z) \leq 1$ such that*

$$\int_{\mathbb{R}} \gamma(t, z)\theta(t, z)\nu(dz) = \alpha(t) \qquad \text{for a.a. } (t, \omega) \qquad (1.4.12)$$

*and such that the process $Z(t)$ defined by*

$$Z(t) = \exp\Big\{ \int_0^t \int_{\mathbb{R}} \ln(1 - \theta(s, z))\widetilde{N}(ds, dz)$$

$$+ \int_0^t \int_{\mathbb{R}} \{\ln(1 - \theta(s, z)) + \theta(s, z)\}\nu(dz)ds \Big\} \qquad (1.4.13)$$

*exists for $0 \leq t \leq T$. Define a measure $Q$ on $\mathcal{F}_T$ by*

$$dQ(\omega) = Z(T)dP(\omega). \qquad (1.4.14)$$

*Assume that*

$$E_P[Z(T)] = 1. \qquad (1.4.15)$$

*Then $Q$ is an equivalent local martingale measure for $X(t)$.*

*Proof.* By Lemma 1.33 and (1.4.11) we have

$$dX(t) = \alpha(t)dt + \int_{\mathbb{R}} \gamma(t,z)N(dt,dz) - \int_{\mathbb{R}} \gamma(t,z)\nu(dz)dt$$

$$= \alpha(t)dt + \int_{\mathbb{R}} \gamma(t,z)\{\widetilde{N}^Q(dt,dz) + (1 - \theta(t,z))\nu(dz)dt\}$$

$$- \int_{\mathbb{R}} \gamma(t,z)\nu(dz)dt$$

$$= \int_{\mathbb{R}} \gamma(t,z)\widetilde{N}^Q(dt,dz) + \Big[\alpha(t) - \int_{\mathbb{R}} \gamma(t,z)\theta(t,z)\nu(dz)\Big]dt$$

$$= \int_{\mathbb{R}} \gamma(t,z)\widetilde{N}^Q(dt,dz),$$

which is a local $Q$-martingale.                                                                                                        □

Similarly, in the $n$-dimensional case we get:

**Theorem 1.35 (Girsanov theorem II for jump processes).** *Let $X(t)$ be an $n$-dimensional Itô-Lévy process of the form*

$$dX(t) = \alpha(t)dt + \int_{\mathbb{R}^n} \gamma(t,z)\widetilde{N}(dt,dz),$$

*where $\alpha(t) = \alpha(t,\omega) \in \mathbb{R}^n$, $\gamma(t,z) \in \mathbb{R}^{n\times\ell}$ and $\widetilde{N}(dt,dz) = (\widetilde{N}_1(dt,dz_1),\ldots,$ $\widetilde{N}_\ell(dt,dz_\ell))$ is $\ell$-dimensional. Assume that there exists a process $\theta(t,z) = (\theta_1(t,z),\ldots,\theta_\ell(t,z))^T \in \mathbb{R}^\ell$ such that $\theta_j(s,z) \leq 1$ and*

$$\sum_{j=1}^{\ell} \int_{\mathbb{R}} \gamma_{ij}(t,z_j)\theta_j(t,z_j)\nu_j(dz_j) = \alpha_i(t); \qquad i = 1,\ldots,n, \quad t \in [0,T] \quad (1.4.16)$$

*and such that the process*

$$Z(t) := \exp\Big\{\sum_{j=1}^{\ell} \int_0^t \int_{\mathbb{R}} \big[\ln(1 - \theta_j(s,z_j))N_j(ds,dz_j) + \theta_j(s,z_j)\nu_j(dz_j)ds\big]\Big\}$$

*exists for $0 \leq t \leq T$. Define a measure $Q$ on $\mathcal{F}_T$ by*

$$dQ(\omega) = Z(T)dP(\omega).$$

*Assume that*

$$E[Z(T)] = 1.$$

*Then $Q$ is an equivalent local martingale measure for $X(t)$.*

**Example 1.36.**

**(i)** Suppose $X(t) = \pi(t) := \int_{\mathbb{R}} z\widetilde{N}(t, dz) \in \mathbb{R}$ is a Poisson process. Then

$$\nu(dz) = \delta_{z_1}(dz) \qquad \text{for some } z_1 \in \mathbb{R} \setminus \{0\}$$

and condition (1.4.12) gets the form

$$\gamma(t, z_1)\theta(t, z_1) = \alpha(t); \qquad t \in [0, T]. \tag{1.4.17}$$

This corresponds to the equation (1.4.3) in the Brownian motion case. Note that there is *at most one* solution $\theta(t, z_1)$ of (1.4.17) (unless $\alpha(t) = \gamma(t, z_1) = 0$).

**(ii)** Next, suppose $\nu$ is supported on $n$ points $z_1, \ldots, z_n$. Then (1.4.16) gets the form

$$\sum_{j=1}^{n} \gamma(t, z_j)\theta(t, z_j)\nu(\{z_j\}) = \alpha(t); \qquad t \in [0, T].$$

For each $t \in [0, T]$ this is one linear equation in the $n$ unknowns $\theta(t, z_1), \ldots, \theta(t, z_n)$. So unless in degenerate cases this equation will have infinitely many solutions and – under some conditions on $\{\gamma(t, z_j)\}_{j=1}^{n}$ – also infinitely many solutions satisfying $\theta(t, z_j) \leq 1$. This corresponds to infinitely many equivalent local martingale measures, which again is equivalent to incompleteness of the associated financial market, according to the Second Fundamental Theorem of Asset Pricing (see e.g. [DS], [LS]).

## 1.5 Application to finance

It has been argued (see e.g. [EK], [B-N], [Sc], [Eb] and [CT]) that Lévy processes are relevant in mathematical finance, in particular in the modelling of stock prices.

Consider the following Lévy version of the Black-Scholes market:

(Bond price)   $dS_0(t) = rS_0(t)dt;$                    $S_0(0) = 1$

(Stock price)   $dS_1(t) = S_1(t^-)[\mu\, dt + \gamma\, d\eta(t)];$     $S_1(0) = x > 0$

where $r, \mu$ and $\gamma \neq 0$ are constants and

$$\eta(t) = \int_0^t \int_{\mathbb{R}} z\widetilde{N}(dt, dz)$$

is a pure jump Lévy martingale. To ensure that $S_1(t) \geq 0$ for all $t \geq 0$ we assume as before that $\gamma z \geq -1$ a.s. $\nu$. Assume in addition that

$$\int_{\mathbb{R}} \nu(dz) > \frac{\mu - r}{\gamma} \,, \tag{1.5.1}$$

i.e. that the total mass of the jump measure (Lévy measure) $\nu$ exceeds $\frac{\mu-r}{\gamma}$ .

The *normalized* stock price $\bar{S}_1(t)$ is given by

$$\bar{S}_1(t) = \frac{1}{S_0(t)} S_1(t) = e^{-\rho t} S_1(t).$$

Note that

$$d\bar{S}_1(t) = \bar{S}_1(t^-)[(\mu - r)dt + \gamma \, d\eta(t)]; \qquad \bar{S}_1(0) = x.$$

We seek an equivalent local martingale measure $Q$ of the process $\bar{S}_1(t)$.

To this end we apply Theorem 1.34 and try to find a solution $\theta(z) \le 1$ of the equation (1.4.12), which in this case gets the form

$$\int_{\mathbb{R}} \theta(z)\nu(dz) = \frac{\mu - r}{\gamma} \,. \tag{1.5.2}$$

By (1.5.1) we see that if $A \subset \mathbb{R}$ with $\frac{\mu-r}{\gamma} < \nu(A) < \infty$ then

$$\theta(z) = \frac{\mu - r}{\gamma \nu(A)} \mathcal{X}_A(z)$$

is a possible solution. If $\nu$ is concentrated on one point $z_0$, i.e. if

$$\nu(\mathbb{R}) = \nu(\{z_0\})$$

(which means that $\eta(t)$ is a Poisson process multiplied by $z_0$) then this is the only solution. On the other hand, if there exist two sets $A, B \subset (-1, \infty)$ such that $A \cap B = \emptyset$ and

$$\nu(A) > 0, \quad \nu(B) > 0 \tag{1.5.3}$$

then we see that there are infinitely many solutions $\theta(z)$ of (1.5.2) such that $\theta(z) < 1$.

Fix a solution $\theta(z) < 1$ of (1.5.2) and define

$$Z(t) = Z^\theta(t) = \exp\left\{ \int_0^t \int_{\mathbb{R}} \ln(1 - \theta(z)) N(ds, dz) + \frac{\mu - r}{\gamma} t \right\}; \quad 0 \le t \le T$$

and

$$dQ = dQ^\theta = Z^\theta(T) dP \quad \text{on } \mathcal{F}_T.$$

Then by Lemma 1.33 $\bar{S}_1(t)$ is a local martingale with respect to $Q$.

We now discuss the concept of *arbitrage* in this market. For more information on the mathematics of finance see e.g. [KS] or [Ø1, Chapter 12] and the references therein.

A *portfolio* in this market is a predictable process $\phi(t) = (\phi_0(t), \phi_1(t)) \in \mathbb{R}^2$ such that

$$\int_0^T \{\phi_0^2(s) + \phi_1^2(s)\}ds < \infty \quad \text{a.s. } P. \tag{1.5.4}$$

The corresponding *wealth process* $V^\phi(t)$ is defined by

$$V^\phi(t) = \phi_0(t)S_0(t) + \phi_1(t)S_1(t); \qquad 0 \le t \le T. \tag{1.5.5}$$

We say that $(\phi_0, \phi_1)$ is *self-financing* if $V^\phi(t)$ is also given by

$$V^\phi(t) = V^\phi(0) + \int_0^t \phi_0(s)dS_0(s) + \int_0^t \phi_1(s)dS_1(s). \tag{1.5.6}$$

If, in addition,

$$\left\{V^\phi(t)\right\}_{t \in [0,T]} \quad \text{is lower bounded} \tag{1.5.7}$$

we say that $\phi$ is *admissible* and write $\phi \in \mathcal{A}_0$. A portfolio $\phi \in \mathcal{A}_0$ is called an *arbitrage* if

$$V^\phi(0) = 0, \quad V^\phi(T) \ge 0 \quad \text{and} \quad P[V^\phi(T) > 0] > 0. \tag{1.5.8}$$

Does this market have an arbitrage? To answer this we combine (1.5.5) and (1.5.6) to get

$$\phi_0(t) = e^{-rt}(V^\phi(t) - \phi_1(t)S_1(t))$$

and

$$dV^\phi(t) = rV^\phi(t) + \phi_1(t)S_1(t^-)\left[(\mu - r)dt + \gamma \int_{\mathbb{R}} z\widetilde{N}(dt, dz)\right].$$

From this we obtain

$$d(e^{-rt}V^\phi(t)) = e^{-rt}\phi_1(t)S_1(t^-)\left[(\mu - r)dt + \gamma \int_{\mathbb{R}} z\widetilde{N}(dt, dz)\right]$$

or

$$e^{-rt}V^\phi(t) = V^\phi(0) + \int_0^t e^{-rs}\phi_1(s)S_1(s^-)\left[(\mu - r)ds + \gamma \int_{\mathbb{R}} z\widetilde{N}(ds, dz)\right].$$

Therefore $e^{-rt}V^\phi(t)$ is a lower bounded local martingale, and hence a supermartingale, with respect to $Q$. But then

$$0 = E_Q[V^\phi(0)] \ge E_Q[e^{-rT}V^\phi(T)],$$

which shows that (1.5.8) cannot hold.

We conclude that *there is no arbitrage in this market* (if (1.5.1) holds).

This example illustrates the First Fundamental Theorem of Asset Pricing, which states the connection between

(i) the existence of an equivalent local martingale measure and
(ii) the nonexistence of arbitrage, or No Free Lunch with Vanishing Risk
  (NFLVR) to be more precise. See e.g. [DS], [LS].

## 1.6 Exercises

**Exercise\* 1.1.** Suppose

$$dX(t) = \alpha dt + \sigma dB(t) + \int_{\mathbb{R}} \gamma(z)\bar{N}(dt, dz), \qquad X(0) = x \in \mathbb{R},$$

where $\alpha, \sigma$ are constants, $\gamma : \mathbb{R} \to \mathbb{R}$ is a given function.

**(i)** Use Itô's formula to find $dY(t)$ when

$$Y(t) = \exp(X(t)) .$$

**(ii)** How do we choose $\alpha, \sigma$ and $\gamma(z)$ if we want $Y(t)$ to solve the SDE

$$dY(t) = Y(t^-)\Big[\beta dt + \theta dB(t) + \lambda \int_{\mathbb{R}} z\bar{N}(dt, dz)\Big],$$

for given constants, $\beta, \theta$ and $\lambda$?

**Exercise\* 1.2.** Solve the following Lévy SDEs:

**(i)** $dX(t) = (m - X(t))dt + \sigma dB(t) + \gamma \int_{\mathbb{R}} z\bar{N}(dt, dz); X(0) = x \in \mathbb{R}$

$\quad$ ($m, \sigma, \gamma$ constants) (the mean-reverting Lévy-Ornstein-Uhlenbeck process)

**(ii)** $dX(t) = \alpha dt + \gamma X(t^-) \int_{\mathbb{R}} z\bar{N}(dt, dz); X(0) = x \in \mathbb{R}$

$\quad$ ($\alpha, \gamma$ constants, $\gamma z > -1$ a.s. $\nu$).

$\big[$Hint: Try to multiply the equation by

$$F(t) := \exp\Big\{ -\int_0^t \int_{\mathbb{R}} \theta(z)\bar{N}(dt, dz) + \int_{|z| < R} (e^{\theta(z)} - 1 - \theta(z))\nu(dz) \cdot t \Big\},$$

for suitable $\theta(z)$.$\big]$

**Exercise 1.3 (Geometric Lévy martingales).** Let $h \in L^2(\mathbb{R})$ be deterministic and define

$$Y(t) = \exp\Big\{ \int_0^t \int_{\mathbb{R}} h(s)z\widetilde{N}(ds, dz) - \int_0^t \int_{\mathbb{R}} \Big(e^{h(s)z} - 1 - h(s)z\Big) \nu(dz)ds \Big\}.$$

Show that

$$dY(t) = Y(t^-) \int_{\mathbb{R}} \left( e^{h(t)z} - 1 \right) \tilde{N}(dt, dz).$$

**Exercise 1.4.** Find the generator $A$ of the following jump diffusions:

a) (The geometric Lévy process )

$$dX(t) = X(t^-) \left[ \mu dt + \sigma dB(t) + \gamma \int_{\mathbb{R}} z \tilde{N}(dt, dz) \right].$$

b) (The mean-reverting Lévy-Ornstein-Uhlenbeck process )

$$dX(t) = (m - X(t))dt + \sigma dB(t) + \gamma \int_{\mathbb{R}} z \tilde{N}(dt, dz).$$

c) (The graph of the geometric Lévy process )

$$dY(t) = \begin{bmatrix} dt \\ dX(t) \end{bmatrix}, \quad \text{where } X(t) \text{ is as in a)}.$$

d) (The $n$ dimensional geometric Lévy process )

$$X(t) = \begin{bmatrix} X(t) \\ \vdots \\ X_n(t) \end{bmatrix},$$

where

$$dX_i(t) = X_i(t^-) \left[ \mu_i dt + \sum_{j=1}^{n} \sigma_{ij} dB_j(t) + \sum_{j=1}^{n} \gamma_{ij} \int_{\mathbb{R}} z_j \tilde{N}_j(dt, dz_j) \right];$$

$$1 \le i \le n.$$

**Exercise 1.5 (The first exit time from a ball).**
Let $K = \{x \in \mathbb{R}^n \ ; \ |x| \le R\}$ be the open ball of radius $R$ in $\mathbb{R}^n$ and let

$$\eta(t) = (\eta_1(t), \dots, \eta_n(t)),$$

where

$$\eta_i(t) = a_i + \int_0^t \int_{\mathbb{R}} z_i \tilde{N}_i(dt, dz_i) \ ; \ 1 \le i \le n$$

are independent 1-dimensional pure jump Lévy processes and $a = (a_1, \dots, a_n)$ $\in K$ is constant. We assume that $0 < E \left[ \sum_{i=1}^{n} (\eta_i(t) - a_i)^2 \right] < \infty$ for all $t > 0$.

a) Find the generator $A$ of $\eta(\cdot)$ and show that if $f(x) = |x|^2$ then

$$Af(x) = \sum_{i=1}^{n} \int_{\mathbb{R}} |\zeta|^2 \nu_i(d\zeta) := \rho(n) \in (0, \infty) \quad \text{(constant)}.$$

b) Let
$$\tau = \inf\{t > 0 \; ; \; \eta(t) \notin K\} \leq \infty$$

and put
$$\tau_k = \tau \wedge k \; ; \; k = 1, 2, \ldots$$

Let $\{f_m\}_{m=1}^{\infty}$ be a sequence of functions in $C^2(\mathbb{R}^n)$ such that
$f_m(x) = |x|^2$ for $|x| \leq R$, $0 \leq f_m(x) \leq 2R^2$ for all $x \in \mathbb{R}$,
$\operatorname{supp} f_m \subset \{x \in \mathbb{R}^n \; ; \; |x| \leq R + \frac{1}{m}\}$ for all $m$ and
$f_m(x) \to |x|^2 \cdot \chi_K(x)$ as $m \to \infty$, for all $x \in \mathbb{R}^n$.
Use the Dynkin formula I to show that

$$E^a\left[f_m\left(\eta(\tau_k)\right)\right] = |a|^2 + \rho(n) \cdot E^a[\tau_k] \qquad \text{for all } m, k.$$

c) Show that

$$E^a[\tau] = \frac{1}{\rho(n)}\left(R^2 P^a[\eta(\tau) \in K] - |a|^2\right) \leq \frac{1}{\rho(n)}\left(R^2 - |a|^2\right).$$

In particular, $\tau < \infty$, a.s.

**Remark.** If we replace $\eta(\cdot)$ by an $n$-dimensional Brownian motion $B(\cdot)$, then the corresponding exit time $\tau^{(B)}$ satisfies

$$E^a\left[\tau^{(B)}\right] = \frac{1}{n}\left(R^2 - |a|^2\right)$$

(see e.g. [Ø1], Example 7.4.2).

**Exercise\* 1.6.** Show that

$$E\left[\exp\left(\int_0^t \int_{\mathbb{R}} \gamma(s, z)\widetilde{N}(ds, dz)\right)\right] = \exp\left(\int_0^t \int_{\mathbb{R}} \{e^{\gamma(s,z)} - 1 - \gamma(s, z)\}\nu(dz)ds\right),$$

provided that the right hand side is finite.

**Exercise\* 1.7.** Let

$$dX_i(t) = \int_{\mathbb{R}} \gamma_i(t, z)\widetilde{N}(dt, dz) \; ; \qquad i = 1, 2$$

be two 1-dimensional Itô-Lévy processes. Use the 2-dimensional Itô formula (Theorem 1.16) to prove the following *integration by parts formula* :

$$X_1(t)X_2(t) = X_1(0)X_2(0) + \int_0^t X_1(s^-)dX_2(s) + \int_0^t X_2(s^-)dX_1(s)$$

$$+ \int_0^t \int_{\mathbb{R}} \gamma_1(s, z)\gamma_2(s, z)N(ds, dz) . \qquad (1.6.1)$$

**Remark.** The process

$$[X_1, X_2](t) := \int\limits_0^t \int\limits_{\mathbb{R}} \gamma_1(s, z)\gamma_2(s, z)N(ds, dz)$$

$$= \int\limits_0^t \int\limits_{\mathbb{R}} \gamma_1(s, z)\gamma_2(s, z)\nu(dz)ds + \int\limits_0^t \int\limits_{\mathbb{R}} \gamma_1(s, z)\gamma_2(s, z)\widetilde{N}(ds, dz)$$

(1.6.2)

is called the *quadratic covariation* of $X_1$ and $X_2$. See Definition 1.27.

**Exercise\* 1.8.** Consider the following market

(Bond price)    $dS_0(t) = 0;$                                                    $S_0(0) = 0$
(Stock price 1) $dS_1(t) = S_1(t^-)[\mu_1 dt + \gamma_{11} d\eta_1(t) + \gamma_{12} d\eta_2(t)];$ $S_1(0) = x_1 > 0$
(Stock price 2) $dS_2(t) = S_2(t^-)[\mu_2 dt + \gamma_{21} d\eta_1(t) + \gamma_{22} d\eta_2(t)];$ $S_2(0) = x_2 > 0$

where $\mu_i$ and $\mu_{ij}$ are constants and $\eta_1(t), \eta_2(t)$ are independent Lévy martingales of the form

$$d\eta_i(t) = \int\limits_{\mathbb{R}} z\widetilde{N}_i(dt, dz); \qquad i = 1, 2.$$

Assume that the matrix $\gamma := [\gamma_{ij}]_{1 \le i,j \le 2} \in \mathbb{R}^2$ is invertible, with inverse

$$\gamma^{-1} = \lambda = [\lambda_{ij}]_{1 \le i,j \le 2}$$

and assume that

$$\nu_i(\mathbb{R}) > \lambda_{i1}\mu_1 + \lambda_{i2}\mu_2 \qquad \text{for} \quad i = 1, 2. \tag{1.6.3}$$

Find an equivalent local martingale measure $Q$ for $(S_1(t), S_2(t))$ and use this to deduce that there is no arbitrage in this market.

**2**

# Optimal Stopping of Jump Diffusions

## 2.1 A general formulation and a verification theorem

Fix an open set $\mathcal{S} \subset \mathbb{R}^k$ (the *solvency region*) and let $Y(t)$ be a jump diffusion in $\mathbb{R}^k$ given by

$$dY(t) = b(Y(t))dt + \sigma(Y(t))dB(t) + \int_{\mathbb{R}^k} \gamma(Y(t^-), z)\bar{N}(dt, dz), \quad Y(0) = y \in \mathbb{R}^k$$

where $b : \mathbb{R}^k \to \mathbb{R}^k$, $\sigma : \mathbb{R}^k \to \mathbb{R}^{k \times m}$ and $\gamma : \mathbb{R}^k \times \mathbb{R}^k \to \mathbb{R}^{k \times \ell}$ are given functions such that a unique solution $Y(t)$ exists (see Theorem 1.19). Let

$$\tau_{\mathcal{S}} = \tau_{\mathcal{S}}(y, \omega) = \inf\{t > 0; Y(t) \notin \mathcal{S}\} \tag{2.1.1}$$

be the *bankruptcy time* and let $\mathcal{T}$ denote the set of all stopping times $\tau \leq \tau_{\mathcal{S}}$.

The results below remain valid, with the natural modifications, if we allow $\mathcal{S}$ to be any Borel set such that $\mathcal{S} \subset \overline{\mathcal{S}^0}$ where $\mathcal{S}^0$ denotes the interior of $\mathcal{S}$, $\overline{\mathcal{S}^0}$ its closure.

Let $f : \mathbb{R}^k \to \mathbb{R}$ and $g : \mathbb{R}^k \to \mathbb{R}$ be continuous functions satisfying the conditions

$$E^y\left[\int_0^{\tau_{\mathcal{S}}} f^-(Y(t))dt\right] < \infty \qquad \text{for all } y \in \mathbb{R}^k \tag{2.1.2}$$

The family $\{g^-(Y(\tau)) \cdot \mathcal{X}_{\{\tau < \infty\}}; \tau \in \mathcal{T}\}$ is uniformly integrable, for all $y \in \mathbb{R}^k$. $\tag{2.1.3}$

(If $x$ is a real number, then $x^- := \max(-x, 0)$ denotes the *negative part* of $x$.)

The general *optimal stopping problem* is the following:

Find $\Phi(y)$ and $\tau^* \in \mathcal{T}$ such that

$$\Phi(y) = \sup_{\tau \in \mathcal{T}} J^\tau(y) = J^{\tau^*}(y) ; \qquad y \in \mathbb{R}^k$$

where

$$J^\tau(y) = E^y\left[\int\limits_0^\tau f(Y(t))dt + g(Y(\tau)) \cdot \mathcal{X}_{\{\tau < \infty\}}\right] ; \qquad \tau \in \mathcal{T}$$

is the *performance criterion* .

The function $\Phi$ is called the *value function* and the stopping time $\tau^*$ (if it exists) is called an *optimal stopping time* .

In the following we let $A$ be the integrodifferential operator which coincides with the generator of $Y(t)$ on $C_0^2(\mathbb{R}^k)$, i.e.

$$A\phi(y) = \sum_{i=1}^k b_i(y)\frac{\partial \phi}{\partial y_i}(y) + \tfrac{1}{2}\sum_{i,j=1}^k (\sigma\sigma^T)_{ij}(y)\frac{\partial^2 \phi}{\partial y_i \partial y_j}(y)$$

$$+ \sum_{j=1}^\ell \int\limits_\mathbb{R} \{\phi(y + \gamma^{(j)}(y, z_j)) - \phi(y) - \nabla\phi(y) \cdot \gamma^{(j)}(y, z_j)\}\nu_j(dz_j) \quad (2.1.4)$$

for all $\phi : \mathbb{R}^k \to \mathbb{R}$ and $y \in \mathbb{R}^k$ such that (2.1.4) exists. (See Theorem 1.22 and Theorem 1.24.)

We will need the following result. A proof of a related result (in the no jump case) can be found in [Ø1].

**Theorem 2.1 (Approximation theorem).**
*Let $D$ be an open set, $D \subset \mathcal{S}$. Assume that*

$$\partial D \text{ is a Lipschitz surface} \qquad (2.1.5)$$

*(i.e. $\partial D$ is locally the graph of a Lipschitz continuous function) and let $\varphi : \bar{\mathcal{S}} \to \mathbb{R}$ be a function with the following properties:*

$$\varphi \in C^1(\mathcal{S}) \cap C(\bar{\mathcal{S}}) \qquad (2.1.6)$$

*and*

$$\varphi \in C^2(\mathcal{S} \backslash \partial D) \qquad (2.1.7)$$

*and the second order derivatives of $\varphi$ are locally bounded near $\partial D$.*

*Then there exists a sequence $\{\varphi_m\}_{m=1}^\infty \subset C^2(\mathcal{S}) \cap C(\bar{\mathcal{S}})$ such that, with $A$ as in (2.1.4),*

$$\varphi_m \to \varphi \text{ pointwise dominatedly in } \bar{\mathcal{S}} \text{ as } m \to \infty \qquad (2.1.8)$$

$$\frac{\partial \varphi_m}{\partial x_i} \to \frac{\partial \varphi}{\partial x_i} \text{ pointwise dominatedly in } \mathcal{S} \text{ as } m \to \infty \qquad (2.1.9)$$

$$\frac{\partial^2 \varphi_m}{\partial x_i \partial x_j} \to \frac{\partial^2 \varphi}{\partial x_i \partial x_j} \text{ and } A\varphi_m \to A\varphi$$

$$\text{pointwise dominatedly in } \mathcal{S} \backslash \partial D \text{ as } m \to \infty. \qquad (2.1.10)$$

We now formulate a set of sufficient conditions that a given function $\phi$ actually coincides with the value function $\Phi$ and that a corresponding stopping time, $\tau_D$, actually is optimal. The result is analogous to the variational inequality verification theorem for optimal stopping of continuous diffusions. See e.g. [Ø1, Theorem 10.4.1].

**Theorem 2.2 (Integro-variational inequalities for optimal stopping).**

**a)** *Suppose we can find a function* $\phi : \bar{\mathcal{S}} \to \mathbb{R}$ *such that*

(i) $\phi \in C^1(\mathcal{S}) \cap C(\bar{\mathcal{S}})$
(ii) $\phi \geq g$ *on* $\mathcal{S}$.
  *Define*

$$D = \{y \in \mathcal{S}; \phi(y) > g(y)\} \quad \text{(the continuation region)}.$$

  *Suppose*

(iii) $E^y \left[ \int\limits_0^{\tau_{\mathcal{S}}} \mathcal{X}_{\partial D}(Y(t)) dt \right] = 0$

(iv) $\partial D$ *is a Lipschitz surface*
(v) $\phi \in C^2(\mathcal{S} \setminus \partial D)$ *with locally bounded derivatives near* $\partial D$
(vi) $A\phi + f \leq 0$ *on* $\mathcal{S} \setminus \partial D$
(vii) $Y(\tau_{\mathcal{S}}) \in \partial \mathcal{S}$ *a.s. on* $\{\tau_{\mathcal{S}} < \infty\}$ *and* $\lim\limits_{t \to \tau_{\mathcal{S}}^-} \phi(Y(t)) = g(Y(\tau_{\mathcal{S}})) \cdot \mathcal{X}_{\{\tau_{\mathcal{S}} < \infty\}}$

  *and*

(viii) $E^y \Big[ |\phi(Y(\tau))| + \int\limits_0^{\tau_{\mathcal{S}}} \big\{ |A\phi(Y(t))| + |\sigma^T(Y(t))\nabla\phi(Y(t))|^2$

$$+ \sum_{j=1}^{\ell} \big[ \int_{\mathbb{R}} |\phi(Y(t) + \gamma^{(j)}(Y(t), z)) - \phi(Y(t))|^2 \nu_j(dz_j) \big] \big\} dt \Big] < \infty$$

  *for all* $\tau \in \mathcal{T}$.

*Then* $\phi(y) \geq \Phi(y)$ *for all* $y \in \bar{\mathcal{S}}$.

**b)** *Moreover, assume*

(ix) $A\phi + f = 0$ *on* $D$
(x) $\tau_D := \inf\{t > 0; Y(t) \notin D\} < \infty$ *a.s. for all* $y$
(xi) $\{\phi(Y(\tau)); \tau \in \mathcal{T}\}$ *is uniformly integrable, for all* $y$.
  *Then*

$$\phi(y) = \Phi(y)$$

  *and*

$$\tau^* = \tau_D \quad \text{is an optimal stopping time}.$$

PROOF OF THEOREM 2.2 (Sketch)

**a)** Let $\tau \leq \tau_{\mathcal{S}}$ be a stopping time. By Theorem 2.1 we can assume that $\phi \in C^2(\mathcal{S})$. Then by (vii) and (viii) and the Dynkin formula (Theorem 1.24) applied to $\tau_m := \min(\tau, m); \; m = 1, 2, \ldots$ we have, by (vi),

$$E^y[\phi(Y(\tau_m))] = \phi(y) + E^y\left[\int_0^{\tau_m} A\phi(Y(t))dt\right] \leq \phi(y) - E^y\left[\int_0^{\tau_m} f(Y(t))dt\right].$$

Hence by (ii) and the Fatou lemma

$$\phi(y) \geq \liminf_{m \to \infty} E^y\left[\int_0^{\tau_m} f(Y(t))dt + \phi(Y(\tau_m))\right]$$

$$\geq E^y\left[\int_0^{\tau} f(Y(t))dt + g(Y(\tau))\chi_{\{\tau < \infty\}}\right] = J^\tau(y).$$

Hence

$$\phi(y) \geq \Phi(y) . \tag{2.1.11}$$

**b)** Moreover, if we apply the above argument to $\tau = \tau_D$ then by (ix), (x) and (xi) and the definition of $D$ we get *equality* in (2.1.11), so that

$$\phi(y) = J^{\tau_D}(y) \leq \Phi(y) . \tag{2.1.12}$$

From (2.1.11) and (2.1.12) we conclude that $\phi(y) = \Phi(y)$ and $\tau_D$ is optimal. ∎

The following result is sometimes helpful.

**Proposition 2.3.** *Suppose the conditions of Theorem 2.2 hold. Suppose $g \in C^2(\mathbb{R}^k)$ and that $\phi = g$ satisfies (viii). Define*

$$U = \{y \in \mathcal{S}; Ag(y) + f(y) > 0\} .$$

*Suppose that for all $y \in U$ there exists a neighbourhood $W_y$ of $y$ such that $\tau_{W_y} := \inf\{t > 0; Y(t) \notin W_y\} < \infty$ a.s. Then*

$$U \subset \{y \in \mathcal{S}; \Phi(y) > g(y)\} = D .$$

*Hence it is never optimal to stop while $Y(t) \in U$.*

*Proof.* Choose $y_0 \in U$ and let $W \subset U$ be a neighbourhood of $y_0$ with $\tau_W < \infty$ a.s. Then by the Dynkin formula (Theorem 1.24)

$$E^y[g(Y(\tau_W))] = g(y) + E^y\left[\int_0^{\tau_W} Ag(Y(t))dt\right] > g(y) - E^y\left[\int_0^{\tau_W} f(Y(t))dt\right].$$

Hence

$$g(y) < E^y \left[ \int_0^{\tau_W} f(Y(t))dt + g(Y_{\tau_W}) \right] \leq \Phi(y) \, ,$$

as claimed. □

Another useful observation is

**Proposition 2.4.** *Let $U$ be as in Proposition 2.3. Suppose $U = \emptyset$. Then*

$$\Phi(y) = g(y) \quad and \quad \tau^* = 0 \quad is \ optimal.$$

*Proof.* If $U = \emptyset$ then $Ag(y) + f(y) \leq 0$ for all $y \in \mathcal{S}$. Hence the function $\phi = g$ satisfies all the conditions of Theorem 2.2. Therefore $D = \emptyset$, $g(y) = \Phi(y)$ and $\tau^* = 0$ is optimal. □

## 2.2 Applications and examples

**Example 2.5 (The optimal time to sell).** Suppose the price $X(t)$ at time $t$ of an asset (a property, a stock ...) is a geometric Lévy process given by

$$dX(t) = X(t^-) \left[ \alpha dt + \beta dB(t) + \gamma \int_{\mathbb{R}} z \widetilde{N}(dt, dz) \right], \qquad X(0) = x > 0 \, , \ (2.2.1)$$

where $\alpha, \beta$ and $\gamma$ are constants, $\gamma z > -1$ a.s. $\nu$. If we sell the asset at time $s + \tau$ we get the expected discounted net payoff

$$J^\tau(s, x) := E^{s,x} \left[ e^{-\rho(s+\tau)} (X(\tau) - a) \mathcal{X}_{\{\tau < \infty\}} \right]$$

where $\rho > 0$ (the discounting exponent) and $a > 0$ (the transaction cost) are constants.

We seek the value function $\Phi(s, x)$ and an optimal stopping time $\tau^* \leq \infty$ such that

$$\Phi(s, x) = \sup_{\tau \leq \infty} J^\tau(s, x) = J^{\tau^*}(s, x) \, . \tag{2.2.2}$$

We apply Theorem 2.2 to solve this problem as follows: Put $\mathcal{S} = \mathbb{R} \times (0, \infty)$ and

$$Y(t) = \begin{bmatrix} s + t \\ X(t) \end{bmatrix}; \qquad t \geq 0$$

Then

$$dY(t) = \begin{bmatrix} 1 \\ \alpha X(t) \end{bmatrix} dt + \begin{bmatrix} 0 \\ \beta X(t) \end{bmatrix} dB(t) + \begin{bmatrix} 0 \\ \gamma X(t^-) \int_{\mathbb{R}} z \widetilde{N}(dt, dz) \end{bmatrix}; \quad Y(0) = \begin{bmatrix} s \\ x \end{bmatrix}$$

and the generator $A$ of $Y(t)$ is

$$A\phi(s,x) = \frac{\partial\phi}{\partial s} + \alpha x \frac{\partial\phi}{\partial x} + \tfrac{1}{2}\beta^2 x^2 \frac{\partial^2\phi}{\partial x^2} + \int\limits_{\mathbb{R}} \left\{ \phi(s, x+\gamma xz) - \phi(s,x) - \gamma xz \frac{\partial\phi}{\partial x} \right\} \nu(dz) \; .$$
(2.2.3)

If we try a function $\phi$ of the form

$$\phi(s,x) = e^{-\rho s} x^\lambda \qquad \text{for some constant } \lambda \in \mathbb{R}$$

we get

$$\begin{aligned}
A\phi(s,x) &= e^{-\rho s}\Big[ -\rho x^\lambda + \alpha x \lambda x^{\lambda-1} + \tfrac{1}{2}\beta^2 x^2 \lambda(\lambda-1)x^{\lambda-2} \\
&\qquad + \int\limits_{\mathbb{R}} \{(x+\gamma xz)^\lambda - x^\lambda - \gamma xz\lambda x^{\lambda-1}\}\nu(dz)\Big] \\
&= e^{-\rho s} x^\lambda h(\lambda) \; ,
\end{aligned}$$

where

$$h(\lambda) = -\rho + \alpha\lambda + \tfrac{1}{2}\beta^2 \lambda(\lambda-1) + \int\limits_{\mathbb{R}} \{(1+\gamma z)^\lambda - 1 - \lambda\gamma z\}\nu(dz) \; .$$

Note that

$$h(1) = \alpha - \rho \qquad \text{and} \qquad \lim_{\lambda\to\infty} h(\lambda) = \infty \; .$$

Therefore, if we assume that

$$\alpha < \rho \; ,$$
(2.2.4)

then we get that there exists $\lambda_1 > 1$ such that

$$h(\lambda_1) = 0 \; .$$
(2.2.5)

With this value of $\lambda_1$ we put

$$\phi(s,x) = \begin{cases} e^{-\rho s} C x^{\lambda_1} & \text{for} \quad (s,x) \in D \\ e^{-\rho s}(x-a) & \text{for} \quad (s,x) \notin D \end{cases}$$
(2.2.6)

for some constant $C$, to be determined.

To find a reasonable guess for the continuation region $D$ we use Proposition 2.3: In this case we have $f = 0$ and $g(s,x) = e^{-\rho s}(x-a)$ and hence by (2.2.3)

$$Ag + f = e^{-\rho s}(-\rho(x-a) + \alpha x) = e^{-\rho s}((\alpha-\rho)x + \rho a) \; .$$

Therefore

$$U = \{(s,x); (\alpha-\rho)x + \rho a > 0\} \; .$$

**Case 1:  $\alpha \geq \rho$.**  In this case $U = \mathbb{R}^2$ and it is easily seen that $\Phi = \infty$: We can get as high expected payoff as we wish by waiting long enough before stopping.

**Case 2:  $\alpha < \rho$.**  In this case

$$U = \left\{ (s, x); x < \frac{\rho a}{\rho - \alpha} \right\}. \tag{2.2.7}$$

Therefore, in view of Proposition 2.3 we now guess that the continuation region $D$ has the form

$$D = \{(s, x); 0 < x < x^*\} \tag{2.2.8}$$

for some $x^*$ such that $U \subseteq D$, i.e.

$$x^* \geq \frac{\rho a}{\rho - \alpha}. \tag{2.2.9}$$

Hence, by (2.2.6) we now put

$$\phi(s, x) = \begin{cases} e^{-\rho s} C x^{\lambda_1} & \text{for } 0 < x < x^* \\ e^{-\rho s}(x - a) & \text{for } x^* \leq x, \end{cases} \tag{2.2.10}$$

for some constant $C > 0$ (to be determined). We guess that the value function is $C^1$ at $x = x^*$ and this gives the following "high contact"- conditions :

$$C(x^*)^{\lambda_1} = x^* - a \qquad \text{(continuity at } x = x^*)$$

and

$$C\lambda_1(x^*)^{\lambda_1 - 1} = 1 \qquad \text{(differentiability at } x = x^*).$$

It is easy to see that the solution of these equations is

$$x^* = \frac{\lambda_1 a}{\lambda_1 - 1}, \qquad C = \frac{1}{\lambda_1}(x^*)^{1 - \lambda_1}. \tag{2.2.11}$$

It remains to verify that with these values of $x^*$ and $C$ the function $\phi$ given by (2.2.10) satisfies all the conditions (i)–(xi) of Theorem 2.2.

To this end, first note that (i) and (ix) hold by construction of $\phi$. Moreover, $\phi = g$ outside $D$. Therefore, to verify (ii) we only need to prove that $\phi \geq g$ on $D$, i.e. that

$$C x^{\lambda_1} \geq x - a \qquad \text{for } 0 < x < x^*. \tag{2.2.12}$$

Define $k(x) = C x^{\lambda_1} - x + a$. By our chosen values of $C$ and $x^*$ we have $k(x^*) = k'(x^*) = 0$. Moreover, $k''(x) = C\lambda_1(\lambda_1 - 1)x^{\lambda_1 - 2} > 0$ for $x < x^*$. Therefore $k(x) > 0$ for $0 < x < x^*$ and (2.2.12) holds and hence (ii) is proved.

(iii): In this case $\partial D = \{(s, x); x = x^*\}$ and hence

$$E^y\left[\int_0^\infty \mathcal{X}_{\partial D}(Y(t))dt\right] = \int_0^\infty P^x[X(t) = x^*]dt = 0 \ .$$

(iv) and (v) are trivial.

(vi): Outside $D$ we have $\phi(s, x) = e^{-\rho s}(x - a)$ and therefore

$$A\phi + f(s, x) = e^{-\rho s}(-\rho(x - a) + \alpha x) = e^{-\rho s}((\alpha - \rho)x + \rho a)$$

So by (2.2.4) we get that

$$A\phi + f(s, x) \leq 0 \qquad \text{for all} \quad x \geq x^*$$
$$\Updownarrow$$
$$(\alpha - \rho)x + \rho a \leq 0 \qquad \text{for all} \quad x \geq x^*$$
$$\Updownarrow$$
$$(\alpha - \rho)x^* + \rho a \leq 0$$
$$\Updownarrow$$
$$x^* \geq \frac{\rho a}{\rho - \alpha} \ ,$$

which holds by (2.2.9).

(x): To check if $\tau_D < \infty$ a.s. we consider the solution $X(t)$ of (2.2.1), which by (1.2.5) is given by

$$X(t) = x \exp\left\{\left(\alpha - \tfrac{1}{2}\beta^2 - \gamma \int_{\mathbb{R}} z\nu(dz)\right)t + \int_0^t\int_{\mathbb{R}} \ln(1 + \gamma z)N(dt, dz) + \beta B(t)\right\}.$$

$$(2.2.13)$$

By the law of iterated logarithm for Brownian motion (see the argument in [Ø, Chapter 5]) we see that if

$$\alpha > \tfrac{1}{2}\beta^2 + \gamma \int_{\mathbb{R}} z\nu(dz) \qquad\qquad (2.2.14)$$

and

$$z \geq 0 \quad \text{a.s. } \nu \qquad\qquad (2.2.15)$$

then

$$\lim_{t\to\infty} X(t) = \infty \quad \text{a.s.}$$

and in particular $\tau_D < \infty$ a.s.

(xi): Since $\phi$ is bounded on $[0, x^*]$ it suffices to check that

$$\left\{e^{-\rho\tau}X(\tau)\right\}_{\tau\in\mathcal{T}} \qquad \text{is uniformly integrable.}$$

For this to hold it suffices that there exists a constant $K$ such that

$$E[e^{-2\rho\tau}X^2(\tau)] \leq K \qquad \text{for all } \tau \in \mathcal{T} . \tag{2.2.16}$$

By (2.2.13) and Exercise 1.6 we have

$$E[e^{-2\rho T}X^2(\tau)] = x^2 E\Big[\exp\Big\{\Big(2\alpha - 2\rho - \beta^2 - 2\gamma \int_{\mathbb{R}} z\nu(dz)\Big)\tau$$

$$+ 2\int_0^\tau \int_{\mathbb{R}} \ln(1+\gamma z)N(dt,dz) + 2\beta B(\tau)\Big\}\Big]$$

$$= x^2 E\Big[\exp\Big\{\Big(2\alpha - 2\rho + \beta^2 + 2\int_{\mathbb{R}} (\ln(1+\gamma z) - \gamma z)\nu(dz)\Big)\tau$$

$$+ 2\int_0^\tau \int_{\mathbb{R}} \ln(1+\gamma z)\widetilde{N}(dt,dz)\Big\}\Big]$$

$$= x^2 E\Big[\exp\Big\{\Big(2\alpha - 2\rho + \beta^2 + \int_{\mathbb{R}} [2\ln(1+\gamma z) - 2\gamma z + (1+\gamma z)^2$$

$$- 1 - 2\ln(1+\gamma z)]\nu(dz)\Big)\tau\Big\}\Big]$$

$$= x^2 E\Big[\exp\Big\{\Big(2\alpha - 2\rho + \beta^2 + \int_{\mathbb{R}} [(1+\gamma z)^2 - 1 - 2\gamma z]\nu(dz)\Big)\tau\Big\}\Big]$$

We conclude that if

$$2\alpha - 2\rho + \beta^2 + \gamma^2 \int_{\mathbb{R}} z^2\nu(dz) \leq 0$$

then (2.2.16) holds and hence (xi) holds also.

(vii): holds since we have assumed that $z > -1$ a.s. $\nu$.

Finally, for (viii) to hold it suffices that

$$E^x\Big[\int_0^\infty e^{-2\rho t}\Big\{X^2(t) + \gamma^2 \int_{\mathbb{R}} z^2\nu(dz)t\Big\}dt\Big] < \infty .$$

By the above this holds if

$$2\alpha - 2\rho + \beta^2 + \gamma^2 \int_{\mathbb{R}} z^2\nu(dz) < 0 . \tag{2.2.17}$$

We summarize what we have proved:

**Theorem 2.6.** *Suppose that $\alpha < \rho$ and that (2.2.14) and (2.2.17) hold. Then, with $\lambda_1, C$ and $x^*$ given by (2.2.5) and (2.2.11), the function $\phi$ given by (2.2.10) coincides with the value function $\Phi$ of problem (2.2.2) and $\tau^* = \tau_D$ is an optimal stopping time, where $D$ is given by (2.2.8).*

**Remark 2.7.** For other applications of optimal stopping to jump diffusions we refer to [Ma].

## 2.3 Exercises

**Exercise\* 2.1.** Solve the optimal stopping problem

$$\Phi(s, x) = \sup_{\tau \geq 0} E^{(s,x)}\left[e^{-\rho(s+\tau)}(X(\tau) - a)\right]$$

where

$$dX(t) = dB(t) + \gamma \int_{\mathbb{R}} z\bar{N}(dt, dz) ; \qquad X(0) = x \in \mathbb{R}$$

and $\rho > 0$, $a > 0$, $\sigma$ and $\gamma$ are constants.

**Exercise\* 2.2 (An optimal resource extraction stopping problem).**
Suppose the price $P(t)$ per unit of a resource (oil, gas ...) at time $t$ is given by

(i)   $dP(t) = \alpha P(t)dt + \beta P(t)dB(t) + \gamma P(t^-) \int_{\mathbb{R}} z\widetilde{N}(dt, dz); \quad P(0) = p > 0$

and the remaining amount of resources $Q(t)$ at time $t$ is

(ii)   $dQ(t) = -\lambda Q(t)dt ; \quad Q(0) = q > 0$

where $\lambda > 0$ is the (constant) relative extraction rate and $\alpha, \beta, \gamma \geq 0$ are constants. We assume that $\gamma z > -1$ a.s. $\nu$.
    If we decide to stop extraction and close the field at a stopping time $\tau \geq 0$, the expected discounted total net profit $J^\tau(s, p, q)$ is given by

$$J^\tau(s, p, q) = E^{(s,p,q)}\left[\int_0^\tau e^{-\rho(s+t)}(\lambda P(t)Q(t) - K)dt + \theta e^{-\rho(s+\tau)}P(\tau)Q(\tau)\right],$$

where $K > 0$ is the (constant) running cost rate, $\theta > 0$ another constant. Find $\Phi$ and $\tau^*$ such that

$$\Phi(s, p, q) = \sup_{\tau \geq 0} J^\tau(s, p, q) = J^{\tau^*}(s, p, q) .$$

[*Hint*: Try $\phi(s, p, q) = e^{-\rho s}\psi(p \cdot q)$ for some function $\psi : \mathbb{R} \to \mathbb{R}$.]

**Exercise\* 2.3.** Solve the optimal stopping problem

$$\Phi(s, x) = \sup_{\tau \geq 0} E^x \left[ e^{-\rho(s+\tau)} |X(\tau)| \right]$$

where

$$dX(t) = dB(t) + \int_{\mathbb{R}} z \tilde{N}(dt, dz)$$

and $\rho > 0$ is a constant. Assume that the Lévy measure $\nu$ of $X$ is symmetric, i.e.

$$\nu(G) = \nu(-G) \text{ for all measurable } G \subset \mathbb{R} \backslash \{0\}.$$

# 3

## Stochastic Control of Jump Diffusions

### 3.1 Dynamic programming

Fix a domain $\mathcal{S} \subset \mathbb{R}^k$ (our solvency region ) and let $Y(t) = Y^{(u)}(t)$ be a stochastic process of the form

$$
\begin{aligned}
dY(t) =& b(Y(t), u(t))dt + \sigma(Y(t), u(t))dB(t) \\
& + \int_{\mathbb{R}^k} \gamma(Y(t^-), u(t^-), z)\bar{N}(dt, dz) , \quad Y(0) = y \in \mathbb{R}^k , \qquad (3.1.1)
\end{aligned}
$$

where

$$
b : \mathbb{R}^k \times U \to \mathbb{R}^k, \quad \sigma : \mathbb{R}^k \times U \to \mathbb{R}^{k \times m} \quad \text{and} \quad \gamma : \mathbb{R}^k \times U \times \mathbb{R}^k \to \mathbb{R}^{k \times \ell}
$$

are given functions, $U \subset \mathbb{R}^k$ is a given set. The process $u(t) = u(t, \omega) :$ $[0, \infty) \times \Omega \to U$ is our *control process* , assumed to be cadlag and adapted. We call $Y(t) = Y^{(u)}(t)$ a *controlled jump diffusion* .

We consider a performance criterion $J = J^{(u)}(y)$ of the form

$$
J^{(u)}(y) = E^y \Big[ \int_0^{\tau_{\mathcal{S}}} f(Y(t), u(t))dt + g(Y(\tau_{\mathcal{S}})) \cdot \mathcal{X}_{\{\tau_{\mathcal{S}} < \infty\}} \Big]
$$

where

$$
\tau_{\mathcal{S}} = \inf\{t > 0; Y^{(u)}(t) \notin \mathcal{S}\} \qquad \text{(the bankruptcy time )}
$$

and $f : \mathcal{S} \to \mathbb{R}$ and $g : \mathbb{R}^k \to \mathbb{R}$ are given continuous functions.

We say that the control process $u$ is *admissible* and write $u \in \mathcal{A}$ if (3.1.1) has a unique, strong solution $Y(t)$ for all $y \in \mathcal{S}$ and

$$
E^y \Big[ \int_0^{\tau_{\mathcal{S}}} f^-(Y(t), u(t))dt + g^-(Y(\tau_{\mathcal{S}})) \cdot \mathcal{X}_{\{\tau_{\mathcal{S}} < \infty\}} \Big] < \infty .
$$

The *stochastic control problem* is to find the *value function* $\Phi(y)$ and an *optimal control* $u^* \in \mathcal{A}$ defined by

$$\Phi(y) = \sup_{u \in \mathcal{A}} J^{(u)}(y) = J^{(u^*)}(y) . \qquad (3.1.2)$$

It turns out that – under mild conditions (see e.g. [Ø1, Theorem 11.2.3])– it suffices to consider *Markov* controls , i.e. controls $u(t)$ of the form

$$u(t) = u_0(Y(t^-))$$

for some function $u_0 : \mathbb{R}^k \to U$. Therefore, from now on we will only consider Markov controls and we will, with a slight abuse of notation, write $u(t) = u(Y(t^-))$.

Note that if $u = u(y)$ is a Markov control then $Y(t) = Y^{(u)}(t)$ is a Lévy diffusion with generator

$$A\phi(y) = A^u \phi(y) = \sum_{i=1}^k b_i(y, u(y)) \frac{\partial \phi}{\partial y_i}(y) + \tfrac{1}{2} \sum_{i,j=1}^k (\sigma\sigma^T)_{ij}(y, u(y)) \cdot \frac{\partial^2 \phi}{\partial y_i \partial y_j}(y)$$

$$+ \sum_{j=1}^\ell \int_{\mathbb{R}} \{\phi(y + \gamma^{(j)}(y, u(y), z_j)) - \phi(y) - \nabla\phi(y) \cdot \gamma^{(j)}(y, u(y), z_j)\} \nu_j(dz_j) .$$

We now formulate a verification theorem for the optimal control problem (3.1.2), analogous to the classical Hamilton-Jacobi-Bellman (HJB) for (continuous) Itô diffusions:

**Theorem 3.1 (HJB for optimal control of jump diffusions).**
**a)** *Suppose* $\phi \in C^2(\mathcal{S}) \cap C(\bar{\mathcal{S}})$ *satisfies the following:*

(i)     $A^v \phi(y) + f(y, v) \leq 0$   *for all* $y \in \mathcal{S}$, $v \in U$
(ii)     $Y(\tau_{\mathcal{S}}) \in \partial\mathcal{S}$ *a.s. on* $\{\tau_{\mathcal{S}} < \infty\}$ *and*
         $\lim_{t \to \tau_{\mathcal{S}}^-} \phi(Y(t)) = g(Y(\tau_{\mathcal{S}})) \cdot \mathcal{X}_{\{\tau_{\mathcal{S}} < \infty\}}$ *a.s., for all* $u \in \mathcal{A}$

(iii)    $E^y \Big[ |\phi(Y(\tau))| + \int_0^{\tau_{\mathcal{S}}} \{|A\phi(Y(t))| + |\sigma^T(Y(t)) \nabla\phi(Y(t))|^2$

         $+ \sum_{j=1}^\ell \int_{\mathbb{R}} |\phi(Y(t) + \gamma^{(j)}(Y(t), u(t), z_j)) - \phi(Y(t))|^2 \nu_j(dz_j)\} dt \Big] < \infty,$
    *for all* $u \in \mathcal{A}$ *and all* $\tau \in \mathcal{T}$.
(iv)    $\{\phi^-(Y(\tau))\}_{\tau \leq \tau_{\mathcal{S}}}$ *is uniformly integrable for all* $u \in \mathcal{A}$ *and* $y \in \mathcal{S}$.

*Then*

$$\phi(y) \geq \Phi(y) \qquad \text{for all } y \in \mathcal{S} . \qquad (3.1.3)$$

**b)** *Moreover, suppose that for each* $y \in \mathcal{S}$ *there exists* $v = \hat{u}(y) \in U$ *such that*

(v)     $A^{\hat{u}(y)} \phi(y) + f(y, \hat{u}(y)) = 0$

*and*

(vi)    $\{\phi(Y^{(\hat{u})}(\tau))\}_{\tau \leq \tau_S}$ *is uniformly integrable.*

*Suppose $u^*(t) := \hat{u}(Y(t^-)) \in \mathcal{A}$. Then $u^*$ is an optimal control and*

$$\phi(y) = \Phi(y) = J^{(u^*)}(y) \qquad for \ all \ y \in \mathcal{S} . \tag{3.1.4}$$

*Proof.* **a)** Let $u \in \mathcal{A}$. For $n = 1, 2, \ldots$ put $\tau_n = \min(n, \tau_S)$. Then by the Dynkin formula (Theorem 1.24) we have

$$E^y[\phi(Y(\tau_n))] = \phi(y) + E^y\left[\int_0^{\tau_n} A^u\phi(Y(t))dt\right] \leq \phi(y) - E\left[\int_0^{\tau_n} f(Y(t), u(t))dt\right].$$

Hence

$$\phi(y) \geq \liminf_{n\to\infty} E^y\left[\int_0^{\tau_n} f(Y(t), u(t))dt + \phi(Y(\tau_n))\right]$$

$$\geq E^y\left[\int_0^{\tau_S} f(Y(t), u(t))dt + g(Y(\tau_S)) \cdot \mathcal{X}_{\{\tau_S < \infty\}}\right] = J^{(u)}(y) . \tag{3.1.5}$$

Since $u \in \mathcal{A}$ was arbitrary we conclude that

$$\phi(y) \geq \Phi(y) \qquad for \ all \ y \in \mathcal{S} . \tag{3.1.6}$$

**b)** Now apply the above argument to $u(t) = \hat{u}(Y(t))$, where $\hat{u}$ is as in (v). Then we get *equality* in (3.1.5) and hence

$$\phi(y) = J^{(\hat{u})}(y) \leq \Phi(y) \qquad for \ all \ y \in \mathcal{S} . \tag{3.1.7}$$

Combining (3.1.6) and (3.1.7) we get (3.1.4).                        $\square$

**Example 3.2 (Optimal consumption and portfolio in a Lévy type Black-Scholes market [Aa], [FØS1]).**
   Suppose we have a market with two possible investments:

(i)   a *safe* investment (bond, bank account) with price dynamics

$$dP_1(t) = rP_1(t)dt ; \qquad P_1(0) = p_1 > 0$$

(ii)   a *risky* investment (stock) with price dynamics

$$dP_2(t) = P_2(t^-)\left[\mu dt + \sigma dB(t) + \int_{-1}^{\infty} z\tilde{N}(dt, dz)\right], \qquad P_2(0) = p_2 > 0$$

where $r > 0$, $\mu > 0$ and $\sigma \in \mathbb{R}$ are constants. We assume that

$$\int_{-1}^{\infty} |z| d\nu(z) < \infty \quad \text{and} \quad \mu > r .$$

Assume that at any time $t$ the investor can choose a consumption rate $c(t) \geq 0$ (adapted, cadlag) and is also free to transfer money from one investment to the other without transaction cost. Let $X_1(t), X_2(t)$ be the amounts of money invested in the bonds and the stocks, respectively. Let

$$\theta(t) = \frac{X_2(t)}{X_1(t) + X_2(t)}$$

be the fraction of the total wealth invested in stocks at time $t$. Define the *performance criterion* by

$$J^{(c,\theta)}(s, x_1, x_2) = E^{x_1, x_2}\left[\int_0^{\infty} e^{-\delta(s+t)} \frac{c^{\gamma}(t)}{\gamma} dt\right]$$

where $\delta > 0$, $\gamma \in (0,1)$ are constants and $E^{x_1,x_2}$ is the expectation w.r.t. the probability law $P^{x_1,x_2}$ of $(X_1(t), X_2(t))$ when $X_1(0^-) = x_1$, $X_2(0^-) = x_2$. Call the control $u(t) = (c(t), \theta(t)) \in [0, \infty) \times [0, 1]$ *admissible* and write $u \in \mathcal{A}$ if the corresponding *total wealth*

$$W(t) = W^{(u)}(t) = X_1^{(u)}(t) + X_2^{(u)}(t)$$

is *nonnegative* for all $t \geq 0$.

The problem is to find $\Phi(s, x_1, x_2)$ and $u^*(c^*, \theta^*) \in \mathcal{A}$ such that

$$\Phi(s, x_1, x_2) = \sup_{u \in \mathcal{A}} J^{(u)}(s, x_1, x_2) = J^{(u^*)}(s, x_1, x_2) .$$

*Case 1:* $\boldsymbol{\nu = 0}$ .

In this case the problem was solved by Merton [M]. He proved that if

$$\delta > \gamma\left[r + \frac{(\mu - r)^2}{2\sigma^2(1 - \gamma)}\right] \tag{3.1.8}$$

then the value function is

$$\Phi_0(s, x_1, x_2) = K_0 e^{-\delta s}(x_1 + x_2)^{\gamma} \tag{3.1.9}$$

where

$$K_0 = \frac{1}{\gamma}\left[\frac{1}{1-\gamma}\left(\delta - \gamma r - \frac{\gamma(\mu - r)^2}{2\sigma^2(1 - \gamma)}\right)\right]^{\gamma - 1} . \tag{3.1.10}$$

Moreover, the optimal consumption rate $c_0^*(t)$ is given by

$$c_0^*(t) = (K_0\gamma)^{\frac{1}{\gamma-1}}(X_1(t) + X_2(t)) \tag{3.1.11}$$

and the optimal portfolio $\theta_0^*(t)$ is (the constant)

$$\theta_0^*(t) = \frac{\mu - r}{\sigma^2(1 - \gamma)} \qquad \text{for all } t \in [0, \infty) . \tag{3.1.12}$$

In other words, it is optimal to keep the state $(X_1(t), X_2(t))$ on the line

$$x_2 = \frac{\theta_0^*}{1 - \theta_0^*} x_1 \tag{3.1.13}$$

in the $(x_1, x_2)$-plane at all times (the Merton line). See figure 3.1.
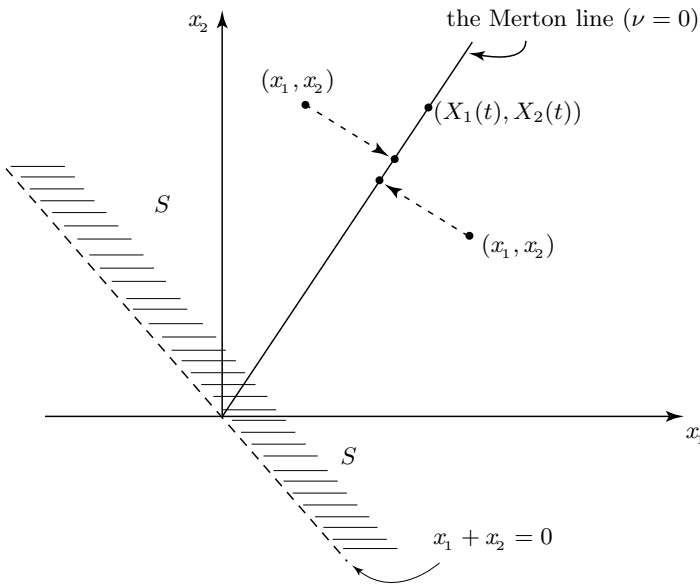


**Fig. 3.1.** The Merton line

*Case 2:* $\boldsymbol{\nu \neq 0}$

We now ask: How does the presence of jumps influence the optimal strategy?

As in [M] we reduce the dimension by introducing

$$W(t) = X_1(t) + X_2(t) .$$

Then we see that

$$dW(t) = \big([r(1 - \theta(t)) + \mu\theta(t)]W(t) - c(t)\big)dt + \sigma\theta(t)W(t)dB(t)$$

$$+ \theta(t)W(t^-) \int\limits_{-1}^{\infty} z\widetilde{N}(dt, dz) ; \quad W(0^-) = x_1 + x_2 = w \geq 0 .$$

The generator $A^{(u)}$ of the controlled process

$$Y(t) = \begin{bmatrix} s + t \\ W(t) \end{bmatrix} ; \quad t \geq 0 , \quad Y(0^-) = y = \begin{bmatrix} s \\ w \end{bmatrix}$$

is

$$A^{(u)}\phi(y) = \frac{\partial\phi}{\partial s} + \big([r(1 - \theta) + \mu\theta]w - c\big)\frac{\partial\phi}{\partial w} + \tfrac{1}{2}\sigma^2\theta^2 w^2 \frac{\partial^2\phi}{\partial w^2}$$

$$+ \int\limits_{-1}^{\infty} \Big\{\phi(s, w + \theta wz) - \phi(s, w) - \frac{\partial\phi}{\partial w}(s, w)\theta wz\Big\}\nu(dz) .$$

If we try

$$\phi(y) = \phi(s, w) = e^{-\delta s}\psi(w)$$

we get

$$A^{(u)}\phi(y) = e^{-\delta s}A_0^{(u)}\psi(w), \quad \text{where}$$

$$A_0^{(u)}\psi(w) = -\rho\psi(w) + \big([r(1 - \theta) + \mu\theta]w - c\big)\psi'(w) + \tfrac{1}{2}\sigma^2\theta^2 w^2\psi''(w)$$

$$+ \int\limits_{-1}^{\infty} \{\psi((1 + \theta z)w) - \psi(w) - \psi'(w)\theta wz\}\nu(dz) .$$

In particular, if we try

$$\psi(w) = Kw^\gamma$$

we get

$$A_0^{(u)}\psi(w) + f(w, u) = -\rho Kw^\gamma + \big([r(1 - \theta) + \mu\theta]w - c\big)K\gamma w^{\gamma-1}$$

$$+ K \cdot \tfrac{1}{2}\sigma^2\theta^2 w^2\gamma(\gamma - 1)w^{\gamma-2} + Kw^\gamma \int\limits_{-1}^{\infty} \{(1 + \theta z)^\gamma - 1 - \gamma\theta z\}\nu(dz) + \frac{c^\gamma}{\gamma} .$$

Let $h(c, \theta)$ be the expression on the right hand side. Then $h$ is concave in $(c, \theta)$ and the maximum of $h$ is attained at the critical points, i.e. when

$$\frac{\partial h}{\partial c} = -K\gamma w^{\gamma-1} + c^{\gamma-1} = 0 \qquad\qquad (3.1.14)$$

and

$$\frac{\partial h}{\partial \theta} = (\mu-r)K\gamma w^\gamma + K\sigma^2\theta\gamma(\gamma-1)w^\gamma + Kw^\gamma \int\limits_{-1}^{\infty} \{\gamma(1+\theta z)^{\gamma-1}z - \gamma z\}\nu(dz) = 0 \; .$$

$$(3.1.15)$$

From (3.1.14) we get

$$c = \hat{c} = (K\gamma)^{\frac{1}{\gamma-1}} w \tag{3.1.16}$$

and from (3.1.15) we get that $\theta = \hat{\theta}$ should solve the equation

$$\Lambda(\theta) := \mu - r - \sigma^2\theta(1-\gamma) - \int\limits_{-1}^{\infty} \{1 - (1+\theta z)^{\gamma-1}\}z\nu(dz) = 0 \; . \tag{3.1.17}$$

Since $\Lambda(0) = \mu - r > 0$ we see that if

$$\sigma^2(1-\gamma) + \int\limits_{-1}^{\infty} \{1 - (1+z)^{\gamma-1}\}z\nu(dz) \geq \mu - r \tag{3.1.18}$$

then there exists an optimal $\theta = \hat{\theta} \in (0,1]$.

With this choice of $c = \hat{c} = (K\gamma)^{\frac{1}{\gamma-1}} w$ and $\theta = \hat{\theta}$ (constant) we require that

$$A_0^{(\hat{u})}\psi(w) + f(w, \hat{u}) = 0 \qquad \text{i.e.}$$

$$-\rho K + ([r(1-\hat{\theta}) + \mu\hat{\theta}] - (K\gamma)^{\frac{1}{\gamma-1}})K\gamma$$

$$+ K\tfrac{1}{2}\sigma^2\hat{\theta}^2\gamma(\gamma-1) + K\int\limits_{-1}^{\infty} \{(1+\hat{\theta}z)^\gamma - 1 - \gamma\hat{\theta}z\}\nu(dz) + (K\gamma)^{\frac{\gamma}{\gamma-1}}\frac{1}{\gamma} = 0$$

or

$$-\delta + \gamma[r(1-\hat{\theta}) + \mu\hat{\theta}] - (K\gamma)^{\frac{1}{\gamma-1}}\gamma$$

$$-\tfrac{1}{2}\sigma^2\hat{\theta}^2(1-\gamma)\gamma + \int\limits_{-1}^{\infty} \{(1+\hat{\theta}z)^\gamma - 1 - \gamma\hat{\theta}z\}\nu(dz) + K^{\frac{1}{\gamma-1}} \cdot \gamma^{\frac{\gamma}{\gamma-1}} \cdot \frac{1}{\gamma}$$

or

$$(K\gamma)^{\frac{1}{\gamma-1}}[1-\gamma] = \delta - \gamma[r(1-\hat{\theta}) + \mu\hat{\theta}] + \tfrac{1}{2}\sigma^2\hat{\theta}^2(1-\gamma)\gamma - \int\limits_{-1}^{\infty} \{(1+\hat{\theta}z)^\gamma - 1 - \gamma\hat{\theta}z\}\nu(dz)$$

or

$$K = \frac{1}{\gamma}\Big[\frac{1}{1-\gamma}\Big(\delta - \gamma\{r(1-\hat{\theta}) + \mu\hat{\theta}\} + \tfrac{1}{2}\sigma^2\hat{\theta}^2(1-\gamma)\gamma$$

$$- \int\limits_{\mathbb{R}} \{(1+\hat{\theta}z)^\gamma - 1 - \gamma\hat{\theta}z\}\nu(dz)\Big)\Big]^{\gamma-1} . \tag{3.1.19}$$

We now study condition (iii):

Here $\sigma^T \nabla \phi(y) = e^{-\delta s} \sigma \theta w K \gamma w^{\gamma-1} = e^{-\delta s} \sigma \theta K w^\gamma$ and

$$\phi(Y(t) + \gamma(Y(t), u(t))) - \phi(Y(t)) = KW(t)^\gamma e^{-\delta s}[(1 + \theta z)^\gamma - 1].$$

So (iii) holds if

$$E\Big[ \int_0^T e^{-2\delta t} W^{2\gamma}(t) dt \Big] + \int_{\mathbb{R}} [(1 + \theta z)^\gamma - 1] \nu(dz) < \infty \qquad (3.1.20)$$

We refer to [FØS1] for sufficient conditions on the parameters for (3.1.20) to hold.

We conclude that the value function is

$$\Phi(s, w) = \Phi(s, x_1, x_2) = e^{-\delta s}(x_1 + x_2)^\gamma \qquad (3.1.21)$$

with optimal control $u^*(t) = (c^*(t), \theta^*(t))$ where $c^* = \hat{c} = (K\gamma)^{\frac{1}{\gamma-1}}(x_1 + x_2)$ is given by (3.1.16) and $\theta^* = \hat{\theta}$ is given by (3.1.17), with $K$ given by (3.1.19).

Finally we compare the solution in the jump case ($\nu \neq 0$) with Merton's solution in the no jump case ($\nu = 0$):

As before let $\Phi_0, c_0^*$ and $\theta_0^*$ be the solution when there are no jumps ($\nu = 0$). Then it can be seen that

$$K < K_0 \quad \text{and hence} \quad \Phi(s, w) = e^{-\delta s} K w^\gamma < e^{-\delta s} K_0 w^\gamma = \Phi_0(s, w)$$

$$c^*(s, w) \geq c_0^*(s, w)$$

$$\theta^* \leq \theta_0^* .$$

So with jumps it is optimal to place a *smaller* wealth fraction in the risky investment, consume *more* relative to the current wealth and the resulting value is *smaller* than in the no-jump case.

For more details we refer to [FØS1].

**Remark 3.3.** For more information and other applications of stochastic control of jump diffusions see [GS], [BKR], [Ma] and the references therein.

## 3.2 The maximum principle

Suppose the state $X(t) = X^{(u)}(t)$ of a controlled jump diffusion in $\mathbb{R}^n$ is given by

$$dX(t) = b(t, X(t), u(t))dt + \sigma(t, X(t), u(t))dB(t)$$
$$+ \int_{\mathbb{R}^n} \gamma(t, X(t^-), u(t^-), z) \widetilde{N}(dt, dz). \qquad (3.2.1)$$
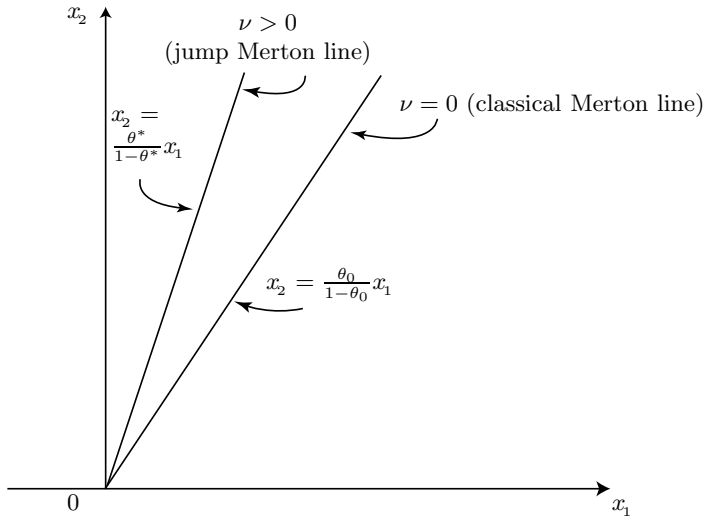
**Fig. 3.2.** The Merton line for $\nu = 0$ and $\nu > 0$

As before $\widetilde{N}(dt, dz) = (\widetilde{N}_1(dt, dz_1), \ldots, \widetilde{N}_\ell(dt, dz_\ell))^T$ where

$$\widetilde{N}_j(dt, dz_j) = N_j(dt, dz_j) - \nu_j(dz_j)dt \;; \qquad 1 \le j \le \ell$$

(see the notation of Theorem 1.16).

   The process $u(t) = u(t, \omega) \in U \subset \mathbb{R}^k$ is our *control*. We assume that $u$ is adapted and cadlag, and that the corresponding equation (3.2.1) has a unique strong solution $X^{(u)}(t); t \in [0, T]$. Such controls are called *admissible*. The set of admissible controls is denoted by $\mathcal{A}$.

   Suppose the performance criterion has the form

$$J(u) = E\Big[ \int\limits_0^T f(t, X(t), u(t))dt + g(X(T))\Big]; \qquad u \in \mathcal{A}$$

where $f : [0, T] \times \mathbb{R}^n \times U \to \mathbb{R}$ is continuous, $g : \mathbb{R}^n \to \mathbb{R}$ is $C^1$, $T < \infty$ is a fixed deterministic time and

$$E\Big[ \int\limits_0^T f^-(t, X(t), u(t))dt + g^-(X(T))\Big] < \infty \qquad \text{for all } u \in \mathcal{A} \,.$$

Consider the problem to find $u^* \in \mathcal{A}$ such that

$$J(u^*) = \sup_{u \in \mathcal{A}} J(u) \,. \tag{3.2.2}$$

In the previous chapter we saw how to solve such a problem using dynamic programming and the associated HJB equation. Here we present an alternative approach, based on what is called *the maximum principle* . In the deterministic case this principle was first introduced by Pontryagin and his group [PBGM]. A corresponding maximum principle for Itô diffusions was formulated by Kushner [Ku], Bismut [Bi] and subsequently further developed by Bensoussan [Ben], Haussmann [H] and others. For *jump diffusions* a sufficient maximum principle has recently been formulated in [FØS3] and it is this approach that is presented here, in a somewhat simplified version.

Define the *Hamiltonian* $H : [0,T] \times \mathbb{R}^n \times U \times \mathbb{R}^n \times \mathbb{R}^{n \times m} \times \mathcal{R} \to \mathbb{R}$ by

$$H(t,x,u,p,q,r) = f(t,x,u) + b^T(t,x,u)p + \text{tr}(\sigma^T(t,x,u)q)$$

$$+ \sum_{j=1}^{\ell} \sum_{i=1}^{n} \int_{\mathbb{R}} \gamma_{ij}(t,x,u,z_j) r_{ij}(t,z) \nu_j(dz_j) \qquad (3.2.3)$$

where $\mathcal{R}$ is the set of functions $r : \mathbb{R}^{n+1} \to \mathbb{R}^{n \times \ell}$ such that the integrals in (3.2.3) converge. From now on we assume that $H$ is differentiable with respect to $x$.

The *adjoint* equation (corresponding to $u$ and $X^{(u)}$) in the unknown processes $p(t) \in \mathbb{R}^n$, $q(t) \in \mathbb{R}^{n \times m}$ and $r(t,z) \in \mathbb{R}^{n \times \ell}$ is the backward stochastic differential equation

$$\begin{cases} dp(t) = -\nabla_x H(t, X(t), u(t), p(t), q(t), r(t,\cdot)) dt \\ \qquad\qquad + q(t) dB(t) + \displaystyle\int_{\mathbb{R}^n} r(t^-, z) \widetilde{N}(dt, dz) ; \qquad t < T \\ p(T) = \nabla g(X(T)) \end{cases} \qquad (3.2.4)$$

We assume from now on that

$$E\Big[ \int_0^T \Big\{ \sigma\sigma^T(t, X(t), u(t)) $$

$$+ \sum_{k=1}^{\ell} \int_{\mathbb{R}} \big| \gamma^{(k)}(t, X(t), u(t), z_k) \big|^2 \nu_k(dz_k) \Big\} dt \Big] < \infty \qquad \text{for all } u \in \mathcal{A} .$$

**Theorem 3.4 (A sufficient maximum principle [FØS3]).** *Let $\hat{u} \in \mathcal{A}$ with corresponding solution $\hat{X} = X^{(\hat{u})}$ and suppose there exists a solution $(\hat{p}(t), \hat{q}(t), \hat{r}(t,z))$ of the corresponding adjoint equation (3.2.4) satisfying*

$$E\Big[ \int_0^T \Big\{ \hat{q}\hat{q}^T(t) + \sum_{k=1}^{n} \int_{\mathbb{R}} |r^{(k)}(t,z_k)|^2 \nu_k(dz_k) \Big\} dt \Big] < \infty . \qquad (3.2.5)$$

*Moreover, suppose that*

$$H(t, \hat{X}(t), \hat{u}(t), \hat{p}(t), \hat{q}(t, ), \hat{r}(t, \cdot)) = \sup_{v \in U} H(t, \hat{X}(t), v, \hat{p}(t), \hat{q}(t), \hat{r}(t, \cdot))$$

*for all $t$, that $g(x)$ is a concave function of $x$ and that*

$$\widehat{H}(x) := \max_{v \in U} H(t, x, v, \hat{p}(t), \hat{q}(t), \hat{r}(t, \cdot)) \ \textit{exists and is}$$

*a concave function of $x$, for all $t \in [0, T]$ (the Arrow condition)*

(3.2.6)

*Then $\hat{u}$ is an optimal control.*

**Remark 3.5.** For (3.2.6) to hold it suffices that the function

$$(x, v) \rightarrow H(t, x, v, \hat{p}(t), \hat{q}(t), \hat{r}(t, \cdot)) \qquad \text{is concave, for all } t \in [0, T] \ . \quad (3.2.7)$$

To prove Theorem 3.4 we first establish the following:

**Lemma 3.6 (Integration by parts).** *Suppose $E[(Y^{(j)}(T)^2] < \infty$ for $j = 1, 2$, where*

$$dY^{(j)}(t) = b^{(j)}(t, \omega)dt + \sigma^{(j)}(t, \omega)dB(t) + \int_{\mathbb{R}^n} \gamma^{(j)}(t, z, \omega)\widetilde{N}(dt, dz)$$

$$Y^{(j)}(0) = y^{(j)} \in \mathbb{R}^n ; \qquad j = 1, 2$$

*where $b^{(j)} \in \mathbb{R}^n$, $\sigma^{(j)} \in \mathbb{R}^{n \times m}$ and $\gamma^{(j)} \in \mathbb{R}^{n \times \ell}$. Then*

$$E[Y^{(1)}(T) \cdot Y^{(2)}(T)] = y_1 \cdot y_2 + E\Big[\int_0^T Y^{(1)}(t^-)dY^{(2)}(t) + \int_0^T Y^{(2)}(t^-)dY^{(1)}(t)$$

$$+ \int_0^T \operatorname{tr}[\sigma^{(1)^T}\sigma^{(2)}](t)dt + \int_0^T \Big[ \sum_{j=1}^{\ell} \Big( \sum_{i=1}^{n} \int_{\mathbb{R}} \gamma_{ij}^{(1)}(t, z_j)\gamma_{ij}^{(2)}(t, x) \Big) \nu_j(dz_j) \Big] dt.$$

*Proof.* The Itô formula (Theorem 1.16). (See also Exercise 1.7.)     □

PROOF OF THEOREM 3.4     Let $u \in \mathcal{A}$ be an admissible control with corresponding state process $X(t) = X^{(u)}(t)$. Then

$$J(\hat{u}) - J(u) = E\Big[ \int_0^T \{f(t, \hat{X}(t), \hat{u}(t)) - f(t, X(t), u(t))\}dt + g(\hat{X}(T)) - g(X(T)) \Big].$$

Since $g$ is concave we get by Lemma 3.6

$$E[g(\hat{X}(T)) - g(X(T))] \geq E[(\hat{X}(T) - X(T))^T \nabla g(\hat{X}(T))]$$

$$= E[(\hat{X}(T) - X(T))^T \hat{p}(T)]$$

$$= E\Big[ \int_0^T (\hat{X}(t^-) - X(t^-))^T d\hat{p}(t) + \int_0^T \hat{p}(t^-)^T (d\hat{X}(t) - dX(t))$$

$$+ \int_0^T \mathrm{tr}[\{\sigma(t, \hat{X}(t), \hat{u}(t)) - \sigma(t, X(t), u(t))\}^T \hat{q}(t)] dt$$

$$+ \int_0^T \Big( \sum_{j=1}^{\ell} \Big( \sum_{i=1}^{n} \int_{\mathbb{R}} \{\gamma_{ij}(t, \hat{X}(t), \hat{u}(t), z_j)$$

$$- \gamma_{ij}(t, X(t), u(t), z_j)\} \hat{r}_{ij}(t, z_j) \Big) \nu_j(dz_j) \Big) dt \Big]$$

$$= E\Big[ \int_0^T (\hat{X}(t) - X(t))^T \big( -\nabla_x H(t, \hat{X}(t), \hat{u}(t), \hat{p}(t), \hat{q}(t), \hat{r}(t, \cdot)) \big) dt$$

$$+ \int_0^T \hat{p}^T(t^-) \{b(t, \hat{X}(t), \hat{u}(t)) - b(t, X(t), u(t))\} dt$$

$$+ \int_0^T \mathrm{tr}[\{\sigma(t, \hat{X}(t), \hat{u}(t)) - \sigma(t, X(t), u(t))\}^T \hat{q}(t)] dt$$

$$+ \int_0^T \Big( \sum_{j=1}^{\ell} \Big( \sum_{i=1}^{n} \int_{\mathbb{R}} \{\gamma_{ij}(t, \hat{X}(t), \hat{u}(t), z_j)$$

$$- \gamma_{ij}(t, X(t), u(t), z_j)\} \hat{r}_{ij}(t, z_j) \Big) \nu_j(dz_j) \Big) dt \Big] \qquad (3.2.8)$$

By the definition of $H$ we find

$$E\Big[ \int_0^T \{f(t, \hat{X}(t), \hat{u}(t)) - f(t, X(t), u(t))\} dt \Big]$$

$$= E\Big[ \int_0^T \{H(t, \hat{X}(t), \hat{u}(t), \hat{p}(t), \hat{q}(t), \hat{r}(t, \cdot))$$

$$- H(t, X(t), u(t), p(t), q(t), r(t, \cdot))\} dt$$

$$- \int_0^T \{b(t, \hat{X}(t), \hat{u}(t)) - b(t, X(t), u(t))\}^T \hat{p}(t) dt$$

$$- \int_0^T \mathrm{tr}[\{\sigma(t, \hat{X}(t), \hat{u}(t)) - \sigma(t, X(t), u(t))\}^T \hat{q}(t)] dt$$

$$-\int\limits_0^T \Big( \sum_{j=1}^{\ell} \Big( \sum_{i=1}^{n} \int_{\mathbb{R}} \{ \gamma_{ij}(t, \hat{X}(t), \hat{u}(t), z_j)$$

$$- \gamma_{ij}(t, X(t), u(t), z_j) \} \hat{r}_{ij}(t, z_j) \Big) \nu_j(dz_j) \Big) dt \Big]. \tag{3.2.9}$$

Adding (3.2.8) and (3.2.9) we get

$$J(\hat{u}) - J(u) \geq E^x \Big[ \int\limits_0^T \{ H(t, \hat{X}(t), \hat{u}(t), \hat{p}(t), \hat{q}(t), \hat{r}(t, \cdot))$$

$$- H(t, X(t), u(t), \hat{p}(t), \hat{q}(t), \hat{r}(t, \cdot))$$

$$- (\hat{X}(t) - X(t))^T \nabla_x H(t, \hat{X}(t), \hat{u}(t), \hat{p}(t), \hat{q}(t), \hat{r}(t, \cdot)) \} dt \Big].$$

If (3.2.6) (or (3.2.7)) holds then $J(\hat{u}) - J(u) \geq 0$. This follows from the proof in [SeSy, p. 108]. For details we refer to [FØS3].

We mention briefly the relation to dynamic programming : Define

$$J^{(u)}(s, x) = E \Big[ \int\limits_0^{T-s} f(s + t, X^x(t), u(t)) dt + g(X^x(T - s)) \Big]; \qquad u \in \mathcal{A}$$

where $X^x(t)$ is the solution of (3.2.1) for $t \geq 0$ with initial value $X(0) = x$.
Then put

$$V(s, x) = \sup_{u \in \mathcal{A}} J^{(u)}(s, x) . \tag{3.2.10}$$

**Theorem 3.7 (FØS3).** *Assume that $V(s, x) \in C^{1,3}(\mathbb{R} \times \mathbb{R}^n)$ and that there exists an optimal Markov control $u^*(t, x)$ for problem (3.2.2), with corresponding solution $X^*(t)$ of (3.2.1). Define*

$$p_i(t) = \frac{\partial V}{\partial x_i}(t, X^*(t)) ; \quad 1 \leq i \leq n$$

$$q_{jk}(t) = \sum_{i=1}^{n} \sigma_{ik}(t, X^*(t), u^*(t)) \frac{\partial^2 V}{\partial x_i \partial x_j}(t, X^*(t)) ; \quad 1 \leq j \leq n, \ 1 \leq k \leq m$$

$$r_{ik}(t, z) = \frac{\partial V}{\partial x_i}(t, X^*(t) + \gamma^{(k)}(t, X^*(t), u^*(t), z_k)) - \frac{\partial V}{\partial x_i}(t, X^*(t)) ;$$

$$1 \leq i \leq n, \quad 1 \leq k \leq \ell .$$

*Then $p(t), q(t), r(t, \cdot)$ solve the adjoint equation (3.2.4).*

For a proof see [FØS3].

**Remark 3.8.** A general discussion of impulse control for jump diffusions can be found in [F]. A study with vanishing impulse costs is given in [ØUZ].

## 3.3 Application to finance

The following example is from [FØS3].

Consider a financial market with two investment possibilities, a risk free (e.g. a bond or bank account) and risky (e.g. a stock), whose prices $S_0(t), S_1(t)$ at time $t \in [0, T]$ are given by

(bond)   $dS_0(t) = \rho_t S_0(t) dt$ ;    $S_0(0) = 1$                    (3.3.1)

(stock)   $dS_1(t) = S_1(t^-)\Big[\mu_t dt + \sigma_t dB(t) + \displaystyle\int_{\mathbb{R}} \gamma(t, z)\widetilde{N}(dt, dz)\Big]$;    $S_1(0) > 0$

(3.3.2)

where $\rho_t > 0$, $\mu_t, \sigma_t$ and $\gamma(t, z) \geq -1$ are given bounded deterministic functions. We assume that the function

$$t \to \int_{\mathbb{R}} \gamma^2(t, z)\nu(dz) \qquad \text{is locally bounded .} \qquad (3.3.3)$$

We may regard this market as a jump diffusion extension of the classical Black-Scholes market (see section 1.5).

A *portfolio* in this market is a two-dimensional cadlag, adapted process $\theta(t) = (\theta_0(t), \theta_1(t))$ giving the number of units of bonds and stocks, respectively, held at time $t$ by an agent.

The corresponding *wealth* process $X(t) = X^{(\theta)}(t)$ is defined by

$$X(t) = \theta_0(t)S_0(t) + \theta_1(t)S_1(t) ; \qquad t \in [0, T] . \qquad (3.3.4)$$

The portfolio $\theta$ is called *self-financing* if

$$X(t) = X(0) + \int_0^t \theta_0(s)dS_0(s) + \int_0^t \theta_1(s)dS_1(s) \qquad (3.3.5)$$

or, in short hand notation,

$$dX(t) = \theta_0(t)dS_0(t) + \theta_1(t)dS_1(t) . \qquad (3.3.6)$$

Alternatively, the portfolio can also be expressed in terms of the *amounts* $w_0(t), w_1(t)$ invested in the bond and stock, respectively. They are given by

$$w_i(t) = \theta_i(t)S_i(t) ; \qquad i = 0, 1. \qquad (3.3.7)$$

Now put

$$u(t) = w_1(t) . \qquad (3.3.8)$$

Then $w_0(t) = X(t) - u(t)$ and (3.3.6) gets the form

$$dX(t) = [\rho_t X(t) + (\mu_t - \rho_t)u(t)]dt + \sigma_t u(t)dB(t) + u(t^-) \int_{\mathbb{R}} \gamma(t,z)\widetilde{N}(dt,dz) \; .$$

(3.3.9)

We call $u(t)$ *admissible* and write $u(t) \in \mathcal{A}$ if (3.3.9) has a unique solution $X(t) = X^{(u)}(t)$ such that $E[(X^{(u)}(T))^2] < \infty$.

The *mean-variance portfolio selection problem* is to find $u(t)$ which minimizes

$$\mathrm{Var}[X(T)] := E\big[(X(T) - E[X(T)])^2\big]$$

(3.3.10)

under the condition that

$$E[X(T)] = A \; , \quad \text{a given constant} \; .$$

(3.3.11)

By the Lagrange multiplier method the problem can be reduced to minimizing, for a given constant $a \in \mathbb{R}$,

$$E[(X(T) - a)^2]$$

without constraints. To see this, consider

$$
\begin{aligned}
E[(X(T) - A)^2 &- \lambda([X(T)] - A)] \\
&= E[X^2(T) - 2(A + \tfrac{\lambda}{2})X(T) + A^2 + \lambda A] \\
&= E[(X(T) - (A + \tfrac{\lambda}{2}))^2] + \tfrac{\lambda^2}{4} \; , \quad \text{where } \lambda \in \mathbb{R} \text{ is constant.}
\end{aligned}
$$

We will consider the equivalent problem

$$\sup_{u \in \mathcal{A}} E[-\tfrac{1}{2}(X^{(u)}(T) - a)^2].$$

(3.3.12)

In this case the Hamiltonian (3.2.3) gets the form

$$H(t,x,u,p,q,r) = \{\rho_t x + (\mu_t - \rho_t)u\}p + \sigma_t uq + u \int_{\mathbb{R}} \gamma(t,z)r(t,z)\nu(dz) \; .$$

Hence the adjoint equations (3.2.4) are

$$
\begin{cases}
dp(t) = -\rho_t p(t)dt + q(t)dB(t) + \displaystyle\int_{\mathbb{R}} r(t^-,z)\widetilde{N}(dt,dz) \; ; \quad t < T \\[2mm]
p(T) = -(X(T) - a).
\end{cases}
$$

(3.3.13)

We try a solution of the form

$$p(t) = \phi_t X(t) + \psi_t \; ,$$

(3.3.14)

where $\phi_t, \psi_t$ are deterministic $C^1$ functions. Substituting in (3.3.13) and using (3.3.9) we get

$$
\begin{aligned}
dp(t) = \phi_t &\Big[ \{\rho_t X(t) + (\mu_t - \rho_t)u(t)\} dt + \sigma_t u(t) dB(t) \\
&+ u(t^-) \int_{\mathbb{R}} \gamma(t,z)\widetilde{N}(dt,dz) \Big] + X(t)\phi'_t \, dt + \psi'_t \, dt \\
= [\phi_t \rho_t X(t) &+ \phi_t(\mu_t - \rho_t)u(t) + X(t)\phi'_t + \psi'_t] dt \\
&+ \phi_t \sigma_t u(t) dB(t) + \phi_t u(t^-) \int_{\mathbb{R}} \gamma(t,z)\widetilde{N}(dt,dz) \ . \qquad (3.3.15)
\end{aligned}
$$

Comparing with (3.3.13) we get

$$\phi_t \rho_t X(t) + \phi_t(\mu_t - \rho_t)u(t) + X(t)\phi'_t + \psi'_t = -\rho_t(\phi_t X(t) + \psi_t) \qquad (3.3.16)$$
$$q(t) = \phi_t \sigma_t u(t) \qquad (3.3.17)$$
$$r(t,z) = \phi_t u(t)\gamma(t,z). \qquad (3.3.18)$$

Let $\hat{u} \in \mathcal{A}$ be a candidate for the optimal control with corresponding $\hat{X}$ and $\hat{p}, \hat{q}, \hat{r}$. Then

$$
\begin{aligned}
H(t, &\hat{X}(t), u, \hat{p}(t), \hat{q}(t), \hat{r}(t, \cdot)) \\
&= \rho_t \hat{X}(t)\hat{p}(t) + u\Big[ (\mu_t - \rho_t)\hat{p}(t) + \sigma_t \hat{q}(t) + \int_{\mathbb{R}} \gamma(t,z)\hat{r}(t,z)\nu(dz) \Big].
\end{aligned}
$$

Since this is a linear expression in $u$, it is natural to guess that the coefficient of $u$ vanishes, i.e.:

$$(\mu_t - \rho_t)\hat{p}(t) + \sigma_t \hat{q}(t) + \int_{\mathbb{R}} \gamma(t,z)\hat{r}(t,z)\nu(dz) = 0. \qquad (3.3.19)$$

Using that by (3.3.17) and (3.3.18) we have

$$\hat{q}(t) = \phi_t \sigma_t \hat{u}(t) \ , \qquad \hat{r}(t,z) = \phi_t \hat{u}(t)\gamma(t,z)$$

we get from (3.3.19) that

$$\hat{u}(t) = \frac{(\rho_t - \mu_t)\hat{p}(t)}{\phi_t \Lambda_t} = \frac{(\rho_t - \mu_t)(\phi_t \hat{X}(t) + \psi_t)}{\phi_t \Lambda_t} \qquad (3.3.20)$$

where

$$\Lambda_t = \sigma_t^2 + \int_{\mathbb{R}} \gamma^2(t,z)\nu(dz). \qquad (3.3.21)$$

On the other hand, from (3.3.16) we have

$$\hat{u}(t) = \frac{(\phi_t \rho_t + \phi_t')\hat{X}(t) + \rho_t(\phi_t \hat{X}(t) + \psi_t) + \psi_t'}{\phi_t(\rho_t - \mu_t)} \ . \tag{3.3.22}$$

Combining (3.3.20) and (3.3.22) we get the equations

$$(\rho_t - \mu_t)^2 \phi_t - [2\rho_t \phi_t + \phi_t']\Lambda_t = 0 \ ; \qquad \phi(T) = -1$$
$$(\rho_t - \mu_t)^2 \psi_t - [\rho_t \psi_t + \psi_t']\Lambda_t = 0 \ ; \qquad \psi(T) = a$$

which have the solutions

$$\phi_t = -\exp\Big(\int_t^T \Big\{ \frac{(\rho_s - \mu_s)^2}{\Lambda_s} - 2\rho_s \Big\}ds\Big); \qquad 0 \le t \le T \tag{3.3.23}$$

$$\psi_t = a\exp\Big(\int_t^T \Big\{ \frac{(\rho_s - \mu_s)^2}{\Lambda_s} - \rho_s \Big\}ds\Big); \qquad 0 \le t \le T \ . \tag{3.3.24}$$

With this choice of $\phi_t$ and $\psi_t$ the processes

$$\hat{p}(t) := \phi_t \hat{X}(t) + \psi_t, \quad \hat{q}(t) := \sigma_t \sigma_t \hat{u}(t) \quad \text{and} \quad \hat{r}(t, z) := \phi_t \hat{u}(t)\gamma(t, z)$$

solve the adjoint equation, and by (3.3.19) we see that all the conditions of the sufficient maximum principle (Theorem 3.4) are satisfied. We conclude that $\hat{u}(t)$ given by (3.3.20) is an optimal control. In feedback form the control can be written

$$\hat{u}(t, x) = \frac{(\rho_t - \mu_t)(\phi_t x + \psi_t)}{\phi_t \Lambda_t} \ . \tag{3.3.25}$$

## 3.4 Exercises

**Exercise* 3.1.** Suppose the wealth $X(t) = X^{(u)}(t)$ of a person with consumption rate $u(t) \ge 0$ satisfies the following Lévy type mean reverting Ornstein-Uhlenbeck SDE

$$dX(t) = (\mu - \rho X(t) - u(t))dt + \sigma dB(t) + \theta\int_{\mathbb{R}} z\widetilde{N}(dt, dz) \ ; \qquad t > 0$$

$$X(0) = x > 0$$

Fix $T > 0$ and define

$$J^{(u)}(s, x) = E^{s,x}\Big[ \int_0^{T-s} e^{-\delta(s+t)} \frac{u^\gamma(t)}{\gamma}dt + \lambda X(T - s)\Big].$$

Use dynamic programming to find the value function $\Phi(s, x)$ and the optimal consumption rate (control) $u^*(t)$ such that

$$\Phi(s, x) = \sup_{u(\cdot)} J^{(u)}(s, x) = J^{(u^*)}(s, x) \ .$$

In the above $\mu, \rho, \sigma, \theta, T, \ \delta > 0, \ \gamma \in (0, 1)$ and $\lambda > 0$ are constants.

**Exercise\* 3.2.** Solve the problem of Exercise 3.1 by using the stochastic maximum principle.

**Exercise\* 3.3.** Define

$$dX^{(u)}(t) = dX(t) = \begin{bmatrix} dX_1(t) \\ dX_2(t) \end{bmatrix} = \begin{bmatrix} u(t, \omega) \int\limits_{\mathbb{R}} z \widetilde{N}(dt, dz) \\ \int\limits_{\mathbb{R}} z^2 \widetilde{N}(dt, dz) \end{bmatrix} \in \mathbb{R}^2$$

and, for fixed $T > 0$ (deterministic)

$$J(u) = E\big[ - (X_1(T) - X_2(T))^2 \big] \ .$$

Use the stochastic maximum principle to find $u^*$ such that

$$J(u^*) = \sup_u J(u) \ .$$

**Interpretation:** Put $F(\omega) = \int\limits_{\mathbb{R}} z^2 \widetilde{N}(T, dz)$. We may regard $F$ as a given $T$-claim in the normalized market with the two investment possibilities bond and stock, whose prices are

(bond)      $dS_0(t) = 0$ ;      $S_0(0) = 1$

(stock)      $dS_1(t) = \int\limits_{\mathbb{R}} z \widetilde{N}(dt, dz),$   a Lévy martingale.

Then $-J(u)$ is the variance of the difference between $F = X_2(T)$ and the wealth $X_1(T)$ generated by a self-financing portfolio $u(t, \omega)$. See [BDLØP] for more information on minimal variance hedging in markets driven by Lévy martingales.

**Exercise\* 3.4.** Solve the stochastic control problem

$$\Phi_1(s, x) = \inf_{u \geq 0} E^{s,x} \left[ \int_0^\infty e^{-\rho(s+t)} (X^2(t) + \theta u^2(t) dt \right]$$

where

$$dX(t) = u(t)dt + \sigma dB(t) + \int_{\mathbb{R}} z \tilde{N}(dt, dz) \ ; \ X(0) = x$$

where $\rho > 0$, $\theta > 0$ and $\sigma > 0$ are constants.

    The interpretation of this problem is that we want to push the process $X(t)$ as close as possible to 0 by using a minimum of energy, its rate being measured by $\theta u^2(t)$.

[*Hint:* Try $\varphi(s, x) = e^{-\rho s}(ax^2 + b)$ for some constants $a, b$.]

**Exercise\* 3.5 (The stochastic linear regulator problem).**
Solve the stochastic control problem

$$\Phi_0(x) = \inf_u E^x \left[ \int_0^T (X^2(t) + \theta u(t)^2)dt + \lambda X^2(T) \right]$$

where

$$dX(t) = u(t)dt + \sigma dB(t) + \int_{\mathbb{R}} z\tilde{N}(dt, dz) \; ; \; X(0) = x$$

and

$$T > 0 \text{ is a constant.}$$

**a)** by using dynamic programming (Theorem 3.1).
**b)** by using the stochastic maximum principle (Theorem 3.4).

**Exercise\* 3.6.** Solve the stochastic control problem

$$\Phi(s, x) = \sup_{c(t) \geq 0} E^{s,x} \left[ \int_0^{\tau_0} e^{-\delta(s+t)} \ln c(t)dt \right],$$

where the supremum is taken over all $\mathcal{F}_t$-adapted processes $c(t) \geq 0$ and

$$\tau_0 = \inf\{t > 0 \; ; \; X(t) \leq 0\},$$

where

$$dX(t) = X(t^-) \left[ \mu dt + \sigma dB(t) + \theta \int_{\mathbb{R}} z\tilde{N}(dt, dz) \right] - c(t)dt, \; X(0) = x > 0,$$

where $\delta > 0$, $\mu, \sigma$ and $\theta$ are constants, and

$$\theta z > -1 \text{ for a.a. } z \text{ w.r.t. } \nu.$$

We may interpret $c(t)$ as the consumption rate, $X(t)$ as the corresponding wealth and $\tau_0$ as the bankrupty time. Thus $\Phi$ represents the maximal expected total discounted logarithmic utility of the consumption up to bankrupty time.

[*Hint:* Try $\varphi(s, x) = e^{-\delta s}(a \ln x + b)$ as a candidate for $\Phi(s, x)$, where $a$ and $b$ are suitable constants.]

**Exercise 3.7.** Use the stochastic maximum principle (Theorem 3.4) to solve the problem

$$\sup_{c(t) \geq 0} E \left[ \int_0^T e^{-\delta t} \ln c(t)dt + \lambda e^{-\delta T} \ln X(T) \right],$$

where

$$dX(t) = X(t^-) \left[ \mu dt + \sigma dB(t) + \theta \int_{\mathbb{R}} z \tilde{N}(dt, dz) \right] - c(t)dt, \quad X(0) > 0.$$

Here $\delta > 0$, $\lambda > 0$, $\mu$, $\sigma$ and $\theta$ are constants, and

$$\theta z > -1 \text{ for a.a. } z(\nu).$$

(See Exercise 3.6 for an interpretation of this problem).

[*Hint:* Try $p(t) = ae^{-\delta t} X^{-1}(t)$ and $c(t) = \frac{X(t)}{a}$, for some constant $a > 0$.]

# 4

# Combined Optimal Stopping
# and Stochastic Control of Jump Diffusions

## 4.1 Introduction

In this chapter we discuss combined optimal stopping and stochastic control problems and their associated HJB variational inequalities. This is a subject which deserves to be better known because of its many applications. A thorough treatment of such problems (but without the associated HJB variational inequalities) can be found in Krylov [K].

This chapter may also serve as a brief review of the theory of optimal stopping and their variational inequalities on one hand, and the theory of stochastic control and their HJB equations on the other. An introduction to these topics separately can be found in [Ø1].

As an illustration of how combined optimal stopping and stochastic control problems may appear in economics, let us consider the following example which is an extension of Exercise 2.2.

**Example 4.1 (An optimal resource extraction control and stopping problem).** Suppose the price $P_t = P(t)$ of one unit of a resource (e.g. gas or oil) at time $t$ is varying like a geometric Lévy process, i.e.

$$dP(t) = P(t^-)(\alpha dt + \beta dB(t) + \gamma \int_{\mathbb{R}} z \bar{N}(dt, dz)) \; ; \qquad P_0 = p \geq 0 \quad (4.1.1)$$

where $\alpha, \beta \neq 0, \gamma$ are constants and $\gamma z \geq -1$ a.s. $\nu$.

Let $Q_t$ denote the amount of remaining resources at time $t$. If we extract the resources at the "intensity" $u_t = u_t(\omega) \in [0, m]$ at time $t$, then the dynamics of $Q_t$ is

$$dQ_t = -u_t Q_t dt \; ; \qquad Q_0 = q \geq 0 \; . \quad (4.1.2)$$

($m$ is a constant giving the maximal intensity).

We assume as before that our *control* $u_t(\omega)$ is adapted to the filtration $\{\mathcal{F}\}_{t \geq 0}$. If the running cost is given by $K_0 + K_1 u_t$ (with $K_0, K_1 \geq 0$ constants)

as long as the field is open and if we decide to stop the extraction for good at time $\tau(\omega) \geq 0$ let us assume that the expected total discounted profit is

$$J^{(u,\tau)}(p,q) = E^{(p,q)}\left[\int_0^\tau e^{-\rho t}(u_t(P_tQ_t - K_1) - K_0)dt + e^{-\rho\tau}(\theta P_\tau Q_\tau - a)\right]$$

(4.1.3)

where $\rho > 0$ is the discounting exponent and $\theta > 0$, $a \geq 0$ are constants. Thus $e^{-\rho t}(u_t(P_tQ_t - K_1) - K_0)$ gives the discounted net profit rate when the field is in operation, while $e^{-\rho\tau}(\theta P_\tau Q_\tau - a)$ gives the discounted net value of the remaining resources at time $\tau$. (We may interpret $a \geq 0$ as a transaction cost.) We assume that the closing time $\tau$ is a *stopping time* with respect to the filtration $\{\mathcal{F}_t\}_{t\geq 0}$, i.e. that

$$\{\omega; \tau(\omega) \leq t\} \in \mathcal{F}_t \qquad \text{for all } t .$$

Thus both the extraction intensity $u_t$ and the decision whether to close before or at time $t$ must be based on the information $\mathcal{F}_t$ only, not on any future information.

The problem is to find the *value function* $\Phi(p,q)$ and the *optimal control* $u_t^* \in [0,m]$ and the *optimal stopping time* $\tau^*$ such that

$$\Phi(p,q) = \sup_{u_t,\tau} J^{(u,\tau)}(p,q) = J^{(u^*,\tau^*)}(p,q) .$$

(4.1.4)

This problem is an example of a combined optimal stopping and stochastic control problem. It is a modification of a problem discussed in [BØ1] and [DZ].

We will return to this and other examples after presenting a general theory for problems of this type.

## 4.2 A general mathematical formulation

Consider a controlled stochastic system of the same type as in Chapter 3, where the state $Y^{(u)}(t) = Y(t) \in \mathbb{R}^k$ at time $t$ is given by

$$dY(t) = b(Y(t), u(t))dt + \sigma(Y(t), u(t))dB(t) + \int_{\mathbb{R}^k} \gamma(Y(t^-), u(t^-), z)\bar{N}(dt, dz)$$

$$Y(0) = y \in \mathbb{R}^k .$$

(4.2.1)

Here $b : \mathbb{R}^k \times U \to \mathbb{R}^k$, $\sigma : \mathbb{R}^k \times U \to \mathbb{R}^{k\times m}$ and $\gamma : \mathbb{R}^k \times U \times \mathbb{R}^k \to \mathbb{R}^{k\times\ell}$ are given continuous functions and $u(t) = u(t,\omega)$ is our *control*, assumed to be $\mathcal{F}_t$-adapted and with values in a given closed, convex set $U \subset \mathbb{R}^\ell$.

Associated to a control $u = u(t,\omega)$ and an $\mathcal{F}_t$-stopping time $\tau = \tau(\omega)$ belonging to a given set $\mathcal{T}$ of *admissible* stopping times we assume there is a *performance criterion* of the form

$$J^{(u,\tau)}(y) = E^y\left[\int_0^{\tau} f(Y(t), u(t))dt + g(Y(\tau))\chi_{\{\tau<\infty\}}\right] \qquad (4.2.2)$$

where $f : \mathbb{R}^k \times U \to \mathbb{R}$ (the profit rate) and $g : \mathbb{R}^k \to \mathbb{R}$ (the bequest function) are given functions.

We assume that we are given a set $\mathcal{U} = \mathcal{U}(y)$ of admissible controls $u$ which is contained in the set of controls $u$ such that a unique strong solution $Y(t) = Y^{(u)}(t)$ of (4.2.1) exists and the following, (4.2.3)–(4.2.4), hold:

- $$E^y\left[\int_0^{\tau_{\mathcal{S}}} |f(Y(t), u(t))|dt\right] < \infty \qquad \text{for all } y \in \mathcal{S} \qquad (4.2.3)$$

    where $\tau_{\mathcal{S}} = \tau_{\mathcal{S}}(y, u) = \inf\{t > 0; Y^{(u)}(t) \notin \mathcal{S}\}$,
- the family $\{g^-(Y^{(u)}(\tau)); \tau \in \mathcal{T}\}$ is uniformly $P^y$-integrable for all $y$
    $\in \mathcal{S}$, where $g^-(y) = \max(0, -g(y))$. $\qquad (4.2.4)$

We interpret $g(Y(\tau(\omega)))$ as 0 if $\tau(\omega) = \infty$. Here, and in the following, $E^y$ denotes expectation with respect to $P$ when $Y(0) = y$ and $\mathcal{S} \subset \mathbb{R}^k$ is a fixed Borel set such that
$$\mathcal{S} \subset \overline{\mathcal{S}^0}.$$

We can think of $\mathcal{S}$ as the "universe" or "solvency set" of our system, in the sense that we are only interested in the system up to time $T$, which may be interpreted as the time of bankruptcy.

We now consider the following *combined optimal stopping and control-problem* :

Let $\mathcal{T}$ be the set of $\mathcal{F}_t$-stopping times $\tau \leq \tau_{\mathcal{S}}$. Find $\Phi(y)$ and $u^* \in \mathcal{U}$, $\tau^* \in \mathcal{T}$ such that
$$\Phi(y) = \sup\{J^{(u,\tau)}(y); u \in \mathcal{U}, \tau \in \mathcal{T}\} = J^{(u^*,\tau^*)}(y) . \qquad (4.2.5)$$

We will prove a verification theorem for this problem. The theorem can be regarded as a combination of the variational inequalities for optimal stopping (Theorem 2.2) and the Hamilton-Jacobi-Bellman (HJB) equation for stochastic control (Theorem 3.1).

We say that the control $u$ is *Markov* or *Markovian* if it has the form
$$u(t) = u_0(Y(t))$$

for some function $u_0 : \bar{\mathcal{S}} \to U$. If this is the case we usually do not distinguish notationally between $u$ and $u_0$ and write (with abuse of notation)
$$u(t) = u(Y(t)) .$$

If $u \in \mathcal{U}$ is Markovian then $Y^{(u)}(t)$ is a Markov process whose generator coincides on $C_0^2(\mathbb{R}^k)$ with the differential operator $L = L^u$ defined for $y \in \mathbb{R}^k$ by

$$L^u \psi(y) = \sum_{i=1}^{k} b_i(y, u(y)) \frac{\partial \psi}{\partial y_i} + \frac{1}{2} \sum_{i,j=1}^{k} (\sigma \sigma^T)_{ij}(y, u(y)) \frac{\partial^2 \psi}{\partial y_i \partial y_j}$$

$$+ \sum_{j=1}^{\ell} \int_{\mathbb{R}} \{\psi(y + \gamma^{(j)}(y, u(y), z_j)) - \psi(y) - \nabla \psi(y).\gamma^{(j)}(y, u(y), z_j)\} \nu_j(dz_j)$$

$$(4.2.6)$$

for all functions $\psi : \mathbb{R}^k \to \mathbb{R}$ which are twice differentiable at $y$.

Typically the value function $\Phi$ will be $C^2$ outside the boundary $\partial D$ of the continuation region $D$ (see (ii) below) and it will satisfy a Hamilton-Jacobi-Bellman (HJB) equation in $D$ and an HJB inequality outside $\bar{D}$. Across $\partial D$ the function $\Phi$ will not be $C^2$, but it will usually be $C^1$, and this feature is often referred to as the "*high contact*" – or "*smooth fit*" – principle. This is the background for the verification theorem given below, Theorem 4.2. Note however, that there are cases when $\Phi$ is not even $C^1$ at $\partial D$. To handle such cases one can use a verification theorem based on the viscosity solution concept. See Chapter 9 and in particular Section 9.2.

**Theorem 4.2 (HJB-variational inequalities for optimal stopping and control).**
**a)** *Suppose we can find a function $\varphi : \bar{S} \to \mathbb{R}$ such that*

(i)  $\varphi \in C^1(S^0) \cap C(\bar{S})$
(ii) $\varphi \geq g$ *on* $S^0$ .
     *Define*

$$D = \{y \in S; \varphi(y) > g(y)\} \qquad \text{(the continuation region ) .}$$

*Suppose $Y^{(u)}(t)$ spends 0 time on $\partial D$ a.s., i.e.*

(iii)  $E^y \left[ \int_0^{\tau_S} \mathcal{X}_{\partial D}(Y^{(u)}(t)) dt \right] = 0$   *for all $y \in S$, $u \in \mathcal{U}$*

     *and suppose that*

(iv)   $\partial D$ *is a Lipschitz surface*
(v)    $\varphi \in C^2(S^0 \backslash \partial D)$ *and the second order derivatives of $\varphi$ are locally bounded near $\partial D$*
(vi)   $L^v \varphi(y) + f(y, v) \leq 0$ *on $S^0 \setminus \partial D$ for all $v \in U$*
(vii)  $Y^{(u)}(\tau_S) \in \partial S$ *a.s. on $\{\tau_S < \infty\}$ and*

$$\lim_{t \to \tau_S^-} \varphi(Y^{(u)}(t)) = g(Y^{(u)}(\tau_S)) \chi_{\{\tau_S < \infty\}} \quad a.s.$$

(viii) $E^y \Big[ |\varphi(Y^{(u)}(\tau))| + \int_0^{\tau_S} \{|A\varphi(Y^{(u)}(t))| + |\sigma^T(Y(t))\nabla\varphi(Y(t))|^2$

$$+ \sum_{j=1}^{\ell} \int_{\mathbb{R}} |\varphi(Y(t) + \gamma^{(j)}(Y(t), u(t), z_j)) - \varphi(Y(t))|^2 \nu_j(dz_j)\} dt \Big] < \infty.$$

*Then*

$$\varphi(y) \geq \Phi(y) \qquad \text{for all } y \in \mathcal{S} \ .$$

**b)** *Suppose, in addition to (i)–(viii) above, that*

(ix)  *for each $y \in D$ there exists $\hat{u}(y) \in U$ such that*

$$L^{\hat{u}(y)}\varphi(y) + f(y, \hat{u}(y)) = 0,$$

(x)  $\tau_D := \inf\{t > 0; Y^{(\hat{u})}(t) \notin D\} < \infty$ *a.s. for all $y \in \mathcal{S}$,*
     *and*

(xi)  *the family $\{\varphi(Y^{(\hat{u})}(\tau)); \ \tau \in \mathcal{T}\}$ is uniformly integrable with respect to $P^y$ for all $y \in D$.*
     *Suppose $\hat{u} \in \mathcal{U}$. Then*

$$\varphi(y) = \Phi(y) \qquad \text{for all } y \in \mathcal{S} \ .$$

*Moreover, $u^* := \hat{u}$ and $\tau^* := \tau_D$ are optimal control and stopping times, respectively.*

*Proof.* The proof is a synthesis of the proofs of Theorem 2.2 and Theorem 3.1. For completeness we give some details:

**a)** By Theorem 2.1 we may assume that $\varphi \in C^2(\mathcal{S}^0) \cap C(\bar{\mathcal{S}})$. Choose $u \in \mathcal{U}$ and put $Y(t) = Y^{(u)}(t)$. Let $\tau \leq \tau_\mathcal{S}$ be a stopping time. Then by Dynkin's formula (Theorem 1.24)

$$E^y\big[\varphi(Y(\tau \wedge m))\big] = \varphi(y) + E^y\Big[\int_0^{\tau \wedge m} L^u\varphi(Y(t))dt\Big]. \qquad (4.2.7)$$

Hence by (vii) and the Fatou lemma

$$\varphi(y) = \lim_{m \to \infty} E^y\Big[\int_0^{\tau \wedge m} -L^u\varphi(Y(t))dt + \varphi(Y(\tau \wedge m))\Big]$$

$$\geq E^y\Big[\int_0^{\tau} -L^u\varphi(Y(t))dt + g(Y(\tau))\chi_{\{\tau<\infty\}}\Big]. \qquad (4.2.8)$$

If we now use (vi) we can conclude that

$$\varphi(y) \geq E^y\Big[\int_0^{\tau} f(Y(t), u(t))dt + g(Y(\tau))\chi_{\{\tau<\infty\}}\Big] = J^{u,\tau}(y). \qquad (4.2.9)$$

Since $u \in \mathcal{U}$ and $\tau \leq \tau_\mathcal{S}$ was arbitrary we conclude that

$$\varphi(y) \geq \sup_{u,\tau} J^{(u,\tau)}(y) = \Phi(y) \ , \qquad (4.2.10)$$

which proves a).

To prove the opposite inequality, assume that (ix)-(xi) hold. Choose a point $y \in D$. Then apply the argument above to the Markovian control $\hat{u}(t) := \hat{u}(Y(t))$ and the stopping time $\hat{\tau} = \tau_D = \inf\{t > 0; \hat{Y}(t) \notin D\}$, where $\hat{Y}(t) = Y^{(\hat{u})}(t)$: We get

$$E^y\big[\varphi(\hat{Y}(\hat{\tau}))\big] = \varphi(y) + E^y\bigg[\int_0^{\hat{\tau}} L^{\hat{u}}\varphi(\hat{Y}(t))dt\bigg], \qquad (4.2.11)$$

which implies that

$$\varphi(y) = E^y\bigg[\int_0^{\hat{\tau}} f(\hat{Y}(t), \hat{u}(t))dt + g(\hat{Y}(\hat{\tau}))\bigg] = J^{(\hat{u},\hat{\tau})}(y).$$

Combined with a) this shows that $\varphi(y) = \Phi(y)$ and that $(\hat{u}, \hat{\tau})$ is optimal if $y \in D$.

Finally, if $y \notin D$ then $\varphi(y) = g(y) \leq \Phi(y)$ and hence by a) we have $\varphi(y) = \Phi(y)$ also if $y \notin D$. In this case $\tau^* = 0$ is an optimal stopping time.   □

**Remark 4.3.** If we neglect all the technical conditions of Theorem 4.2 and concentrate on conditions (ii), (vi) and (ix), then we can write the conditions of Theorem 4.2 in the following condensed form

$$\max\bigg(\sup_{v \in U}\{L^v\phi(y) + f(y,v)\}, g(y) - \phi(y)\bigg) = 0 ; \qquad y \in \mathcal{S}^0 . \qquad (4.2.12)$$

Since this is a combination of the Hamilton-Jacobi-Bellman (HJB) equation of stochastic control and the variational inequality (VI) of optimal stopping, we call (4.2.12) a HJBVI.

One can prove that, under some conditions, the value function $\phi$ is indeed a solution of (4.2.12) in the weak sense of *viscosity*. See Chapter 9 for a discussion of this concept.

**Remark 4.4.** Note that the problem (4.2.5) contains the general optimal stopping problem as a special case.

More precisely, if the functions $b, \sigma$ and $f$ do not depend on $u$, then the problem reduces to the optimal stopping problem

$$\Phi(y) = \sup_{\tau \in \mathcal{T}} E^y\bigg[\int_0^\tau f(Y(t))dt + g(Y(\tau))\chi_{\{\tau < \infty\}}\bigg]$$

discussed in Chapter 2 and the HJBVI (4.2.12) becomes the VI

$$\max(L\Phi(y) + f(y), g(y) - \Phi(y)) = 0 ; \qquad y \in \mathcal{S}^0 . \qquad (4.2.13)$$

The problem (4.2.5) is also closely related to the general *stochastic control problem* discussed in Chapter 3. In such a problem the stopping time $\tau$ is

fixed to $\tau = \tau_\mathcal{S}$ and hence the problem is to find $\Phi(y)$ and $u^* \in \mathcal{U}$ such that

$$\Phi(y) = \sup_{u \in \mathcal{U}} J^{(u)}(y) = J^{(u^*)}(y) \tag{4.2.14}$$

where

$$J^{(u)}(y) = E^y \left[ \int_0^{\tau_\mathcal{S}} f(Y(t), u(t)) dt + g(Y(\tau_\mathcal{S})) \chi_{\{\tau_\mathcal{S} < \infty\}} \right].$$

## 4.3 Applications

To illustrate Theorem 4.2 let us apply it to the problem of Example 4.1:
    In this case the generator $L^u$ of the time-space state process

$$dY(t) = (dt, dP_t, dQ_t) ; \qquad Y(0) = (s, p, q) \in [0, \infty)^3$$

is given by (see Theorem 1.22)

$$L^u \psi(s, p, q) = \frac{\partial \psi}{\partial s} + \alpha p \frac{\partial \psi}{\partial p} - uq \frac{\partial \psi}{\partial q} + \tfrac{1}{2}\beta^2 p^2 \frac{\partial^2 \psi}{\partial p^2}$$

$$+ \int_{\mathbb{R}} \{\psi(s, p + \gamma pz, q) - \psi(s, p, q) - \frac{\partial \psi}{\partial p}(s, p, q) \cdot \gamma zp\} \nu(dz) .$$

In view of Theorem 4.2 we are looking for a subset $D$ of $\mathcal{S} = [0, \infty)^3$ and a function $\varphi(s, p, q) : \mathcal{S} \to \mathbb{R}$ such that

$$\varphi(s, p, q) = e^{-\rho s}(\theta pq - a) \quad \text{for all } (s, p, q) \notin D \tag{4.3.1}$$

$$\varphi(s, p, q) \geq e^{-\rho s}(\theta pq - a) \quad \text{for all } (s, p, q) \in \mathcal{S} \tag{4.3.2}$$

$$L^v \varphi(s, p, q) + e^{-\rho s}(v(pq - K_1) - K_0) \leq 0 \quad \text{for all } (s, p, q) \in \mathcal{S}^0 \setminus \bar{D}$$

$$\text{and all } v \in [0, m] \tag{4.3.3}$$

$$\sup_{v \in [0, m]} \{L^v \varphi(s, p, q) + e^{-\rho s}(v(pq - K_1) - K_0)\} = 0 \quad \text{for all } (s, p, q) \in D .$$

$$\tag{4.3.4}$$

Let us try a function $\varphi$ of the form

$$\varphi(s, p, q) = e^{-\rho s} F(w) \qquad \text{where } w = pq \tag{4.3.5}$$

and a continuation region $D$ of the form

$$D = \{(s, p, q); pq > w_0\} \qquad \text{for some } w_0 > 0 .$$

Then (4.3.1)–(4.3.4) get the form

$$F(w) = \theta w - a \qquad \text{for all } w \ge w_0 \tag{4.3.6}$$

$$F(w) \ge \theta w - a \qquad \text{for all } w < w_0 \tag{4.3.7}$$

$$- \rho F(w) + (\alpha - v)w F'(w) + \tfrac{1}{2}\beta^2 w^2 F''(w)$$
$$+ \int_{\mathbb{R}} \{F(w + \gamma z w) - F(w) - F'(w)\gamma z w\}\nu(dz) + v(w - K_1) - K_0 \le 0$$
$$\text{for all } w < w_0, \ v \in [0, m] \tag{4.3.8}$$

$$\sup_{v \in [0,m]} \{ -\rho F(w) + (\alpha - v)w F'(w) + \tfrac{1}{2}\beta^2 w^2 F''(w)$$
$$+ \int_{\mathbb{R}} \{F(w + \gamma z w) - F(w) - F'(w)\gamma z w\}\nu(dz) + v(w - K_1) - K_0 \} = 0$$
$$\text{for all } w > w_0 \ . \tag{4.3.9}$$

From (4.3.9) and (xi) of Theorem 4.2 we get the following candidate $\hat{u}$ for the optimal control:

$$v = \hat{u}(w) = \underset{v \in [0,m]}{\text{Argmax}} \left\{ v(w(1 - F'(w)) - K_1) \right\}$$

$$\tag{4.3.10}$$

$$= \begin{cases} m & \text{if } F'(w) < 1 - \frac{K_1}{w} \\ 0 & \text{if } F'(w) > 1 - \frac{K_1}{w} \end{cases}$$

Let $F_m(w)$ be the solution of (4.3.9) with $v = m$, i.e. the solution of

$$-\rho F_m(w) + (\alpha - m)w F'_m(w) + \tfrac{1}{2}\beta^2 w^2 F''_m(w)$$
$$+ \int_{\mathbb{R}} \{F(w + \gamma z w) - F(w) - F'(w)\gamma z w\}\nu(dz) = K_0 + mK_1 - mw \ . \tag{4.3.11}$$

A solution of (4.3.11) is

$$F_m(w) = C_1 w^{\lambda_1} + C_2 w^{\lambda_2} + \frac{mw}{\rho + m - \alpha} - \frac{K_0 + mK_1}{\rho} \tag{4.3.12}$$

where $C_1, C_2$ are constants and $\lambda_1 > 0, \lambda_2 < 0$ are roots of the equation

$$h(\lambda) = 0 \tag{4.3.13}$$

with

$$h(\lambda) = -\rho + (\alpha - m)\lambda + \tfrac{1}{2}\beta^2 \lambda(\lambda - 1) + \int_{\mathbb{R}} \{(1 + \gamma z)^{\lambda} - 1 - \lambda \gamma z\}\nu(dz). \tag{4.3.14}$$

(Note that $h(0) = -\rho < 0$ and $\lim_{|\lambda| \to \infty} h(\lambda) = \infty$). The solution will depend on the relation between the parameters involved and we will not give a complete discussion, but only consider some special cases.

**Case 1**

Let us assume that

$$\alpha \leq \rho \,, \qquad K_1 = a = 0 \qquad \text{and} \ \ 0 < \theta < \tfrac{m}{\rho + m - \alpha} \,. \qquad (4.3.15)$$

It is easy to see that

$$\lambda_1 > 1 \Longleftrightarrow \rho + m > \alpha \,. \qquad (4.3.16)$$

Let us try (guess) that

$$C_1 = 0 \qquad (4.3.17)$$

and that the continuation region $D = \{(s, p, q); pq > w_0\}$ is such that (see (4.3.10))

$$F'_m(w) < 1 \qquad \text{for all} \ \ w > w_0 \,. \qquad (4.3.18)$$

The intuitive motivation for trying this is the belief that it is optimal to use the maximal extraction intensity $m$ all the time until closure, at least if $\theta$ is small enough.

These guesses lead to the following candidate for the value function $F(w)$:

$$F(w) = \begin{cases} \theta w & \text{if} \ \ 0 \leq w \leq w_0 \\ F_m(w) = C_2 w^{\lambda_2} + \frac{mw}{\rho + m - \alpha} - \frac{K_0}{\rho} & \text{if} \ \ w > w_0 \,. \end{cases} \qquad (4.3.19)$$

We now use continuity and differentiability at $w = w_0$ to determine $w_0$ and $C_2$:

$$\text{(Continuity)} \qquad C_2 w_0^{\lambda_2} + \frac{mw_0}{\rho + m - \alpha} - \frac{K_0}{\rho} = \theta w_0 \qquad (4.3.20)$$

$$\text{(Differentiability)} \qquad C_2 \lambda_2 w_0^{\lambda_2 - 1} + \frac{m}{\rho + m - \alpha} = \theta \,. \qquad (4.3.21)$$

Easy calculations show that the unique solution of (4.3.20), (4.3.21) is

$$w_0 = \frac{(-\lambda_2)K_0(\rho + m - \alpha)}{(1 - \lambda_2)\rho[m - \theta(\rho + m - \alpha)]} \qquad (> 0 \quad \text{by (4.3.15))} \qquad (4.3.22)$$

and

$$C_2 = \frac{[m - \theta(\rho + m - \alpha)]w_0^{1 - \lambda_2}}{(-\lambda_2)(\rho + m - \alpha)} \qquad (> 0 \quad \text{by (4.3.15)).} \qquad (4.3.23)$$

It remains to verify that with these values of $w_0$ and $C_2$ the set $D = \{(s, p, q); pq > w_0\}$ and the function $F(w)$ given by (4.3.19) satisfies (4.3.6)–(4.3.9), as well as all the other conditions of Theorem 4.2:

To verify (4.3.6) we have to check that (4.3.18) holds, i.e. that

$$F'_m(w) = C_2 \lambda_2 w^{\lambda_2 - 1} + \frac{m}{\rho + m - \alpha} < 1 \qquad \text{for all} \ \ w > w_0 \,.$$

Since $\lambda_2 < 0$ and we have assumed $\alpha \leq \rho$ (in (4.3.15)) this is clear. So (4.3.9) holds. If we substitute $F(w) = \theta w$ in (4.3.8) we get

$$-\rho\theta w + (\alpha - m)w\theta + mw - K_0 = w[m - \theta(\rho + m - \alpha)] - K_0 .$$

We know that this is 0 for $w = w_0$ by (4.3.20) and (4.3.21). Hence it is less than 0 for $w < w_0$. So (4.3.8) holds. Condition (4.3.6) holds by definition of $D$ and $F$. Finally, since $F(w_0) = \theta w_0$, $F'(w_0) = \theta$ and

$$F''(w) = F_m''(w) = C_2 \lambda_2 (\lambda_2 - 1) w^{\lambda_2 - 2} > 0$$

we must have $F(w) > \theta w$ for $w > w_0$. Hence (4.3.7) holds. Similarly one can verify all the other conditions of Theorem 4.2. We have proved:

**Theorem 4.5.** *Suppose* (4.3.15) *holds. Then the optimal strategy* $(u^*, \tau^*)$ *for problem* (4.1.3)–(4.1.4) *is*

$$u^* = m , \qquad \tau^* = \inf\{t > 0; \ P_t Q_t \leq w_0\} \tag{4.3.24}$$

*where $w_0$ is given by* (4.3.22). *The corresponding value function is* $\Phi(s, p, q) = e^{-\rho s} F(p \cdot q)$, *where $F$ is given by* (4.3.19) *with $\lambda_2 < 0$ as in* (4.3.11) *and $C_2 > 0$ as in* (4.3.23).

For other values of the parameters it might be optimal not to produce at all but just wait for the best closing/sellout time. For example, we mention without proof the following cases (see Exercise 4.2):

**Case 2**

Assume that

$$\theta = 1 \qquad \text{and} \qquad \rho \leq \alpha . \tag{4.3.25}$$

Then $u^* = 0$ and

$$\Phi = \infty .$$

**Case 3**

Assume that

$$\theta = 1 , \quad \rho > \alpha \qquad \text{and} \qquad K_0 < \rho a < K_0 + \rho K_1 . \tag{4.3.26}$$

Then

$$u^* = 0 \qquad \text{and} \qquad \tau^* = \inf\{t > 0; \ P_t Q_t \geq w_1\} ,$$

for some $w_1 > 0$.

## 4.4 Exercises

**Exercise\* 4.1. a)** Solve the following stochastic control problem:

$$\Phi(s, x) = \sup_{u(t) \geq 0} J^{(u)}(s, x) = J^{(u^*)}(s, x) \,,$$

where

$$J^{(u)}(s, x) = E^x \Big[ \int_0^{\tau_S} e^{-\delta(s+t)} \frac{u^\gamma(t)}{\gamma} dt \Big].$$

Here

$$\tau_S = \tau_S(\omega) = \inf\{t > 0; X(t) \leq 0\} \quad \text{(the time of bankruptcy)}$$

and

$$dX(t) = (\mu X(t) - u(t))dt + \sigma X(t)dB(t) + \theta X(t^-) \int_{\mathbb{R}} z \bar{N}(dt, dz) \,; \quad X_0 = x > 0$$

with $\gamma \in (0, 1)$, $\delta > 0$, $\mu, \sigma \neq 0$, $\theta$ constants, $\theta z > -1$ a.s. $\nu$. The interpretation of this is the following: $X(t)$ represents the total wealth at time $t$, $u(t) = u(t, \omega) \geq 0$ represents the chosen consumption rate (the control). We want to find the consumption rate $u^*(t)$ which maximizes the expected total discounted utility of the consumption up to the time of bankruptcy, $\tau_S$.

[*Hint:* Try a value function of the form

$$\phi(s, x) = Ke^{-\delta s}x^\gamma$$

for a suitable value of the constant $K$.]

**b)** Consider the following combined stochastic control and optimal stopping problem:

$$\Phi(s, x) = \sup_{u, \tau} J^{(u, \tau)}(s, x) = J^{(u^*, \tau^*)}(s, x)$$

where

$$J^{(u, \tau)}(s, x) = E^x \Big[ \int_0^\tau e^{-\delta(s+t)} \frac{u^\gamma(t)}{\gamma} dt + \lambda X^\gamma(\tau) \Big]$$

with $X(t)$ as in a), $\lambda > 0$ a given constant.

Now the supremum is taken over all $\mathcal{F}_t$-adapted controls $u(t) \geq 0$ and all $\mathcal{F}_t$-stopping times $\tau \leq \tau_S$.

Let $K$ be the constant found in a). Show that
**(i)**   if $\lambda \geq K$ then it is optimal to stop immediately
**(ii)**  if $\lambda < K$ then it is optimal never to stop.

**Exercise 4.2. a)** Verify the statements in Case 2 and Case 3 at the end of Section 4.3.

**b)**  What happens in the cases
Case 4: $\theta = 1$, $\rho > \alpha$ and $\rho a \leq K_0$ ?
Case 5: $\theta = 1$, $\rho > \alpha$ and $K_0 + \rho K_1 \leq \rho a$ ?

**Exercise 4.3 (A stochastic linear regulator problem with optimal stopping).**

Solve the stochastic linear regulator problem in Exercise 3.5, with the additional option of stopping, i.e. solve the problem

$$\Phi(s, x) = \inf \left\{ J^{(u,\tau)}(s, x) \ ; \ u \in \mathcal{U}, \ \tau \in \mathcal{T} \right\},$$

where

$$J^{(u,\tau)}(s, x) = E^{s,x} \left[ \int_0^\tau e^{-\rho(s+t)} \left( X^2(t) + \theta u^2(t) \right) dt + \lambda e^{-\rho(s+\tau)} X^2(\tau) \right],$$

($\rho > 0$ constant), where the state process is

$$Y(t) = \begin{bmatrix} s+t \\ X(t) \end{bmatrix} \ ; \ t \geq 0, \ Y(0) = \begin{bmatrix} s \\ x \end{bmatrix} = y \in \mathbb{R}^2$$

with

$$dX(t) = dX^{(u)}(t) = u(t)dt + \sigma dB(t) + \int_\mathbb{R} z\widetilde{N}(dt, dz) \ ; \ X(0) = x.$$

[Hint: As a candidate for the value function $\Phi$ try a function $\varphi$ of the form

$$\varphi(s, x) = \begin{cases} e^{-\rho s}(ax^2 + b) & ; \ |x| \geq \delta \\ e^{-\rho s}\lambda x^2 & ; \ |x| < \delta \end{cases}$$

for suitable values of the constants $a, b$ and $\delta$.]

# 5
# Singular Control for Jump Diffusions

## 5.1 An illustrating example

We illustrate singular control problems by the following example, studied in [FØS2]:

**Example 5.1 (Optimal consumption rate under proportional transaction costs).** Consider again a financial market of the form (3.3.1)–(3.3.2), where we have two investment possibilities:

(i) A bank account/bond where the value/price $P_1(t)$ at time $t$ grows with interest rate $r$, i.e.,

$$dP_1(t) = rP_1(t)dt , \qquad P_1(0) = 1 .$$

(ii) A stock, with price $P_2(t)$ satisfying the equation

$$dP_2(t) = P_2(t^-)\Big[\mu dt + \beta dB(t) + \int_{\mathbb{R}} z\widetilde{N}(dt, dz)\Big]; \qquad P_2(0) = p_2 > 0 .$$

Here $r > 0$, $\mu$ and $\beta > 0$ are given constants, and we assume that $z > -1$ a.s. $\nu$.

Assume that at any time $t$ the investor can choose an adapted, cadlag consumption rate process $c(t) = c(t, \omega) \geq 0$, taken from the bank account. Moreover, the investor can at any time transfer money from one investment to another with a transaction cost which is proportional to the size of the transaction. Let $X_1(t)$ and $X_2(t)$ denote the amounts of money invested in the bank and in the stocks, respectively. Then the evolution equations for $X_1(t)$ and $X_2(t)$ are

$$dX_1(t) = dX_1^{x_1,c,\xi}(t)$$
$$= (rX_1(t) - c(t))dt - (1 + \lambda)d\xi_1(t) + (1 - \mu)d\xi_2(t); \quad X_1(0^-) = x_1 \in \mathbb{R}$$

$$dX_2(t) = dX_2^{x,c,\xi}(t)$$
$$= X_2(t^-)\left[\mu dt + \beta dB(t) + \int_{\mathbb{R}} z\widetilde{N}(dt, dz)\right] + d\xi_1(t) - d\xi_2(t); \ X_2(0^-) = x_2 \in \mathbb{R}.$$

Here $\xi = (\xi_1, \xi_2)$, where $\xi_1(t), \xi_2(t)$ represent cumulative purchase and sale, respectively, of stocks up to time $t$. The constants $\lambda \geq 0$, $\mu \in [0, 1]$ represent the constants of proportionality of the transaction costs.

Define the *solvency region* $\mathcal{S}$ to be the set of states $(x_1, x_2)$ such that the net wealth is non-negative, i.e. (see Figure 5.1)

$$\mathcal{S} = \{(x_1, x_2) \in \mathbb{R}^2; \ x_1 + (1 + \lambda)x_2 \geq 0 \quad \text{and} \quad x_1 + (1 - \mu)x_2 \geq 0 \quad (5.1.1)$$
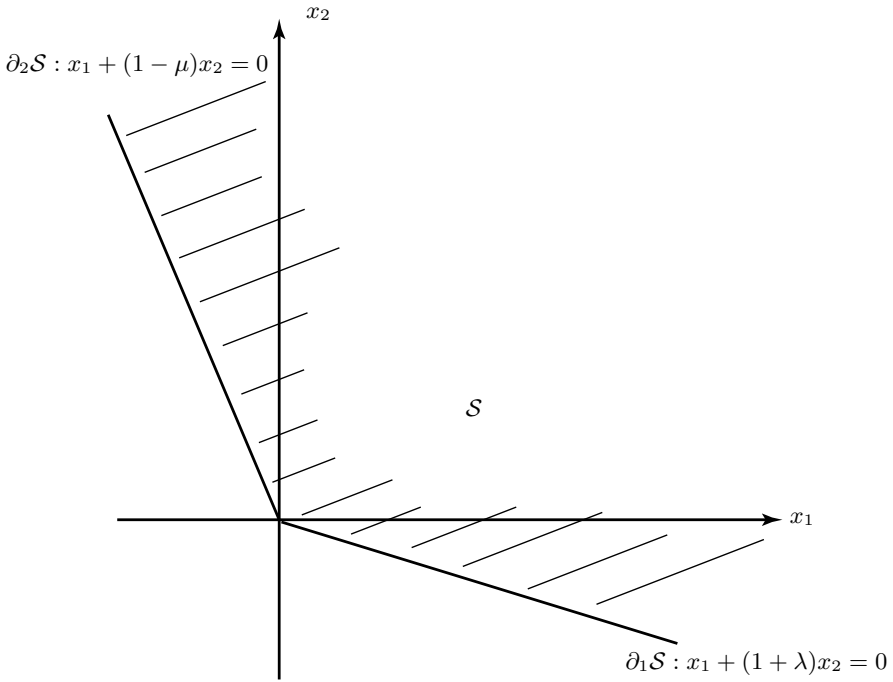


**Fig. 5.1.** The solvency region $\mathcal{S}$

We define the set $\mathcal{A}$ of *admissible* controls as the set of predictable consumption-investment policies $(c, \xi)$ such that $\xi = (\xi_1, \xi_2)$ where each $\xi_i(t)$ ; $i = 1, 2$ is right-continuous, non-decreasing, $\xi(0^-) = 0$ and such that

$$(x_1, x_2) \in \mathcal{S} \Rightarrow (X_1(t), X_2(t)) \in \mathcal{S} \qquad \text{for all } t \geq 0 \,.$$

The performance criterion is defined by

$$J^{c,\xi}(s, x_1, x_2) = E^{s,x_1,x_2} \Big[ \int_0^\infty e^{-\delta(s+t)} \frac{c^\gamma(t)}{\gamma} dt \Big], \qquad (5.1.2)$$

where $\delta > 0$, $\gamma \in (0, 1)$ are constants. We seek $(c^*, \xi^*) \in \mathcal{A}$ and $\Phi(s, x_1, x_2)$ such that

$$\Phi(s, x_1, x_2) = \sup_{(c,\xi) \in \mathcal{A}} J^{c,\xi}(s, x_1, x_2) = J^{c^*, \xi^*}(s, x_1, x_2) \,. \qquad (5.1.3)$$

This is an example of a *singular* stochastic control problem. It is called singular because the investment control measure $d\xi(t)$ is allowed to be singular with respect to Lebesgue measure $dt$. In fact, as we shall see, the optimal control measure $d\xi^*(t)$ turns out to be singular.

We now give a general theory of singular control of jump diffusions and return to the above example afterwards.

## 5.2 A general formulation

Let $\kappa = [\kappa_{ij}] \in \mathbb{R}^{k \times p}$ be a constant matrix and $\theta = [\theta_i] \in \mathbb{R}^p$ be a constant vector. Suppose the state $Y(t) = Y^{u,\xi}(t) \in \mathbb{R}^k$ is described by the equation

$$dY(t) = b(Y(t), u(t))dt + \sigma(Y(t), u(t))dB(t)$$
$$+ \int_{\mathbb{R}^k} \gamma(Y(t^-), u(t^-), z)\widetilde{N}(dt, dz) + \kappa d\xi(t) \,; \qquad Y(0^-) = y \in \mathbb{R}^k.$$

Here $\xi(t) \in \mathbb{R}^p$ is an adapted cadlag process with increasing components and $\xi(0^-) = 0$. Since $d\xi(t)$ may be singular with respect to Lebesgue measure $dt$, we call $\xi$ our *singular control* or our *intervention control*. The process $u(t)$ is an adapted cadlag process with values in a given set $U$ (our absolutely continuous control). Suppose we are given a performance functional $J^{u,\xi}(y)$ of the form

$$J^{u,\xi}(y) = E^y \Big[ \int_0^{\tau_{\mathcal{S}}} f(Y(t), u(t))dt + g(Y(\tau_{\mathcal{S}})) \cdot \mathcal{X}_{\{\tau_{\mathcal{S}} < \infty\}} + \int_0^{\tau_{\mathcal{S}}} \theta^T d\xi(t) \Big],$$

where $f : \mathbb{R}^k \times U \to \mathbb{R}$, $g : \mathbb{R}^k \to \mathbb{R}$ are continuous functions and

$$\tau_{\mathcal{S}} = \inf\{t > 0; Y^{u,\xi}(t) \notin \mathcal{S}\} \leq \infty \quad \text{is the time of bankruptcy,}$$

where $\mathcal{S} \subset \mathbb{R}^k$ is a given *solvency set*, assumed to satisfy $\mathcal{S} \subset \overline{\mathcal{S}^0}$.

Let $\mathcal{A}$ be a given family of *admissible controls* $(u, \xi)$, contained in the set of $(u, \xi)$ such that

$$E^y \left[ \int_0^{\tau_S} |f(Y(t), u(t))| dt + |g(Y(\tau_S))| \cdot \mathcal{X}_{\{\tau_S < \infty\}} + \int_0^{\tau_S} \sum_{j=1}^p |\theta_j| d\xi_j(t) \right] < \infty .$$

The problem is to find the value function $\Phi(y)$ and an optimal control $(u^*, \xi^*) \in \mathcal{A}$ such that

$$\Phi(y) = \sup_{(u, \xi) \in \mathcal{A}} J^{u, \xi}(y) = J^{u^*, \xi^*}(y) . \tag{5.2.1}$$

Note that if we apply a Markov control $u(t) = u(Y(t)) \in U$ and $d\xi = 0$, then $Y(t) = Y^{u, 0}(t)$ has the generator $A^u$ given by

$$A^u \phi(y) = \sum_{i=1}^k b_i(y, u(y)) \frac{\partial \phi}{\partial y_i} + \tfrac{1}{2} \sum_{i,j=1}^k (\sigma\sigma^T)_{ij}(y, u(y)) \frac{\partial^2 \phi}{\partial y_i \partial y_j}$$

$$+ \sum_{j=1}^\ell \int_{\mathbb{R}} \left\{ \phi(y + \gamma^{(j)}(y, u(y), z_j)) - \phi(y) - \nabla\phi(y)^T \gamma^{(j)}(y, u(y), z_j) \right\} \nu_j(dz_j) .$$

Note that we distinguish between the jumps of $Y^{u, \xi}(t)$ caused by the jump of $N(t, z)$, denoted by $\Delta_N Y(t)$, and the jump caused by the singular control $\xi$, denoted by $\Delta_\xi Y(t)$. Thus

$$\Delta_N Y(t) = \int_{\mathbb{R}^k} \gamma(Y(t^-), u(t^-), z) \widetilde{N}(\{t\}, dz), \quad \text{while} \quad \Delta_\xi Y(t) = \kappa \Delta\xi(t) .$$
$$\tag{5.2.2}$$

We let $t_1, t_2, \ldots$ denote the jumping times of $\xi(t)$ and we let

$$\Delta_\xi \phi(Y(t_n)) = \phi(Y(t_n)) - \phi(Y(t_n^-) + \Delta_N Y(t_n)) \tag{5.2.3}$$

be the increase in $\phi$ due to the jump of $\xi(t)$ at $t = t_n$.

We give now a verification theorem for singular control problems.

**Theorem 5.2 (Integro-variational inequalities for singular control).**
**a)** *Suppose there exists a function $\phi \in C^2(\mathcal{S}^0) \cap C(\bar{\mathcal{S}})$ such that*

(i)     $A^v \phi(y) + f(y, v) \leq 0$ *for all (constant)* $v \in U$ *and* $y \in \mathcal{S}$

(ii)    $\sum_{i=1}^k \kappa_{ij} \dfrac{\partial \phi}{\partial y_i}(y) + \theta_j \leq 0$ *for all* $y \in \mathcal{S}$, $j = 1, \ldots, p$.

(iii)   $E^y \left[ |\phi(Y(\tau))| + \int_0^{\tau_S} \{ |A\phi(Y(t))| + |\sigma^T(Y(t), u(t)) \nabla\phi(Y(t))|^2 \right.$

$$\left. + \sum_{m=1}^\ell \int_{\mathbb{R}^k} |\phi(Y(t) + \gamma^{(m)}) - \phi(Y(t))|^2 \nu_m(dz_m) \} dt \right] < \infty$$

*for all* $(u, \xi) \in \mathcal{A}$, $\tau \in \mathcal{T}$

(iv)   $\lim_{t \to \tau_{\mathcal{S}}^-} \phi(Y(t)) = g(Y(\tau_{\mathcal{S}}))\mathcal{X}_{\{\tau_{\mathcal{S}} < \infty\}}$ *a.s., for all* $(u, \xi) \in \mathcal{A}$

(v)   $Y(\tau_{\mathcal{S}}) \in \partial\mathcal{S}$ *a.s. on* $\{\tau_{\mathcal{S}} < \infty\}$ *and* $\{\phi^-(Y(\tau))\}_{\tau \le \tau_{\mathcal{S}}}$ *is uniformly integrable for all* $(u, \xi) \in \mathcal{A}$, $y \in \mathcal{S}$.

*Then*

$$\phi(y) \ge \Phi(y) \qquad \text{for all } y \in \mathcal{S} . \tag{5.2.4}$$

**b)**  *Define the non-intervention region $D$ by*

$$D = \Big\{ y \in \mathcal{S}; \max_{1 \le j \le p} \Big\{ \sum_{i=1}^{k} \kappa_{ij} \frac{\partial \phi}{\partial y_i}(y) + \theta_j \Big\} < 0 \Big\}. \tag{5.2.5}$$

*Suppose, in addition to (i)-(v) above, that for all $y \in \bar{D}$ there exists $v = \hat{u}(y)$ such that*

(vi)   $A^{\hat{u}(y)}\phi(y) + f(y, \hat{u}(y)) = 0$

*Moreover, suppose there exists $\hat{\xi}$ such that $(\hat{u}, \hat{\xi}) \in \mathcal{A}$ and*

(vii)   $Y^{\hat{u}, \hat{\xi}}(t) \in \bar{D}$   *for all $t$*

(viii)   $d\hat{\xi}(t) = 0$   *if*   $Y(t) \in D$

(ix)   $\sum_{j=1}^{p} \Big\{ \sum_{i=1}^{k} \kappa_{ij} \frac{\partial \phi}{\partial y_i}(Y(t^-)) + \theta_j \Big\} d\hat{\xi}_j^{(c)} = 0$   *for all $t$; $1 \le j \le p$*

*where $\xi^{(c)}(t)$ is the continuous part of $\xi(t)$, i.e. the process obtained by removing the jumps of $\xi(t)$*

(x)   $\Delta_{\hat{\xi}}\phi(Y(t_n)) + \sum_{j=1}^{p} \theta_j \Delta\hat{\xi}_j(t_n) = 0$   *for all jumping times $t_n$ of $\hat{\xi}(t)$*

*and*

(xi)   $\lim_{R \to \infty} E^y \big[ \phi\big(Y^{\hat{u}, \hat{\xi}}(T_R)\big) \big] = E^y \big[ g\big(Y^{\hat{u}, \hat{\xi}}(\tau_{\mathcal{S}})\big) \cdot \mathcal{X}_{\{\tau_{\mathcal{S}} < \infty\}} \big]$

*where*   $T_R = \min(\tau_{\mathcal{S}}, R)$   *for $R < \infty$ .*

*Then*

$$\phi(y) = \Phi(y)$$

*and*

$$(\hat{u}, \hat{\xi}) \quad \text{is an optimal control} .$$

*Proof.* **a)** Choose $(c, \xi) \in \mathcal{A}$. Then by the Itô formula for the semimartingale $Y(t) = Y^{u, \xi}(t)$ we have (see [P, Theorem II.33] and Theorem 1.24)

$$E^y[\phi(Y(T_R))] = \phi(y) + E^y \Big[ \int_0^{T_R} A^u \phi(Y(t)) dt$$

$$+ \int_0^{T_R} \sum_{i=1}^{k} \frac{\partial \phi}{\partial y_i}(Y(t^-)) \sum_{j=1}^{p} \kappa_{ij} d\xi_j^{(c)}(t) + \sum_{0 < t_n \le T_R} \Delta_{\xi}\phi(Y(t_n)) \Big]. \tag{5.2.6}$$

By the mean value theorem we have

$$\Delta_\xi \phi(Y(t_n)) = \nabla \phi(\hat{Y}^{(n)})^T \Delta_\xi Y(t_n) = \sum_{i=1}^{k} \sum_{j=1}^{p} \frac{\partial \phi}{\partial y_i}(\hat{Y}^{(n)}) \kappa_{ij} \Delta \xi_j(t_n) \quad (5.2.7)$$

where $\hat{Y}^{(n)}$ is some point on the straight line between $Y(t_n)$ and $Y(t_n^-) + \Delta_N Y(t_n)$ (see (5.2.3)).

Hence, by (i) and (ii) and (5.2.6), (5.2.7),

$$\phi(y) \geq E^y \Big[ \int_0^{T_R} f(Y(t), u(t))dt + \phi(Y(T_R))$$

$$- \sum_{j=1}^{p} \sum_{i=1}^{k} \Big\{ \int_0^{T_R} \frac{\partial \phi}{\partial y_i}(Y(t^-)) \kappa_{ij} d\xi_j^{(c)}(t) + \sum_{0 < t_n \leq T_R} \frac{\partial \phi}{\partial y_i}(\hat{Y}^{(n)}) \kappa_{ij} \Delta \xi_j(t_n) \Big\} \Big]$$

$$\geq E^y \Big[ \int_0^{T_R} f(Y(t), u(t))dt + \phi(Y(T_R)) + \sum_{j=1}^{p} \int_0^{T_R} \theta_j d\xi_j(t) \Big].$$

Letting $R \to \infty$ and applying (iv) and (v) we obtain from (5.2.7) that

$$\phi(y) \geq J^{u,\xi}(y).$$

Since $(u, \xi) \in \mathcal{A}$ was arbitrary, this proves (5.2.4).

**b)** Now apply the above argument to $(\hat{u}, \hat{\xi}) \in \mathcal{A}$, as given by (vi)-(xi). Then we get *equality* everywhere in a) and we end up with

$$\phi(y) = J^{\hat{u}, \hat{\xi}}(y)$$

and the proof is complete.    □

**Remark 5.3.** In many applications the process $Y^{u,\xi}(t)$ will have the form

$$Y^{u,\xi}(t) = \begin{bmatrix} s+t \\ X^{u,\xi}(t) \end{bmatrix} \in \mathbb{R}^{n+1} ; \qquad Y(0) = y = \begin{bmatrix} s \\ x \end{bmatrix}.$$

In this case we see by inspecting the proof that Theorem 5.2 still holds even if $\kappa = [\kappa_{ij}]$ and $\theta = [\theta_i]$ are not constant, as long as they depend on $s$ only: $\kappa = \kappa(s)$, $\theta = \theta(s)$.

## 5.3 Application to portfolio optimization with transaction costs

We now apply this theorem to Example 5.1.

In this case our state process is

$$dY(t) = \begin{bmatrix} dt \\ dX_1(t) \\ dX_2(t) \end{bmatrix} = \begin{bmatrix} 1 \\ rX_1(t) - c(t) \\ \mu X_2(t) \end{bmatrix} dt + \begin{bmatrix} 0 \\ 0 \\ \beta X_2(t) \end{bmatrix} dB(t)$$

$$+ \begin{bmatrix} 0 \\ 0 \\ X_2(t^-) \int\limits_{\mathbb{R}} z\widetilde{N}(dt, dz) \end{bmatrix} + \begin{bmatrix} 0 & 0 \\ -(1+\lambda) & (1-\mu) \\ 1 & -1 \end{bmatrix} \begin{bmatrix} d\xi_1(t) \\ d\xi_2(t) \end{bmatrix}$$

The generator of $Y(t)$ when there are no interventions is

$$A^c \phi(y) = \frac{\partial \phi}{\partial s} + (rx_1 - c)\frac{\partial \phi}{\partial x_1} + \mu x_2 \frac{\partial \phi}{\partial x_2} + \tfrac{1}{2}\beta^2 x_2^2 \frac{\partial^2 \phi}{\partial x_2^2}$$

$$+ \int\limits_{\mathbb{R}} \left\{ \phi(s, x_1, x_2 + x_2 z) - \phi(s, x_1, x_2) - x_2 z \frac{\partial \phi}{\partial x_2}(s, x_1, x_2) \right\} \nu(dz) .$$

Or, if

$$\phi(s, x_1, x_2) = \psi(x_1, x_2)$$

we have

$$A^c \phi(s, x_1, x_2) = e^{-\delta s} A_0^c \psi(x_1, x_2)$$

where

$$A_0^c \psi(x_1, x_2) = -\delta \psi + (rx_1 - c)\frac{\partial \psi}{\partial x_1} + \mu x_2 \frac{\partial \psi}{\partial x_2} + \tfrac{1}{2}\beta^2 x_2^2 \frac{\partial^2 \psi}{\partial x_2^2}$$

$$+ \int\limits_{\mathbb{R}} \left\{ \psi(x_1, x_2 + x_2 z) - \psi(x_1, x_2) - x_2 z \frac{\partial \psi}{\partial x_2}(x_1, x_2) \right\} \nu(dz) .$$

Here

$$\theta = g = 0, \quad u(t) = c(t), \quad f(y, u) = f(s, x_1, x_2, c) = e^{-\delta s} \frac{c^\gamma}{\gamma} .$$

Condition (ii) of Theorem 5.2 gets the form

$$-(1+\lambda)\frac{\partial \psi}{\partial x_1} + \frac{\partial \psi}{\partial x_2} \leq 0$$

and

$$(1-\mu)\frac{\partial \psi}{\partial x_1} - \frac{\partial \psi}{\partial x_2} \leq 0 .$$

The non-intervention region $D$ in (5.2.5) therefore becomes

$$D = \left\{ (s, x_1, x_2) \in \mathcal{S}; -(1+\lambda)\frac{\partial \psi}{\partial x_1} + \frac{\partial \psi}{\partial x_2} < 0 \text{ and } (1-\mu)\frac{\partial \psi}{\partial x_1} - \frac{\partial \psi}{\partial x_2} < 0 \right\}.$$

Conditions (i), (vi) become, respectively,

$$A_0^c \psi(x_1, x_2) + \frac{c^\gamma}{\gamma} \leq 0 \qquad \text{for all } c \geq 0$$

$$A_0^{\hat{c}} \psi(x_1, x_2) + \frac{\hat{c}^\gamma}{\gamma} = 0 \qquad \text{on } D .$$

Using these integro-variational inequalities together with the remaining conditions of Theorem 5.2 it is possible to prove the following about the optimal consumption rate $c^*(t)$ and the optimal portfolio $\xi^*(t)$:

**Theorem 5.4.** *[FØS2]. Suppose*

$$\delta > \gamma\mu - \tfrac{1}{2}\sigma^2\gamma(1-\gamma) - \gamma\|\nu\| + \int\limits_{-1}^{\infty} \{(1+z)\gamma - 1\}\nu(dz)$$

*where*

$$\|\nu\| = \nu((-1,\infty)) < \infty .$$

*Then*

$$c^*(x_1, x_2) = \left(\frac{\partial\phi}{\partial x_1}\right)^{\frac{1}{\gamma-1}}$$

*and there exist $\hat{\theta}_1, \hat{\theta}_2 \in \left[0, \frac{\pi}{2}\right]$ with $\hat{\theta}_1 < \hat{\theta}_2$ such that*

$$D = \{re^{i\theta}; \hat{\theta}_1 < \theta < \hat{\theta}_2\} \qquad (i = \sqrt{-1})$$

*is the non-intervention region and the optimal intervention strategy (portfolio) $\xi^*(t)$ is the local time of the process $(X_1(t), X_2(t))$ reflected back into $D$ in the direction parallel to $\partial_1 S$ at $\theta = \theta_1$ and in the direction parallel to $\partial_2 S$ at $\theta = \theta_2$. See Figure 5.2.*

For proofs and more details we refer to [FØS2].

**Remark 5.5.** For other applications of singular control theory of jump diffusions see [Ma].

## 5.4 Exercises

**Exercise* 5.1 (Optimal dividend policy under proportional transaction costs).**
Suppose the cash flow $X(t) = X^{(\xi)}(t)$ of a firm at time $t$ is given by (with $\alpha, \sigma, \beta, \lambda > 0$ constants)

$$dX(t) = \alpha dt + \sigma dB(t) + \beta \int\limits_{\mathbb{R}} z\widetilde{N}(dt, dz) - (1+\lambda)d\xi(t) ; \qquad X(0^-) = x > 0 ,$$

$x_1 + (1 - \mu)x_2 = 0$

$x_2$

$(1 - \mu)\frac{\partial \psi}{\partial x_1} - \frac{\partial \psi}{\partial x_2} = 0$

$\theta = \hat{\theta}_2$

(Sell region)

$\partial_2 S$

$(x_1, x_2)$

sale

$D =$

non-intervention region

jump

$(x_1', x_2')$

$\theta = \hat{\theta}_1$

jump

$(x_1', x_2')$

purchase    $(x_1, x_2)$

$-(1 + \lambda)\frac{\partial \psi}{\partial x_1} + \frac{\partial \psi}{\partial x_2} = 0$ (Buy region)

$x_1$

$\partial_1 S$

$x_1 + (1 + \lambda)x_2 = 0$
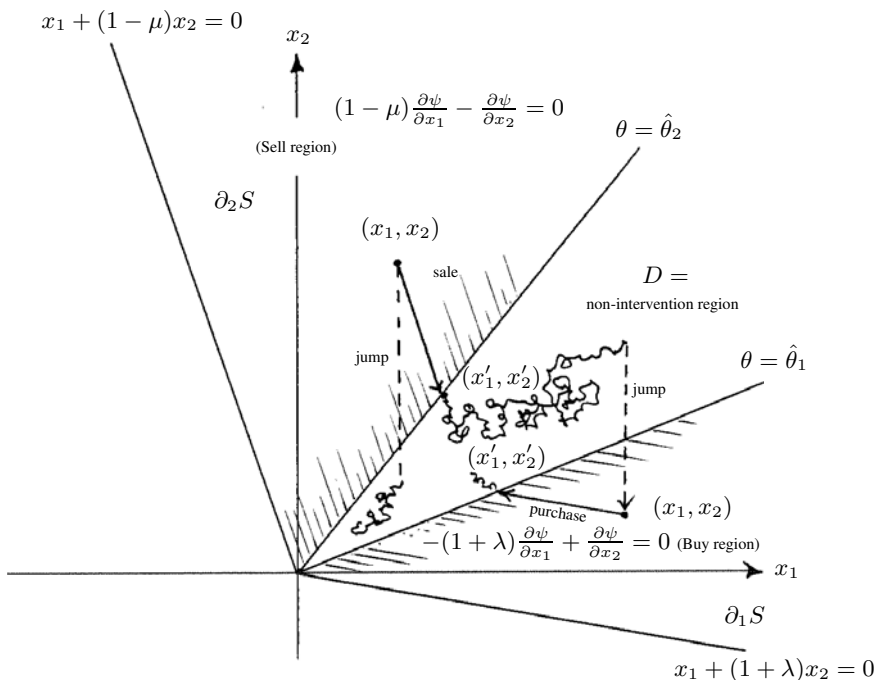
$\ell$ = purchase amount, $m$ = sale amount (at transaction)
$x_1' = x_1 - (1 + \lambda)\ell + (1 - \mu)m$ (new value of $x_1$ after transaction)
$x_2' = x_2 + \ell - m$             (new value of $x_2$ after transaction)

**Fig. 5.2.** The optimal portfolio $\xi^*(t)$.

where $\xi(t)$ is an increasing, adapted cadlag process representing the total dividend taken out up to time $t$ (our control process). Let

$$\tau_S = \inf\{t > 0; X^{(\xi)}(t) \leq 0\}$$

be the time of bankruptcy of the firm and let

$$J^\xi(s, x) = E^{s,x}\left[ \int_0^{\tau_S} e^{-\rho(s+t)} d\xi(t) \right], \qquad \rho > 0 \text{ constant,}$$

be the expected total discounted amount taken out up to bankruptcy time.
Find $\Phi(s, x)$ and a dividend policy $\xi^*$ such that

$$\Phi(s, x) = \sup_\xi J^{(\xi)}(s, x) = J^{\xi^*}(s, x) \ .$$

**Exercise* 5.2.** Let $\Phi(s, x_1, x_2)$ be the value function of the optimal consumption problem (5.1.2) with proportional transaction costs and let $\Phi_0(s, x_1, x_2) =$

$Ke^{-\delta s}(x_1+x_2)^\gamma$ be the corresponding value function when there are no trans-
action costs, i.e. $\mu = \lambda = 0$) (Example 3.2). Use Theorem 5.2a) to prove that

$$\Phi(s, x_1, x_2) \le Ke^{-\delta s}(x_1 + x_2)^\gamma.$$

**Exercise 5.3 (Optimal harvesting).**
Suppose the size $X(t)$ at time $t$ of a certain fish population is modeled by
a geometric Lévy process, i.e.

$$dX(t) = dX^{(\xi)}(t) = X(t^-)\left[\mu dt + \int_{\mathbb{R}} z\widetilde{N}(dt, dz)\right] - d\xi(t) \; ; \; t > 0$$

$$X(0^-) = x > 0$$

where $\mu > 0$ is a constant, $z > -1$ a.s. $\nu(dz)$ and $\xi(t)$ is an increasing adapted
process giving the amount harvested from the population from time 0 up to
time $t$. We assume that $\xi(t)$ is right-continuous. Consider the *optimal harvest-
ing problem*

$$\Phi(s, x) = \sup_{\xi} J^{(\xi)}(s, x),$$

where

$$J^{(\xi)}(s, x) = E^{s,x}\left[\int_0^{\tau_s} \theta e^{-\rho(s+t)}d\xi(t)\right],$$

with $\theta > 0$, $\rho > 0$ constants and

$$\tau_S = \inf\{t > 0 \; ; \; X^{(\xi)}(t) \le 0\} \quad \text{(extinction time).}$$

If we interpret $\theta$ as the price per unit harvested, then $J^{(\xi)}(s, x)$ represents
the expected total discounted value of the harvested amount up to extinction
time.

**a)** Write down the integro-variational inequalities (i), (ii), (vi), and (ix) of
Theorem 5.2 in this case, with the state process

$$Y(t) = \begin{bmatrix} s+t \\ X(t) \end{bmatrix} \; ; \; t \ge 0, \quad Y(0) = y = \begin{bmatrix} s \\ x \end{bmatrix} \in \mathbb{R}^+ \times \mathbb{R}^+.$$

**b)** Suppose $\mu \le \rho$.
   Show that in this case it is optimal to harvest all the population immedi-
   ately, i.e. it is optimal to choose the harvesting strategy $\hat{\xi}$ defined by

$$\hat{\xi}(t) = x \text{ for all } t \ge 0$$

   (sometimes called the "take the money and run"- strategy).
   This gives the value function

$$\Phi(s, x) = \theta e^{-\rho s}x.$$

**c)** Suppose $\mu > \rho$.
   Show that in this case $\Phi(s, x) = \infty$.

# 6

# Impulse Control of Jump Diffusions

## 6.1 A general formulation and a verification theorem

Suppose that – if there are no interventions – the state $Y(t) \in \mathbb{R}^k$ of the system we consider is a jump diffusion of the form

$$dY(t) = b(Y(t))dt + \sigma(Y(t))dB(t) + \int_{\mathbb{R}^k} \gamma(Y(t^-), z)\widetilde{N}(dt, dz) \qquad (6.1.1)$$

where $b : \mathbb{R}^k \to \mathbb{R}^k$, $\sigma : \mathbb{R}^k \to \mathbb{R}^{k \times m}$ and $\gamma : \mathbb{R}^k \times \mathbb{R}^k \to \mathbb{R}^{k \times \ell}$ are given functions satisfying the conditions for the existence and uniqueness of a solution $Y(t)$ (see Theorem 1.19).

The generator $A$ of $Y(t)$ is

$$A\phi(y) = \sum_{i=1}^{k} b_i(y) \frac{\partial \phi}{\partial y_i} + \tfrac{1}{2} \sum_{i,j=1}^{k} (\sigma \sigma^T)_{ij}(y) \frac{\partial^2 \phi}{\partial y_i \partial y_j}$$

$$+ \sum_{j=1}^{\ell} \int_{\mathbb{R}} \{\phi(y + \gamma^{(j)}(y, z_j)) - \phi(y) - \nabla\phi(y) \cdot \gamma^{(j)}(y, z_j)\}\nu_j(dz_j) \, .$$

Now suppose that at any time $t$ and any state $y$ we are free to intervene and give the system an impulse $\zeta \in \mathcal{Z} \subset \mathbb{R}^p$, where $\mathcal{Z}$ is a given set (the set of admissible impulse values). Suppose the result of giving the impulse $\zeta$ when the state is $y$ is that the state jumps immediately from $y = Y(t^-)$ to $Y(t) = \Gamma(y, \zeta) \in \mathbb{R}^k$, where $\Gamma : \mathbb{R}^k \times \mathcal{Z} \to \mathbb{R}^k$ is a given function.

An *impulse control* for this system is a double (possibly finite) sequence

$$v = (\tau_1, \tau_2, \ldots, \tau_j, \ldots; \zeta_1, \zeta_2, \ldots, \zeta_j, \ldots)_{j \leq M} \, ; \qquad M \leq \infty$$

where $\tau_1 < \tau_2 < \cdots$ are $\mathcal{F}_t$-stopping times (the *intervention times*) and $\zeta_1, \zeta_2, \ldots$ are the corresponding *impulses* at these times. We assume that $\zeta_j$ is $\mathcal{F}_{\tau_j}$-measurable for all $j$.

If $v = (\tau_1, \tau_2, \ldots; \zeta_1, \zeta_2, \ldots)$ is an impulse control, the corresponding state process $Y^{(v)}(t)$ is defined by

$$Y^{(v)}(t) = Y(t) ; \qquad 0 \le t \le \tau_1 \tag{6.1.2}$$

$$Y^{(v)}(\tau_j) = \Gamma(Y^{(v)}(\tau_j^-) + \Delta_N Y(\tau_j), \zeta_j) ; \qquad j = 1, 2, \ldots \tag{6.1.3}$$

$$dY^{(v)}(t) = b(Y^{(v)}(t))dt + \sigma(Y^{(v)}(t))dB(t)$$
$$+ \int_{\mathbb{R}^k} \gamma(Y^{(v)}(t), z)\tilde{N}(dt, dz) \qquad \text{for } \tau_j \le t < \tau_{j+1} \wedge \tau^* \tag{6.1.4}$$

where, as in (5.2.2), $\Delta_N Y(t)$ is the jump of $Y$ stemming from the jump of the random measure $N(t, \cdot)$ only and

$$\tau^* = \tau^*(\omega) = \lim_{R \to \infty} (\inf\{t > 0; |Y^{(v)}(t)| \ge R\}) \le \infty \tag{6.1.5}$$

is the explosion time of $Y^{(v)}(t)$.

Note that here we must distinguish between the (possible) jump of $Y^{(v)}(\tau_j)$ stemming from $N$, denoted by $\Delta_N Y(\tau_j)$ and the jump caused by the intervention $v$, given by

$$\Delta_v Y(\tau_j) := \Gamma(\check{Y}(\tau_j^-), \zeta) - \check{Y}(\tau_j^-) , \tag{6.1.6}$$

where

$$\check{Y}(\tau_j^-) = Y(\tau_j^-) + \Delta_N Y(\tau_j) . \tag{6.1.7}$$

Let $\mathcal{S} \subset \mathbb{R}^k$ be a fixed open set (the solvency region). Define

$$\tau_\mathcal{S} = \inf\{t \in (0, \tau^*); Y^{(v)}(t) \notin \mathcal{S}\} . \tag{6.1.8}$$

Suppose we are given a *profit function* $f : \mathcal{S} \to \mathbb{R}$ and a *bequest function* $g : \mathbb{R}^k \to \mathbb{R}$. Moreover, suppose the profit/utility of making an intervention with impulse $\zeta \in \mathcal{Z}$ when the state is $y$ is $K(y, \zeta)$, where $K : \mathcal{S} \times \mathcal{Z} \to \mathbb{R}$ is a given function.

We assume we are given a set $\mathcal{V}$ of *admissible impulse controls* which is included in the set of $v = (\tau_1, \tau_2, \ldots; \zeta_1, \zeta_2, \ldots)$ such that a unique solution $Y^{(v)}$ of (6.1.2)–(6.1.4) exists and

$$\tau^* = \infty \quad \text{a.s} \tag{6.1.9}$$

and

$$\lim_{j \to \infty} \tau_j = \tau_\mathcal{S} \quad \text{a.s.} \qquad (\text{if } M < \infty \text{ we assume } \tau_M = \tau_\mathcal{S} \text{ a.s.}) \tag{6.1.10}$$

We also assume that

$$E^y\left[\int_0^{\tau_\mathcal{S}} f^-(Y^{(v)}(t))dt\right] < \infty \qquad \text{for all } y \in \mathbb{R}^k, \ v \in \mathcal{V} \tag{6.1.11}$$

and

$$E\big[g^-(Y^{(v)}(\tau_{\mathcal{S}}))\mathcal{X}_{\{\tau_{\mathcal{S}}<\infty\}}\big] < \infty \qquad \text{for all } y \in \mathbb{R}^k, \ v \in \mathcal{V} \qquad (6.1.12)$$

and

$$E\Big[\sum_{\tau_j \le \tau_{\mathcal{S}}} K^-(\check{Y}^{(v)}(\tau_j^-),\zeta_j)\Big] < \infty \qquad \text{for all } y \in \mathbb{R}^k, \ v \in \mathcal{V}. \qquad (6.1.13)$$

Now define the performance criterion

$$J^{(v)}(y) = E^y\Big[ \int\limits_0^{\tau_{\mathcal{S}}} f(Y^{(v)}(t))dt + g(Y^{(v)}(\tau_{\mathcal{S}}))\mathcal{X}_{\{\tau_{\mathcal{S}}<\infty\}} + \sum_{\tau_j \le \tau_{\mathcal{S}}} K(\check{Y}^{(v)}(\tau_j^-),\zeta_j)\Big].$$

The *impulse control* problem is the following:
   Find $\Phi(y)$ and $v^* \in \mathcal{V}$ such that

$$\Phi(y) = \sup\{J^{(v)}(y); v \in \mathcal{V}\} = J^{(v^*)}(y) . \qquad (6.1.14)$$

The following concept is crucial:

**Definition 6.1.** *Let $\mathcal{H}$ be the space of all measurable functions $h : \mathcal{S} \to \mathbb{R}$. The* intervention operator $\mathcal{M} : \mathcal{H} \to \mathcal{H}$ *is defined by*

$$\mathcal{M}h(y) = \sup\{h(\Gamma(y,\zeta)) + K(y,\zeta); \ \zeta \in \mathcal{Z} \ \text{ and } \ \Gamma(y,\zeta) \in \mathcal{S}\} . \qquad (6.1.15)$$

   As in Chapter 2 we put

$$\mathcal{T} = \{\tau; \tau \text{ stopping time, } 0 \le \tau \le \tau_{\mathcal{S}} \text{ a.s.}\}.$$

We can now state the main result of this chapter, a verification theorem for impulse control problems:

**Theorem 6.2 (Quasi-integrovariational inequalities for impulse control).**
**a)** *Suppose we can find $\phi : \bar{\mathcal{S}} \to \mathbb{R}$ such that*

(i) $\phi \in C^1(\mathcal{S}) \cap C(\bar{\mathcal{S}})$,
(ii) $\phi \ge \mathcal{M}\phi$ *on $\mathcal{S}$.*
   *Define*

$$D = \{y \in \mathcal{S}; \phi(y) > \mathcal{M}\phi(y)\} \qquad \text{(the continuation region).}$$

   *Assume*
(iii) $E^y\Big[ \int\limits_0^{\tau_{\mathcal{S}}} \mathcal{X}_{\partial D}(Y(t))dt \Big] = 0 \quad \text{for all } y \in \mathcal{S}, v \in \mathcal{V},$

(iv) $\partial D$ *is a Lipschitz surface,*
(v) $\phi \in C^2(\mathcal{S} \setminus \partial D)$ *with locally bounded derivatives near $\partial D$,*

(vi) $A\phi + f \leq 0$ on $\mathcal{S} \setminus \partial D$,

(vii) $\phi(Y(t)) \to g(Y(\tau_{\mathcal{S}})) \cdot \mathcal{X}_{\{\tau_{\mathcal{S}} < \infty\}}$ as $t \to \tau_{\mathcal{S}}^-$ a.s., for all $y \in \mathcal{S}$, $v \in \mathcal{V}$,

(viii) $\{\phi^-(Y(\tau)); \tau \in \mathcal{T}\}$ is uniformly integrable, for all $y \in \mathcal{S}$, $v \in \mathcal{V}$,

(ix) $E^y \Big[ |\phi(Y(\tau))| + \int\limits_0^{\tau_{\mathcal{S}}} \Big\{ |A\phi(Y(t))| + |\sigma^T(Y(t))\nabla\phi(Y(t))|^2$

$$+ \sum_{j=1}^{\ell} \int\limits_{\mathbb{R}} |\phi(Y(t) + \gamma^{(j)}(Y(t), z_j)) - \phi(Y(t))|^2 \nu_j(dz_j) \Big\} dt \Big] < \infty$$

for all $\tau \in \mathcal{T}$, $v \in \mathcal{V}$, $y \in \mathcal{S}$.

Then

$$\phi(y) \geq \Phi(y) \qquad \text{for all } y \in \mathcal{S}. \tag{6.1.16}$$

**b)** *Suppose in addition that*

(x) $A\phi + f = 0$ in $D$,

(xi) $\hat{\zeta}(y) \in \text{Argmax}\{\phi(\Gamma(y, \cdot)) + K(y, \cdot)\} \in \mathcal{Z}$ *exists for all* $y \in \mathcal{S}$ *and* $\hat{\zeta}(\cdot)$ *is a Borel measurable selection.*

Put $\hat{\tau}_0 = 0$ and define $\hat{v} = (\hat{\tau}_1, \hat{\tau}_2, \ldots; \hat{\zeta}_1, \hat{\zeta}_2, \ldots)$ inductively by
$\hat{\tau}_{j+1} = \inf\{t > \hat{\tau}_j; Y^{(\hat{v}_j)}(t) \notin D\} \wedge \tau_{\mathcal{S}}$ and $\hat{\zeta}_{j+1} = \hat{\zeta}(Y^{(\hat{v}_j)}(\hat{\tau}_{j+1}^-))$
if $\hat{\tau}_{j+1} < \tau_{\mathcal{S}}$, where $Y^{(\hat{v}_j)}$ is the result of applying
$\hat{v}_j := (\hat{\tau}_1, \ldots, \hat{\tau}_j; \hat{\zeta}_1, \ldots, \hat{\zeta}_j)$ to $Y$.

Suppose

(xii) $\hat{v} \in \mathcal{V}$ and $\{\phi(Y^{(\hat{v})}(\tau)); \tau \in \mathcal{T}\}$ is uniformly integrable.

Then

$$\phi(y) = \Phi(y) \qquad \text{and } \hat{v} \text{ is an optimal impulse control}. \tag{6.1.17}$$

SKETCH OF PROOF. **a)** By Theorem 2.1 and (iii)–(v), we may assume that $\phi \in C^2(\mathcal{S}) \cap C(\bar{\mathcal{S}})$. Choose $v = (\tau_1, \tau_2, \ldots; \zeta_1, \zeta_2, \ldots) \in \mathcal{V}$ and set $\tau_0 = 0$. By another approximation argument we may assume that we can apply the Dynkin formula to the stopping times $\tau_j$. Then for $j = 0, 1, 2, \ldots$, with $Y = Y^{(v)}$,

$$E^y[\phi(Y(\tau_j))] - E^y[\phi(\check{Y}(\tau_{j+1}^-))] = -E^y \Big[ \int\limits_{\tau_j}^{\tau_{j+1}} A\phi(Y(t)) dt \Big], \tag{6.1.18}$$

where $\check{Y}(\tau_{j+1}^-) = Y(\tau_{j+1}^-) + \Delta_N Y(\tau_{j+1})$, as before. Summing this from $j = 0$ to $j = m$ we get

$$\phi(y) + \sum_{j=1}^m E^y[\phi(Y(\tau_j)) - \phi(\check{Y}(\tau_j^-))] - E^y[\phi(\check{Y}(\tau_{m+1}^-))]$$

$$= -E^y \Big[ \int\limits_0^{\tau_{m+1}} A\phi(Y(t)) dt \Big] \geq E^y \Big[ \int\limits_0^{\tau_{m+1}} f(Y(t)) dt \Big]. \tag{6.1.19}$$

Now

$$\phi(Y(\tau_j)) = \phi(\Gamma(\check{Y}(\tau_j^-), \zeta_j))$$
$$\leq \mathcal{M}\phi(\check{Y}(\tau_j^-)) - K(\check{Y}(\tau_j^-), \zeta_j) \qquad \text{if } \tau_j < \tau_{\mathcal{S}} \text{ by (6.1.15)}$$

and

$$\phi(Y(\tau_j)) = \phi(\check{Y}(\tau_j^-)) \qquad \text{if } \tau_j = \tau_{\mathcal{S}} \text{ by (vii).}$$

Therefore

$$\mathcal{M}\phi(\check{Y}(\tau_j^-)) - \phi(\check{Y}(\tau_j^-)) \geq \phi(Y(\tau_j)) - \phi(\check{Y}(\tau_j^-)) + K(\check{Y}(\tau_j^-), \zeta_j)$$

and

$$\phi(y) + \sum_{j=1}^{m} E^y[\{\mathcal{M}\phi(\check{Y}(\tau_j^-)) - \phi(\check{Y}(\tau_j^-))\} \cdot \mathcal{X}_{\{\tau_j < \tau_{\mathcal{S}}\}}]$$

$$\geq E^y\Big[ \int_0^{\tau_{m+1}} f(Y(t))dt + \phi(\check{Y}(\tau_{m+1}^-)) + \sum_{j=1}^{m} K(\check{Y}(\tau_j^-), \zeta_j)\Big].$$

Letting $m \to M$ we get

$$\phi(y) \geq E^y\Big[ \int_0^{\tau_{\mathcal{S}}} f(Y(t))dt + g(Y(\tau_{\mathcal{S}}))\mathcal{X}_{\{\tau_{\mathcal{S}} < \infty\}} + \sum_{j=1}^{M} K(\check{Y}(\tau_j^-), \zeta_j)\Big] = J^{(v)}(y) .$$

$$(6.1.20)$$

Hence    $\phi(y) \geq \Phi(y)$.

**b)**  Next assume (x)–(xii) also hold. Apply the above argument to $\hat{v} = (\hat{\tau}_1, \hat{\tau}_2, \ldots; \hat{\zeta}_1, \hat{\zeta}_2, \ldots)$. Then by (x) we get *equality* in (6.1.19) and by our choice of $\zeta_j = \hat{\zeta}_j$ we have *equality* in (6.1.20). Hence

$$\phi(y) = J^{(\hat{v})}(y) ,$$

which combined with a) completes the proof.

**Remark 6.3.** In the case of a pure diffusion process, the same verification theorem holds ; just skip condition (ix).

## 6.2 Examples

**Example 6.4 (Optimal stream of dividends under transaction costs).**
This example is an extension to the jump diffusion case of a problem studied

in [J-PS]. Suppose that if we make no interventions the amount $X(t)$ available (cash flow) is given by

$$dX(t) = \mu dt + \sigma dB(t) + \theta \int_{\mathbb{R}} z\widetilde{N}(dt, dz) ; \qquad X(0) = x > 0 \qquad (6.2.1)$$

where $\mu, \sigma > 0$, $\theta \geq 0$ are constants. Suppose that at any time $t$ we are free to take out an amount $\zeta > 0$ from $X(t)$ by applying the transaction cost

$$k(\zeta) = c + \lambda \zeta \qquad (6.2.2)$$

where $c > 0$, $\lambda \geq 0$ are constants. The constant $c$ is called the *fixed* part and the quantity $\lambda \zeta$ is called the *proportional* part, respectively, of the transaction cost. The resulting cash flow $X^{(v)}(t)$ is given by

$$X^{(v)}(t) = X(t) \qquad \text{if } 0 \leq t < \tau_1, \qquad (6.2.3)$$

$$X^{(v)}(\tau_j) = \Gamma(X^{(v)}(\tau_j^-) + \Delta_N X(\tau_j), \zeta_j) = \check{X}^{(v)}(\tau_j^-) - (1 + \lambda)\zeta_j - c \quad (6.2.4)$$

and

$$dX^{(v)}(t) = \mu dt + \sigma dB(t) + \theta \int_{\mathbb{R}} z\widetilde{N}(dt, dz) \qquad \text{if } \tau_j \leq t < \tau_{j+1} . \qquad (6.2.5)$$

Put

$$\tau_{\mathcal{S}} = \inf\{t > 0; X^{(v)}(t) \leq 0\} \qquad \text{(time of bankruptcy)} \qquad (6.2.6)$$

and

$$J^{(v)}(s, x) = E^{s,x}\left[ \sum_{\tau_j \leq \tau_{\mathcal{S}}} e^{-\rho(s+\tau_j)}\zeta_j \right] \qquad (6.2.7)$$

where $\rho > 0$ is constant (the discounting exponent).

We seek $\Phi(s, x)$ and $v^* = (\tau_1^*, \tau_2^*, \ldots; \zeta_1^*, \zeta_2^*, \ldots) \in \mathcal{V}$ such that

$$\Phi(s, x) = \sup_{v \in \mathcal{V}} J^{(v)}(s, x) = J^{(v^*)}(s, x) , \qquad (6.2.8)$$

where $\mathcal{V}$ is the set of impulse controls s.t. $X^{(v)}(t) \geq 0$ for all $t \leq \tau_{\mathcal{S}}$. This is a problem of the type (6.1.14), with

$$Y^{(v)}(t) = \begin{bmatrix} s+t \\ X^{(v)}(t) \end{bmatrix}; \quad t \geq 0 \qquad Y^{(v)}(0^-) = \begin{bmatrix} s \\ x \end{bmatrix} = y$$

$$\Gamma(y, \zeta) = \Gamma(s, x, \zeta) = \begin{bmatrix} s \\ x - c - (1 + \lambda)\zeta \end{bmatrix}$$

$$K(y, \zeta) = K(s, x, \zeta) = e^{-\rho s}\zeta , \qquad f = g = 0$$

and

$$\mathcal{S} = \{(s, x); x > 0\} .$$

As a candidate for the value function $\Phi$ we try

$$\phi(s, x) = e^{-\rho s}\psi(x) . \tag{6.2.9}$$

Then

$$\mathcal{M}\psi(x) = \sup\left\{\psi(x - c - (1 + \lambda)\zeta) + \zeta \;;\; 0 < \zeta < \tfrac{x-c}{1+\lambda}\right\}.$$

We now guess that the continuation region has the form

$$D = \{(s, x); 0 < x < x^*\} \qquad \text{for some } x^* > 0 . \tag{6.2.10}$$

Then (x) of Theorem 6.2 gives

$$-\rho\psi(x) + \mu\psi'(x) + \tfrac{1}{2}\sigma^2\psi''(x) + \int_{\mathbb{R}}\{\psi(x + \theta z) - \psi(x) - \psi'(x)\theta z\}\nu(dz) = 0 .$$

To solve this equation we try a function of the form

$$\psi(x) = e^{rx}$$

for some constant $r \in \mathbb{R}$. Then $r$ must solve the equation

$$h(r) := -\rho + \mu r + \tfrac{1}{2}\sigma^2 r^2 + \int_{\mathbb{R}}\{e^{r\theta z} - 1 - r\theta z\}\nu(dz) = 0 . \tag{6.2.11}$$

Since $h(0) = -\rho < 0$ and $\lim_{|r| \to \infty} h(r) = \infty$, we see that there exist two solutions $r_1, r_2$ of $h(r) = 0$ such that

$$r_2 < 0 < r_1 .$$

Moreover, since $e^{r\theta z} - 1 - r\theta z \geq 0$ for all $r, z$ we have

$$|r_2| > r_1 .$$

With such a choice of $r_1, r_2$ we try

$$\psi(x) = A_1 e^{r_1 x} + A_2 e^{r_2 x}; \qquad A_i \text{ constants.}$$

Since

$$\psi(0) = 0 \quad \text{we have} \quad A_1 + A_2 = 0$$

so we write $A_1 = A = -A_2 > 0$ and

$$\psi(x) = A\big(e^{r_1 x} - e^{r_2 x}\big) ; \qquad 0 < x < x^* .$$

Define

$$\psi_0(x) = A\big(e^{r_1 x} - e^{r_2 x}\big) \qquad \text{for } all \ x > 0 . \tag{6.2.12}$$

To study $\mathcal{M}\psi$ we first consider

$$g(\zeta) := \psi_0(x - c - (1 + \lambda)\zeta) + \zeta \, ; \qquad \zeta > 0 \, .$$

The first order condition for a maximum point $\hat{\zeta} = \hat{\zeta}(x)$ for $g(\zeta)$ is that

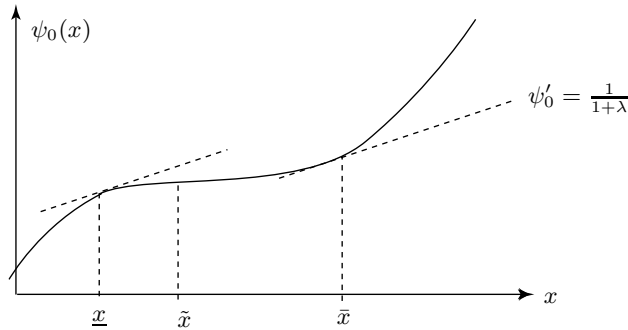$$\psi_0'(x - c - (1 + \lambda)\hat{\zeta}) = \frac{1}{1 + \lambda} \, .$$

Now

$$\psi_0'(x) > 0 \qquad \text{for all } x \text{ and}$$

$$\psi_0''(x) < 0 \qquad \text{iff} \quad x < \tilde{x} := \frac{2(\ln |r_2| - \ln r_1)}{r_1 - r_2} \, .$$

Therefore the equation $\psi_0'(x) = \frac{1}{1+\lambda}$ has exactly two solutions $x = \underline{x}$, $x = \bar{x}$ where

$$0 < \underline{x} < \tilde{x} < \bar{x}$$

(provided that $\psi_0'(\tilde{x}) < \frac{1}{1+\lambda} < \psi_0'(0)$). See Figure 6.1.



**Fig. 6.1.** The function $\psi_0(x)$

Choose

$$x^* = \bar{x} \quad \text{and put} \quad \hat{x} = \underline{x} \, . \tag{6.2.13}$$

If we require that $\psi(x) = \mathcal{M}\psi_0(x)$ for $x \geq x^*$ we get

$$\psi(x) = \psi_0(\hat{x}) + \hat{\zeta}(x) \qquad \text{for} \quad x \geq x^*$$

where

$$x - c - (1 + \lambda)\hat{\zeta}(x) = \hat{x}$$

i.e.

$$\hat{\zeta}(x) = \frac{x - \hat{x} - c}{1 + \lambda} \qquad \text{for} \quad x \geq x^* \, . \tag{6.2.14}$$

Hence we propose that $\psi$ has the form

$$\psi(x) = \begin{cases} \psi_0(x) = A(e^{r_1 x} - e^{r_2 x}) \,; & 0 < x < x^* \\ \psi_0(\hat{x}) + \frac{x - \hat{x} - c}{1 + \lambda} \,; & x \geq x^* \end{cases} \tag{6.2.15}$$

Now choose $A$ such that $\psi$ is continuous at $x = x^*$. This gives

$$A = (1 + \lambda)^{-1} \left[ e^{r_1 x^*} - e^{r_2 x^*} - e^{r_1 \hat{x}} + e^{r_2 \hat{x}} \right]^{-1} (x^* - \hat{x} - c) \,. \tag{6.2.16}$$

By our choice of $x^*$ we then have that $\psi$ is also differentiable at $x = x^*$.

We can now check that, with these values of $x^*, \hat{x}$ and $A$, our choice of $\phi(s, x) = e^{-\rho s} \psi(x)$ satisfies all the requirements of Theorem 6.2, provided that some conditions on the parameters are satisfied. We leave this verification to the reader.

Thus the solution of the impulse control problem (6.2.8) can be described as follows: As long as $X(t) < x^*$ we do nothing. If $X(t)$ reaches the value $x^*$ or jumps above this value, then immediately we make an intervention to bring $X(t)$ down to the value $\hat{x}$. See Figure 6.2.
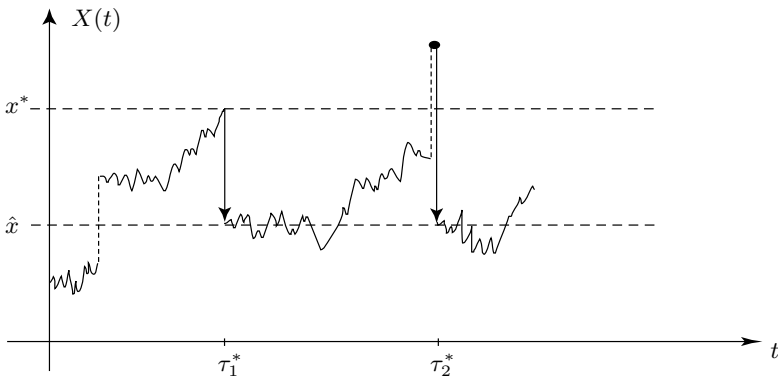


**Fig. 6.2.** The optimal impulse control of Example 6.4

**Example 6.5.** As another illustration of how to apply Theorem 6.2 we consider the following example, which is a jump diffusion version of the example in [Ø2] studied in connection with questions involving vanishing fixed costs. Variations of this problem have been studied by many authors. See e.g. [HST], [J], [MØ], [ØS], [ØUZ] and [V]. One possible economic interpretation is that the given process represents the exchange rate of a given currency and the impulses represent the interventions taken in order to keep the exchange rate in a given "target zone". See e.g. [J] and [MØ].

Suppose that without interventions the system has the form

$$Y(t) = \begin{bmatrix} s+t \\ X(t) \end{bmatrix} \in \mathbb{R}^2 \; ; \; Y(0) = y = (s, x) \tag{6.2.17}$$

where $X(t) = x + B(t) + \int_0^t \int_{\mathbb{R}} z\tilde{N}(ds, dz)$ and $B(0) = 0$. Suppose that we are only allowed to give the system impulses $\zeta$ with values in $\mathcal{Z} := (0, \infty)$ and that if we apply an impulse control $v = (\tau_1, \tau_2, \ldots \; ; \; \zeta_1, \zeta_2, \ldots)$ to $Y(t)$ it gets the form

$$Y^{(v)}(t) = \begin{bmatrix} s+t \\ X(t) - \sum_{\tau_k \leq t} \zeta_k \end{bmatrix} = \begin{bmatrix} s+t \\ X^{(v)}(t) \end{bmatrix}. \tag{6.2.18}$$

Suppose that the *cost rate* $f(t, \xi)$ if $X^{(v)}(t) = \xi$ at time $t$ is given by

$$f(t, \xi) = e^{-\rho t}\xi^2 \tag{6.2.19}$$

where $\rho > 0$ is constant. In an effort to reduce the cost one can apply the impulse control $v$ in order to reduce the value of $X^{(v)}(t)$. However, suppose the cost of an intervention of size $\zeta > 0$ at time $t$ is

$$K(t, \xi, \zeta) = K(\zeta) = c + \lambda\zeta, \tag{6.2.20}$$

where $c > 0$, $\lambda \geq 0$ are constants. Then the expected total discounted cost associated to a given impulse control is

$$J^{(v)}(s, x) = E^x \left[ \int_0^\infty e^{-\rho(s+t)}(X^{(v)}(t))^2 dt + \sum_{k=1}^N e^{-\rho(s+\tau_k)}(c + \lambda\zeta_k) \right]. \tag{6.2.21}$$

We seek $\Phi(s, x)$ and $v^* = (\tau_1^*, \tau_2^*, \ldots \; ; \; \zeta_1^*, \zeta_2^*, \ldots)$ such that

$$\Phi(s, x) = \inf_v J^{(v)}(s, x) = J^{(v^*)}(s, x). \tag{6.2.22}$$

This is an impulse control problem of the type described above, except that it is a minimum problem rather than a maximum problem. Theorem 6.2 still applies, with the corresponding changes.

Note that it is not optimal to move $X(t)$ downwards if $X(t)$ is already below 0. Hence we may restrict ourselves to consider impulse controls $v = (\tau_1, \tau_2, \ldots \; ; \; \zeta_1, \zeta_2, \ldots)$ such that

$$\sum_{j=1}^{\tau_k} \zeta_j \leq X(\tau_k) \text{ for all } k. \tag{6.2.23}$$

We let $\mathcal{V}$ denote the set of such impulse controls.

We guess that the optimal strategy is to wait until the level of $X(t)$ reaches an (unknown) value $x^* > 0$. At this time, $\tau_1$, we intervene and give $X(t)$ an

impulse $\zeta_1$, which brings it down to a lower value $\hat{x} > 0$. Then we do nothing until the next time, $\tau_2$, that $X(t)$ reaches the level $x^*$ etc. This suggests that the continuation region $D$ in Theorem 6.2 has the form

$$D = \{(s, x) \; ; \; x < x^*\}. \tag{6.2.24}$$

See Figure 6.3.

Let us try a value function $\varphi$ of the form

$$\varphi(s, x) = e^{-\rho s}\psi(x) \tag{6.2.25}$$

where $\psi$ remains to be determined.



**Fig. 6.3.** The optimal impulse control of Example 6.5

Condition (x) of Theorem 6.2 gives that for $x < x^*$ we should have

$$A\varphi + f = e^{-\rho s}\left(-\rho\psi(x) + \frac{1}{2}\psi''(x) + \int_{\mathbb{R}} \{\psi(x + z) - \psi(x) - z\psi'(x)\}\nu(dz)\right)$$
$$+ e^{-\rho s}x^2 = 0.$$

So for $x < x^*$ we let $\psi$ be a solution $h(x)$ of the equation

$$\int_{\mathbb{R}} \{h(x + z) - h(x) - zh'(x)\}\nu(dz) + \frac{1}{2}h''(x) - \rho h(x) + x^2 = 0. \tag{6.2.26}$$

We see that any function $h(x)$ of the form

$$h(x) = C_1 e^{r_1 x} + C_2 e^{r_2 x} + \frac{1}{\rho} x^2 + \frac{1 + \int_{\mathbb{R}} z^2 \nu(dz)}{\rho^2} \tag{6.2.27}$$

where $C_1, C_2$ are arbitrary constants, is a solution of (6.2.26), provided that $r_1 > 0, r_2 < 0$ are roots of the equation

$$K(r) := \int_{\mathbb{R}} \{e^{rz} - 1 - rz\} \nu(dz) + \frac{1}{2} r^2 - \rho = 0.$$

Note that if we make no interventions at all, then the cost is

$$J^{(v)}(s, x) = e^{-\rho s} E^x \left[ \int_0^\infty e^{-\rho t} (X(t))^2 dt \right]$$

$$= e^{-\rho s} \int_0^\infty e^{-\rho t} (x^2 + tb) dt = e^{-\rho s} \left( \frac{1}{\rho} x^2 + \frac{b}{\rho^2} \right), \tag{6.2.28}$$

where $b = 1 + \int_{\mathbb{R}} z^2 \nu(dz)$. Hence we must have

$$0 \le \psi(x) \le \frac{1}{\rho} x^2 + \frac{b}{\rho^2} \quad \text{for all } x. \tag{6.2.29}$$

Comparing this with (6.2.27) we see that we must have $C_2 = 0$. Hence $C_1 \le 0$. So we put

$$\psi(x) = \psi_0(x) := \frac{1}{\rho} x^2 + \frac{b}{\rho^2} - a e^{r_1 x} \text{ for } x \le x^* \tag{6.2.30}$$

where $a = -C_1$ remains to be determined.
We guess that $a > 0$.

To determine $a$ we first find $\psi$ for $x > x^*$ and then require $\psi$ to be $C^1$ at $x = x^*$.
By (ii) and (6.2.24) we know that for $x > x^*$ we have

$$\psi(x) = \mathcal{M}\psi(x) := \inf\{\psi(x - \zeta) + c + \lambda \zeta \; ; \; \zeta > 0\}. \tag{6.2.31}$$

The first order condition for a minimum $\hat{\zeta} = \hat{\zeta}(x)$ of the function

$$G(\zeta) := \psi(x - \zeta) + c + \lambda \zeta \; ; \; \zeta > 0$$

is

$$\psi'(x - \hat{\zeta}) = \lambda.$$

Suppose there is a unique point $\hat{x} \in (0, x^*)$ such that

$$\psi'(\hat{x}) = \lambda. \tag{6.2.32}$$

Then
$$\hat{x} = x - \hat{\zeta}(x) \text{ i.e. } \hat{\zeta}(x) = x - \hat{x}$$

and from (6.2.31) we deduce that
$$\psi(x) = \psi_0(\hat{x}) + c + \lambda(x - \hat{x}) \text{ for } x \geq x^*.$$

In particular,
$$\psi'(x^*) = \lambda \tag{6.2.33}$$

and
$$\psi(x^*) = \psi_0(\hat{x}) + c + \lambda(x^* - \hat{x}). \tag{6.2.34}$$

To summarize we put
$$\psi(x) = \begin{cases} \frac{1}{\rho}x^2 + \frac{b}{\rho^2} - ae^{r_1 x} \text{ for } x \leq x^* \\ \psi_0(\hat{x}) + c + \lambda(x - \hat{x}) \text{ for } x > x^* \end{cases} \tag{6.2.35}$$

where $\hat{x}, x^*$ and $a$ are determined by (6.2.32), (6.2.33) and (6.2.34), i.e.

$$ar_1 e^{r_1 \hat{x}} = \frac{2}{\rho}\hat{x} - \lambda \qquad (\text{i.e. } \psi'(\hat{x}) = \lambda) \tag{6.2.36}$$

$$ar_1 e^{r_1 x^*} = \frac{2}{\rho}x^* - \lambda \qquad (\text{i.e. } \psi'(x^*) = \lambda) \tag{6.2.37}$$

$$ae^{r_1 x^*} - ae^{r_1 \hat{x}} = \frac{1}{\rho}((x^*)^2 - (\hat{x})^2) - c - \lambda(x^* - \hat{x}). \tag{6.2.38}$$

One can now prove (see [Ø2], Theorem 2.5) :

For each $c > 0$ there exists $a = a^*(c) > 0$, $\hat{x} = \hat{x}(c) > 0$ and $x^* = x^*(c) > \hat{x}$ such that (6.2.36)-(6.2.38) hold. With this choice of $a$, $\hat{x}$, $x^*$, the function $\varphi(s, x) = e^{-\rho s}\psi(x)$ with $\psi$ given by (6.2.35) coincides with the value function $\Phi$ defined in (6.2.22). Moreover, the optimal impulse control $v^* = (\tau_1^*, \tau_2^*, \ldots ; \zeta_1^*, \zeta_2^*, \ldots)$ is to do nothing while $X(t) < x^*$, then move $X(t)$ from $x^*$ down to $\hat{x}$ (i.e., apply $\zeta_1^* = x^* - \hat{x}$) at the first time $\tau_1^*$ when $X(t)$ reaches a value $\geq x^*$, then wait until the next time, $\tau_2^*$, $X(t)$ again reaches the value $x^*$ etc.

**Remark 6.6.** In [Ø2] this result is used to study how the value function $\Phi(s, x) = \Phi_c(s, x)$ depends on the fixed part $c > 0$ of the intervention cost. It is proved that the function
$$c \to \Phi_c(s, x)$$

is continuous but not differentiable at $c = 0$. In fact, we have

$$\frac{\partial}{\partial c}\Phi_c(s, x) \to \infty \text{ as } c \to 0^+.$$

Subsequently this high $c$-sensitivity of the value function for $c$ close to 0 was proved for other processes as well. See [ØUZ].

**Remark 6.7.** For applications of impulse control theory in inventory control see e.g. [S], [S2] and the references therein.

## 6.3 Exercices

**Exercise\* 6.1.** Solve the impulse control problem

$$\Phi(s, x) = \inf_{v} J^{(v)}(s, x) = J^{(v^*)}(s, x)$$

where

$$J^{(v)}(s, x) = E\left[\int_0^{\infty} e^{-\rho(s+t)}(X^{(v)}(t))^2 dt + \sum_{k=1}^{N} e^{-\rho(s+\tau_k)}(c + \lambda|\zeta_k|)\right].$$

The inf is taken over all impulse controls $v = (\tau_1, \tau_2, \dots ; \zeta_1, \zeta_2, \dots)$ with $\zeta_i \in \mathbb{R}$ and the corresponding process $X^{(v)}(t)$ is given by

$$X^{(v)}(t) = x + B(t) + \int_0^t \int_{\mathbb{R}} z\tilde{N}(ds, dz) + \sum_{\tau_k \leq t} \zeta_k,$$

where $B(0) = 0$, $x \in \mathbb{R}$, and we assume that the corresponding Lévy measure $\nu$ is *symmetric*, i.e. $\nu(G) = \nu(-G)$ for all $G \subset \mathbb{R}\backslash\{0\}$.

**Exercise\* 6.2 (Optimal stream of dividends with transaction costs from a geometric Lévy process).**
    For $v = (\tau_1, \tau_2, \dots ; \zeta_1, \zeta_2, \dots)$ with $\zeta_i \in \mathbb{R}_+$ define $X^{(v)}(t)$ by

$$
\begin{aligned}
dX^{(v)}(t) &= \mu X^{(v)}(t)dt + \sigma X^{(v)}(t)dB(t) \\
&\quad + \theta X^{(v)}(t^-)\int_{\mathbb{R}} z\tilde{N}(ds, dz) \;;\; \tau_i \leq t \leq \tau_{i+1} \\
X^{(v)}(\tau_{i+1}) &= \check{X}^{(v)}(\tau_{i+1}^-) - (1+\lambda)\zeta_{i+1} - c \;;\; i = 0, 1, 2, \dots \\
X^{(v)}(0^-) &= x > 0
\end{aligned}
$$

where $\mu, \sigma \neq 0$, $\theta, \lambda \geq 0$ and $c > 0$ are constants (see (6.1.7)), $\theta z \geq -1$ a.s. $\nu$.
    Find $\Phi$ and $v^*$ such that

$$\Phi(s, x) = \sup_{v} J^{(v)}(s, x) = J^{(v^*)}(s, x).$$

Here

$$J^{(v)}(s, x) = E^x\left[\sum_{\tau_k < \tau_{\mathcal{S}}} e^{-\rho(s+\tau_k)}\zeta_k\right] \qquad (\rho > 0 \text{ constant})$$

is the expected discounted total dividend up to time $\tau_{\mathcal{S}}$, where

$$\tau_{\mathcal{S}} = \tau_{\mathcal{S}}(\omega) = \inf\{t > 0 \;;\; X^{(v)}(t) \leq 0\}$$

is the time of bankruptcy. (See also Exercise 7.2).

**Exercise\* 6.3 (Optimal forest management (inspired by Y. Willassen [W])).**
Suppose the biomass of a forest at time $t$ is given by

$$X(t) = x + \mu t + \sigma B(t) + \theta \int_{\mathbb{R}} z \widetilde{N}(t, dz) ,$$

where $\mu > 0$, $\sigma > 0$, $\theta > 0$ are constants. At times $0 \leq \tau_1 < \tau_2 < \cdots$ we decide to cut down the forest and replant it, with the cost

$$c + \lambda \check{X}(\tau_k^-), \quad \text{with} \quad \check{X}(\tau_k^-) = X(\tau_k^-) + \Delta_N X(\tau_k) ,$$

where $c > 0$, $\lambda \in [0, 1)$ are constants and $\Delta_N X(t)$ is the (possible) jump in $X$ at $t$ coming from the jump in $N(t, \cdot)$ only, not from the intervention.
Find the sequence of stopping times $v = (\tau_1, \tau_2, \ldots)$ which maximizes the expected total discounted net profit $J^{(v)}(s, x)$ given by

$$J^{(v)}(s, x) = E^x \left[ \sum_{k=1}^{\infty} e^{-\rho(s+\tau_k)} (\check{X}(\tau_k^-) - c - \lambda \check{X}(\tau_k^-)) \right],$$

where $\rho > 0$ is a given discounting exponent.

# 7

# Approximating Impulse Control of Diffusions by Iterated Optimal Stopping

## 7.1 Iterative scheme

In general it is not possible to reduce impulse control to optimal stopping, because the choice of the first intervention time $\tau_1$ and the first impulse $\zeta_1$ will influence the next and so on. However, if we allow only (up to) a fixed finite number $n$ of interventions, then the corresponding impulse control problem can be solved by solving iteratively $n$ optimal stopping problems. Moreover, if we restrict the number of interventions to (at most) $n$ in a given impulse control problem, then the value function of this restricted problem will converge to the value function of the original problem as $n \to \infty$. Thus it is possible to reduce a given impulse control problem to a sequence of iterated optimal stopping problems. This is useful both for theoretical purposes and numerical applications.

We now make this more precise:

Using the notation of Chapter 6 consider the impulse control problem

$$\Phi(y) = \sup\{J^{(v)}(y) \; ; \; v \in \mathcal{V}\} = J^{(v^*)}(y) \; ; \; y \in \mathcal{S} \tag{7.1.1}$$

where, with $\tau_{\mathcal{S}} = \tau_{\mathcal{S}}^{(v)} = \inf\{t > 0 \; ; \; Y^{(v)}(t) \notin \mathcal{S}\}$,

$$J^{(v)}(y)$$

$$= E^y \left[ \int_0^{\tau_{\mathcal{S}}} f(Y^{(v)}(t))dt + g(Y^{(v)}(\tau_{\mathcal{S}}))\chi_{\{\tau_{\mathcal{S}}<\infty\}} + \sum_{\tau_j<\tau_{\mathcal{S}}} K(\check{Y}^{(v)}(\tau_j^-), \zeta_j) \right].$$

$$\tag{7.1.2}$$

Here $\mathcal{V}$ denotes the set of admissible controls $v = (\tau_1, \tau_2, \dots \; ; \; \zeta_1, \zeta_2, \dots)$. See (6.1.9)–(6.1.13).

For $n = 1, 2, \ldots$ let $\mathcal{V}_n$ denote the set of all $v \in \mathcal{V}$ such that $v = (\tau_1, \tau_2, \ldots, \tau_n, \tau_{n+1}; \zeta_1, \zeta_2, \ldots, \zeta_n)$ with $\tau_{n+1} = \tau_{\mathcal{S}}$ a.s. In other words, $\mathcal{V}_n$ is the set of all admissible controls *with at most $n$ interventions*. Then

$$\mathcal{V}_n \subseteq \mathcal{V}_{n+1} \subseteq \mathcal{V} \qquad \text{for all } n . \tag{7.1.3}$$

Define

$$\Phi_n(y) = \sup\{J^{(v)}(y) \; ; \; v \in \mathcal{V}_n\} \; ; \; n = 1, 2, \ldots . \tag{7.1.4}$$

Then $\Phi_n(y) \le \Phi_{n+1}(y) \le \Phi(y)$ because $\mathcal{V}_n \subseteq \mathcal{V}_{n+1} \subseteq \mathcal{V}$. Moreover, we have

**Lemma 7.1.** *Suppose $g \ge 0$. Then*

$$\lim_{n \to \infty} \Phi_n(y) = \Phi(y) \text{ for all } y \in \mathcal{S}.$$

*Proof.* We have already seen that

$$\lim_{n \to \infty} \Phi_n(y) \le \Phi(y).$$

To get the opposite inequality let us first assume $\Phi(y) < \infty$. Then for each $\varepsilon > 0$ there exists $v = (\tau_1, \tau_2, \ldots ; \zeta_1, \zeta_2, \ldots) \in \mathcal{V}$ such that

$$
\begin{aligned}
J^{(v)}(y) = E^y \Bigg[ &\int_0^{\tau_{\mathcal{S}}^{(v)}} f(Y^{(v)}(t))dt + g(Y^{(v)}(\tau_{\mathcal{S}}^{(v)}))\chi_{\{\tau_{\mathcal{S}}^{(v)} < \infty\}} \\
&+ \sum_{\tau_j < \tau_{\mathcal{S}}^{(v)}} K(\check{Y}^{(v)}(\tau_j^-), \zeta_j) \Bigg] \ge \Phi(y) - \varepsilon.
\end{aligned}
\tag{7.1.5}
$$

For $n = 1, 2, \ldots$ define $v_n = (\tau_1, \tau_2, \ldots, \tau_n, \tau_{\mathcal{S}} \; ; \; \zeta_1, \zeta_2, \ldots, \zeta_n)$, i.e. $v_n$ is obtained by truncating the $v$ sequence after $n$ steps. Then

$$Y^{(v_n)}(t) = Y^{(v)}(t) \text{ for all } t \le \tau_n. \tag{7.1.6}$$

Since $\tau_j \to \tau_{\mathcal{S}}$ a.s. when $j \to \infty$, we get by assumptions (6.1.11) and (6.1.12) that there exists $n$ such that

$$E^y \left[ \int_{\tau_n}^{\tau_{\mathcal{S}}^{(v)}} f^-(Y^{(v)}(t))dt + \int_{\tau_n}^{\tau_{\mathcal{S}}^{(v_n)}} f^-(Y^{(v_n)}(t))dt \right] < \varepsilon \tag{7.1.7}$$

and

$$E^y \left[ \sum_{\substack{j > n \\ \tau_j < \tau_{\mathcal{S}}^{(v)}}} K^-(\check{Y}^{(v)}(\tau_j^-), \zeta_j) \right] < \varepsilon. \tag{7.1.8}$$

Moreover, by (7.1.6) we have

$$\chi_{\{\tau_{\mathcal{S}}^{(v)}<\infty\}} \leq \liminf_{n\to\infty} \chi_{\{\tau_{\mathcal{S}}^{(v_n)}<\infty\}}. \tag{7.1.9}$$

Combining (7.1.6)-(7.1.9) we get

$$\liminf_{n\to\infty} J^{(v_n)}(y) = \liminf_{n\to\infty} \left\{ E^y \left[ \int_0^{\tau_n} + \int_{\tau_n}^{\tau_{\mathcal{S}}^{(v_n)}} f(Y^{(v_n)}(t)dt \right] \right.$$

$$\left. + E^y \left[ g(Y^{(v_n)}(\tau_{\mathcal{S}}^{(v_n)}))\chi_{\{\tau_{\mathcal{S}}^{(v_n)}<\infty\}} \right] + E^y \left[ \sum_{j=1}^n K(\check{Y}^{(v_n)}(\tau_j^-),\zeta_j) \right] \right\}$$

$$\geq J^{(v)}(y) - 2\varepsilon + \liminf_{n\to\infty} E^y [g(Y^{(v_n)}(\tau_{\mathcal{S}}^{(v_n)}))\chi_{\{\tau_{\mathcal{S}}^{(v_n)}<\infty\}}$$

$$- g(Y^{(v)}(\tau_{\mathcal{S}}^{(v)}))\chi_{\{\tau_{\mathcal{S}}^{(v)}<\infty\}}]$$

$$\geq J^{(v)}(y) - 2\varepsilon + E^y \left[ \liminf_{n\to\infty} (g(Y^{(v_n)}(\tau_{\mathcal{S}}^{(v_n)})) - g(Y^{(v)}(\tau_{\mathcal{S}}^{(v)})))\chi_{\{\tau_{\mathcal{S}}^{(v_n)}<\infty\}} \right]$$

$$= J^{(v)}(y) - 2\varepsilon.$$

Hence by (7.1.5) $\liminf_{n\to\infty} \Phi_n(y) \geq \liminf_{n\to\infty} J^{(v_n)}(y) \geq \Phi(y) - 3\varepsilon$.

Since $\varepsilon > 0$ was arbitrary this proves Lemma 7.1 in the case when $\Phi(y) < \infty$. If $\Phi(y) = \infty$ the proof is similar, except that now we use that for each $M < \infty$ there exists $v \in \mathcal{V}$ such that $J^{(v)}(y) \geq M$. Choosing $v_n$ as before and using (7.1.6)-(7.1.9) with $\varepsilon = 1$ we get $J^{(v_n)}(y) \geq M - 2$. Since $M$ was arbitrary this shows that

$$\lim_{n\to\infty} \Phi_n(y) \geq \lim_{n\to\infty} J^{(v_n)}(y) = \infty.$$

Let

$$\mathcal{M}h(y) = \sup_{\zeta\in\mathcal{Z}}\{h(\Gamma(y,\zeta)) + K(y,\zeta)\} \ ; \ h \in \mathcal{H}, \ y \in \mathbb{R}^k \tag{7.1.10}$$

be the intervention operator (Definition 6.1). The iterative procedure is the following :

Let $Y(t) = Y^{(0)}(t)$ be the process (6.1.1) without interventions. Define

$$\varphi_0(y) = E^y \left[ \int_0^{\tau_{\mathcal{S}}} f(Y(t))dt + g(Y(\tau_{\mathcal{S}}))\chi_{\{\tau_{\mathcal{S}}<\infty\}} \right] \tag{7.1.11}$$

and inductively, for $j = 1, 2, \ldots, n$,

$$\varphi_j(y) = \sup_{\tau\in\mathcal{T}} E^y \left[ \int_0^\tau f(Y(t))dt + \mathcal{M}\varphi_{j-1}(Y(\tau)) \right], \tag{7.1.12}$$

where, as before $\mathcal{T}$ denotes the set of stopping times $\tau \leq \tau_{\mathcal{S}}$, with

$$\tau_{\mathcal{S}} = \inf\{t > 0 \ ; \ Y(t) \notin \mathcal{S}\}.$$

Let $\mathcal{P}(\mathbb{R}^k)$ denote the set of functions $h : \mathbb{R}^k \to \mathbb{R}$ of at most *polynomial growth* , i.e. with the property that there exists constants $C$ and $m$ (depending on $h$) such that

$$|h(y)| \leq C(1 + |y|^m) \text{ for all } y \in \mathbb{R}^k. \tag{7.1.13}$$

The main result of this chapter is the following :

**Theorem 7.2.** *Suppose*

$$f, g \text{ and } \mathcal{M}\varphi_{j-1} \in \mathcal{P}(\mathbb{R}^k) \text{ for } j = 1, 2, \ldots, n. \tag{7.1.14}$$

*Then*

$$\varphi_n = \Phi_n.$$

To prove this we need a *dynamic programming principle* (or *Bellman* principle). This principle is due to Krylov ([K], Theorem 9 and Theorem 11, p. 134) when there is no jumps. The proof of the dynamic programming principle for jump processes can be found in Ishikawa ([Ish], Section 4).

**Lemma 7.3.** *Suppose $G \in \mathcal{P}(\mathbb{R}^k)$. Define*

$$\psi(y) = \sup_{\tau \in \mathcal{T}} E^y \left[ \int_0^\tau f(Y(s))ds + G(Y(\tau)) \right]. \tag{7.1.15}$$

**a)** *Then for all stopping times $\beta$ we have*

$$\psi(y) = \sup_{\tau \in \mathcal{T}} E^y \left[ \int_0^{\tau \wedge \beta} f(Y(t))dt + G(Y(\tau))\chi_{\{\tau \leq \beta\}} + \psi(Y(\beta))\chi_{\{\tau > \beta\}} \right]. \tag{7.1.16}$$

**b)** *For $\varepsilon > 0$ define*

$$D^{(\varepsilon)} = \{y \in \mathcal{S} ; \psi(y) > G(y) + \varepsilon\}$$

*and put*

$$\tau^{(\varepsilon)} = \inf\{t > 0 ; Y(t) \notin D^{(\varepsilon)}\}.$$

*Then if $\beta$ is a stopping time such that $\beta \leq \tau^{(\varepsilon)}$ for some $\varepsilon > 0$ we have*

$$\psi(y) = E^y \left[ \int_0^\beta f(Y(t))dt + \psi(Y(\beta)) \right]. \tag{7.1.17}$$

**Corollary 7.4. a)** *For each given $\tau \in \mathcal{T}$ the process*

$$U(t) := \int_0^{t \wedge \tau} f(Y(s))ds + \psi(Y(t \wedge \tau)) ; \ t \geq 0$$

*is a* supermartingale. *In particular, if $\tau_1 \leq \tau_2 \leq \tau_{\mathcal{S}}$ are stopping times then*

$$E^y[\psi(Y(\tau_1))] \geq E^y \left[ \int_{\tau_1}^{\tau_2} f(Y(s))ds + \psi(Y(\tau_2)) \right].$$

**b)** *For $\varepsilon > 0$ let $\tau^{(\varepsilon)}$ be as in Lemma 7.3a) and let $\beta_1 \leq \beta_2 \leq \tau^{(\varepsilon)}$ be stopping times. Then*

$$E^y[\psi(Y(\beta_1))] = E^y\left[\int_{\beta_1}^{\beta_2} f(Y(t))dt + \psi(Y(\beta_2))\right].$$

*Proof of Corollary 7.4.*

**a)** Define

$$R(t) = \left(Y(t \wedge \tau), \int_0^{t \wedge \tau} f(Y(r))dr\right) \in \mathbb{R}^{k+1}$$

and put

$$H(y, u) = \psi(y) + u \; ; \; (y, u) \in \mathbb{R}^k \times \mathbb{R}.$$

Then if $t > s$ we have by the Markov property

$$E^y[U(t)|\mathcal{F}_s] = E^y[H(R(t))|\mathcal{F}_s] = E^{R(s)}[H(R(t-s))].$$

Then by (7.1.16) applied to $\beta = (t - s) \wedge \tau$ and $(y, u) = R(s)$

$$E^{R(s)}[H(R(t-s))] = E^y\left[\psi(Y((t-s) \wedge \tau)) + u + \int_0^{(t-s)\wedge\tau} f(Y(r))dr\right]$$

$$= E^y\left[\psi(Y(\beta)) + u + \int_0^\beta f(Y(r))dr\right]$$

$$\leq u + \sup_{\tau \geq \beta} E^y\left[\int_0^{\tau \wedge \beta} f(Y(r))dr + \psi(Y(\tau \wedge \beta))\right]$$

$$\leq u + \psi(y)$$

$$= H(R(s)) = U(s).$$

Hence $E^y[U(t)|\mathcal{F}_s] \geq U(s)$ for all $t > s$. This proves that $U(t)$ is a supermartingale. The second statement follows from Doob's optional sampling theorem.

**b)** By Lemma 7.3b) we have

$$\psi(y) = E^y\left[\int_0^{\beta_1} f(Y(t))dt + \psi(Y(\beta_1))\right]$$

$$+ E^y\left[\int_{\beta_1}^{\beta_2} f(Y(t))dt + \psi(Y(\beta_2)) - \psi(Y(\beta_1))\right]$$

$$= \psi(y) + E^y\left[\int_{\beta_1}^{\beta_2} f(Y(t))dt + \psi(Y(\beta_2)) - \psi(Y(\beta_1))\right],$$

from which b) follows.

*Proof of Theorem 7.2.*

Choose $v_n = (\tau_1, \tau_2, \ldots, \tau_n \; ; \; \zeta_1, \zeta_2, \ldots, \zeta_n)$ with $\tau_n \leq \tau_S$ and set $\tau_{n+1} = \tau_S$. Let $Y(t) = Y^{(v_n)}(t)$. Then by Corollary 7.4a) we have

$$E^y[\varphi_{n-j}(Y(\tau_j))] \geq E^y \left[ \int_{\tau_j}^{\tau_{j+1}} f(Y(t))dt + \varphi_{n-j}(\check{Y}(\tau_{j+1}^-)) \right]. \qquad (7.1.18)$$

Choosing $\tau = 0$ in (7.1.12) we obtain that

$$\varphi_{n-j} \geq \mathcal{M}\varphi_{n-j-1} \text{ if } n - j \geq 1. \qquad (7.1.19)$$

Moreover, from the definition of $\mathcal{M}$ it follows that

$$\mathcal{M}\varphi_{n-j-1}(\check{Y}(\tau_{j+1}^-)) \geq \varphi_{n-j-1}(Y(\tau_{j+1})) + K(\check{Y}(\tau_{j+1}^-), \zeta_{j+1}). \qquad (7.1.20)$$

Combining (7.1.18)-(7.1.20) we get

$$E^y[\varphi_{n-j}(Y(\tau_j))]$$
$$\geq E^y \left[ \int_{\tau_j}^{\tau_{j+1}} f(Y(t))dt + \varphi_{n-j-1}(Y(\tau_{j+1})) + K(\check{Y}(\tau_{j+1}^-), \zeta_{j+1})) \right]$$
$$\qquad (7.1.21)$$

for $j = 0, 1, \ldots, n - 1$, where we have put $\tau_0 = 0$. Summing (7.1.21) from $j = 0$ to $j = n - 1$ we get

$$\sum_{j=0}^{n-1} E^y[\varphi_{n-j}(Y(\tau_j)) - \varphi_{n-j-1}(Y(\tau_{j+1}))]$$
$$\geq E^y \left[ \int_0^{\tau_n} f(Y(t))dt + \sum_{j=1}^n K(\check{Y}(\tau_j^-), \zeta_j) \right]$$

or

$$E^y[\varphi_n(Y(0)) - \varphi_0(Y(\tau_n))] \geq E^y \left[ \int_0^{\tau_n} f(Y(t))dt + \sum_{j=1}^n K(\check{Y}(\tau_j^-), \zeta_j) \right]. \qquad (7.1.22)$$

Now

$$E^y[\varphi_n(Y(0))] = \varphi_n(y) \qquad (7.1.23)$$

and by the strong Markov property

$$E^y[\varphi_0(Y(\tau_n))] = E^y \left[ E^{Y(\tau_n)} \left[ \int_0^{\tau_S} f(Y(t))dt + g(Y(\tau_S))\chi_{\{\tau_S < \infty\}} \right] \right]$$
$$= E^y \left[ \int_{\tau_n}^{\tau_S} f(Y(t))dt + g(Y(\tau_S))\chi_{\{\tau_S < \infty\}} \right]. \qquad (7.1.24)$$

Combining (7.1.22)-(7.1.24) we obtain

$$\varphi_n(y) \geq E^y \left[ \int_0^{\tau_S} f(Y(t))dt + g(Y(\tau_S))\chi_{\{\tau_S < \infty\}} + \sum_{j=1}^{n} K(\check{Y}(\tau_j^-), \zeta_j) \right]$$

$$= J^{(v)}(y).$$

$$(7.1.25)$$

Since $v \in \mathcal{V}_n$ was arbitrary we conclude that

$$\varphi_n(y) \geq \Phi_n(y). \qquad (7.1.26)$$

To get the opposite inequality choose $\varepsilon > 0$ and define an increasing sequence of stopping times $0 = \hat{\tau}_0 < \hat{\tau}_1 < \cdots < \hat{\tau}_n$ as follows :

For $j = 1, 2, \ldots, n$ let

$$D_j^{(\varepsilon)} = \{y \; ; \; \varphi_j(y) > \mathcal{M}\varphi_{j-1}(y) + \varepsilon\}. \qquad (7.1.27)$$

Define

$$\hat{\tau}_1 = \inf\{t > 0 \; ; \; Y^{(0)}(t) \notin D_n^{(\varepsilon)}\}, \qquad (7.1.28)$$

where $Y^{(0)}(t) = Y(t)$ is the process without interventions. Then choose $\hat{\zeta}_1 = \bar{\zeta}_1(Y(\hat{\tau}_1^-))$, where $\bar{\zeta}_1 = \bar{\zeta}_1(y) \in \mathcal{Z}$ is $\varepsilon$-optimal for $\varphi_{n-1}$, in the sense that

$$\mathcal{M}\varphi_{n-1}(y) \leq \varphi_{n-1}(\Gamma(y, \bar{\zeta}_1)) + K(y, \bar{\zeta}_1) + \varepsilon. \qquad (7.1.29)$$

Inductively, if $0 = \hat{\tau}_0, \ldots, \hat{\tau}_j \; ; \; \hat{\zeta}_1, \ldots, \hat{\zeta}_j$ have been chosen, where $j \leq n-1$, we let $Y^{(j)}(t)$ be the process obtained by applying $\hat{v}_j = (\hat{\tau}_1, \ldots, \hat{\tau}_j \; ; \; \hat{\zeta}_1, \ldots, \hat{\zeta}_j)$ to $Y(t)$. Define

$$\hat{\tau}_{j+1} = \inf\{t > \hat{\tau}_j \; ; \; Y^{(j)}(t) \notin D_{n-j}^{(\varepsilon)}\} \qquad (7.1.30)$$

and choose $\bar{\zeta}_{j+1}(Y(\hat{\tau}_{j+1}^-))$, where $\bar{\zeta}_{j+1} = \bar{\zeta}_{j+1}(y) \in \mathcal{Z}$ is $\varepsilon$-optimal for $\varphi_{n-j-1}$, in the sense that

$$\mathcal{M}\varphi_{n-j-1}(y) \leq \varphi_{n-j-1}(\Gamma(y, \bar{\zeta}_{j+1})) + K(y, \bar{\zeta}_{j+1}) + \varepsilon. \qquad (7.1.31)$$

Finally put $\hat{\tau}_{n+1} = \tau_S$ and define

$$\hat{v} = (\hat{\tau}_1, \ldots \hat{\tau}_n \; ; \; \hat{\zeta}_1, \ldots, \hat{\zeta}_n) \in \mathcal{V}_n.$$

Now apply the argument (7.1.18)–(7.1.25) to $\hat{v}$ :

By Corollary 7.4 b) we have

$$E^y[\varphi_{n-j}(Y(\hat{\tau}_j))] = E^y \left[ \int_{\hat{\tau}_j}^{\hat{\tau}_{j+1}} f(Y(t))dt + \varphi_{n-j}(\check{Y}(\hat{\tau}_{j+1}^-)) \right]. \qquad (7.1.32)$$

Since $\check{Y}(\hat{\tau}_{j+1}^-) \notin D_{n-j}^{(\varepsilon)}$ we deduce that

$$\varphi_{n-j}(\check{Y}(\hat{\tau}_{j+1}^-)) \leq \mathcal{M}\varphi_{n-j-1}(\check{Y}(\hat{\tau}_{j+1}^-)) + \varepsilon \tag{7.1.33}$$

and by (7.1.31), we get, with $Y = Y^{(\hat{v})}$,

$$\mathcal{M}\varphi_{n-j-1}(\check{Y}(\hat{\tau}_{j+1}^-)) \leq \varphi_{n-j-1}(Y(\hat{\tau}_{j+1})) + K(\check{Y}(\hat{\tau}_{j+1}^-), \hat{\zeta}_j) + \varepsilon. \tag{7.1.34}$$

Combining (7.1.32)-(7.1.34) we obtain

$$
E^y[\varphi_{n-j}(Y(\hat{\tau}))]
$$
$$
\leq E^y \left[ \int_{\hat{\tau}_j}^{\hat{\tau}_{j+1}} f(Y(t))dt + \varphi_{n-j-1}(Y(\hat{\tau}_{j+1})) + K(\check{Y}(\hat{\tau}_{j+1}^-), \hat{\zeta}_{j+1}) \right] + 2\varepsilon.
$$
$$\tag{7.1.35}$$

Now sum (7.1.35) from $j = 0$ to $j = n - 1$. The result is

$$
\varphi_n(y) \leq E^y \left[ \int_0^{\hat{\tau}_n} f(Y(t))dt + \varphi_0(Y(\hat{\tau}_n)) + \sum_{j=1}^n K(\check{Y}(\hat{\tau}_j^-), \hat{\zeta}_j) \right] + 2n\varepsilon.
$$

Therefore by (7.1.24),

$$
\varphi_n(y) \leq E^y \left[ \int_0^{\hat{\tau}_S} f(Y(t))dt + g(Y(\hat{\tau}_S))\chi_{\{\hat{\tau}_S < \infty\}} + \sum_{j=1}^n K(\check{Y}(\hat{\tau}_j^-), \hat{\zeta}_j) \right] + 2n\varepsilon
$$
$$
= J^{(\hat{v})}(y) + 2n\varepsilon.
$$

Since $\varepsilon$ was arbitrary we deduce that

$$\varphi_n(y) \leq \sup\{J^{(v)}(y) \; ; \; v \in \mathcal{V}_n\} = \Phi_n(y).$$

Combined with (7.1.26) this proves Theorem 7.2.

**Remark 7.5.** Note that the proof of Theorem 7.2 actually also gives a $2n\varepsilon$-optimal impulse control $\hat{v} = (\hat{\tau}_1, \ldots, \hat{\tau}_n \; ; \; \hat{\zeta}_1, \ldots, \hat{\zeta}_n)$ : It is defined inductively by (7.1.27)–(7.1.31).

In particular, if it is possible to choose $\hat{\zeta}_j = \zeta_j^*$ to be optimal (i.e. (7.1.31) holds with $\varepsilon = 0$), then $v^* = (\hat{\tau}_1, \ldots, \hat{\tau}_n \; ; \; \zeta_1^*, \ldots, \zeta_n^*)$ will be an *optimal impulse* control for $\Phi_n$ given by the following procedure :

For $j = 1, 2, \ldots, n$ let

$$D_j = \{y \; ; \; \varphi_j(y) > \mathcal{M}\varphi_{j-1}(y)\}. \tag{7.1.36}$$

Define

$$\hat{\tau}_1 = \inf\{t > 0 \; ; \; Y^{(0)}(t) \notin D_n\} \tag{7.1.37}$$

and

$$\hat{\zeta}_1 = \bar{\zeta}_1(Y^{(0)}(\hat{\tau}_1^-)) \tag{7.1.38}$$

where $\bar{\zeta}_1 = \bar{\zeta}_1(y)$ is a Borel measurable function such that

$$\mathcal{M}\varphi_{n-1}(y) = \varphi_{n-1}(\Gamma(y, \bar{\zeta}_1)) + K(y, \bar{\zeta}_1) \; ; \; y \in \mathcal{S}. \tag{7.1.39}$$

Then if $(\hat{\tau}_1, \ldots, \hat{\tau}_j \; ; \; \hat{\zeta}_1, \ldots, \hat{\zeta}_j)$ is defined, put

$$\hat{\tau}_{j+1} = \inf\{t > \hat{\tau}_j \; ; \; Y^{(j)}(t) \notin D_{n-j}\} \tag{7.1.40}$$

and

$$\hat{\zeta}_{j+1} = \bar{\zeta}_{j+1}(Y^{(j)}(\hat{\tau}_{j+1}^-)) \tag{7.1.41}$$

where $\bar{\zeta}_{j+1} = \bar{\zeta}_{j+1}(y)$ is a Borel measurable function such that

$$\mathcal{M}\varphi_{n-(j+1)}(y) = \varphi_{n-(j+1)}(\Gamma(y, \bar{\zeta}_{j+1})) + K(y, \bar{\zeta}_{j+1}) \; ; \; y \in \mathcal{S}, \; j+1 \le n. \tag{7.1.42}$$

As before $Y^{(j)}(t)$ denotes the result of applying the impulse control $\hat{v}_j = (\hat{\tau}_1, \ldots, \hat{\tau}_j \; ; \; \hat{\zeta}_1, \ldots, \hat{\zeta}_j)$ to $Y$.

**Corollary 7.6.** *Assume that $g \ge 0$ and that $f, g, \mathcal{M}\varphi_n \in \mathcal{P}(\mathbb{R}^k)$ for $n = 0, 1, 2, \ldots$, where $\varphi_n$ is as defined in (7.1.11)-(7.1.12). Then*

$$\varphi_n(y) \uparrow \Phi(y) \text{ as } n \to \infty, \text{ for all } y.$$

**Corollary 7.7.** *Suppose $g \ge 0$ and $f, g, \mathcal{M}\varphi_n \in \mathcal{P}(\mathbb{R}^k)$ for $n = 0, 1, 2, \ldots$. Then $\Phi$ is a solution of the following non-linear optimal stopping problem*

$$\Phi(y) = \sup_{\tau \in \mathcal{T}} E^y \left[ \int_0^\tau f(Y(t)) dt + \mathcal{M}\Phi(Y(\tau)) \right] \tag{7.1.43}$$

*Proof.* Let $\{\varphi_n\}_{n=0}^\infty$ be as in (7.1.11)-(7.1.12). Then $\varphi_n \uparrow \Phi$ as $n \to \infty$, by Corollary 7.6. Therefore

$$\varphi_n(y) \le \sup_{\tau \in \mathcal{T}} E^y \left[ \int_0^\tau f(Y(t)) dt + \mathcal{M}\Phi(Y(\tau)) \right]$$

for all $n$ and hence

$$\Phi(y) \le \sup_{\tau \in \mathcal{T}} E^y \left[ \int_0^T f(Y(t)) dt + \mathcal{M}\Phi(Y(\tau)) \right].$$

To get the opposite inequality, choose $\varepsilon > 0$ and let $\hat{\tau} \in \mathcal{T}$ be a stopping time such that

$$E^y \left[ \int_0^{\hat{\tau}} f(Y(t)) dt + \mathcal{M}\Phi(Y(\hat{\tau})) \right] \ge \sup_{\tau \in \mathcal{T}} E^y \left[ \int_0^\tau f(Y(t)) dt + \mathcal{M}\Phi(Y(\tau)) \right] - \varepsilon.$$

Then by monotone convergence

$$\Phi(y) = \lim_{n\to\infty} \varphi_n(y) \geq \lim_{n\to\infty} E^y\left[\int_0^{\hat\tau} f(Y(t))dt + \mathcal{M}\varphi_{n-1}(Y(\hat\tau))\right]$$

$$= E^y\left[\int_0^{\hat\tau} f(Y(t))dt + \mathcal{M}\Phi(Y(\hat\tau))\right] \tag{7.1.44}$$

$$\geq \sup_{\tau\in\mathcal{T}} E^y\left[\int_0^{\tau} f(Y(t))dt + \mathcal{M}\Phi(Y(\tau))\right] - \varepsilon.$$

**Corollary 7.8.** *Suppose $g \geq 0$ and $f, g, \mathcal{M}\varphi_n \in \mathcal{P}(\mathbb{R}^k)$ for all $n = 0, 1, 2, \ldots$. Morover, suppose that*

$$\Phi, \mathcal{M}\Phi, \varphi_n \text{ and } \mathcal{M}\varphi_n \text{ are continuous for all } n, \tag{7.1.45}$$

*where $\Phi, \varphi_n$ are defined in (7.1.1) and (7.1.11)-(7.1.12), respectively. Define*

$$D = \{y \;;\; \Phi(y) > \mathcal{M}\Phi(y)\} \tag{7.1.46}$$

*and*

$$D_n = \{y \;;\; \varphi_n(y) > \mathcal{M}\varphi_{n-1}(y)\} \;;\; n = 1, 2, \ldots. \tag{7.1.47}$$

*Then*

$$D \subseteq D_{n+1} \subseteq D_n \text{ for all } n. \tag{7.1.48}$$

*Proof.* Suppose there exists a point $y \in D_{n+1}\backslash D_n$. Then

$$\varphi_{n+1}(y) > \mathcal{M}\varphi_n(y)$$

and

$$\varphi_n(y) = \mathcal{M}\varphi_{n-1}(y).$$

Then by Lemma 7.3a) we have

$$\mathcal{M}\varphi_n(y) < \varphi_{n+1}(y) = \sup_{\tau\in\mathcal{T}} E^y\left[\int_0^T f(Y(t))dt + \mathcal{M}\varphi_n(Y(\tau))\right]$$

$$\leq \sup_{\tau\in\mathcal{T}} E^y\left[\int_0^T f(Y(t))dt + \varphi_n(Y(\tau))\right] = \varphi_n(y) = \mathcal{M}\varphi_{n-1}(y).$$

This is a contradiction, because $\mathcal{M}\varphi_n \geq \mathcal{M}\varphi_{n-1}$. This contradiction shows that $D_{n+1} \subseteq D_n$ for all $n$. A similar argument, based on Corollary 7.7, shows that $D \subseteq D_n$ for all $n$. □

Combining the results above we get the following general picture of the optimal impulse control $(\hat\tau_1, \ldots, \hat\tau_n \;;\; \hat\zeta_1, \ldots, \hat\zeta_n)$ for $\Phi_n$ (see (7.1.11)-(7.1.12) and (7.1.36)-(7.1.42)):

Make the first intervention the first time $t = \hat\tau_1$ that $Y(t) \notin D_n$. Then give the system the impulse $\hat\zeta_1$ according to (7.1.38). Now we have only $n-1$ interventions left, so we wait until $Y(t)$ exits from the larger set $D_{n-1}$ before making the next intervention, and so on. The last intervention time $\hat\tau_n$ is the first time after $\hat\tau_{n-1}$ that $Y(t) \notin D_1$. See Figure 7.1.
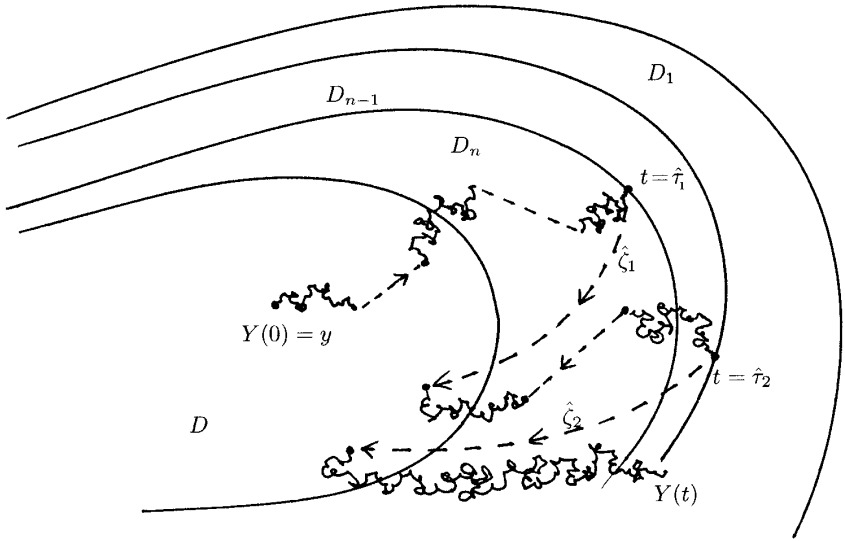
**Fig. 7.1.** Optimal impulse control for $\Phi_n$

## 7.2 Examples

**Example 7.9.** Consider the impulse control problems

$$\Phi_n(s,x) = \inf_{v \in \mathcal{V}_n} J^{(v)}(s,x) = J^{(v_n^*)}(s,x) \; ; \; n = 1,2,\ldots \qquad (7.2.1)$$

$$\Phi(s,x) = \inf_{v \in \mathcal{V}} J^{(v)}(s,x) = J^{(v^*)}(s,x) \qquad (7.2.2)$$

where

$$J^{(v)}(s,x) = E^x \left[ \int_0^\infty e^{-\rho(s+t)} (X^{(v)}(t))^2 dt + c \sum_j e^{-\rho(s+\tau_j)} \right],$$

$c > 0$, $\rho > 0$ are constants, and

$$X^{(v)}(t) = x + B(t) + \int_0^t \int_R z \tilde{N}(ds,dz) + \sum_j \zeta_j \chi_{\{\tau_j \leq t\}} \; ; \; B(0) = 0,$$

when

$$v = (\tau_1, \tau_2, \ldots, \tau_n \; ; \; \zeta_1, \zeta_2, \ldots, \zeta_n) \in \mathcal{V}_n$$

or

$$v = (\tau_1, \tau_2, \ldots \; ; \; \zeta_1, \zeta_2, \ldots) \in \mathcal{V}, \; \zeta_j \in \mathbb{R}.$$

This is related to Example 6.5, except that now we allow the impulses $\zeta_j$ to be arbitrary real numbers, so that $X(t)$ can be moved both up and down. Moreover, to simplify matters we have put $\lambda = 0$.

First, let us find $\Psi_n(x) := \Phi_n(0, x)$ by using the iterative procedure (7.1.11)–(7.1.12):

In this case we get from (7.1.11) (see (6.2.29))

$$\Psi_0(x) = E^x \left[ \int_0^\infty e^{-\rho t} (x + B(t) + \int_0^t \int_R z\tilde{N}(ds, dz))^2 dt \right] = \frac{1}{\rho} x^2 + \frac{b}{\rho^2},$$
(7.2.3)

where $b = 1 + \int_{\mathbb{R}} z^2 \nu(dz)$. Hence

$$\mathcal{M}\Psi_0(x) = \inf\{\Psi_0(x + \zeta) + c \; ; \; \zeta \in \mathbb{R}\backslash\{0\}\} = \Psi(0) + c = \frac{b}{\rho^2} + c. \quad (7.2.4)$$

Therefore, by (7.1.12)

$$\Psi_1(x) = \inf_{\tau \geq 0} E^x \left[ \int_0^\tau e^{-\rho t} (x + B(t) + \int_0^t \int_R z\tilde{N}(ds, dz))^2 dt + e^{-\rho\tau} \left( \frac{b}{\rho^2} + c \right) \right].$$
(7.2.5)

To solve this optimal stopping problem we consider the three basic associated variational inequalities

$$-\rho\psi_1(x) + \frac{1}{2}\psi_1''(x) + \int_{\mathbb{R}} \{\psi_1(x + z) - \psi_1(x) - z\psi_1'(x)\}\nu(dz) + x^2 \geq 0 \text{ for all } x$$
(7.2.6)

$$\psi_1(x) \leq \frac{b}{\rho^2} + c \text{ for all } x \quad (7.2.7)$$

$$-\rho\psi_1(x) + \frac{1}{2}\psi_1''(x) + \int_{\mathbb{R}} \{\psi_1(x + z) - \psi_1(x) - z\psi_1'(x)\}\nu(dz) + x^2 = 0$$

$$\text{on} \qquad D_1 := \left\{ x \; ; \; \psi_1(x) < \frac{b}{\rho^2} + c \right\}.$$

If we guess that $D_1$ has the form

$$D_1 = (-x_1^*, x_1^*) \text{ for some } x_1^* > 0$$

then we are led to the following candidate for $\psi_1$:

$$\psi_1(x) = \begin{cases} \frac{1}{\rho} x^2 + \frac{b}{\rho^2} - a_1 \cosh(r_1 x) \; ; \; |x| < x_1^* \\ \frac{b}{\rho^2} + c \; ; \; |x| \geq x_1^*, \end{cases}$$
(7.2.8)

where $a_1$ is a constant to determine and $r_1 > 0$ is a root of the equation
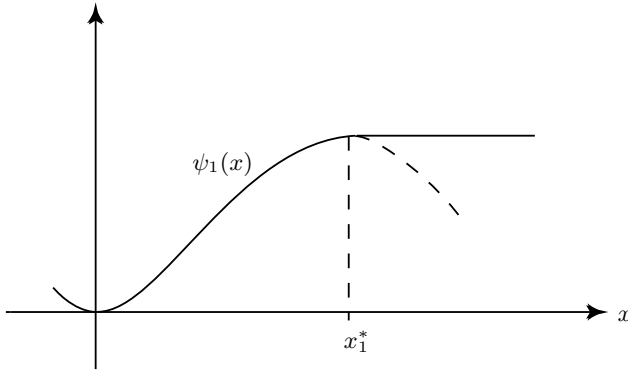
**Fig. 7.2.** The graph of $\psi_1(x)$

$$K(r) := \int_{\mathbb{R}} \{e^{rz} - 1 - rz\}\nu(dz) + \frac{1}{2}r^2 - \rho = 0.$$

(see Example 6.5).

If we require $\psi_1$ to be continuous and differentiable at $x = \pm x_1^*$, we get the following two equations to determine $x_1^*$ and $a_1$:

$$\frac{1}{\rho}(x_1^*)^2 + \frac{b}{\rho^2} - a_1 \cosh(r_1 x_1^*) = \frac{b}{\rho^2} + c \tag{7.2.9}$$

$$\frac{2}{\rho}x_1^* - a_1 r_1 \sinh(r_1 x_1^*) = 0. \tag{7.2.10}$$

Combining these two equations we get

$$\mathrm{tgh}(z_1^*) = \frac{2z_1^*}{(z_1^*)^2 - \rho r_1^2 c} \tag{7.2.11}$$

and

$$a_1 = \frac{2z_1^*}{\rho r_1^2 \sinh(z_1^*)}, \tag{7.2.12}$$

where

$$z_1^* = r_1 x_1^*. \tag{7.2.13}$$

It is easy to see that (7.2.11) has a unique solution $z_1^* > r_1 \sqrt{\rho c}$ i.e. $x_1^* > \sqrt{\rho c}$. If we choose this value of $x_1^*$ and let $a_1 > 0$ be the corresponding value given by (7.2.12) we can now verify that the candidate $\psi_1$ given by (7.2.8) satisfies all the conditions of the verification theorem and we conclude that

$$\psi_1 = \Psi_1.$$

We now repeat the procedure to find $\Psi_2$:

First note that

$$\mathcal{M}\Psi_1(x) = \inf\{\Psi_1(x + \zeta) + c \; ; \; \zeta \in \mathbb{R}\backslash\{0\}\} = \Psi_1(0) + c$$
$$= \frac{b}{\rho^2} + c - a_1. \tag{7.2.14}$$

Hence

$$\Psi_2(x) = \inf_{\tau \geq 0} E^x \left[ \int_0^\tau e^{-\rho t} \left( x + B(t) + \int_0^t \int_R z\tilde{N}(ds, dz) \right)^2 dt \right.$$
$$\left. + e^{-\rho\tau} \left( \frac{b}{\rho^2} + c - a_1 \right) \right].$$

The same procedure as above leads us to the candidate

$$\psi_2(x) = \begin{cases} \frac{1}{\rho}x^2 + \frac{b}{\rho^2} - a_2 \cosh(r_1 x) \; ; \; |x| < x_2^* \\ \frac{b}{\rho^2} + c - a_1 \; ; \; |x| \geq x_2^* \end{cases} \tag{7.2.15}$$

where $x_2^*, a_2$ solve the equations

$$\text{tgh}(z_2^*) = \frac{2z_2^*}{(z_2^*)^2 - \rho r_1^2(c - a_1)} \tag{7.2.16}$$

$$a_2 = \frac{2z_2^*}{\rho r_1^2 \sinh(z_2^*)}, \qquad \text{where } z_2^* = r_1 x_2^*. \tag{7.2.17}$$

As above, we see that (7.2.16) has a unique solution $z_2^* > 0$ i.e. $x_2^* > 0$. If we choose this value of $x_2^*$ and let $a_2 > 0$ be the corresponding value given by (7.2.17) we can verify that the candidate $\psi_2$ given by (7.2.15) coincides with $\Psi_2$. It is easy to see that $x_2^* < x_1^*$ and $a_2 > a_1$.

Continuing this inductively we get a sequence of functions $\psi_n$ of the form

$$\psi_n(x) = \begin{cases} \frac{1}{\rho}x^2 + \frac{b}{\rho^2} - a_n \cosh(r_1 x) \; ; \; |x| < x_n^* \\ \frac{b}{\rho^2} + c - a_{n-1} \; ; \qquad\qquad |x| \geq x_n^* \end{cases} \tag{7.2.18}$$

where $x_n^*, a_n$ solve the equations

$$\text{tgh}(z_n^*) = \frac{2z_n^*}{(z_n^*)^2 - \rho r_1^2(c - a_{n-1})} \tag{7.2.19}$$

and

$$a_n = \frac{2z_n^*}{\rho r_1^2 \sinh(z_n^*)}, \quad \text{where } z_n^* = r_1 x_n^*. \tag{7.2.20}$$

We find that $x_n^* < x_{n-1}^*$ and $a_n > a_{n-1}$. Using the verification theorem we conclude that $\psi_n = \Psi_n$. In the limiting case when there is no bound on the number of interventions the corresponding value function will be

$$\Psi(x) = \begin{cases} \frac{1}{\rho}x^2 + \frac{b}{\rho^2} - a\cosh(\sqrt{2\rho}x) \; ; \; |x| < x^* \\ \frac{b}{\rho^2} + c - a \; ; \; |x| \geq x^* \end{cases} \qquad (7.2.21)$$

where $x^*, a$ solve the coupled system of equations

$$\text{tgh}(z^*) = \frac{2z^*}{(z^*)^2 - \rho r_1^2(c-a)} \qquad (7.2.22)$$

$$a = \frac{2z^*}{\rho r_1^2 \sinh(z^*)}, \quad \text{where } z^* = r_1 x^*. \qquad (7.2.23)$$

It can be proved that this system has a unique solution $z^* = r_1 x^* > 0$, $a > 0$ and that $x^* < x_n^*$ and $a > a_n$ for all $n$. The situation in the case $n = 3$ is shown on Figure 7.3. Note that with only 3 interventions allowed the optimal strategy is first to wait until the first time $\tau_1^*$ when $X(t) \geq x_3^*$, then move $X(t)$ down to 0, next wait until the first time $\tau_2^* > \tau_1^*$ when $X(t) \geq x_2^*$, then move $X(t)$ back to 0 and finally wait until the first time $t = \tau_3^* > \tau_2^*$ when $X(t) \geq x_1^*$ before making the last intervention.
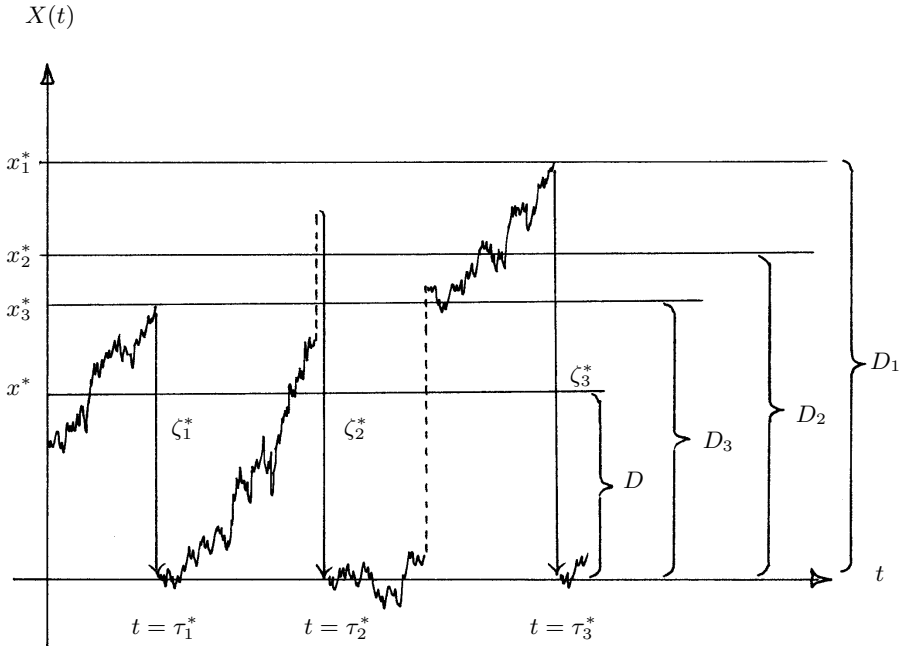


**Fig. 7.3.** The optimal impulse control for $\Psi_3$ (Example 7.9)

## 7.3 Exercices

**Exercise\* 7.1 (Optimal forest management revisited).** Using the notation of Exercise 6.3 let

$$\Phi(x) = \sup \left\{ J^{(v)}(x); v = (\tau_1, \tau_2, \ldots) \right\}$$

be the value function when there are no restrictions on the number of interventions. For $n = 1, 2, \ldots$ let

$$\Phi_n(x) = \sup \left\{ J^{(v)}(x); v = (\tau_1, \tau_2, \ldots, \tau_n) \right\}$$

be the value function when up to $n$ interventions are allowed. Use Theorem 7.2 to find $\Phi_1$ and $\Phi_2$.

**Exercise\* 7.2 (Optimal stream of dividends with transaction costs from a geometric Lévy process).** This is an addition to Exercise 6.2 Suppose that we at times $0 \leq \tau_1 < \tau_2 < \ldots$ decide to take out dividends of sizes $\zeta_1, \zeta_2, \ldots \in (0, \infty)$ from an economic quantity growing like a geometric Lévy process. If we let $X^{(v)}(t)$ denote the size at time $t$ of this quantity when the dividend policy $v = (\tau_1, \tau_2, \ldots; \zeta_1, \zeta_2, \ldots)$ is applied, we assume that $X^{(v)}(t)$ is described by (see (6.2.4))

$$dX^{(v)}(t) =$$

$$\mu X^{(v)}(t)dt + \sigma X^{(v)}(t)dB(t) + \theta X^{(v)}(t^-) \int_{\mathbb{R}} z \tilde{N}(dt, dz) \, ; \, \tau_i \leq t < \tau_{i+1}$$

$$X^{(v)}(\tau_{i+1}) = \hat{X}^{(v)}(\tau_{i+1}^-) - (1+\lambda)\zeta_{i+1} - c \, ; \qquad i = 0, 1, 2, \ldots \quad \text{(see (6.1.7))}$$

where $\mu, \sigma \neq 0$, $\theta, \lambda \geq 0$ and $c > 0$ are constants, $\theta z \geq -1$, a.s. $(\nu)$. Let

$$\Phi(x) = \sup \left\{ J^{(v)}(x); v = (\tau_1, \tau_2, \ldots; \zeta_1, \zeta_2, \ldots) \right\}$$

and

$$\Phi_n(x) = \sup \left\{ J^{(v)}(x); v = (\tau_1, \tau_2, \ldots, \tau_n; \zeta_1, \zeta_2, \ldots, \zeta_n) \right\},$$

be the value function with no restrictions on the number of interventions and with at most $n$ interventions, respectively, where

$$J^{(v)}(x) = E^x \left[ \sum_{\tau_k < \tau_{\mathcal{S}}} e^{-\rho \tau_k} \zeta_k \right] \qquad (\rho > 0 \text{ constant})$$

is the expected total discounted dividend and

$$\tau_{\mathcal{S}} = \inf \left\{ t > 0; X^{(v)}(t) \leq 0 \right\}$$

is the time of bankruptcy. Show that

$$\Phi(x) = \Phi_n(x) = \Phi_1(x) \qquad \text{for all } n.$$

Thus in this case we achieve the optimal result with just one intervention.

# 8

# Combined Stochastic Control
# and Impulse Control of Jump Diffusions

## 8.1 A verification theorem

Consider the general situation in Chapter 6, except that now we assume that we, in addition, are free at any state $y \in \mathbb{R}^k$ to choose a Markov control $u(y) \in U$, where $U$ is a given closed convex set in $\mathbb{R}^\ell$. If, as before, $v = (\tau_1, \tau_2, \ldots; \zeta_1, \zeta_2, \ldots) \in \mathcal{V}$ denotes a given impulse control we call $w := (u, v)$ a *combined control*. If $w = (u, v)$ is applied, we assume that the corresponding state $Y(t) = Y^{(w)}(t)$ at time $t$ is given by (see (3.1.1) and (6.1.2)-(6.1.7))

$$dY(t) = b(Y(t), u(t)dt + \sigma(Y(t), u(t))dB(t)$$
$$+ \int_{\mathbb{R}} \gamma(Y(t^-), u(t^-), z)\tilde{N}(dt, dz) ; \qquad \tau_j < t < \tau_{j+1} \qquad (8.1.1)$$

$$Y(\tau_{j+1}) = \Gamma(\check{Y}(\tau_{j+1}^-), \zeta_{j+1}) ; \qquad j = 0, 1, 2, \ldots \qquad (8.1.2)$$

where $u(t) = u(Y(t))$ and $b : \mathbb{R}^k \times U \to \mathbb{R}^k$, $\sigma : \mathbb{R}^k \times U \to \mathbb{R}^{k \times d}$ and $\gamma : \mathbb{R}^k \times U \times \mathbb{R}^k \to \mathbb{R}^{k \times \ell}$ are given continuous functions, $\tau_0 = 0$.

As before we let our "universe" $\mathcal{S}$ be a fixed Borel set in $\mathbb{R}^k$ such that $\mathcal{S} \subset \overline{\mathcal{S}^0}$ and we define

$$\tau^* = \lim_{R \to \infty} \inf\{t > 0; |Y^{(w)}(t)| \geq R\} \leq \infty \qquad (8.1.3)$$

and

$$\tau_{\mathcal{S}} = \inf\{t \in (0, T^*(\omega)); Y^{(w)}(t, \omega) \notin \mathcal{S}\} . \qquad (8.1.4)$$

If $Y^{(w)}(t, \omega) \in \mathcal{S}$ for all $t < \tau^*$ we put $\tau_{\mathcal{S}} = \tau^*$.

We assume that we are given a set $\mathcal{W}$ of *admissible* combined controls $w = (u, v)$ which includes the combined controls $w = (u, v)$ such that a unique strong solution $Y^{(w)}(t)$ of (8.1.1), (8.1.2) exists and

$$\tau^* = \infty \text{ and } \lim_{j \to \infty} \tau_j = \tau_{\mathcal{S}} \text{ a.s. } Q^y \text{ for all } y \in \mathbb{R}^\varepsilon. \qquad (8.1.5)$$

Define the *performance* or *total expected profit/utility* of $w = (u,v) \in \mathcal{W}$, $v = (\tau_1, \tau_2, \ldots; \zeta_1, \zeta_2, \ldots)$, by

$$J^{(w)}(y) = E^y \Big[ \int_0^{\tau_S} f(Y^{(w)}(t), u(t))dt + g(Y^{(w)}(\tau_S))\mathcal{X}_{\{\tau_S < \infty\}}$$

$$+ \sum_{\tau_j \leq \tau_S} K(\check{Y}^{(w)}(\tau_j^-), \zeta_j) \Big] \qquad (8.1.6)$$

where $f : \mathcal{S} \times U \to \mathbb{R}$, $g : \mathbb{R}^k \to \mathbb{R}$ and $K : \bar{\mathcal{S}} \times \mathcal{Z} \to \mathbb{R}$ are given functions statisfying the conditions similar to (6.1.11)–(6.1.13).

The *combined stochastic control and impulse control problem* is the following:

Find the value function $\Phi(y)$ and an optimal control $w^* = (u^*, v^*) \in \mathcal{W}$ such that

$$\Phi(y) = \sup\{J^{(w)}(y); w \in \mathcal{W}\} = J^{(w^*)}(y). \qquad (8.1.7)$$

We now state a verification theorem for this problem. It is a combination of the HJB equation of control theory and the QVI for impulse control:

Define

$$L^\alpha h(y) = \sum_{i=1}^k b_i(y, \alpha) \frac{\partial h}{\partial y_i} + \frac{1}{2} \sum_{i,j=1}^k (\sigma\sigma^T)_{ij}(y, \alpha) \frac{\partial^2 h}{\partial y_i \partial y_j}$$

$$+ \int_{\mathbb{R}} \sum_{j=1}^\ell \{h(y + \gamma^{(j)}(y, \alpha, z)) - h(y) - \nabla h(y).\gamma^{(j)}(y, \alpha, z)\}\nu_j(dz_j) \qquad (8.1.8)$$

for each $\alpha \in U$ and for each twice differentiable function $h$. This is the generator of $Y^{(w)}(t)$ if we apply the constant control $\alpha$ and no (impulse) interventions.

As in Chapter 6 we let

$$\mathcal{M}h(y) = \sup\{h(\Gamma(y, \zeta)) + K(y, \zeta); \zeta \in \mathcal{Z} \quad \text{and} \quad \Gamma(y, \zeta) \in \mathcal{S}\} \qquad (8.1.9)$$

be the intervention operator.

Then the verification theorem is the following (compare with Theorem 3.1 and Theorem 6.2):

**Theorem 8.1 (HJBQVI verification theorem for combined stochastic and impulse control).**
**a)** *Suppose we can find a function $\phi : \bar{\mathcal{S}} \to \mathbb{R}$ such that*
  (i)  $\phi \in C^1(\mathcal{S}^0) \cap C(\bar{\mathcal{S}})$
  (ii)  $\phi \geq \mathcal{M}\phi$ *on $\mathcal{S}$.*
*Define*

$$D = \{y \in \mathcal{S}; \phi(y) > \mathcal{M}\phi(y)\} \quad \text{(the continuation region)}. \qquad (8.1.10)$$

*Suppose that $Y^{(w)}(t)$ spends 0 time on $\partial D$ a.s., i.e.*

(iii)   $E^y \left[ \int_0^{\tau_{\mathcal{S}}} \mathcal{X}_{\partial D}(Y^{(w)}(t)) dt \right] = 0$  *for all $y \in \mathcal{S}$, $w \in \mathcal{W}$ and suppose that*

(iv)   $\partial D$ *is a Lipschitz surface*

(v)   $\phi \in C^2(S^0 \backslash \partial D)$ *and the second order derivatives of $\phi$ are locally bounded near $\partial D$*

(vi)   $L^\alpha \phi(y) + f(y, \alpha) \leq 0$ *for all $\alpha \in U$, $y \in S^0 \setminus \partial D$*

(vii)   $Y(\tau_{\mathcal{S}}) \in \partial \mathcal{S}$ *a.s. on $\{\tau_{\mathcal{S}} < \infty\}$ and*
$\phi(Y^{(w)}(t)) \to g(Y^{(w)}(\tau_{\mathcal{S}})) \cdot \mathcal{X}_{\{\tau_{\mathcal{S}} < \infty\}}$ *as $t \to \tau_{\mathcal{S}}^-$ a.s. $Q^y$, for all $y \in \mathcal{S}$, $w \in \mathcal{W}$*

(viii)   *the family $\{\phi^-(Y^{(w)}(\tau)); \tau \in \mathcal{T}\}$ is uniformly $Q^y$-integrable for all $y \in \mathcal{S}$, $w \in \mathcal{W}$*

(ix) $E^y \Big[ |\phi(Y(\tau))| + \int_0^{\tau_{\mathcal{S}}} \Big\{ |A\phi(Y(t))| + |\sigma^T(Y(t))\nabla\phi(Y(t))|^2$

$+ \sum_{j=1}^{\ell} \int_{\mathbb{R}} |\phi(Y(t) + \gamma^{(j)}(Y(t), z_j)) - \phi(Y(t))|^2 \nu_j(dz_j) \Big\} dt < \infty$ *for all $\tau \in \mathcal{T}$, $w \in \mathcal{W}$, $y \in \mathcal{S}$.*

*Then*
$$\phi(y) \geq \Phi(y) \qquad \text{for all } y \in \mathcal{S} \ .$$

**b)** *Suppose in addition that*

(x) *there exists a function $\hat{u} : D \to \mathbb{R}$ such that*
$$L^{\hat{u}(y)}\phi(y) + f(y, \hat{u}(y)) = 0 \qquad \text{for all } y \in D$$

   *and*

(xi)
$$\hat{\zeta}(y) \in \text{Argmax}\{\phi(\Gamma(y, \cdot)) + K(y, \cdot)\}$$

   *exists for all $y \in \mathcal{S}$ and $\hat{\zeta}(\cdot)$ is a Borel measurable selection.*

*Define an impulse control $\hat{v} = (\hat{\tau}_1, \hat{\tau}_2, \ldots; \hat{\zeta}_1, \hat{\zeta}_2, \ldots)$ as follows:*
   *Put $\tau_0 = 0$ and inductively*

$$\hat{\tau}_{k+1} = \inf\{t > \hat{\tau}_k; Y^{(k)}(t) \in D\} \wedge T$$

$$\hat{\zeta}_{k+1} = \hat{\zeta}(Y^{(k)}(\hat{\tau}_{k+1}^-)) \qquad \text{if } \hat{\tau}_{k+1} < T; \ k = 0, 1, \ldots$$

*where $Y^{(k)}(t)$ is the result of applying the combined control*

$$\widehat{w}_k := (\hat{u}, (\hat{\tau}_1, \ldots, \hat{\tau}_k; \hat{\zeta}_1, \ldots, \hat{\zeta})) \ .$$

*Put $\widehat{w} = (\hat{u}, \hat{v})$. Suppose $\widehat{w} \in \mathcal{W}$ and that*

(xii) $\{\phi(Y^{(\widehat{w})}(\tau)); \tau \in \mathcal{T}\}$  *is $Q^y$-uniformly integrable for all $y \in \mathcal{S}$.*

*Then*
$$\phi(y) = \Phi(y) \qquad \text{for all } y \in \mathcal{S}$$

*and*

$$\widehat{w} \in \mathcal{W} \quad \text{is an optimal combined control.}$$

*Proof.* The proof is similar to the proof of Theorem 6.2 and is omitted.     □

## 8.2 Examples

**Example 8.2 (Optimal combined control of the exchange rate).** This example is a simplification of a model studied in [MØ].

Suppose a government has two means of influencing the foreign exchange rate of its own currency:

(i) at all times the government can choose the domestic interest rate $r$
(ii) at times selected by the government it can intervene in the foreign exchange market by buying or selling large amounts of foreign currency.

Let $r(t)$ denote the interest rate chosen and let $\tau_1, \tau_2, \ldots$ be the (stopping) times when it is decided to intervene, with corresponding amounts $\zeta_1, \zeta_2, \ldots$ If $\zeta > 0$ the government buys foreign currency, if $\zeta < 0$ it sells. Let $v = (\tau_1, \tau_2, \ldots; \zeta_1, \zeta_2, \ldots)$ be corresponding impulse control.

If the combined control $w = (r, v)$ is applied, we assume that the corresponding exchange rate $X(t)$ (measured in the number of domestic monetary units it takes to buy one average foreign monetary unit) is given by

$$X(t) = x - \int_0^t F(r(s) - \bar{r}(s))ds + \sigma B(t) + \sum_{j:\tau_j \leq t} \gamma(\zeta_j) ; \qquad t \geq 0 \quad (8.2.1)$$

where $\sigma > 0$ is a constant, $\bar{r}(s)$ is the (average) foreign interest rate and $F : \mathbb{R} \to \mathbb{R}$ and $\gamma : \mathbb{R} \to \mathbb{R}$ are known functions which give the effects on the exchange rate by the interest rate differential $r(s) - \bar{r}(s)$ and the amount $\zeta_j$, respectively.

The total expected cost of applying the combined control $w = (r, v)$ is assumed to be of the form

$$J^{(w)}(s, x) = E^x \left[ \int_s^T e^{-\rho t} \{M(X(t) - m) + N(r(t) - \bar{r}(t))\}dt + \sum_{j;\tau_j \leq T} L(\zeta_j)e^{-\rho\tau_j} \right]$$
$$(8.2.2)$$

where $M(X(t) - m)$ and $N(r(t) - \bar{r}(t))$ give the costs incurred by the difference $X(t) - m$ between $X(t)$ and an optimal value $m$ and by the difference $r(t) - \bar{r}(t)$ between the domestic and the average foreign interest rate $\bar{r}(t)$, respectively. The cost of buying/selling the amount $\zeta_j$ is $L(\zeta_j)$ and $\rho > 0$ is a constant discounting exponent.

The problem is to find $\Phi(s, x)$ and $w^* = (r^*, v^*)$ such that

$$\Phi(s, x) = \inf_w J^{(w)}(s, x) = J^{(w^*)}(s, x) . \qquad (8.2.3)$$

Since this is a minimum problem, the corresponding HJBQVIs of Theorem 8.1 are changed to minima also and they get the form

$$\min \left( \inf_{r \in \mathbb{R}} \left\{ e^{-\rho s}(M(x - m) + N(r - \bar{r}(s)) + \frac{\partial \phi}{\partial s} - F(r - \bar{r}(s))\frac{\partial \phi}{\partial x} \right.\right.$$
$$\left.\left. + \tfrac{1}{2}\sigma^2 \frac{\partial^2 \phi}{\partial x^2} \right\}, \mathcal{M}\phi(s, x) - \phi(s, x) \right) = 0 \qquad (8.2.4)$$

where

$$\mathcal{M}\phi(s, x) = \inf_{\zeta \in \mathbb{R}} \left\{ \phi(s, x + \gamma(\zeta)) + e^{-\rho s} L(\zeta) \right\} . \qquad (8.2.5)$$

In general this is difficult to solve for $\phi$, even for simple choices of the functions $M, N$ and $F$. A detailed discussion on a special case can be found in [MØ].

**Example 8.3 (Optimal consumption and portfolio with both fixed and proportional transaction costs (1)).** This application is studied in [ØS].

Suppose there are two investment possibilities, say a bank account and a stock. Let $X_1(t), X_2(t)$ denote the amount of money invested in these two assets, respectively, at time $t$. In the absence of consumption and transactions suppose that

$$dX_1(t) = rX_1(t)dt \qquad (8.2.6)$$

and

$$dX_2(t) = \mu X_2(t)dt + \sigma X_2(t)dB(t) \qquad (8.2.7)$$

where $r, \mu$ and $\sigma \neq 0$ are constants and

$$\mu > r > 0 . \qquad (8.2.8)$$

Suppose that at any time $t$ the investor is free to choose a *consumption rate* $u(t) \geq 0$. This consumption is automatically drawn from the bank account holding with no extra costs. In addition the investor may at any time transfer money from the bank to the stock and conversely. Suppose that such a transaction of size $\zeta$ incurs a transaction cost given by

$$c + \lambda|\zeta| \qquad (8.2.9)$$

where $c > 0$ an $\lambda \geq 0$ are constants. (If $\zeta > 0$ we buy stocks and if $\zeta < 0$ we sell stocks.) Thus the control of the investor consists of a combination of a stochastic control $u(t)$ and an impulse control $v = (\tau_1, \tau_2, \ldots; \zeta_1, \zeta_2, \ldots)$, where $\tau_1, \tau_2, \ldots$ are the chosen transaction times and $\zeta_1, \zeta_2, \ldots$ are corresponding transaction amounts.

If such a combined control $w = (u, v)$ is applied, the corresponding system $(X_1(t), X_2(t)) = (X_1^{(w)}(t), X_2^{(w)}(t))$ gets the form

$$dX_1(t) = (rX_1(t) - u(t))dt ; \qquad \tau_i < t < \tau_{i+1} \qquad (8.2.10)$$
$$dX_2(t) = \mu X_2(t)dt + \sigma X_2(t)dB(t) ; \qquad \tau_i < t < \tau_{i+1} \qquad (8.2.11)$$
$$X_1(\tau_{i+1}) = X_1(\tau_{i+1}^-) - \zeta_{i+1} - c - \lambda|\zeta_{i+1}| \qquad (8.2.12)$$
$$X_2(\tau_{i+1}) = X_2(\tau_{i+1}^-) + \zeta_{i+1}. \qquad (8.2.13)$$

If we do not allow any negative amounts held in the bank account or in the stock, the *solvency region* $\mathcal{S}$ is given by

$$\mathcal{S} = [0, \infty) \times [0, \infty) . \qquad (8.2.14)$$

We call $w = (u, v)$ *admissible* if $(X_1^{(w)}(t), X_2^{(w)}(t)) \in \mathcal{S}$ for all $t$. The set of admissible controls is denoted by $\mathcal{W}$.

The investor's objective is to maximize

$$J^{(w)}(y) = E^y \left[ \int_0^\infty e^{-\delta(s+t)} \frac{u^\gamma(t)}{\gamma} dt \right] \qquad (8.2.15)$$

where $\delta > 0$, $\gamma \in (0, 1)$ are constants and $E^y$ with $y = (s, x_1, x_2)$ denotes the expectation when $X_1(0^-) = x_1 \geq 0$, $X_2(0^-) = x_2 \geq 0$. Thus we seek the value function $\Phi(y)$ and an optimal control $w^* = (u^*, v^*) \in \mathcal{W}$ such that

$$\Phi(y) = \sup_{w \in \mathcal{W}} J^{(w)}(y) = J^{(w^*)}(y) . \qquad (8.2.16)$$

This problem may be regarded as a generalization of optimal consumption and portfolio problems studied by Merton [M] and Davis and Norman [DN]. See also Shreve and Soner [SS]. [M] considers the case with no transaction costs ($c = \lambda = 0$). The problem reduces then to an ordinary stochastic control problem and it is it is optimal to keep the positions $(X_1(t), X_2(t))$ on the line $y = \frac{\pi^*}{1-\pi^*} x$ in the $(x, y)$-plane at all times (the Merton line), where $\pi^* = \frac{\mu - r}{(1-\gamma)\sigma^2}$ (see Example 3.2).

[DN] and [SS] consider the case when the cost is proportional ($\lambda > 0$), with no fixed component ($c = 0$). In this case the problem can be formulated as a singular stochastic control problem and under some conditions it is proved that there exists a *no-transaction cone NT* bounded by two straight lines $\Gamma_1, \Gamma_2$ such that it is optimal to make no transactions if $(X_1(t), X_2(t)) \in NT$ and make transactions corresponding to local time at $\partial(NT)$, resulting in reflections back to $NT$ every time $(X_1(t), X_2(t)) \in \partial(NT)$. See Figure 8.1. These results have subsequently been extended to jump diffusion markets by [FØS2]. (See Example 5.1).

In the general combined control case numerical results indicate (see [CØS]) that the optimal control $w^* = (u^*, v^*)$ has the following form:
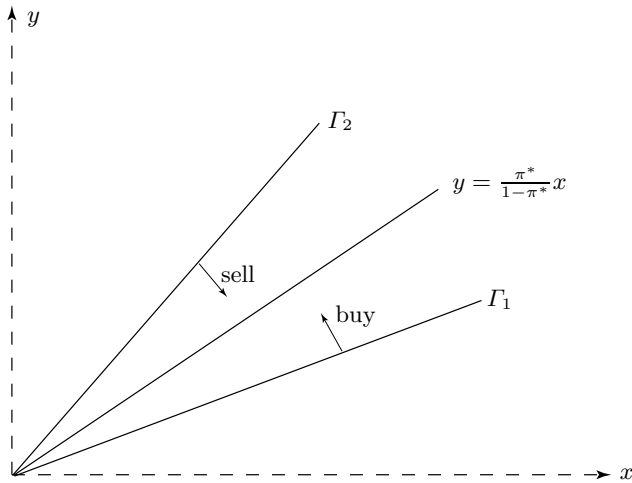
**Fig. 8.1.** The no-transaction cone (no fixed cost: $c = 0$)

There exist two pairs of lines, $(\Gamma_1, \hat{\Gamma}_1)$ and $(\Gamma_2, \hat{\Gamma}_2)$ from the origin such that the following is optimal: Make no transactions (only consume at the rate $u^*(t)$) while $(X_1(t), X_2(t))$ belongs to the region $D$ bounded by the outer curves $\Gamma_1, \Gamma_2$, and if $(X_1(t), X_2(t))$ hits $\partial D = \Gamma_1 \cup \Gamma_2$ then sell or buy so as to bring $(X_1(t), X_2(t))$ to the curve $\hat{\Gamma}_1$ or $\hat{\Gamma}_2$. See Figure 8.2.
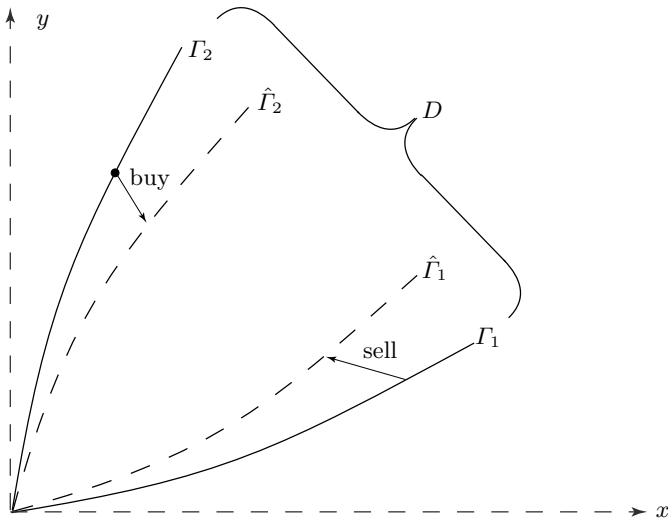


**Fig. 8.2.** The no-transaction region $D$ $(c > 0)$

Note that if we *sell* stocks $(\zeta < 0)$ then $(X_1(t), X_2(t)) = (x_1, x_2)$ moves to a point $(x_1', x_2') = (x_1'(\zeta), x_2'(\zeta))$ on the line

$$x_1' + (1-\lambda)x_2' = x_1 + (1-\lambda)x_2 - c \ . \tag{8.2.17}$$

Similarly, if we *buy* stocks $(\zeta > 0)$ then the new position $(x_1', x_2') = (x_1'(\zeta), x_2'(\zeta))$ is on the line

$$x_1' + (1+\lambda)x_2' = x_1 + (1+\lambda)x_2 - c \ . \tag{8.2.18}$$

If there are no interventions then the process

$$Y(t) = \begin{bmatrix} s+t \\ X_1(t) \\ X_2(t) \end{bmatrix} \tag{8.2.19}$$

has the generator

$$L^u \phi(s, x_1, x_2) = \frac{\partial \phi}{\partial s} + (rx_1 - u)\frac{\partial \phi}{\partial x_1} + \mu x_2 \frac{\partial \phi}{\partial x_2} + \tfrac{1}{2}\sigma^2 x_2^2 \frac{\partial^2 \phi}{\partial x_2^2} \ . \tag{8.2.20}$$

Therefore, if we put $\phi(s, x_1, x_2) = e^{-\delta s}\psi(x_1, x_2)$ the corresponding HJBQVI is

$$\max\left(\sup_{u \geq 0}\left\{\frac{u^\gamma}{\gamma} - \rho\psi(x_1, x_2) + (rx_1 - u)\frac{\partial \psi}{\partial x_1} + \mu x_2 \frac{\partial \psi}{\partial x_2} + \tfrac{1}{2}\sigma^2 x_2^2 \frac{\partial^2 \psi}{\partial x_2^2}\right\},\right.$$

$$\left. \psi(x_1, x_2) - \mathcal{M}\psi(x_1, x_2)\right) = 0 \quad \text{for all } (x_1, x_2) \in \mathcal{S}, \tag{8.2.21}$$

where (see (8.2.17)–(8.2.18))

$$\mathcal{M}\psi(x_1, x_2) = \sup\{\psi(x_1'(\zeta), x_2'(\zeta)); \zeta \in \mathbb{R}\backslash\{0\}, (x_1'(\zeta), x_2'(\zeta)) \in \mathcal{S}\} \ . \tag{8.2.22}$$

See Example 9.12 for a further discussion of this.

## 8.3 Iterative methods

In Chapter 7 we saw that an impulse control problem can be regarded as a limit of iterated optimal stopping problems. A similar result holds for combined control problems. More precisely, a combined stochastic control and impulse control problem can be regarded as a limit of iterated combined stochastic control and optimal stopping problems.

We now describe this in more detail. The presentation is similar to the approach in Chapter 7.

For $n = 1, 2, \ldots$ let $\mathcal{W}_n$ denote the set of all admissible combined controls $w = (u, v) \in \mathcal{W}$ with $v \in \mathcal{V}_n$, where $\mathcal{V}_n$ is the set of impulse controls $v =$

$(\tau_1, \ldots, \tau_n, \tau_{n+1}; \zeta_1, \zeta_2, \ldots, \zeta_n)$ with *at most $n$ interventions* (i.e. $\tau_{n+1} = \infty$). Then

$$\mathcal{W}_n \subseteq \mathcal{W}_{n+1} \subseteq \mathcal{W} \qquad \text{for all } n. \tag{8.3.1}$$

Define, with $J^{(w)}(y)$ as in (8.1.6),

$$\Phi_n(y) = \sup\left\{J^{(w)}(y); w \in \mathcal{W}_n\right\}; \qquad n = 1, 2, \ldots \tag{8.3.2}$$

Then

$$\Phi_n(y) \leq \Phi_{n+1}(y) \leq \Phi(y) \qquad \text{because } \mathcal{W}_n \subseteq \mathcal{W}_{n+1} \subseteq \mathcal{W}.$$

Moreover, we have

**Lemma 8.4.** *Suppose $g \geq 0$. Then*

$$\lim_{n \to \infty} \Phi_n(y) = \Phi(y) \qquad \text{for all } y \in \mathcal{S}.$$

*Proof.* The proof is similar to the proof of Lemma 7.1 and is omitted. $\quad\square$

The iterative procedure is the following:
Let $Y(t) = Y^{(u,0)}(t)$ be the process in (8.1.1) obtained by using the control $u$ and no interventions. Define

$$\phi_0(y) = \sup_{u \in \mathcal{U}} E^y\left[\int_0^{\tau_{\mathcal{S}}} f(Y(t), u(t))dt + g(Y(\tau_{\mathcal{S}}))\mathcal{X}_{\{\tau_{\mathcal{S}} < \infty\}}\right] \tag{8.3.3}$$

and inductively, for $j = 1, 2, \ldots, n$,

$$\phi_j(y) = \sup_{u \in \mathcal{U}, \tau \in \mathcal{T}} E^y\left[\int_0^\tau f(Y(t), u(t))dt + \mathcal{M}\phi_{j-1}(Y(\tau))\right]. \tag{8.3.4}$$

As in (7.1.13) we let $\mathcal{P}(\mathbb{R}^k)$ denote the set of functions $h : \mathbb{R}^k \to \mathbb{R}$ with at most polynomial growth. Then we have, as in Chapter 7,

**Theorem 8.5.** *Suppose*

$$f, g \quad \text{and} \quad \mathcal{M}\phi_{j-1} \in \mathcal{P}(\mathbb{R}^k) \tag{8.3.5}$$

*for $j = 1, 2, \ldots, n$. Then*

$$\phi_n = \Phi_n.$$

*Proof.* The proof is basically the same as the proof of Theorem 7.2 and is left to the reader. $\quad\square$

Similarly we obtain combined control versions of the rest of the results of Chapter 7, with obvious modifications. We omit the details.

## 8.4 Exercices

**Exercise\* 8.1.** Let $\Phi(s, X_1, X_2)$ be the value function of the optimal consumption problem (8.2.16) with fixed and proportional transaction costs and let $\Phi_0(s, X_1, X_2) = Ke^{-\delta s}(X_1 + X_2)^\gamma$ be the corresponding value function in the case when there are no transaction costs, i.e. $c = \lambda = 0$. Use Theorem 8.1 a) to prove that

$$\Phi(s, X_1, X_2) \le Ke^{-\delta s}(X_1 + X_2)^\gamma \ .$$

**Exercise 8.2 (A combined impulse linear regulator problem).** (Compare with Exercise 3.5 and Exercise 4.3).
**a)\*** Suppose the state process is

$$Y(t) = \begin{bmatrix} s + t \\ X(t) \end{bmatrix} \ ; \ t \ge 0, \quad Y(0) = \begin{bmatrix} s \\ x \end{bmatrix} = y \in \mathbb{R}^2,$$

with

$$dX(t) = dX^{(w)}(t) \ \text{given by}$$

$$dX(t) = u(t)dt + \sigma dB(t) \ ; \ \tau_i < t < \tau_{i+1}$$
$$X(\tau_{i+1}) = X(\tau_{i+1}^-) + \zeta_{i+1} \ ; \ i = 1, 2, \dots$$

where $w = (u, v)$ is a combined control, $v = (\tau_1, \tau_2 \dots ; \zeta_1, \zeta_2, \dots)$ with $\tau_i \le T$ (constant).

Solve the combined control problem

$$\Phi(y) = \inf_{w \in \mathcal{W}} J^{(w)}(y),$$

where

$$J^{(w)}(y) = E^{(s,x)}\left[ \int_0^{T-s} (X^2(t) + u^2(t))dt - \sum_{\tau_i \le T} c \right],$$

and $c > 0$ is a given constant.

This models the situation where one is trying to keep $X(t)$ close to 0 with a minimum of cost of the two controls, represented by the rate $u^2(t)$ and the intervention cost $c$.

**b)** Extend this to include jumps, i.e. suppose that

$$dX(t) = u(t)dt + \sigma dB(t) + \int_{\mathbb{R}} z \tilde{N}(dt, dz) \ ; \ \tau_i < t < \tau_{i+1}$$
$$X(\tau_{i+1}) = X(\tau_{i+1}^-) + \Delta_N X(\tau_{i+1}) + \zeta_{i+1},$$

and keep $J^{(w)}(y)$ as before.

# 9

# Viscosity Solutions

The main results of Chapters 2, 3, 4, 6, 8 and 5 are all *verification theorems*: Any function $\phi$ which satisfies the given requirements is necessarily the value function $\Phi$ of the corresponding problem. These requirements are made as weak as possible in order to include as many cases as possible. For example, except for the singular control case, we do not require the function $\phi$ to be $C^2$ everywhere (only outside $\partial D$), because except for that case, $\Phi$ will usually not be $C^2$ everywhere. On the other hand, all the above mentioned verification theorems require $\phi$ to be $C^1$ everywhere, because this is often the case for $\Phi$. This $C^1$ assumption on $\Phi$ is usually called the "*hight contact*" – or "*smooth fit*"-principle. As we have seen in the examples and exercises this principle is very convenient, because it provides us with extra information needed to find $\Phi$ and the continuation region $D$.

However, it is important to know that in general $\Phi$ need not be $C^1$. In fact, it need not even be continuous everywhere. Nevertheless, it turns out that $\Phi$ does satisfy the corresponding verification theorems, provided that we interpret these equations in an appropriate weak sense. More precisely, they should be interpreted in the sense of *viscosity solutions*. This weak solution concept was first introduced by Crandall and Lions to handle the HJB equations of stochastic control and later extended by them and others to more general equations. See [CIL], [FS], [BCa] and the references therein.

In case the equation in question is a *linear* partial differential operator, the viscosity solution is the same as the well-known distribution solution. See [Is2]. However, the nice feature of the viscosity solution is that it also applies to the nonlinear equations appearing in control theory.

We now proceed to define viscosity solutions. We will do this in two steps:

First we consider the viscosity solutions of the variational inequalities appearing in the optimal stopping problems of Chapter 2. Then we proceed to discuss more general equations.

## 9.1 Viscosity solutions of variational inequalities

Consider the optimal stopping problem of Chapter 2: The state $Y(t)$ is given by

$$dY(t) = b(Y(t))dt + \sigma(Y(t))dB(t) + \int_{\mathbb{R}^k} \gamma(Y(t^-), z)\bar{N}(dt, dz) ; \quad Y(0) = y \in \mathbb{R}^k$$

(9.1.1)

and the performance criterion is

$$J^\tau(y) = E^y\left[\int_0^\tau f(Y(t))dt + g(Y(\tau))\right]; \qquad \tau \in \mathcal{T} , \qquad (9.1.2)$$

with $f$ and $g$ as in (2.1.2), (2.1.3). The associated variational inequality of the optimal stopping problem

$$\Phi(y) = \sup\{J^\tau(y); \tau \in \mathcal{T}\} \qquad (9.1.3)$$

is (see (4.2.13))

$$\max(L\phi(y) + f(y), g(y) - \phi(y)) = 0 ; \quad y \in \mathcal{S}^0 . \qquad (9.1.4)$$

In addition we have the boundary requirement

$$\phi(y) = g(y) ; \qquad y \in \partial\mathcal{S} \qquad (9.1.5)$$

(see Theorem 2.2).

We will prove that, under some conditions, the function $\phi = \Phi$ is the unique viscosity solution of (9.1.4)–(9.1.5). First we give the definition of a viscosity solution of such equations.

**Definition 9.1.** *Let* $\phi \in C(\bar{\mathcal{S}})$.

(i) *We say that* $\phi$ *is a* viscosity subsolution *of* (9.1.4)–(9.1.5) *if* (9.1.5) *holds and for all* $h \in C^2(\mathbb{R}^k)$ *and all* $y_0 \in \mathcal{S}^0$ *such that*

$$h \geq \phi \quad on\ \mathcal{S} \quad and\ h(y_0) = \phi(y_0)$$

*we have*

$$\max(Lh(y_0) + f(y_0), g(y_0) - \phi(y_0)) \geq 0 . \qquad (9.1.6)$$

(ii) $\phi$ *is a* viscosity supersolution *of* (9.1.4)–(9.1.5) *if* (9.1.5) *holds and for all* $h \in C^2(\mathbb{R}^k)$ *and all* $y_0 \in \mathcal{S}$ *such that*

$$h \leq \phi \quad on\ \mathcal{S} \quad and\ h(y_0) = \phi(y_0)$$

*we have*

$$\max(Lh(y_0) + f(y_0), g(y_0) - \phi(y_0)) \leq 0 . \qquad (9.1.7)$$

*(iii)  $\phi$ is a* viscosity solution *of* (9.1.4)–(9.1.5) *if $\phi$ is both a viscosity subsolution and a viscosity supersolution of* (9.1.4)–(9.1.5).

**Theorem 9.2 ([ØR]).** *Assume that the set $\partial \mathcal{S}$ is* regular *for the process $Y(t)$, i.e.*

$$\tau_{\mathcal{S}^0} = \tau_{\mathcal{S}^0}(y) := \inf\{t > 0; Y(t) \notin \mathcal{S}^0\} = 0 \qquad a.s. \ \ Q^y \ for \ all \ \ y \in \partial\mathcal{S} \ .$$
$$(9.1.8)$$

*Moreover, assume that the value function $\Phi$ of the optimal stopping problem* (9.1.3) *is continuous on $\bar{\mathcal{S}}$. Then $\Phi$ is a viscosity solution of* (9.1.4)–(9.1.5).

*Proof.* First note that (9.1.5) follows directly from (9.1.8). So it remains to consider (9.1.4).

We first prove that $\Phi$ is a subsolution. To this end suppose $h \in C^2(\mathbb{R}^k)$ and $y_0 \in \mathcal{S}$ with $h \geq \Phi$ on $\mathcal{S}$ and $h(y_0) = \Phi(y_0)$. As before let

$$D = \{y \in \mathcal{S} \ ; \ \ \Phi(y) > g(y)\} \ . \qquad (9.1.9)$$

Then if $y_0 \notin D$ we have $\Phi(y_0) = g(y_0)$ and hence (9.1.6) holds trivially.

Next, assume $y_0 \in D$. Then by the dynamic programming principle (Lemma 7.3 b)) we have

$$\Phi(y_0) = E^{y_0}\left[\int_0^\tau f(Y(t))dt + \Phi(Y(\tau))\right] \qquad (9.1.10)$$

for all bounded stopping times $\tau \leq \tau_D = \inf\{t > 0; Y(t) \notin D\}$. Combining this with the Dynkin formula we get

$$\begin{aligned}
\Phi(y_0) &= E^{y_0}\left[\int_0^\tau f(Y(t))dt + \Phi(Y(\tau))\right] \\
&\leq E^{y_0}\left[\int_0^\tau f(Y(t))dt + h(Y(\tau))\right] \\
&= h(y_0) + E^{y_0}\left[\int_0^\tau (Lh(Y(t)) + f(Y(t)))dt\right]
\end{aligned}$$

or

$$E^{y_0}\left[\int_0^\tau (Lh(Y(t)) + f(Y(t)))dt\right] \geq 0 \ . \qquad (9.1.11)$$

If we divide (9.1.11) by $E^{y_0}[\tau]$ and let $\tau \to 0$ we get, by continuity,

$$Lh(y_0) + f(y_0) \geq 0 \ .$$

Hence (9.1.6) holds and we have proved that $\Phi$ is a viscosity subsolution.

Finally we show that $\Phi$ is a viscosity supersolution. So we assume that $h \in C^2(\mathbb{R}^k)$ and $y_0 \in \mathcal{S}$ are such that $h \leq \Phi$ on $\mathcal{S}$ and $h(y_0) = \Phi(y_0)$. Then by the dynamic programming principle (Lemma 7.3 a)) we have

$$\Phi(y_0) \geq E^{y_0}\left[\int_0^\tau f(Y(t))dt + \Phi(Y(\tau))\right] \qquad (9.1.12)$$

for all stopping times $\tau \leq \tau_{\mathcal{S}^0}$. Hence, by the Dynkin formula

$$\Phi(y_0) \geq E^{y_0}\left[\int_0^\tau f(Y(t))dt + \phi(Y(\tau))\right] \geq E^{y_0}\left[\int_0^\tau f(Y(t))dt + h(Y(\tau))\right]$$

$$= h(y_0) + E^{y_0}\left[\int_0^\tau (Lh(Y(t))dt + f(Y(t)))dt\right],$$

for all bounded stopping times $\tau \leq \tau_{\mathcal{S}^0}$. Hence

$$E^{y_0}\left[\int_0^\tau (Lh(Y(t)) + f(Y(t)))dt\right] \leq 0$$

and by dividing by $E^{y_0}[\tau]$ and letting $\tau \to 0$ we get

$$Lh(y_0) + f(y_0) \leq 0 .$$

Hence (9.1.7) holds and we have proved that $\Phi$ is also a viscosity supersolution.
□

## Uniqueness

One important application of the viscosity solution concept is that it can be used as a verification method: In order to verify that a given function $\phi$ is indeed the value function $\Phi$ it suffices to verify that the function is a viscosity solution of the corresponding variational inequality. For this method to work, however, it is necessary that we know that $\Phi$ is the *unique* viscosity solution. Therefore the question of uniqueness is crucial.

In general we need not have uniqueness. The following simple example illustrates this:

**Example 9.3.** Let $Y(t) = B(t) \in \mathbb{R}$ and choose $f = 0$, $\mathcal{S} = \mathbb{R}$ and

$$g(y) = \frac{y^2}{1+y^2} ; \qquad x \in \mathbb{R} . \qquad (9.1.13)$$

Then the value function $\Phi$ of the optimal stopping problem

$$\Phi(y) = \sup_{\tau \in \mathcal{T}} E^y[g(B(\tau))] \qquad (9.1.14)$$

is easily seen to be $\Phi(y) \equiv 1$. The corresponding VI is

$$\max\left(\tfrac{1}{2}\phi''(y), g(y) - \phi(y)\right) = 1 \qquad (9.1.15)$$

and this equation is trivially satisfied by all constant functions

$$\phi(y) \equiv a$$

for any $a \geq 1$.

**Theorem 9.4 (Uniqueness).** *Suppose that*

$$\tau_{\mathcal{S}^0} < \infty \qquad a.s. \ \ P^y \ for \ all \ \ y \in \mathcal{S}^0 \ . \tag{9.1.16}$$

*Let $\phi \in C(\bar{\mathcal{S}})$ be a viscosity solution of* (9.1.4)–(9.1.5) *with the property that*

$$\begin{aligned} &\text{the family } \{\phi(Y(\tau)); \tau \text{ stopping time}, \ \tau \leq \tau_{\mathcal{S}^0}\} \\ &\quad \text{is } P^y\text{-uniformly integrable, for all } y \in \mathcal{S}^0. \end{aligned} \tag{9.1.17}$$

*Then*

$$\phi(y) = \Phi(y) \qquad for \ all \ \ y \in \bar{\mathcal{S}} \ . $$

*Proof.* We refer the reader to [ØR] for the proof in the case where there are no jumps. $\qquad\qquad\square$

## 9.2 The value function is not always $\mathcal{C}^1$

**Example 9.5.** We now give an example of an optimal stopping problem where the value function $\Phi$ is not $C^1$ everywhere. In this case Theorem 2.2 cannot be used to find $\Phi$. However, we can use Theorem 9.4. The example is taken from [ØR]:
Define

$$k(x) = \begin{cases} 1 & \text{for} \quad x \leq 0 \\ 1 - cx & \text{for} \quad 0 < x < a \\ 1 - ca & \text{for} \quad x \geq a \end{cases} \tag{9.2.1}$$

where $c$ and $a$ are constants to be specified more closely later. Consider the optimal stopping problem

$$\Phi(s, x) = \sup_{\tau \in \mathcal{T}} E^{(s,x)}\left[e^{-\rho(s+\tau)}k(B(\tau))\right] \tag{9.2.2}$$

where $B(t)$ is 1-dimensional Brownian motion, $B(0) = x \in \mathbb{R} = \mathcal{S}$, and $\rho > 0$ is a constant. The corresponding variational inequality is (see (9.1.4))

$$\max\left(\frac{\partial \phi}{\partial s} + \tfrac{1}{2}\frac{\partial^2 \phi}{\partial x^2}, e^{-\rho s}k(x) - \phi(s, x)\right) = 0 \ . \tag{9.2.3}$$

If we try a solution of the form

$$\phi(s, x) = e^{-\rho s}\psi(x) \tag{9.2.4}$$

for some function $\psi$, then (9.2.3) becomes

$$\max(-\rho\psi(x) + \tfrac{1}{2}\psi''(x), k(x) - \psi(x)) = 0 \ . \tag{9.2.5}$$

Let us guess that the continuation region $D$ has the form

$$D = \{(s, x); 0 < x < x_1\} \qquad (9.2.6)$$

for some $x_1 > a$. Then (9.2.5) can be split into the 3 equations

$$\begin{aligned}
-\rho\psi(x) + \tfrac{1}{2}\psi''(x) = 0 & \qquad ; \qquad 0 < x < x_1 \\
\psi(x) = 1 & \qquad ; \qquad x \le 0 \\
\psi(x) = 1 - ca & \qquad ; \qquad x \ge x_1
\end{aligned} \qquad (9.2.7)$$

The general solution of (9.2.7) is

$$\psi(x) = C_1 e^{\sqrt{2\rho}\, x} + C_2 e^{-\sqrt{2\rho}\, x} \; ; \qquad 0 < x < x_1$$

where $C_1, C_2$ are arbitrary constants. If we require $\psi$ to be continuous at $x = 0$ and at $x = x_1$ we get the two equations

$$C_1 + C_2 = 1 \qquad (9.2.8)$$

$$C_1 e^{\sqrt{2\rho}\, x_1} + C_2 e^{-\sqrt{2\rho}\, x_1} = 1 - ca \qquad (9.2.9)$$

in the 3 unknowns $C_1, C_2$ and $x_1$. If we also guess that $\psi$ will be $C^1$ at $x = x_1$ we get the third equation

$$C_1 \sqrt{2\rho}\, e^{\sqrt{2\rho}\, x_1} - C_2 \sqrt{2\rho}\, e^{-\sqrt{2\rho}\, x_1} = 0 . \qquad (9.2.10)$$

If we assume that

$$ca < 1 \qquad \text{and} \qquad \sqrt{2\rho} < \frac{1}{a} \ln \left( \frac{1 - ca}{1 - \sqrt{ca(2 - ca)}} \right) \qquad (9.2.11)$$

then the 3 equations (9.2.8), (9.2.9) and (9.2.10) have the unique solution

$$C_1 = \tfrac{1}{2}\left(1 - \sqrt{ca(2 - ca)}\,\right) > 0 , \qquad C_2 = 1 - C_1 > 0 \qquad (9.2.12)$$

and

$$x_1 = \frac{1}{\sqrt{2\rho}} \ln \left( \frac{1 - ca}{2C_1} \right) > a . \qquad (9.2.13)$$

With these values of $C_1, C_2$ and $x_1$ we put

$$\psi(x) = \begin{cases} 1 & \text{if } x \le 0 \\ C_1 e^{\sqrt{2\rho}\, x} + C_2 e^{-\sqrt{2\rho}\, x} & \text{if } 0 < x < x_1 \\ 1 - ca & \text{if } x_1 \le x \end{cases} \qquad (9.2.14)$$

We claim that $\psi$ is a viscosity solution of (9.2.5).

(i) First we verify that $\psi$ is a viscosity *subsolution*: let $h \in C^2(\mathbb{R})$, $h \ge \psi$ and $h(x_0) = \psi(x_0)$. Then if $x_0 \le 0$ or $x_0 \ge x_1$ we have $k(x_0) - \psi(x_0) = 0$. And if $0 < x_0 < x_1$ then $h - \psi$ is $C^2$ at $x = x_0$ and has a local minimum at $x_0$ so
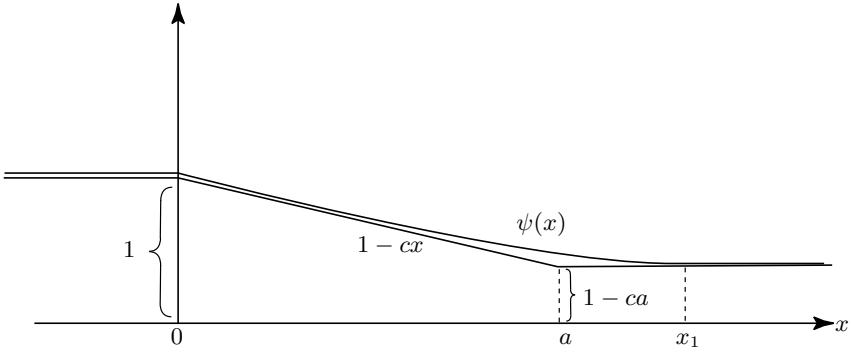
**Fig. 9.1.** The function $\psi$

$$h''(x_0) - \psi''(x_0) \geq 0 \ .$$

Therefore

$$-\rho h(x_0) + \tfrac{1}{2}h''(x_0) \geq -\rho\psi(x_0) + \tfrac{1}{2}\psi''(x_0) = 0 \ .$$

This proves that

$$\max\big(-\rho h(x_0) + \tfrac{1}{2}h''(x_0), k(x_0) - \psi(x_0)\big) \geq 0 \ ,$$

so $\psi$ is a viscosity subsolution of (9.2.5).

(ii) Second, we prove that $\psi$ is a viscosity *super*solution. So let $h \in C^2(\mathbb{R})$, $h \leq \psi$ and $h(x_0) = \psi(x_0)$. Note that we always have

$$k(x_0) - \psi(x_0) \leq 0$$

so in order to prove that

$$\max\big(-\rho h(x_0) + \tfrac{1}{2}h''(x_0), k(x_0) - \psi(x_0)\big) \leq 0$$

it suffices to prove that

$$-\rho h(x_0) + \tfrac{1}{2}h''(x_0) \leq 0 \ .$$

At any point $x_0$ where $\psi$ is $C^2$ this follows in the same way as in (i) above. So it remains only to consider the two cases $x_0 = 0$ and $x_0 = x_1$: If $x_0 = 0$ then no such $h$ exists, so the conclusion trivially holds.

If $x_0 = x_1$ then the function $h - \psi$ has a local maximum at $x = x_0$ and it is $C^2$ to the left of $x_0$ so

$$\lim_{x \to x_0^-} h''(x) - \psi''(x) \leq 0$$

i.e.    $h''(x_0) - \psi''(x_0^-) \leq 0 \ .$

This gives

$$-\rho h(x_0) + \tfrac{1}{2}h''(x_0) \leq -\rho\psi(x_0) + \tfrac{1}{2}\psi''(x_0^-) = 0 \;,$$

and the proof is complete.

We have proved:

Suppose (9.2.11) holds. Then the value function $\Phi(s,x)$ of problem (9.2.2) is given by

$$\Phi(s,x) = e^{-\rho s}\psi(x)$$

with $\psi$ as in (9.2.14), $C_1, C_2$ and $x_1$ as in (9.2.12)–(9.2.13). Note in particular that $\psi(x)$ is not $C^1$ at $x = 0$.

## 9.3 Viscosity solutions of HJBQVI

We now turn to the general combined stochastic control and impulse control problem from Chapter 8. Thus the state $Y(t) = Y^{(w)}(t)$ is

$$
\begin{aligned}
dY(t) &= b(Y(t), u(t))dt + \sigma(Y(t), u(t))dB(t) \\
&\quad + \int_{\mathbb{R}^k} \gamma(Y(t^-), u(t^-), z)\tilde{N}(dt, dz) \qquad \tau_i < t < \tau_{i+1} \\
Y(\tau_{i+1}) &= \Gamma(\check{Y}(\tau_{i+1}^-), \zeta_{i+1}) \;; \qquad i = 0, 1, 2, \ldots
\end{aligned}
\tag{9.3.1}
$$

where $w = (u,v) \in \mathcal{W}$, $u \in \mathcal{U}$, $v = (\tau_1, \tau_2, \ldots; \zeta_1, \zeta_2, \ldots) \in \mathcal{V}$.

The performance is given by

$$J^{(w)}(y) = E^y\left[ \int_0^{\tau_S} f(Y(t), u(t))dt + g(Y(\tau_S))\chi_{\{\tau_S < \infty\}} + \sum_j K(\check{Y}(\tau_j^-), \zeta_j) \right]\tag{9.3.2}$$

and we want to find the value function $\Phi$ defined by

$$\Phi(y) = \sup_{w \in \mathcal{W}} J^{(w)}(y) \;.\tag{9.3.3}$$

To simplify the presentation we will from now on assume that

$$\mathcal{S} \text{ is an open set, i.e. } \mathcal{S} = \mathcal{S}^0 \;,\tag{9.3.4}$$

and that $\partial\mathcal{S}$ is *regular* for $Y^{(w)}(t)$ for all $w \in \mathcal{W}$, i.e.

$$\tau_S = \tau_S(y) = \inf\{t > 0; Y^{(w)}(t) \notin \mathcal{S}\} = 0 \quad \text{for all } y \in \partial\mathcal{S} \text{ and all } w \in \mathcal{W} \;.\tag{9.3.5}$$

These conditions (9.3.4)–(9.3.5) exclude cases where $\Phi$ also satisfies certain HJBQVIs on $\partial\mathcal{S}$ (see e.g. [ØS]), but it is often easy to see how to extend the results to such situations.

Theorem 8.1 associates $\Phi$ to the HJBQVI

$$\max \left( \sup_{\alpha \in U} \{ L^\alpha \Phi(y) + f(y, \alpha) \}, \mathcal{M}\Phi(y) - \Phi(y) \right) = 0 \ , \qquad y \in \mathcal{S} \qquad (9.3.6)$$

with boundary values

$$\Phi(y) = g(y) \ ; \qquad y \in \partial \mathcal{S} \qquad (9.3.7)$$

where

$$L^\alpha \Phi(y) = \sum_{i=1}^k b_i(y, \alpha) \frac{\partial \Phi}{\partial y_i} + \frac{1}{2} \sum_{i,j=1}^k \left( \sigma \sigma^T \right)_{ij} (y, \alpha) \frac{\partial^2 \Phi}{\partial y_i \partial y_j}$$

$$+ \sum_{j=1}^\ell \int_{\mathbb{R}} \left\{ \Phi \left( y + \gamma^{(j)}(y, \alpha, z) \right) - \Phi(y) - \nabla \Phi(y) \cdot \gamma^{(j)}(y, \alpha, z) \right\} \nu_j(dz_j)$$

$$(9.3.8)$$

and

$$\mathcal{M}\Phi(y) = \sup \left\{ \Phi(\Gamma(y, \zeta)) + K(y, \zeta); \zeta \in \mathcal{Z}, \Gamma(y, \zeta) \in \mathcal{S} \right\} . \qquad (9.3.9)$$

Unfortunately, as we have seen already for optimal stopping problems, the value function $\Phi$ need not be $C^1$ everywhere – in general not even continuous! So (9.3.6) is not well-defined, if we interpret the equation in the usual sense. However, it turns out that if we interpret (9.3.6) in the *weak sense of viscosity* then $\Phi$ does indeed solve the equation. In fact, under some assumptions $\Phi$ is the *unique* viscosity solution of (9.3.6)–(9.3.7) (Theorem 9.11 below). This result is an important supplement to Theorem 8.1.

We now define the concept of viscosity solutions of general HJBQVIs of type (9.3.6)–(9.3.7):

**Definition 9.6.** *Let $\varphi \in C(\bar{\mathcal{S}})$.*

*(i) We say that $\varphi$ is a* viscosity subsolution *of*

$$\max \left( \sup_{\alpha \in U} \{ L^\alpha \varphi(y) + f(y, \alpha) \} , \ \mathcal{M}\varphi(y) - \varphi(y) \right) = 0 \ ; \ y \in \mathcal{S} \qquad (9.3.10)$$

$$\varphi(y) = g(y) \ ; \ y \in \partial \mathcal{S} \qquad (9.3.11)$$

*if (9.3.11) holds and for every $h \in C^2(\mathbb{R}^k)$ and every $y_0 \in \mathcal{S}$ such that $h \geq \varphi$ on $\mathcal{S}$ and $h(y_0) = \varphi(y_0)$ we have*

$$\max \left( \sup_{\alpha \in U} \{ L^\alpha h(y_0) + f(y_0, \alpha) \} , \ \mathcal{M}\varphi(y_0) - \varphi(y_0) \right) \geq 0. \qquad (9.3.12)$$

*(ii) We say that $\varphi$ is a* viscosity supersolution *of (9.3.10)-(9.3.11) if (9.3.11) holds and for every $h \in C^2(\mathbb{R}^k)$ and every $y_0 \in \mathcal{S}$ such that $h \leq \varphi$ on $\mathcal{S}$ and $h(y_0) = \varphi(y_0)$ we have*

$$\max\left(\sup_{\alpha\in U}\{L^\alpha h(y_0) + f(y_0, \alpha)\},\ \mathcal{M}\varphi(y_0) - \varphi(y_0)\right) \le 0. \qquad (9.3.13)$$

*(iii) We say that $\varphi$ is a* viscosity solution *of (9.3.10)-(9.3.11) if $\varphi$ is both a viscosity subsolution and a viscosity supersolution of (9.3.10)-(9.3.11).*

**Lemma 9.7.** *Let $\Phi$ be as in (9.3.3). Then $\Phi(y) \ge \mathcal{M}\Phi(y)$ for all $y \in \mathcal{S}$.*

*Proof.* Suppose there exists $y \in \mathcal{S}$ with

$$\Phi(y) < \mathcal{M}\Phi(y),$$

i.e.

$$\Phi(y) < \sup_{\zeta\in\mathcal{Z}}\{\Phi(\Gamma(y,\zeta)) + K(y,\zeta)\}.$$

Then there exists $\hat{\zeta} \in \mathcal{Z}$ such that, with $\hat{y} = \Gamma(y,\hat{\zeta})$,

$$\Phi(y) < \Phi(\hat{y}) + K(y,\hat{\zeta}).$$

Put

$$\epsilon := \frac{1}{2}\left(\Phi(\hat{y}) + K(y,\hat{\zeta}) - \Phi(y)\right),$$

and let $w = (u,v)$, with $v = (\tau_1, \tau_2, \cdots; \zeta_1, \zeta_2, \cdots)$ be $\epsilon$-optimal for $\Phi$ at $\hat{y}$, in the sense that

$$J^{(w)}(\hat{y}) > \Phi(\hat{y}) - \epsilon.$$

Define $\hat{w} := (u, \hat{v})$, where $\hat{v} = (0, \tau_1, \tau_2, \ldots; \hat{\zeta}, \zeta_1, \zeta_2, \cdots)$. Then, with $\tau_0 = 0$, $\zeta_0 = \hat{\zeta}$,

$$\Phi(y) \ge J^{(\hat{w})}(y) = E^y\Big[\int_0^{\tau_\mathcal{S}} f(Y(t), u(t))dt + g(Y(\tau_\mathcal{S}))\chi_{\{\tau_\mathcal{S}<\infty\}}$$

$$+ \sum_{i=0}^{\infty} K(\check{Y}(\tau_i^-), \zeta_i)\Big] = K(y,\hat{\zeta}) + J^{(w)}(\hat{y}).$$

Combining the above we get

$$K(y,\hat{\zeta}) + J^{(w)}(\hat{y}) \le \Phi(y) < \Phi(\hat{y}) + K(y,\hat{\zeta})$$
$$< J^{(w)}(\hat{y}) + \epsilon + K(y,\hat{\zeta}).$$

This implies that

$$\Phi(\hat{y}) + K(y,\hat{\zeta}) - \Phi(y) < \epsilon,$$

a contradiction. $\qquad\square$

Our first main result in this section is the following:

**Theorem 9.8.** *The value function*

$$\Phi(y) = \sup_{w \in \mathcal{W}} J^{(w)}(y)$$

*of the combined stochastic control and impulse control problem* (9.3.3) *is a viscosity solution of* (9.3.6)-(9.3.7).

*Proof.* By (9.3.5) and (9.3.2) we see that $\Phi$ satisfies (9.3.7).

a) We first prove that $\Phi$ is a viscosity subsolution. To this end, choose $h \in C^2(\mathbb{R}^k)$ and $y_0 \in \mathcal{S}$ such that $h \geq \Phi$ on $\mathcal{S}$ and $h(y_0) = \Phi(y_0)$. We must prove that

$$\max\left(\sup_{\alpha \in U}\{L^\alpha h(y_0) + f(y_0, \alpha)\}, \ \mathcal{M}\Phi(y_0) - \Phi(y_0)\right) \geq 0. \qquad (9.3.14)$$

If $\mathcal{M}\Phi(y_0) \geq \Phi(y_0)$ then (9.3.14) holds trivially, so we may assume that

$$\mathcal{M}\Phi(y_0) < \Phi(y_0). \qquad (9.3.15)$$

Choose $\epsilon > 0$ and let $w = (u, v) \in \mathcal{W}$, with $v = (\tau_1, \tau_2, \cdots; \zeta_1, \zeta_2, \cdots) \in \mathcal{V}$, be an $\epsilon$-optimal portfolio, i.e.

$$\Phi(y_0) < J^{(w)}(y_0) + \epsilon.$$

Since $\tau_1$ is a stopping time we know that $\{\omega \, ; \, \tau_1(\omega) = 0\}$ is $\mathcal{F}_0$-measurable and hence either

$$\tau_1(\omega) = 0 \text{ a.s. or } \tau_1(\omega) > 0 \text{ a.s.} \qquad (9.3.16)$$

If $\tau_1 = 0$ a.s. then $Y^{(w)}$ makes an immediate jump from $y_0$ to the point $y' = \Gamma(y_0, \zeta_1) \in \mathcal{S}$ and hence

$$\Phi(y_0) - \epsilon \leq J^{(w')}(y') + K(y, \zeta_1) \leq \Phi(y') + K(y, \zeta_1) \leq \mathcal{M}\Phi(y_0),$$

where $w' = (\tau_2, \tau_3, \cdots; \zeta_2, \zeta_3, \cdots)$.

This is a contradiction if $\epsilon < \Phi(y_0) - \mathcal{M}\Phi(y_0)$. This proves that (9.3.15) implies that it is impossible to have $\tau_1 = 0$ a.s.

So by (9.3.16), we can now assume that $\tau_1 > 0$ a.s.. Choose $R < \infty$, $\rho > 0$ and define

$$\tau := \tau_1 \wedge R \wedge \inf\{t > 0 \, ; \, |\, Y^{(w)}(t) - y_0| \geq \rho\}.$$

By the Dynkin formula we have, with $Y(t) = Y^{(w)}(t)$,

$$E^{y_0}[h(\check{Y}(\tau^-))] = h(y_0) + E^{y_0}\left[\int_0^\tau L^u h(Y(t))dt\right], \qquad (9.3.17)$$

where $\check{Y}(\tau^-) := Y(\tau^-) + \Delta_N Y(\tau)$, with $\Delta_N Y(\tau)$ being the jump of $Y$ at $\tau$ stemming from $N$ only, not from the intervention at $\tau$ (see (6.1.3)-(6.1.7) for details).

By the dynamic programming principle (see lemma 7.3) we have : for each $\varepsilon > 0$, there exists a control $u$ such that

$$\Phi(y_0) \leq E^{y_0}\left[\int_0^\tau f(Y(t), u(t))dt + \Phi(\check{Y}(\tau^-))\right] + \varepsilon. \qquad (9.3.18)$$

Combining (9.3.17) and (9.3.18) and using that $h \geq \Phi$ and $h(y_0) = \Phi(y_0)$, we get

$$E^{y_0}\left[\int_0^\tau \{L^u h(Y(t)) + f(Y(t), u(t))\}dt\right] \geq -\varepsilon.$$

Dividing by $E^{y_0}[\tau]$ and letting $\rho \to 0$ we get

$$L^{\alpha_0} h(y_0) + f(y_0, \alpha_0) \geq -\varepsilon,$$

where

$$\alpha_0 = \lim_{s \to 0^+} u(s).$$

Since $\varepsilon$ is arbitrary, this proves (9.3.14) and hence that $\Phi$ is a viscosity subsolution.

b) Next we prove that $\Phi$ is a viscosity supersolution. So we choose $h \in C^2(\mathbb{R}^k)$ and $y_0 \in \mathcal{S}$ such that $h \leq \Phi$ on $\mathcal{S}$ and $h(y_0) = \Phi(y_0)$. We must prove that

$$\max\left(\sup_{\alpha \in U}\{L^\alpha h(y_0) + f(y_0, \alpha)\}, \mathcal{M}\Phi(y_0) - \Phi(y_0)\right) \leq 0. \qquad (9.3.19)$$

Since $\Phi \geq \mathcal{M}\Phi$ always (Lemma 9.7) it suffices to prove that

$$L^\alpha h(y_0) + f(y_0, \alpha) \leq 0 \text{ for all } \alpha \in U.$$

To this end, fix $\alpha \in U$ and let $w_\alpha = (\alpha, 0)$, i.e. $w_\alpha$ is the combined control $(u_\alpha, v_\alpha) \in \mathcal{W}$ where $u_\alpha = \alpha$ (constant) and $v_\alpha = 0$ (no interventions). Then by the dynamic programming principle and the Dynkin formula we have, with $Y(t) = Y^{(w_\alpha)}(t)$, $\tau = \tau_\mathcal{S} \wedge \rho$,

$$\begin{aligned}
\Phi(y_0) &\geq E^{y_0}\left[\int_0^\tau f(Y(s), \alpha)ds + \Phi(\check{Y}(\tau^-))\right] \\
&\geq E^y\left[\int_0^\tau f(Y(s), \alpha)ds + h(\check{Y}(\tau^-))\right] \\
&= h(y_0) + E^y\left[\int_0^\tau \{L^\alpha h(Y(t)) + f(Y(t), \alpha)\}dt\right].
\end{aligned}$$

Hence

$$E\left[\int_0^\tau \{L^\alpha h(Y(t)) + f(Y(t), \alpha)\}dt\right] \leq 0.$$

Dividing by $E[\tau]$ and letting $\rho \to 0$ we get (9.3.19).

This completes the proof of Theorem 9.8. $\qquad\qquad\qquad\qquad\qquad$ □

Next we turn to the question of *uniqueness* of viscosity solutions of (9.3.10)-(9.3.11). Many types of uniqueness results can be found in the literature. See the references in the end of this section.

Here we give a proof in the case when the process $Y(t)$ has no jumps, i.e. when $N(\cdot, \cdot) = \nu(\cdot) = 0$. The method we use is a generalization of the method in [ØS, Theorem 3.8] .

First we introduce some convenient notation:

Define $\Lambda : \mathbb{R}^{k\times k} \times \mathbb{R}^k \times \mathbb{R}^{\mathcal{S}} \times \mathbb{R}^k \to \mathbb{R}$ by

$$\Lambda(R, r, \varphi, y) := \sup_{\alpha \in U} \left\{ \sum_{i=1}^{k} b_i(y, \alpha)r_i + \frac{1}{2} \sum_{i,j=1}^{k} (\sigma\sigma^T)_{ij}(y, \alpha)R_{ij} \right.$$

$$\left. + \sum_{j=1}^{\ell} \int_{\mathbb{R}} \left\{ \varphi(y + \gamma^{(j)}(y, \alpha, z_j)) - \varphi(y) - r \cdot \gamma^{(j)}(y, \alpha, z_j) \right\} \nu_j(dz_j) + f(y, \alpha) \right\}$$

$$(9.3.20)$$

for $R = [R_{ij}] \in \mathbb{R}^{k\times k}$, $r = (r_i, \ldots, r_k) \in \mathbb{R}^k$, $\varphi : \mathcal{S} \to \mathbb{R}$, $y \in \mathbb{R}^k$, and define $F : \mathbb{R}^{k\times k} \times \mathbb{R}^k \times \mathbb{R}^{\mathcal{S}} \times \mathbb{R}^k \to \mathbb{R}$ by

$$F(R, r, \varphi, y) = \max\{\Lambda(R, r, \varphi, y), \mathcal{M}\varphi(y) - \varphi(y)\}. \qquad (9.3.21)$$

Note that if $\varphi \in C^2(\mathbb{R}^k)$ then

$$\Lambda(D^2\varphi, D\varphi, \varphi, y) = \sup_{\alpha \in U} \left\{ L^\alpha \varphi(y) + f(y, \alpha) \right\},$$

where

$$D^2\varphi = \left[ \frac{\partial^2 \varphi}{\partial y_i \partial y_j} \right](y) \text{ and } D\varphi = \left[ \frac{\partial \varphi}{\partial y_i} \right](y).$$

We recall the concepts of "superjets" $J_{\mathcal{S}}^{2,+}$, $J_{\mathcal{S}}^{2,-}$ and $\bar{J}_{\mathcal{S}}^{2,+}$, $\bar{J}_{\mathcal{S}}^{2,-}$ (see [CIL], section 2):

$$J_{\mathcal{S}}^{2,+}\varphi(y) := \left\{ (r, R) \in \mathbb{R}^{k\times k} \times \mathbb{R}^k \ ; \right.$$

$$\left. \limsup_{\substack{\eta \to y \\ \eta \in \mathcal{S}}} \left[ u(\eta) - u(y) - r(\eta - y) - \frac{1}{2}(\eta - y)^T R(\eta - y) \right] \cdot |\eta - y|^{-2} \le 0 \right\},$$

$$\bar{J}_{\mathcal{S}}^{2,+}\varphi(y) := \left\{ (r, R) \in \mathbb{R}^{k\times k} \times \mathbb{R}^k \ ; \ \text{for all } n \text{ there exists} \right.$$

$(R^{(n)}, r^{(n)}, y^{(n)}) \in \mathbb{R}^{k\times k} \times \mathbb{R}^k \times \mathcal{S}$ such that $(R^{(n)}, r^{(n)}) \in J_{\mathcal{S}}^{2,+}\varphi(y^{(n)})$ and

$(R^{(n)}, r^{(n)}, \varphi(y^{(n)}), y^{(n)}) \to (R, r, \varphi(y), y)$ as $n \to \infty \}$

and

$$J_{\mathcal{S}}^{2,-}\varphi = -J_{\mathcal{S}}^{2,+}(-\varphi), \quad \bar{J}_{\mathcal{S}}^{2,-}\varphi = -\bar{J}_{\mathcal{S}}^{2,+}(-\varphi).$$

In terms of these superjets one can give an equivalent definition of viscosity solutions as follows :

**Theorem 9.9 ([CIL], Section 2).**

*(i) A function $\varphi \in C(\mathcal{S})$ is a viscosity subsolution of (9.3.10)-(9.3.11) if and only if (9.3.11) holds and*

$$\max(\Lambda(R, r, \varphi, y), \mathcal{M}\varphi(y) - \varphi(y)) \geq 0 \text{ for all } (r, R) \in \bar{J}_{\mathcal{S}}^{2,+}\varphi(y), \ y \in \mathcal{S}.$$

*(ii) A function $\varphi \in C(\mathcal{S})$ is a viscosity supersolution of (9.3.10)-(9.3.11) if and only if (9.3.11) holds and*

$$\max(\Lambda(R, r, \varphi, y), \mathcal{M}\varphi(y) - \varphi(y)) \leq 0 \text{ for all } (r, R) \in \bar{J}_{\mathcal{S}}^{2,-}\varphi(y), \ y \in \mathcal{S}.$$

We have now ready for the second main theorem of this section:

**Theorem 9.10 (Comparison theorem).**
    *Assume that*
$$N(\cdot, \cdot) = \nu(\cdot) = 0. \tag{9.3.22}$$

*Suppose that there exists a positive function $\beta \in C^2(\bar{\mathcal{S}})$ which satisfies the strict quasi-variational inequality*

$$\max\left(\sup_{\alpha \in U}\{L^\alpha \beta(y)\}, \sup_{\zeta \in \mathcal{Z}} \beta(\Gamma(y, \zeta)) - \beta(y)\right) \leq -\delta(y) < 0 \ ; \ y \in \mathcal{S}, \tag{9.3.23}$$

*where $\delta(y) > 0$ is bounded away from 0 on compact subsets of $\mathcal{S}$.*

*Let $u$ be a viscosity subsolution and $v$ a viscosity supersolution of (9.3.10)-(9.3.11) and suppose that*

$$\lim_{|y| \to \infty} \left\{\frac{u^+(y)}{\beta(y)} + \frac{v^-(y)}{\beta(y)}\right\} = 0. \tag{9.3.24}$$

*Then*
$$u(y) \leq v(y) \text{ for all } y \in \mathcal{S}.$$

*Proof.* (Sketch) We argue by contradiction. Suppose that

$$\sup_{y \in \mathcal{S}}\{u(y) - v(y)\} > 0.$$

Then by (9.3.24) there exists $\epsilon > 0$ such that if we put

$$v_\epsilon(y) := v(y) + \epsilon \beta(y) \ ; \ y \in \mathcal{S}$$

then
$$M := \sup_{y \in \mathcal{S}}\{u(y) - v_\epsilon(y)\} > 0.$$

For $n = 1, 2, \ldots$ and $(x, y) \in \mathcal{S} \times \mathcal{S}$ define

$$H_n(x, y) := u(x) - v(y) - \frac{n}{2}|x - y|^2 - \frac{\epsilon}{2}(\beta(x) + \beta(y)).$$

and

$$M_n := \sup_{(x,y)\in\mathcal{S}\times\mathcal{S}} H_n(x,y).$$

Then by (9.3.24) we have

$$0 < M_n < \infty \text{ for all } n,$$

and there exists $(x^{(n)}, y^{(n)}) \in \mathcal{S} \times \mathcal{S}$ such that

$$M_n = H_n(x^{(n)}, y^{(n)}).$$

Then by Lemma 3.1 in [CIL] the following holds:

$$\lim_{n\to\infty} n|x^{(n)} - y^{(n)}|^2 = 0$$

and

$$\lim_{n\to\infty} M_n = u(\hat{y}) - v_\epsilon(\hat{y}) = \sup_{y\in\mathcal{S}}\{u(y) - v_\epsilon(y)\} = M,$$

for any limit point $\hat{y}$ of $\{y^{(n)}\}_{n=1}^{\infty}$.

Since $v$ is a supersolution of (9.3.10)-(9.3.11) and (9.3.23) holds, we see that $v_\epsilon$ is a *strict* supersolution of (9.3.10), in the sense that $\varphi = v_\epsilon$ satisfies (9.3.13) in the following strict form:

$$\max\left(\sup_{\alpha\in U}\{L^\alpha h(y_0) + f(y_0, \alpha)\}, \mathcal{M}v_\epsilon(y_0) - v_\epsilon(y_0)\right) \leq -\delta(y_0),$$

with $\delta(\cdot)$ as in (9.3.23).

By [CIL, Theorem 3.2], there exist $k \times k$ matrices $P^{(n)}, Q^{(n)}$ such that, if we put

$$p^{(n)} = q^{(n)} = n(x^{(n)} - y^{(n)})$$

then

$$(p^{(n)}, P^{(n)}) \in \bar{J}^{2,+}u(x^{(n)}) \text{ and } (q^{(n)}, Q^{(n)}) \in \bar{J}^{2,-}v_\epsilon(y^{(n)})$$

and

$$\begin{bmatrix} P^{(n)} & 0 \\ 0 & -Q^{(n)} \end{bmatrix} \leq 3n \begin{bmatrix} I & -I \\ -I & I \end{bmatrix},$$

in the sense that

$$\xi^T P^{(n)} \xi - \eta^T Q^{(n)} \eta \leq 3n|\xi - \eta|^2 \text{ for all } \xi, \eta \in \mathbb{R}^k. \tag{9.3.25}$$

Since $u$ is a subsolution we have, by Theorem 9.9,

$$\max\left(\Lambda(P^{(n)}, p^{(n)}, u, x^{(n)}), \mathcal{M}u(x^{(n)}) - u(x^{(n)})\right) \geq 0 \tag{9.3.26}$$

and since $v_\epsilon$ is a supersolution we have

$$\max\left(\Lambda(Q^{(n)}, q^{(n)}, v_\epsilon, y^{(n)}), \mathcal{M}v_\epsilon(y^{(n)}) - v_\epsilon(y^{(n)})\right) \le 0. \qquad (9.3.27)$$

By (9.3.25) we get

$$\Lambda(P^{(n)}, p^{(n)}, u, x^{(n)}) - \Lambda(Q^{(n)}, q^{(n)}, v_\epsilon, y^{(n)})$$

$$\le \sup_{\alpha \in U}\left\{\sum_{i=1}^{k}(b_i(x^{(n)}, \alpha) - b_i(y^{(n)}, \alpha))(p_i^{(n)} - q_i^{(n)})\right.$$

$$\left. + \frac{1}{2}\sum_{i,j=1}^{k}\left[(\sigma\sigma^T)_{ij}(x^{(n)}, \alpha) - (\sigma\sigma^T)_{ij}(y^{(n)}, \alpha)\right](P_{ij}^{(n)} - Q_{ij}^{(n)})\right\}$$

$$\le 0.$$

Therefore, by (9.3.27),

$$\Lambda(P^{(n)}, p^{(n)}, u, x^{(n)}) \le \Lambda(Q^{(n)}, q^{(n)}, v_\epsilon, y^{(n)}) \le 0$$

and hence, by (9.3.26),

$$\mathcal{M}u(x^{(n)}) - u(x^{(n)}) \ge 0. \qquad (9.3.28)$$

On the other hand, since $v_\epsilon$ is a strict supersolution we have

$$\mathcal{M}v_\epsilon(y^{(n)}) - v_\epsilon(y^{(n)}) < -\delta \text{ for all } n, \qquad (9.3.29)$$

for some constant $\delta > 0$.

Combining the above we get

$$M_n < u(x^{(n)}) - v_\epsilon(y^{(n)}) < \mathcal{M}u(x^{(n)}) - \mathcal{M}v_\epsilon(y^{(n)}) - \delta$$

and hence

$$M = \lim_{n\to\infty} M_n \le \lim_{n\to\infty}(\mathcal{M}u(x^{(n)}) - \mathcal{M}v_\epsilon(y^{(n)}) - \delta)$$

$$\le \mathcal{M}u(\hat{y}) - \mathcal{M}v_\epsilon(\hat{y}) - \delta$$

$$= \sup_{\zeta \in \mathcal{Z}}\{u(\Gamma(\hat{y}, \zeta)) + K(\hat{y}, \zeta)\} - \sup_{\zeta \in \mathcal{Z}}\{v_\epsilon(\Gamma(\hat{y}, \zeta)) + K(\hat{y}, \zeta)\} - \delta$$

$$\le \sup_{\zeta \in \mathcal{Z}}\{u(\Gamma(\hat{y}, \zeta)) - v_\epsilon(\Gamma(\hat{y}, \zeta))\} - \delta \le M - \delta.$$

This contradiction proves Theorem 9.10.     □

**Theorem 9.11 (Uniqueness of viscosity solutions).**
*Suppose that the process $Y(t)$ has no jumps i.e.*

$$N(\cdot, \cdot) = 0$$

and let $\beta \in C^2(\bar{\mathcal{S}})$ be as in Theorem 9.10. Then there is at most one viscosity solution $\varphi$ of (9.3.10)-(9.3.11) with the property that

$$\lim_{|y| \to \infty} \frac{|\varphi(y)|}{\beta(y)} = 0. \tag{9.3.30}$$

*Proof.* Let $\varphi_1, \varphi_2$ be two viscosity solutions satisfying (9.3.30). If we apply Theorem 9.10 to $u = \varphi_1$ and $v = \varphi_2$ we get

$$\varphi_1 \leq \varphi_2.$$

If we apply Theorem 9.10 to $u = \varphi_2$ and $v = \varphi_1$ we get

$$\varphi_2 \leq \varphi_1.$$

Hence $\varphi_1 = \varphi_2$.                                                           $\square$

**Example 9.12 (Optimal consumption and portfolio with both fixed and proportional transaction costs (2)).**

Let us return to Example 8.3. In this case equation (9.3.10) takes the form (8.2.21)-(8.2.22) in $\mathcal{S}^0$. For simplicity we assume Dirichlet boundary conditions, e.g. $\psi = 0$, on $\partial \mathcal{S}$. Fix $\gamma' \in (\gamma, 1)$ such that (see (3.1.8))

$$\delta > \gamma' \left[ r + \frac{(\mu - r)^2}{2\sigma^2(1 - \gamma)} \right]$$

and define

$$\beta(x_1, x_2) = (x_1 + x_2)^{\gamma'}. \tag{9.3.31}$$

Then with $\mathcal{M}$ as in (8.2.22) we have

$$(\mathcal{M}\beta - \beta)(x_1, x_2) \leq (x_1, x_2)^{\gamma_1} \left[ \left(1 - \frac{k}{x_1 + x_2}\right)^{\gamma'} - 1 \right] < 0. \tag{9.3.32}$$

Moreover, with

$$L^u \psi(x_1, x_2) := -\rho \psi(x_1, x_2) + (rx_1 - u)\frac{\partial \psi}{\partial x_1}(x_1, x_2) + \mu x_2 \frac{\partial \psi}{\partial x_2}(x_1, x_2)$$

$$+ \frac{1}{2}\sigma^2 x_2^2 \frac{\partial^2 \psi}{\partial x_2^2}(x_1, x_2) \; ; \; \psi \in C^2(\mathbb{R}^2) \tag{9.3.33}$$

we get

$$\max_{u \geq 0} L^u \beta(x_1, x_2) < 0, \tag{9.3.34}$$

and in both (9.3.32) and (9.3.34) the strict inequality is uniform on compact subsets of $\mathcal{S}^0$. The proofs of these inequalities are left as an exercise (Exercise 9.3).

We conclude that the function $\beta$ in (9.3.31) satisfies the conditions (9.3.23) of Theorem 9.10. Thus by Theorem 9.11 we have in this example uniqueness of viscosity solutions $\varphi$ satisfying the growth condition

$$\lim_{|(x_1,x_2)|\to\infty} (x_1 + x_2)^{-\gamma'} |\varphi(x_1, x_2)| = 0. \qquad (9.3.35)$$

For other results regarding uniqueness of viscosity solutions of equations associated to impulse control, stochastic control and optimal stopping for jump diffusions, we refer to [Am], [AKL], [BKR2], [CIL], [Is1], [Is2], [Ish], [MS], [AT], [Ph], [JK], [FS], [BCa], [BCe] and the references therein.

## 9.4 Numerical analysis of HJBQVI

In this section we give some insights in the numerical solution of HJBQVI. We refer e.g. to [LST] for details on the finite difference approximations and the description of the algorithms to solve dynamic programming equations. Here we focus on the main problem which arises in the case of quasi-variational inequalities, that is the presence of a nonexpansive operator due to the intervention operator.

**Finite difference approximation.**

We want to solve the following HJBQVI numerically

$$\max \left( \sup_{\alpha\in U}\{L^\alpha\Phi(x) + f(x, \alpha)\}, \mathcal{M}\Phi(x) - \Phi(x) \right) = 0 , \qquad x \in \mathcal{S} \qquad (9.4.1)$$

with boundary values

$$\Phi(x) = g(x) ; \qquad x \in \partial\mathcal{S} \qquad (9.4.2)$$

where

$$L^\alpha\Phi(x) = -r\Phi + \sum_{i=1}^{k} b_i(x, \alpha)\frac{\partial\Phi}{\partial x_i} + \tfrac{1}{2}\sum_{i,j=1}^{k} a_{ij}(x, \alpha)\frac{\partial^2\Phi}{\partial x_i\partial x_j} \qquad (9.4.3)$$

and

$$\mathcal{M}\Phi(x) = \sup \left\{ \Phi(\Gamma(x, \zeta)) + K(x, \zeta); \zeta \in \mathcal{Z}, \Gamma(x, \zeta) \in \mathcal{S} \right\} . \qquad (9.4.4)$$

We have denoted here $a_{ij} := (\sigma\sigma^T)_{ij}$. We shall also write $K^\zeta(x)$ for $K(x, \zeta)$.

We assume that $\mathcal{S}$ is bounded, otherwise a change of variable or a localisation procedure has to be performed in order to reduce to a bounded domain. Moreover we assume for simplicity that $\mathcal{S}$ is a box, that is a cartesian product of bounded intervals in $\mathbb{R}^k$. We can also handle Neumann type boundary conditions without additional difficulty.

We discretize (9.4.1) by using a finite difference approximation. Let $\delta_i$ denote the finite difference step in each coordinate direction and set $\delta = (\delta_1, \dots \delta_k)$. Denote by $e_i$ the unit vector in the $i^{th}$ coordinate direction, and consider the grid $\mathcal{S}_\delta = \mathcal{S} \cap \prod_{i=1}^k (\delta_i \mathbb{Z})$. Set $\partial \mathcal{S}_\delta = \partial \mathcal{S} \cap \prod_{i=1}^k (\delta_i \mathbb{Z})$. We use the following approximations:

$$\frac{\partial \Phi}{\partial x_i}(x) \sim \frac{\Phi(x + \delta_i e_i) - \Phi(x - \delta_i e_i)}{2\delta_i} \equiv \partial_i^{\delta_i} \Phi(x) \qquad (9.4.5)$$

or (see (9.4.16))

$$\frac{\partial \Phi}{\partial x_i}(x) \sim \begin{cases} \dfrac{\Phi(x + \delta_i e_i) - \Phi(x)}{\delta_i} \equiv \partial_i^{\delta_i +} \Phi(x) \text{ if } b_i(x) \geq 0 \\[3mm] \dfrac{\Phi(x) - \Phi(x - \delta_i e_i)}{\delta_i} \equiv \partial_i^{\delta_i -} \Phi(x) \text{ if } b_i(x) \leq 0. \end{cases} \qquad (9.4.6)$$

$$\frac{\partial^2 \Phi}{\partial x_i^2}(x) \sim \frac{\Phi(x + \delta_i e_i) - 2\Phi(x) + \Phi(x - \delta_i e_i)}{\delta_i^2} \equiv \partial_{ii}^{\delta_i} \Phi(x). \qquad (9.4.7)$$

If $a_{ij}(x) \geq 0$, $i \neq j$, then

$$\frac{\partial^2 \Phi}{\partial x_i \partial x_j}(x) \sim \frac{2\Phi(x) + \Phi(x + \delta_i e_i + \delta_j e_j) + \Phi(x - \delta_i e_i - \delta_j e_j)}{2\delta_i \delta_j}$$

$$-\left[\frac{\Phi(x + \delta_i e_i) + \Phi(x - \delta_i e_i) + \Phi(x + \delta_j e_j) + \Phi(x - \delta_j e_j)}{2\delta_i \delta_j}\right]$$

$$\equiv \partial_{ij}^{\delta_i \delta_j +} \Phi(x). \quad (9.4.8)$$

If $a_{ij}(x) < 0$, $i \neq j$, then

$$\frac{\partial^2 \Phi}{\partial x_i \partial x_j}(x) \sim -\frac{[2\Phi(x) + \Phi(x + \delta_i e_i - \delta_j e_j) + \Phi(x - \delta_i e_i + \delta_j e_j)]}{2\delta_i \delta_j}$$

$$+\frac{\Phi(x + \delta_i e_i) + \Phi(x - \delta_i e_i) + \Phi(x + \delta_j e_j) + \Phi(x - \delta_j e_j)}{2\delta_i \delta_j}$$

$$\equiv \partial_{ij}^{\delta_i \delta_j -} \Phi(x). \, (9.4.9)$$

These approximations can be justified when the function $\Phi$ is smooth by Taylor expansions. Using approximations (9.4.5), (9.4.7),(9.4.8),(9.4.9), we obtain the following approximation problem:

$$\max \left( \sup_{\alpha \in U} \{L_\delta^\alpha \Phi_\delta(x) + f(x, \alpha)\}, \mathcal{M}_\delta \Phi_\delta(x) - \Phi_\delta(x) \right) = 0 \text{ for all } x \in \mathcal{S}_\delta$$

$$\Phi_\delta(x) = g(x) \qquad\qquad\qquad\qquad \text{for all } x \in \partial \mathcal{S}_\delta$$

$$(9.4.10)$$

where

$$L_\delta^\alpha \Phi(x) = \Phi(x) \left\{ \sum_{i=1}^k \frac{-a_{ii}(x,\alpha)}{\delta_i^2} + \sum_{j \neq i} \frac{|a_{ij}(x,\alpha)|}{2\delta_i\delta_j} - r \right\}$$

$$+ \frac{1}{2} \sum_{i,\kappa=\pm 1} \Phi(x+\kappa\delta_i e_i) \left\{ \frac{a_{ii}(x,\alpha)}{\delta_i^2} - \sum_{j,j\neq i} \frac{|a_{ij}(x,\alpha)|}{\delta_i\delta_j} + \kappa\frac{b_i(x,\alpha)}{\delta_i} \right\} \quad (9.4.11)$$

$$+ \frac{1}{2} \sum_{i\neq j,\kappa=\pm 1,\lambda=\pm 1} \Phi(x+\kappa e_i\delta_i + \lambda e_j\delta_j) \frac{a_{ij}(x,\alpha)^{[\kappa\lambda]}}{\delta_i\delta_j}$$

and

$$\mathcal{M}_\delta\Phi_\delta(x) = \sup\left\{\Phi(\Gamma(x,\zeta)) + K(x,\zeta); \zeta \in \mathcal{Z}_\delta(x)\right\} \quad (9.4.12)$$

with

$$\mathcal{Z}_\delta(x) = \left\{\zeta \in \mathcal{Z}, \Gamma(x,\zeta) \in \mathcal{S}_\delta\right\}. \quad (9.4.13)$$

We have used here the notation

$$a_{ij}^{[\kappa\lambda]}(x,\alpha) = \begin{cases} a_{ij}^+(x,\alpha) \equiv \max(0,a_{ij}(x,\alpha)) & \text{if } \kappa\lambda = 1 \\ a_{ij}^-(x,\alpha) \equiv -\min(0,a_{ij}(x,\alpha)) & \text{if } \kappa\lambda = -1. \end{cases}$$

In (9.4.10), $\Phi_\delta$ denotes an approximation of $\Phi$ at the grid points. This approximation is consistent and stable if the following condition holds: (see [LST] for a proof)

$$|b_i(x,\alpha)| \leq \frac{a_{ii}(x,\alpha)}{\delta_i} - \sum_{j\neq i} \frac{|a_{ij}(x,\alpha)|}{\delta_j} \quad \text{for all } \alpha \text{ in } U, x \text{ in } \mathcal{S}_\delta, i = 1\ldots k.$$

$$(9.4.14)$$

In this case $\phi_\delta$ converges to the viscosity solution of (9.4.1) when the step $\delta$ goes to 0. This can be proved by using techniques introduced by Barles and Souganidis [BS], provided a comparison theorem holds for viscosity sub- and super-solutions of the continuous-time problem.

If (9.4.14) does not hold but only the following weaker condition

$$0 \leq \frac{a_{ii}(x,\alpha)}{\delta_i} - \sum_{j\neq i} \frac{|a_{ij}(x,\alpha)|}{\delta_j} \quad \text{for all } \alpha \text{ in } U, x \text{ in } \mathcal{S}_\delta, i = 1\ldots k. \quad (9.4.15)$$

is satisfied, then it can be shown that we can also a obtain a stable approximation (but of lower order) by using the one sided approximations (9.4.6) for the approximation of the gradient instead of the centered difference (9.4.5). Instead of (9.4.11), the operator $L_\delta^\alpha$ is then equal to

$$L_\delta^\alpha \Phi(x) = \Phi(x) \left\{ \sum_{i=1}^k \frac{-a_{ii}(x,\alpha)}{\delta_i^2} + \sum_{j\neq i} \frac{|a_{ij}(x,\alpha)|}{2\delta_i\delta_j} - \frac{|b_i(x,\alpha)|}{\delta_i} - r \right\}$$

$$+ \frac{1}{2} \sum_{i,\kappa=\pm 1} \Phi(x+\kappa\delta_i e_i) \left\{ \frac{a_{ii}(x,\alpha)}{\delta_i^2} - \sum_{j,j\neq i} \frac{|a_{ij}(x,\alpha)|}{\delta_i\delta_j} + \frac{b_i(x,\alpha)^{[\kappa]}}{\delta_i} \right\}$$

$$+ \frac{1}{2} \sum_{i\neq j,\kappa=\pm 1,\lambda=\pm 1} \Phi(x+\kappa e_i\delta_i + \lambda e_j\delta_j) \frac{a_{ij}(x,\alpha)^{[\kappa\lambda]}}{\delta_i\delta_j}.$$

$$(9.4.16)$$

By replacing the values of the function $\Phi_\delta$ by their known values on the boundary, we obtain the following equation in $\mathcal{S}_\delta$:

$$\max\left( \sup_{\alpha\in U}\{\bar{L}_\delta^\alpha \Phi_\delta(x) + f_\delta(x,\alpha)\}, \mathcal{M}_\delta\Phi_\delta(x) - \Phi_\delta(x) \right) = 0, \quad x\in\mathcal{S}_\delta \quad (9.4.17)$$

where $\bar{L}_\delta^\alpha$ is a square $N_\delta\times N_\delta$ matrix, obtained by retrieving the first and last column from $L_\delta^\alpha$, $N_\delta = \mathrm{Card}(\mathcal{S}_\delta)$, that is the number of points of the grid, and $f_\delta(x,\alpha)$ (which will also be denoted by $f_\delta^\alpha(x)$) takes into acount the boundary values.

### A policy iteration algorithm for HJBQVI.

When the stability conditions (9.4.14) or (9.4.15) hold, then the matrix $\bar{L}_\delta^\alpha$ is diagonally dominant, that is

$$(\bar{L}_\delta^\alpha)_{ij} \geq 0 \text{ for } i\neq j \quad \text{and} \quad \sum_{j=1}^{N_\delta}(\bar{L}_\delta^\alpha)_{ij} \leq -r < 0 \quad \text{for all } i=1\ldots N_\delta.$$

Now let $h$ be a positive number such that

$$h \leq \min_i \frac{1}{|(\bar{L}_\delta^\alpha)_{ii} + r|} \tag{9.4.18}$$

and let $I_\delta$ denote the $N_\delta\times N_\delta$ identity matrix. It is easy to check that the matrix

$$P_\delta^\alpha := I_\delta + h(\bar{L}_\delta^\alpha + rI_\delta)$$

is submarkovian, i.e. $(P_\delta^\alpha)_{ij} \geq 0$ for all $i,j$ and $\sum_{j=1}^{N_\delta}(P_\delta^\alpha)_{ij} \leq 1$ for all $i$. Consequently equation (9.4.17) can be rewritten as

$$\max\left( \sup_{\alpha\in U}\{\frac{1}{h}\left( P_\delta^\alpha \Phi_\delta(x) - (1+rh)\Phi_\delta(x) \right) + f_\delta^\alpha(x)\}, \mathcal{M}_\delta\Phi_\delta(x) - \Phi_\delta(x) \right) = 0,$$

$$(9.4.19)$$

which is equivalent to

$$\Phi_\delta(x) = \max\left( \sup_{\alpha\in U}\mathcal{L}_\delta^\alpha \Phi_\delta(x), \sup_{\zeta\in\mathcal{Z}_\delta(x)} B^\zeta \Phi_\delta(x) \right) \tag{9.4.20}$$

where

$$\mathcal{L}_\delta^\alpha \Phi(x) \ := \ \frac{P_\delta^\alpha \Phi(x) + h f_\delta^\alpha(x)}{1 + rh} \tag{9.4.21}$$

$$\tag{9.4.22}$$

$$B^\zeta \Phi(x) := \Phi(\Gamma(x, \zeta)) + K^\zeta(x). \tag{9.4.23}$$

Let $\mathcal{P}(\mathcal{S}_\delta)$ denote the set of all subsets of $\mathcal{S}_\delta$ and for $(T, \alpha, \zeta)$ in $\mathcal{P}(\mathcal{S}_\delta) \times U \times \mathcal{Z}_\delta$, denote by $\mathcal{O}_{T,\alpha,\zeta}$ the operator :

$$\mathcal{O}_{T,\alpha,\zeta} v(x) := \begin{cases} \mathcal{L}_\delta^\alpha v(x) & \text{if } x \in \mathcal{S}_\delta \backslash T, \\ B^\zeta v(x) & \text{if } x \in T. \end{cases} \tag{9.4.24}$$

Problem (9.4.20) is equivalent to the fixed point problem

$$\Phi_\delta(x) = \sup_{T \in \mathcal{P}(\mathcal{S}_\delta), \alpha \in U, \zeta \in \mathcal{Z}_\delta} \mathcal{O}_{T,\alpha,\zeta} \Phi_\delta(x).$$

We define $\mathbf{T}_{ad}$ as

$$\mathbf{T}_{ad} := \mathcal{P}(\mathcal{S}_\delta) \backslash \mathcal{S}_\delta$$

and restrict ourselves to the following problem

$$\Phi_\delta(x) = \sup_{T \in \mathbf{T}_{ad}, \alpha \in U, \zeta \in \mathcal{Z}_\delta} \mathcal{O}_{T,w,z} \Phi_\delta(x) =: \mathcal{O}\Phi_\delta(x). \tag{9.4.25}$$

In other words, it is not admissible to make interventions at all points of $\mathcal{S}_\delta$ (i.e. the continuation region is never the empty set). We can always assume that we order the points of the grid in such a way that it is not admissible to intervene at $x_1 \in \mathcal{S}_\delta$.

The operator $\mathcal{L}_\delta^\alpha$ is contractive (because $\|P\|_\infty \leq 1$ and $rh > 0$) and satisfies the discrete maximum principle, that is

$$\mathcal{L}_\delta^\alpha v_1 - \mathcal{L}_\delta^\alpha v_2 \leq v_1 - v_2 \Rightarrow v_1 - v_2 \geq 0. \tag{9.4.26}$$

(If $v$ is a function from $S_\delta$ into $\mathbb{R}$, $v \geq 0$ means $v(x) \geq 0$ for all $x \in \mathcal{S}_\delta$).

The operator $B^\zeta$ is nonexpansive and we need some additional hypothesis in order to be able to use a policy iteration algorithm for computing a solution of (9.4.21). We assume

There exists an integer function $\sigma : \{1, 2, \dots N_\delta\} \times \mathcal{Z}_\delta \to \{1, 2, \dots N_\delta\}$
such that for all $\zeta \in \mathcal{Z}_\delta$ and all $i = 1 \dots N_\delta$
$$\Gamma(x_i, \zeta) = x_{\sigma(i,\zeta)} \text{ with } \sigma(i, \zeta) < i. \tag{9.4.27}$$

The operator $B_\zeta$ defined in (9.4.23) can be rewritten as

$$B^\zeta v = \mathbf{B}^\zeta v + K^\zeta$$

where $(\mathbf{B}^\zeta, \zeta \in \mathcal{Z}_\delta)$ is a family of $N_\delta \times N_\delta$ markovian matrices (except for the first row) defined by : $\mathbf{B}^z_{i,j} = 1$ if $j = \sigma(i,z)$ and $i \neq 1$, and 0 elsewhere.

Let $\zeta(.)$ be a feedback Markovian control from $\mathcal{S}_\delta$ into $\mathcal{Z}_\delta$, and define the function $\bar{\sigma}$ on $\mathcal{S}_\delta$ by $\bar{\sigma}(x) := \sigma(x, \zeta(x))$. Condition (9.4.27) implies that the $p$-th composition of $\bar{\sigma}$ starting in $T \in \mathbf{T}_{ad}$ will end up in $\mathcal{S}_\delta \backslash T$ after a finite number of iterations.

We can now consider the following *Howard* or *policy iteration* algorithm to solve problem (9.4.20) in the finite set $\mathcal{S}_\delta$. It consists of constructing two sequences of feedback markovian policies $\{(T_k, \alpha_k, \zeta_k), k \in \mathbb{N}\}$ and functions $\{v_k, k \in \mathbb{N}\}$ as follows: Let $v_0$ be a given initial function in $\mathcal{S}_\delta$. For $k \geq 0$ we do the following iterations :

- (step $2k$) Given $v_k$, compute a feedback markovian admissible policy $(T_{k+1}, \alpha_{k+1}, \zeta_{k+1})$ such that

$$(T_{k+1}, \alpha_{k+1}, \zeta_{k+1}) \in \underset{T,\alpha,\zeta}{\mathrm{Argmax}}\{\mathcal{O}_{T,\alpha,\zeta} v_k\}. \qquad (9.4.28)$$

  In other words

$$\alpha_{k+1}(x) \in \underset{\alpha \in U}{\mathrm{Argmax}} \, \mathcal{L}^\alpha_\delta v_k(x); \text{ for all } x \text{ in } \mathcal{S}_\delta$$

$$\zeta_{k+1}(x) \in \underset{\beta \in \mathcal{Z}_\delta}{\mathrm{Argmax}} \, B^\zeta_\delta v_k(x); \text{ for all } x \text{ in } \mathcal{S}_\delta$$

$$T_{k+1} = \{x \in \mathcal{S}_\delta, \mathcal{L}^{\alpha_{k+1}(x)}_\delta v_k(x) > B^{\zeta_{k+1}(x)}_\delta v_k(x)\}.$$

- (step $2k+1$) Compute $v_{k+1}$ as the solution of

$$v_{k+1} = \mathcal{O}_{T_{k+1}, \alpha_{k+1}, \zeta_{k+1}} v_{k+1}. \qquad (9.4.29)$$

  Set $k \leftarrow k+1$ and go to step $2k$.

It can be proved that if (9.4.15), (9.4.18) and (9.4.27) hold, then the sequence $\{v_k\}$ converges to the solution $\Phi_\delta$ of (9.4.20) and the sequence $\{(T_k, \alpha_k, \zeta_k)\}$ converges to the optimal feedback markovian strategy. See [CMS] for a proof and [BT] for similar problems. For more information on the Howard algorithm, we refer to [Pu] and [LST]. For complements on numerical methods for HJB equations we refer e.g to [KD] and [LST].

**Example 9.13 (Optimal consumption and portfolio with both fixed and proportional transaction costs (3)).** We go back to example 9.12. We want to solve equation (8.2.21) numerically. We assume now that $\mathcal{S} = (0, l) \times (0, l)$ with $l > 0$, and that the following boundary conditions hold:

$$\psi(0, x_2) = \psi(x_1, 0) = 0$$
$$\frac{\partial \psi}{\partial x_1}(l, x_2) = \frac{\partial \psi}{\partial x_2}(x_1, l) = 0 \quad \text{for all } (x_1, x_2) \text{ in } (0, l) \times (0, l).$$

Moreover we assume that the consumption is bounded by $u_{max} > 0$ so that $U = [0, u_{max}]$. Let $\delta > 0$ be a positive step and let $\mathcal{S}_\delta = \{(i\delta, j\delta), i, j \in$

$\{1, \dots N\}\}$ be the finite difference grid (we suppose that $N = l/\delta$ is an integer). We denote by $\psi_\delta$ the approximation of $\psi$ on the grid. We approximate the operator $L^u$ defined in (9.3.33) by the following finite difference operator on $\mathcal{S}_\delta$:

$$L^u_\delta \psi := -r\psi + rx_1 \partial_1^{\delta+}\psi + \mu x_2 \partial_2^{\delta+}\psi - u\partial_1^{\delta-}\psi + \frac{1}{2}\sigma^2 x_2^2 \partial_{22}^{\delta_2+}\psi$$

and set the following boundary values:

$$\psi_\delta(0, x_2) = \psi_\delta(x_1, 0) = 0$$
$$\psi_\delta(l - \delta, x_2) = \psi_\delta(l, x_2)$$
$$\psi_\delta(x_1, l - \delta) = \psi_\delta(x_1, l).$$

We then obtain a stable approximation. Take now

$$h \leq \frac{rx_1}{\delta} + \frac{\mu x_2}{\delta} + \left(\frac{\sigma x_2}{\delta}\right)^2 + \frac{u_{max}}{\delta}.$$

We obtain a problem of the form (9.4.20). In order to be able to apply the Howard algorithm described above, it remains to check that (9.4.27) holds. This is indeed the case since a finite number of transactions brings the state to the continuation region. The details are left as an exercise.

This problem is solved in [CØS] by using another numerical method based on the iterative methods of Chapter 7.

## 9.5 Exercises

**Exercise* 9.1.** Let $k > 0$ be a constant and define

$$h(x) = \begin{cases} k|x| & \text{for} \quad -\frac{1}{k} \leq x \leq \frac{1}{k} \\ 1 & \text{for} \quad |x| > \frac{1}{k} \end{cases}$$

Solve the optimal stopping problem

$$\Phi(s, x) = \sup_{\tau \geq 0} E^x\left[e^{-\rho(s+\tau)}h(B(\tau))\right]$$

where $B(t)$ is a 1-dimensional Brownian motion starting at $x \in \mathbb{R}$. Distinguish between the two cases

**a)** $k \leq \frac{\sqrt{2\rho}}{z}$, where $z > 0$ is the unique positive solution of the equation

$$\text{tgh}(z) = \frac{1}{z},$$

and

$$\text{tgh}(z) = \frac{e^z - e^{-z}}{e^z + e^{-z}}$$

**b)** $k > \frac{\sqrt{2\rho}}{z}$ .

**Exercise\* 9.2.** Assume that the state $X(t) = X^{(w)}(t)$ at time $t$ obtained by using a combined control $w = (u, v)$, where $u = u(t, \omega) \in \mathbb{R}$ and $v = (\tau_1, \tau_2, \cdots ; \zeta_1, \zeta_2, \cdots)$ with $\zeta_i \in \mathbb{R}$ given by

$$dX(t) = u(t)dt + dB(t) + \int_{\mathbb{R}} z\tilde{N}(dt, dz) ; \ \tau_i \leq t < \tau_{i+1}$$

$$X(\tau_{i+1}) = X(\tau_{i+1}^-) + \Delta_N X(\tau_{i+1}) + \zeta_{i+1} ; \ X(0) = x \in \mathbb{R}.$$

Assume that the cost of applying such a control is

$$J^{(w)}(s, x) = E^x \left[ \int_0^\infty e^{-\rho(s+t)} (X^{(w)}(t)^2 + \theta u(t)^2)dt + c \sum_i e^{-\rho(s+\tau_i)} \right]$$

where $\rho, \theta$ and $c$ are positive constants. Consider the problem to find $\Phi(s, x)$ and $w^* = (u^*, v^*)$ such that

$$\Phi(s, x) = \inf_w J^{(w)}(s, x) = J^{(w^*)}(s, x). \tag{9.5.1}$$

Let

$$\Phi_1(s, x) = \sup_u J^{(u,0)}(s, x)$$

be the value function if we de not allow any impulse control (i.e. $v = 0$) and let

$$\Phi_2(s, x) = \sup_v J^{(0,v)}(s, x)$$

be the value function if $u$ is fixed equal to 0, and only impulse controls are allowed. (See Exercice 3.5 and Exercise 6.1, respectively).

   Prove that for $i = 1, 2$, there exists $(s, x) \in \mathbb{R} \times \mathbb{R}$ such that

$$\Phi(s, x) < \Phi_i(s, x).$$

   In other words, no matter how the positive parameter values $\rho, \theta$ and $c$ are chosen it is never optimal for the problem (9.5.1) to choose $u = 0$ or $v = 0$ (compare with Exercise 8.2).

[*Hint*: Use Theorem 9.8].

**Exercise 9.3.** Prove the inequalities (9.3.32) and (9.3.34) and verify that the inequalities hold uniformly on compact subsets of $\mathcal{S}^0$.

# 10

# Solutions of Selected Exercises

## 10.1 Exercises of Chapter 1

**Exercise 1.1.**

Choose $f \in C^2(\mathbb{R})$ and put $Y(t) = f(X(t))$. Then by the Itô formula

$$dY(t) = f'(X(t))[\alpha\, dt + \sigma\, dB(t)] + \tfrac{1}{2}\sigma^2 f''(X(t))dt$$
$$+ \int_{|z|<R} \{f(X(t^-) + \gamma(z)) - f(X(t^-)) - \gamma(z)f'(X(t^-))\}\nu(dz)dt$$
$$+ \int_{\mathbb{R}} \{f(X(t^-) + \gamma(z)) - f(X(t^-))\}\bar{N}(dt, dz). \qquad (10.1.1)$$

**(i)** In particular, if $f(x) = \exp(x)$ this gives

$$dY(t) = Y(t)[\alpha\, dt + \sigma\, dB(t)] + \tfrac{1}{2}\sigma^2 Y(t)dt$$
$$+ \int_{|z|<R} \{\exp(X(t^-) + \gamma(z)) - \exp(X(t^-)) - \gamma(z)\exp(X(t^-))\}\nu(dz)dt$$
$$+ \int_{\mathbb{R}} \{\exp(X(t^-) + \gamma(z)) - \exp(X(t^-))\}\bar{N}(dt, dz)$$
$$= Y(t^-)\Big[\Big(\alpha + \tfrac{1}{2}\sigma^2 + \int_{|z|<R} \{e^{\gamma(z)} - 1 - \gamma(z)\}\nu(dz)\Big)dt$$
$$+ \sigma\, dB(t) + \int_{\mathbb{R}} \{e^{\gamma(z)} - 1\}\bar{N}(dt, dz)\Big] \qquad (10.1.2)$$

**(ii)** By (i) we see that $Y(t)$ solves the equation

$$dY(t) = Y(t^-)\Big[\beta\, dt + \theta\, dB(t) + \lambda \int_{\mathbb{R}} z\bar{N}(dt, dz)\Big]$$

if and only if

$$\alpha + \tfrac{1}{2}\sigma^2 + \int\limits_{|z|<R} \{e^{\gamma(z)} - 1 - \gamma(z)\}\nu(dz) = \alpha, \qquad \sigma = \theta$$

and  $e^{\gamma(z)} - 1 = \lambda z$  (i.e.  $\gamma(z) = \ln(1 + \lambda z))$  a.e.  $\nu$.

## Exercise 1.2.

We first make some general remarks:

Suppose $dX_i(t) = \alpha_i(t,\omega)dt + \sigma_i(t,\omega)dB(t) + \int\limits_{\mathbb{R}} \gamma_i(t,z,\omega)\bar{N}(dt,dz)$ for $i = 1,2$.

Define $Y(t) = X_1(t) \cdot X_2(t)$. Then, by the Itô formula with $f(x_1,x_2) = x_1 \cdot x_2$,

$$
\begin{aligned}
dY(t) &= X_2(t)[\alpha_1 dt + \sigma_1 dB(t)] + X_1(t)[\alpha_2 dt + \sigma_2 dB(t)] + \tfrac{1}{2} \cdot 2\sigma_1\sigma_2 dt \\
&\quad + \int\limits_{|z|<R} \{(X_1(t^-) + \gamma_1(t,z))(X_2(t^-) + \gamma_2(t,z)) - X_1(t^-)X_2(t^-) \\
&\qquad - X_2(t^-)\gamma_1(t,z) - X_1(t^-)\gamma_2(t,z)\}\nu(dz)dt \\
&\quad + \int\limits_{\mathbb{R}} \{(X_1(t^-) + \gamma_1(t,z))(X_2(t^-) + \gamma_2(t,z)) - X_1(t^-)X_2(t^-)\}\bar{N}(dt,dz) \\
&= X_2(t)[\alpha_1 dt + \sigma_1 dB(t)] + X_1(t)[\alpha_2 dt + \sigma_2 dB(t)] + \sigma_1\sigma_2 dt \\
&\quad + \int\limits_{|z|<R} \gamma_1(t,z)\gamma_2(t,z)\nu(dz)dt \\
&\quad + \int\limits_{\mathbb{R}} \{\gamma_1(t,z)\gamma_2(t,z) + X_1(t^-)\gamma_2(t,z) + X_2(t^-)\gamma_1(t,z)\}\bar{N}(dt,dz). \quad (10.1.3)
\end{aligned}
$$

In particular, if $dX(t) = \alpha\,dt + \sigma\,dB(t) + \int\limits_{\mathbb{R}} \gamma(t,z)\bar{N}(dt,dz)$, we get

$$
\begin{aligned}
d(e^{\lambda t}X(t)) &= X(t)\lambda e^{\lambda t}dt + e^{\lambda t}[\alpha\,dt + \sigma\,dB(t)] + \int\limits_{\mathbb{R}} e^{\lambda t}\gamma(t,z)\bar{N}(dt,dz) \\
&= e^{\lambda t}dX(t) + \lambda X(t)e^{\lambda t}dt.
\end{aligned}
$$

**(i)**  Now consider the equation

$$dX(t) = (m - X(t))dt + \sigma\,dB(t) + \gamma \int\limits_{\mathbb{R}} z\widetilde{N}(dt,dz),$$

where we assume that $\gamma z > -1$ for a.a. $z$ ($\nu$). It can be written

$$d(e^t X(t)) = m\,e^t dt + \sigma\,e^t dB(t) + \gamma\,e^t \int\limits_{\mathbb{R}} z\widetilde{N}(dt,dz).$$

This gives the solution

$$X(t) = X(0)e^{-t} + m \int_0^t e^{(s-t)}ds + \sigma \int_0^t e^{(s-t)}dB(s) + \gamma \int_0^t \int_{\mathbb{R}} z e^{(s-t)} \widetilde{N}(dt, dz)$$

or

$$X(t) = m + (X_0 - m)e^{-t} + \sigma \int_0^t e^{s-t}dB(s) + \gamma \int_0^t \int_{\mathbb{R}} z e^{s-t} \widetilde{N}(dt, dz) \qquad (10.1.4)$$

**(ii)** Next consider the equation

$$dX(t) = \alpha \, dt + \gamma \, X(t^-) \int_{\mathbb{R}} z \bar{N}(dt, dz); \qquad X(0) = x \in \mathbb{R}. \qquad (10.1.5)$$

Define, for a given function $\theta(z)$,

$$G(t) = \exp \Big( \int_0^t \int_{\mathbb{R}} \theta(z)\bar{N}(dt, dz) - \int_{|z|<R} \{e^{\theta(z)} - 1 - \theta(z)\}\nu(dz) \cdot t \Big).$$

Then by Itô's formula (see Exercise 1.1)

$$dG(t) = G(t^-) \int_{\mathbb{R}} \{e^{\theta(z)} - 1\}\bar{N}(dt, dz).$$

Hence, if we put

$$\widetilde{X}(t) = X(0)G(t) + \alpha \, G(t) \int_0^t G^{-1}(s)ds \qquad (10.1.6)$$

we have

$$d\widetilde{X}(t) = X(0)dG(t) + \alpha \, G(t)G^{-1}(t)dt + \alpha \int_0^t G^{-1}(s)ds \cdot dG(t)$$

$$= \alpha \, dt + X(0)G(t^-) \int_{\mathbb{R}} \{e^{\theta(z)} - 1\}\bar{N}(dt, dz)$$

$$+ \alpha \cdot \int_0^t G^{-1}(s)ds \cdot \Big[ G(t^-) \int_{\mathbb{R}} \{e^{\theta(z)} - 1\}\bar{N}(dt, dz) \Big]$$

$$= \alpha \, dt + \Big[ X(0)G(t^-) + \alpha \, G(t^-) \int_0^t G^{-1}(s)ds \Big] \int_{\mathbb{R}} \{e^{\theta(z)} - 1\}\bar{N}(dt, dz)$$

$$= \alpha \, dt + \widetilde{X}(t^-) \int_{\mathbb{R}} \{e^{\theta(z)} - 1\}\bar{N}(dt, dz).$$

So $X(t) := \widetilde{X}(t)$ solves (10.1.5) if we choose $\theta(z)$ such that

$$e^{\theta(z)} - 1 = \gamma \, z \quad \text{a.s. } \nu$$

i.e.

$$\theta(z) = \ln(1 + \gamma \, z) \quad \text{a.s. } \nu. \qquad (10.1.7)$$

**Exercise 1.6.**

By (1.2.5) (Example 1.15) we know that the equation

$$dX(t) = X(t^-) \int_{\mathbb{R}} (e^{\gamma(t,z)} - 1)\widetilde{N}(dt, dz); \qquad X(0) = 1 \qquad (10.1.8)$$

has the solution

$$X(t) = \exp\left\{ \int_0^t \int_{\mathbb{R}} \gamma(s, z)N(ds, dz) - \int_0^t \int_{\mathbb{R}} (e^{\gamma(s,z)} - 1)\nu(dz)ds \right\}$$

$$= \exp\left\{ \int_0^t \int_{\mathbb{R}} \gamma(s, z)\widetilde{N}(ds, dz) - \int_0^t \int_{\mathbb{R}} (e^{\gamma(s,z)} - 1 - \gamma(s, t))\nu(dz)ds \right\}. \quad (10.1.9)$$

If we assume that

$$\int_0^t \int_{\mathbb{R}} (e^{\gamma(s,z)} - 1)^2 \nu(dz)ds < \infty \qquad (10.1.10)$$

then by (10.1.8) we see that $E[X(t)] = 1$ and hence by (10.1.9) we get

$$E\left[ \exp\left\{ \int_0^t \int_{\mathbb{R}} \gamma(s, z)\widetilde{N}(ds, dz) \right\} \right] = \exp\left\{ \int_0^t \int_{\mathbb{R}} (e^{\gamma(s,z)} - 1 - \gamma(s, z))\nu(dz)ds \right\}.$$
$$(10.1.11)$$

**Exercise 1.7.**

By (10.1.3) in the solution of Exercise 1.2 we have (with $R = \infty$)

$$d(X_1(t)X_2(t)) = \int_{\mathbb{R}} \gamma_1(t, z)\gamma_2(t, z)\nu(dz)dt$$

$$+ \int_{\mathbb{R}} \{\gamma_1(t, z) + X_1(t^-)\gamma_2(t, z) + X_2(t^-)\gamma_1(t, z)\}\widetilde{N}(dt, dz)$$

$$= X_1(t^-) \int_{\mathbb{R}} \gamma_2(t, z)\widetilde{N}(dt, dz) + X_2(t^-) \int_{\mathbb{R}} \gamma_1(t, z)\widetilde{N}(dt, dz) + \int_{\mathbb{R}} \gamma_1(t, z)\gamma_2(t, z)\nu(dz)$$

$$= X_1(t^-)dX_2(t) + X_2(t^-)dX_1(t) + \int_{\mathbb{R}} \gamma_1(t, z)\gamma_2(t, z)N(dt, dz), \qquad (10.1.12)$$

which is ( 1.6.1).

**Exercise 1.8.**

To find $Q$ we apply Theorem 1.35. So we must find a solution $(\theta_1(z), \theta_2(z))$ of the two equations

(i)  $\gamma_{11} \displaystyle\int_{\mathbb{R}} \theta_1(z)\nu_1(dz) + \gamma_{12} \int_{\mathbb{R}} \theta_2(z)\nu_2(dz) = \alpha_1$

(ii)  $\gamma_{21} \displaystyle\int_{\mathbb{R}} \theta_1(z)\nu_1(dz) + \gamma_{22} \int_{\mathbb{R}} \theta_2(z)\nu_2(dz) = \alpha_2$

and such that $\theta_j(z) < 1$ for $j = 1, 2$.

This system is equivalent to

(iii)  $\displaystyle\int_{\mathbb{R}} \theta_1(z)\nu_1(dz) = \lambda_{11}\alpha_1 + \lambda_{12}\alpha_2$

(iv)  $\displaystyle\int_{\mathbb{R}} \theta_2(z)\nu_2(dz) = \lambda_{21}\alpha_1 + \lambda_{22}\alpha_2$

By our assumption (1.6.3) we see that we can choose $A_i \subset \mathbb{R}$ with

$$\lambda_{i1}\alpha_1 + \lambda_{i2}\alpha_2 < \nu_i(A_i) < \infty$$

and then the functions

$$\theta_i(z) = \frac{\lambda_{i1}\alpha_1 + \lambda_{i2}\alpha_2}{\nu_i(A_i)} \mathcal{X}_{A_i}(z) ; \qquad i = 1, 2$$

solve (i), (ii). With this choice of $\theta_i(z); \; i = 1, 2$ we define

$$Z(t) = \exp\left\{ \sum_{i=1}^{2} \left[ \int_0^t \int_{\mathbb{R}} \ln(1 - \theta_i(z_i)) N_i(ds, dz_i) + (\lambda_{i1}\alpha_1 + \lambda_{i2}\alpha_2)t \right] \right\}; \quad 0 \le t \le T$$

and we put

$$dQ = Z(T)dP \qquad \text{on } \mathcal{F}_T.$$

Then $Q$ is an equivalent local martingale measure for $(S_1(t), S_2(t))$. Just as in Section 1.5 we can now deduce that the market has no arbitrage.

## 10.2 Exercises of Chapter 2

### Exercise 2.1.

We seek

$$\Phi(s, x) = \sup_{\tau \ge 0} E^{(s,x)}[e^{-\rho(s+\tau)}(X(\tau) - a)]$$

where

$$dX(t) = dB(t) + \gamma \int_{\mathbb{R}} z\widetilde{N}(dt, dz); \qquad X(0) = x \in \mathbb{R}.$$

We intend to apply Theorem 2.2 and start by putting

$$Y(t) = \begin{bmatrix} s + t \\ X(t) \end{bmatrix}; \qquad Y(0) = \begin{bmatrix} s \\ x \end{bmatrix} = y \in \mathbb{R}^2 = \mathcal{S}.$$

The generator of $Y$ is

$$A\phi(s,z) = \frac{\partial \phi}{\partial s} + \frac{1}{2}\frac{\partial^2 \phi}{\partial x^2} + \int_{\mathbb{R}} \left\{ \phi(s, x + \gamma z) - \phi(s, x) - \frac{\partial \phi}{\partial x}(s, x)\gamma z \right\} \nu(dz).$$

According to Theorem 2.2 (ix) we should look for a function $\phi$ such that $A\phi(s,x) = 0$ in $D$. We try

$$\phi(s,x) = e^{-\rho s}\psi(x) \quad \text{for some function } \psi.$$

Then

$$A\phi(s,x) = e^{-\rho s}A_0\psi(x),$$

where

$$A_0\psi(x) = -\rho\psi(x) + \frac{1}{2}\psi''(x) + \int_{\mathbb{R}} \{\psi(x + \gamma z) - \psi(x) - \psi'(x)\gamma z\}\nu(dz).$$

Choose

$$\psi(x) = e^{\lambda x} \quad \text{for some constant } \lambda > 0.$$

Then

$$A_0\psi(x) = -\rho e^{\lambda x} + \frac{1}{2}\lambda^2 e^{\lambda x} + \int_{\mathbb{R}} \{e^{\lambda(x+\gamma z)} - e^{\lambda x} - \lambda e^{\lambda x} \cdot \gamma z\}\nu(dz)$$

$$= e^{\lambda x}\left[ -\rho + \frac{1}{2}\lambda^2 + \int_{\mathbb{R}} \{e^{\lambda \gamma z} - 1 - \lambda \gamma z\}\nu(dz) \right].$$

Put

$$h(\lambda) := -\rho + \frac{1}{2}\lambda^2 + \int_{\mathbb{R}} \{e^{\lambda \gamma z} - 1 - \lambda \gamma z\}\nu(dz).$$

Note that $h(0) = -\rho < 0$. Therefore, since

$$e^{\lambda \gamma z} - 1 - \lambda \gamma z \geq 0 \quad \text{for all } x \in \mathbb{R}$$

we see that $\lim_{\lambda \to \infty} h(\lambda) = \infty$.

So the equation $h(\lambda) = 0$ has at least one solution $\lambda_1 > 0$. Define

$$\psi(x) = \begin{cases} x - a & \text{for} \quad x \geq x^* \\ C\, e^{\lambda_1 x} & \text{for} \quad x < x^* \end{cases} \tag{10.2.1}$$

where $C > 0$, $x^* > 0$ are two constants to be determined.

If we require $\psi$ to be continuous at $x = x^*$ we get the equation

$$C\, e^{\lambda_1 x^*} = x^* - a. \tag{10.2.2}$$

If we require $\psi$ to be differentiable at $x = x^*$ we get the additional equation

$$\lambda_1 C\, e^{\lambda_1 x^*} = 1. \tag{10.2.3}$$

Dividing (10.2.1) by (10.2.2) we get

$$x^* = a + \frac{1}{\lambda_1}, \qquad C = \frac{1}{\lambda_1} e^{-(\lambda_1 a + 1)}. \tag{10.2.4}$$

We now propose that the function

$$\phi(s, x) := e^{-\rho s} \psi(x),$$

with $\psi(x)$ given by (10.2.1), (10.2.2) and (10.2.3) satisfies all the requirements of Theorem 2.2 (possibly under some assumptions) and hence that

$$\phi(s, x) = \Phi(s, x)$$

and that

$$\tau^* := \inf\{t > 0; X(t) \geq x^*\}$$

is an optimal stopping time.

We proceed to check if the conditions (i)–(xi) of Theorem 2.2 hold. Many of these conditions are satisfied trivially or by construction of $\phi$. We only discuss the remaining ones:

**(ii):** We know that $\phi = g$ for $x > x^*$, by construction. For $x < x^*$ we must check that

$$C_1 e^{\lambda_1 x} \geq x - a \ .$$

To this end, put

$$k(x) = C_1 e^{\lambda_1 x} - x + a \ ; \qquad x \leq x^*.$$

Then

$$k(x^*) = k'(x^*) = 0 \qquad \text{and}$$
$$k''(x^*) = \lambda_1^2 C_1 e^{\lambda_1 x} > 0 \qquad \text{for } x \leq x^*.$$

Therefore $k'(x) < 0$ for $x < x^*$ and hence $k(x) > 0$ for $x < x^*$. Hence (ii) holds.

**(vi):** We know that $A\phi + f = A\phi = 0$ for $x < x^*$, by construction. For $x > x^*$ we have

$$A\phi = e^{-\rho s} A_0(x - a) = e^{-\rho s}(-\rho(x - a)) < 0.$$

So (vi) holds.

**(viii):** In our case this condition gets the form

$$E\left[ \int_0^\infty \left\{ \sigma^2 e^{-2\rho t} X^2(t) + \int_{\mathbb{R}} e^{-2\rho t} \big|(X(t) + \gamma z)^2 - X^2(t)\big|^2 \nu(dz) \right\} dt \right] < \infty$$

i.e.

$$E\left[ \int_0^\infty e^{-2\rho t} \left\{ \sigma^2 X^2(t) + \int_{\mathbb{R}} \big|2X(t)\gamma z + \gamma^2 z^2\big| \nu(dz) \right\} dt \right] < \infty \ . \qquad (10.2.5)$$

This will hold if

$$z \leq 0 \quad \text{a.s. } \nu \qquad\qquad (10.2.6)$$

or if

$$\sup_{\tau \in \mathcal{T}} E^x \left[ e^{-2\rho \tau} \left( \int_0^\tau \int_{\mathbb{R}} z N(ds, dz) \right)^2 \right] < \infty \ . \qquad (10.2.7)$$

We will not discuss this condition further here.

**(x):** With our proposed solution $\phi$ we have

$$D = \{(s, x) \in \mathbb{R}^2; x < x^*\}.$$

So condition (x) states that

$$\tau_D := \inf\{t > 0; X(t) > x^*\} < \infty \text{ a.s.} \tag{10.2.8}$$

Some conditions are needed on $\sigma, \gamma$ and $\nu$ for (10.2.8) to hold. For example, it suffices that

$$\varlimsup_{t \to \infty} X(t) = \varlimsup_{t \to \infty} \left\{ \sigma B(t) + \int_0^t \int_{\mathbb{R}} \gamma \, z N(ds, dz) \right\} = \infty \quad \text{a.s.} \tag{10.2.9}$$

**(xi):**  For (xi) to hold it suffices that

$$\sup_{\tau \in \mathcal{T}} E^x[e^{-2\rho\tau} X^2(\tau)] < \infty . \tag{10.2.10}$$

Again it suffices to assume that (10.2.7) holds.

*Conclusion.*

Assume that (10.2.7) and (10.2.8) hold. Then the value function is

$$\Phi(s, x) = e^{-\rho s} \psi(x),$$

where $\psi(x)$ is given by (10.2.1) and (10.2.4). An optimal stopping time is

$$\tau^* = \inf\{t > 0; X(t) \geq x^*\}.$$

## Exercise 2.2.

Define

$$dY(t) = \begin{bmatrix} dt \\ dP(t) \\ dQ(t) \end{bmatrix} = \begin{bmatrix} 1 \\ \alpha\,P(t) \\ -\lambda\,Q(t) \end{bmatrix} dt + \begin{bmatrix} 0 \\ \beta\,P(t) \\ 0 \end{bmatrix} dB(t) + \begin{bmatrix} 0 \\ \gamma \int_{\mathbb{R}} P(t^-) z \widetilde{N}(dt, dz) \\ 0 \end{bmatrix}$$

Then the generator $A$ of $Y(t)$ is

$$A\phi(y) = A\phi(s, p, q) = \frac{\partial \phi}{\partial s} + \alpha\, p \frac{\partial \phi}{\partial p} - \lambda q \frac{\partial \phi}{\partial q} + \tfrac{1}{2}\beta^2 p^2 \frac{\partial^2 \phi}{\partial p^2}$$

$$+ \int_{\mathbb{R}} \left\{ \phi(s, p + \gamma\, pz, q) - \phi(s, p, q) - \frac{\partial \phi}{\partial p}(s, p, q) \gamma\, zp \right\} \nu(dz) .$$

If we try

$$\phi(s, p, q) = e^{-\rho s} \psi(w) \qquad \text{with} \quad w = p \cdot q ,$$

then

$$A\phi(s, p, q) = e^{-\rho s} A_0 \psi(w),$$

where

$$A_0 \psi(w) = -\rho\psi(w) + (\alpha - \lambda)w\psi'(w) + \tfrac{1}{2}\beta^2 w^2 \psi''(w)$$

$$+ \int_{\mathbb{R}} \{\psi((1 + \gamma\, z)w) - \psi(w) - \gamma\, wz\psi'(w)\}\nu(dz).$$

Consider the set $U$ defined in Proposition 2.3:

$$U = \{y; Ag(y) + f(y) > 0\} = \{(s, p, q); A_0(\theta\, w) + \lambda\, w - K > 0\}$$

$$= \{(s, p, q); [\theta(\alpha - \rho - \lambda) + \lambda]w - K > 0\}$$

$$= \begin{cases} \left\{(s, p, q) : w > \frac{K}{\theta(\alpha - \rho - \lambda) + \lambda}\right\} & \text{if} \quad \theta(\alpha - \rho - \lambda) + \lambda > 0 \\ \emptyset & \text{if} \quad \theta(\alpha - \rho - \lambda) + \lambda \le 0 \end{cases}$$

By Proposition 2.4 we therefore get:

*Case 1:*

Assume $\lambda \le \theta(\lambda + \rho - \alpha)$.
Then $\tau^* = 0$ is optimal and $\Phi(y) = g(y) = e^{-\rho s} p \cdot q$ for all $y$.

*Case 2:*

Assume $\theta(\lambda + \rho - \alpha) < \lambda$.
Then $U = \left\{(s, w); w > \frac{K}{\lambda - \theta(\lambda + \rho - \alpha)}\right\} \subset D$.
In view of this it is natural to guess that the continuation region $D$ has the form

$$D = \{(s, w); 0 < w < w^*\},$$

for some constant $w^*$; $0 < w^* < \frac{K}{\lambda - \theta(\lambda + \rho - \alpha)}$. In $D$ we try to solve the equation

$$A_0\psi(w) + f(w) = 0.$$

The homogeneous equation $A_0\psi_0(w) = 0$ has a solution $\psi_0(w) = w^r$ if and only if

$$h(r) := -\rho + (\alpha - \lambda)r + \tfrac{1}{2}\beta^2 r(r - 1) + \int_{\mathbb{R}} \{(1 + \gamma\, z)^r - 1 - r\,\gamma\, z\}\nu(dz) = 0.$$

Since $h(0) = -\rho < 0$ and $\lim_{|r| \to \infty} h(r) = \infty$, we see that the equation $h(r) = 0$ has two solutions $r_1, r_2$ such that $r_2 < 0 < r_1$.

Let $r$ be a solution of this equation. To find a particular solution $\psi_1(w)$ of the non-homogeneous equation

$$A_0\psi_1(w) + \lambda w - K = 0$$

we try

$$\psi_1(w) = aw + b$$

and find

$$a = \frac{\lambda}{\lambda + \rho - \alpha}, \qquad b = -\frac{K}{\rho}.$$

This gives that for all constants $C$ the function

$$\psi(w) = C\, w^r + \frac{\lambda}{\lambda + \rho - \alpha} w - \frac{K}{\rho}$$

is a solution of

$$A_0 \psi(w) + \lambda \, w - K = 0.$$

Therefore we try to put

$$\psi(w) = \begin{cases} \theta \, w \; ; & 0 < w \le w^* \\ C \, w^r + \frac{\lambda}{\lambda + \rho - \alpha} w - \frac{K}{\rho} \; ; & w \ge w^* \end{cases} \tag{10.2.11}$$

where $w^* > 0$ and $C$ remain to be determined.

Continuity and differentiability at $w = w^*$ give

$$\theta \, w^* = C(w^*)^r + \frac{\lambda}{\lambda + \rho - \alpha} w^* - \frac{K}{\rho} \tag{10.2.12}$$

$$\theta = C \, r(w^*)^{r-1} + \frac{\lambda}{\lambda + \rho - \alpha} \; . \tag{10.2.13}$$

Combining (10.2.12) and (10.2.13) we get

$$w^* = \frac{(-r)K(\lambda + \rho - \alpha)}{(1-r)\rho(\lambda - \theta(\lambda + \rho - \alpha))} \tag{10.2.14}$$

and

$$C = \frac{\lambda - \theta(\lambda + \rho - \alpha)}{-r} \cdot (w^*)^{1-r}. \tag{10.2.15}$$

Since we need to have $w^* > 0$ we are led to the following condition:

*Case 2a)*

$\theta(\lambda + \rho - \alpha) < \lambda$ and $\lambda + \rho - \alpha > 0$.

Then we choose $r = r_2 < 0$, and with the corresponding values (10.2.14), (10.2.15) of $w^*$ and $C$ the function $\phi(s, p, q) = e^{-\rho s}\psi(p \cdot q)$, with $\psi$ given by (10.2.11), is the value function of the problem. The optimal stopping time $\tau^*$ is

$$\tau^* = \inf\{t > 0; P(t) \cdot Q(t) \le w^*\}, \tag{10.2.16}$$

provided that all the other conditions of Theorem 2.2 are satisfied. (See Remark 10.1).

*Case 2b)*

$\theta(\lambda + \rho - \alpha) < \lambda$  and  $\lambda + \rho - \alpha \le 0$, i.e.

$$\alpha \ge \lambda + \rho \, .$$

In this case we have $\Phi^*(y) = \infty$.

To see this note that since

$$P(t) = p + \int_0^t \alpha \, P(s)ds + \int_0^t \beta \, P(s)dB(s) + \int_0^t \int_{\mathbb{R}} \gamma \, P(s^-)z\tilde{N}(ds, dz),$$

we have

$$E[P(t)] = p + \int_0^t \alpha\, E[P(s)]ds$$

which gives

$$E[P(t)] = p\,e^{\alpha t}.$$

Therefore

$$E[e^{-\rho t}P(t)Q(t)] = E[pq\,e^{-\rho t}e^{-\lambda t}P(t)] = pq\exp\{(\alpha - \lambda - \rho)t\}.$$

Hence

$$\lim_{T\to\infty} E\Big[\int_0^T e^{-\rho t}P(t)Q(t)dt\Big] = \lim_{T\to\infty} pq\int_0^T \exp\{(\alpha - \lambda - \rho)t\}dt = \infty$$

if and only if $\alpha \geq \lambda + \rho$.

**Remark 10.1 (On condition (viii) of Theorem 2.2).** Consider

$$\phi(Y(t)) = e^{-\rho t}\psi(P(t)Q(t)),$$

where

$$P(t) = p\exp\Big\{\Big(\alpha - \tfrac{1}{2}\beta^2 - \gamma\int_{\mathbb{R}} z\,\nu(dz)\Big)t + \int_0^t\int_{\mathbb{R}} \ln(1 + \gamma z)N(dt, dz) + \beta\,B(t)\Big\}$$

and

$$Q(t) = q\exp(-\lambda t).$$

We have

$$P(t)Q(t) = pq\exp\Big\{\Big(\alpha - \lambda - \tfrac{1}{2}\beta^2 - \gamma\int_{\mathbb{R}} z\,\nu(dz)\Big)t + \int_0^t\int_{\mathbb{R}} \ln(1 + \gamma z)N(dt, dz) + \beta\,B(t)\Big\}$$

and

$$e^{-\rho t}P(t)Q(t) = pq\exp\Big\{\Big(\alpha - \lambda - \rho - \tfrac{1}{2}\beta^2 - \gamma\int_{\mathbb{R}} z\,\nu(dz)\Big)t$$

$$+ \int_0^t\int_{\mathbb{R}} \ln(1 + \gamma z)N(ds, dz) + \beta\,B(t)\Big\}.$$

Hence

$$E[(e^{-\rho t}P(t)Q(t))^2] = (pq)^2 E\Big[\exp\Big\{\Big(2\alpha - 2\lambda - 2\rho - \beta^2 - 2\gamma\int_{\mathbb{R}} z\,\nu(dz)\Big)t$$

$$+ 2\int_0^t\int_{\mathbb{R}} \ln(1 + \gamma z)N(ds, dz) + 2\beta\,B(t)\Big\}\Big]$$

$$= (pq)^2 \exp\Big\{\Big(2\alpha - 2\lambda - 2\rho - \beta^2 - 2\gamma\int_{\mathbb{R}} z\,\nu(dz)\Big)t + 2\beta^2 t\Big\}$$

$$\cdot E\Big[\exp\Big(2\int_0^t\int_{\mathbb{R}} \ln(1 + \gamma z)N(dt, dz)\Big)\Big]$$

Using Exercise 1.6 we get

$$E[(e^{-\rho t}P(t)Q(t))^2] = p^2q^2 \exp\left\{\left(2\alpha - 2\lambda - 2\rho + \beta^2 - 2\gamma\int_{\mathbb{R}} z\,\nu(dz)\right.\right.$$
$$\left.\left.+ \int_{\mathbb{R}}\{(1+\gamma z)^2 - 1 - 2\ln(1+\gamma z)\}\nu(dz)\right)t\right\}.$$

So condition (viii) of Theorem 2.2 holds if

$$2\alpha - 2\lambda - 2\rho + \beta^2 + \int_{\mathbb{R}}\{\gamma^2 z^2 - 2\ln(1+\gamma z)\}\nu(dz) < 0.$$

## Exercise 2.3.

In this case we have
$$g(s,x) = e^{-\rho s}|x|$$

and
$$dX(t) = dB(t) + \int_{\mathbb{R}} z\tilde{N}(dt,dz).$$

We look for a solution of the form
$$\phi(s,x) = e^{-\rho s}\psi(x).$$

The continuation region is given by

$$D = \{(s,x) \in \mathbb{R} \times \mathbb{R} : \phi(s,x) > g(s,x)\} = \{(s,x) \in \mathbb{R} \times \mathbb{R} : \psi(x) \geq |x|\}$$

Because of the symmetry we assume that $D$ is of the form

$$D = \{(s,x) \in \mathbb{R} \times \mathbb{R} \;;\; -x^* < x < x^*\}$$

where $x^* > 0$. It is trivial that $D$ is a Lipschitz surface and $X(t)$ spends 0 time on $\partial D$. We must have

$$A\phi \equiv 0 \qquad \text{on } D \tag{10.2.17}$$

where the generator $A$ is given by

$$A\phi = \frac{\partial\phi}{\partial s} + \frac{1}{2}\frac{\partial^2\phi}{\partial x^2} + \int_{\mathbb{R}}\left\{\phi(s,x+z) - \phi(s,x) - \frac{\partial\phi}{\partial x}(s,x)z\right\}\nu(dz).$$

Hence equation (10.2.17) becomes

$$-\rho\psi(x) + \tfrac{1}{2}\psi''(x) + \int_{\mathbb{R}}\{\psi(x+z) - \psi(x) - z\psi'(x)\}\nu(dz) = 0. \tag{10.2.18}$$

Let $\lambda > 0$ and $-\lambda$ be two roots of the equation

$$F(\lambda) := -\lambda + \frac{1}{2}\lambda^2 + \int_{\mathbb{R}}\left\{e^{\lambda z} - 1 - \lambda z\right\}\nu(dz) = 0.$$

Because of the symmetry we guess that

$$\psi(x) = \frac{C}{2}\left(e^{\lambda x} + e^{-\lambda x}\right) = C\cosh(\lambda x) \;;\; x \in D$$

for some constant $C > 0$. Therefore

$$\psi(x) = \begin{cases} C \cosh\left(\lambda x\right) & \text{for } |x| < x^* \\ |x| & \text{for } |x| \geq x^*. \end{cases}$$

In order to find $x^*$ and $C$, we impose the continuity and $C^1$-conditions on $\psi(x)$ at $x = x^*$:

• Continuity:   $1 = |x^*| = C \cosh\left(\lambda x^*\right)$
• $C^1$ :       $1 = C\lambda \sinh(\lambda x^*)$

It follows that

$$C = \frac{x^*}{\cosh(\lambda x^*)} \tag{10.2.19}$$

and $x^*$ is the solution of

$$\operatorname{tgh}\left(\lambda x^*\right) = \frac{1}{\lambda x^*}. \tag{10.2.20}$$

Figure 10.1 illustrates that there exists a unique solution for equation (10.2.20).
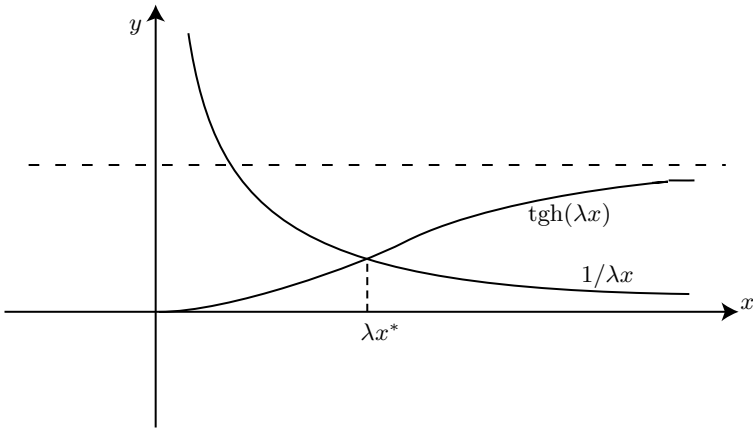


**Fig. 10.1.** The value of $x^*$

Finally we have to verify that the conditions of Theorem 2.2 hold. We check some :

**(ii)** $\psi(x) \geq |x|$ for $(s, x) \in D$.

Define

$$h(x) = C \cosh(\lambda x) - x \; ; \; x > 0.$$

Then $h(x^*) = h'(x^*) = 0$ and $h''(x) = C\lambda^2 \cosh(\lambda x) > 0$ for all $x$. Hence $h(x) > 0$ for $0 < x < x^*$, so (ii) holds. See Figure 10.2.
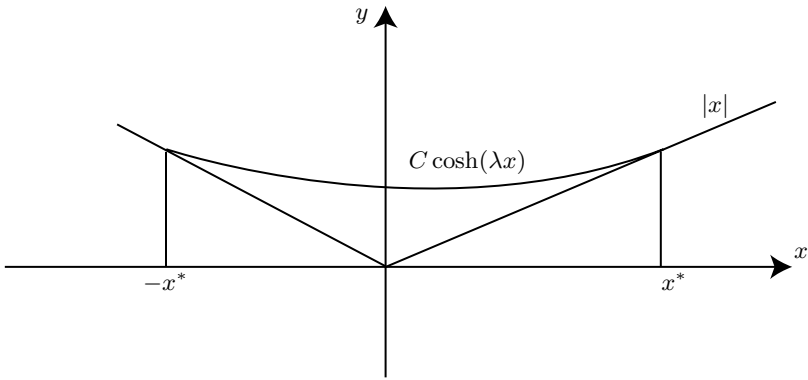
**Fig. 10.2.** The function $\psi$

**(vi)** $A\psi \leq 0$ outside $\bar{D}$.

This holds since

$$A\psi(x) = -\rho|x| + \int_{\mathbb{R}} \{|x+z| - x - z\}\,\nu(dz) \leq 0 \text{ for all } x > x^*.$$

Since all the conditions of Theorem 2.2 are satisfied, we conclude that

$$\phi(s,y) = e^{-\rho s}\psi(y)$$

is the optimal value function and $\tau^* = \inf\{t > 0 \,;\, |B(t)| = x^*\}$.

## 10.3 Exercises of Chapter 3

### Exercise 3.1

Put

$$Y(t) = \begin{bmatrix} s+t \\ X(t) \end{bmatrix}.$$

Then the generator of $Y(t)$ is

$$A^u\phi(y) = A^u\phi(s,x) = \frac{\partial\phi}{\partial s} + (\mu - \rho x - u)\frac{\partial\phi}{\partial x} + \frac{1}{2}\sigma^2\frac{\partial^2\phi}{\partial x^2}$$
$$+ \int_{\mathbb{R}} \left\{\phi(s, x + \theta z) - \phi(s,x) - \frac{\partial\phi}{\partial x}\cdot\theta z\right\}\nu(dz).$$

So the conditions of Theorem 3.1 get the form

**(i)** $A^u\phi(s,x) + e^{-\delta s}\frac{u^\gamma}{\gamma} \leq 0$ for all $u \geq 0$, $s < T$.

(Note: If we put

$$\mathcal{S} = \{(s,x); s < T\}$$

then

$$\tau_{\mathcal{S}} = \inf\{t > 0 \; ; \; Y^{s,x}(t,x) \notin \mathcal{S}\} = T - s.)$$

**(ii)** $\lim_{s \to T^-} \phi(s,x) = \lambda x$

**(iv)** $\{\phi^-(Y)(\tau))\}_{\tau \leq \tau_{\mathcal{S}}}$ is uniformly integrable,

**(v)** $A^{\hat{u}}\phi(s,x) + e^{-\delta s} \frac{\hat{u}^\gamma}{\gamma} = 0$ for $s < T$,

in addition to requirements (iii) and (vi).
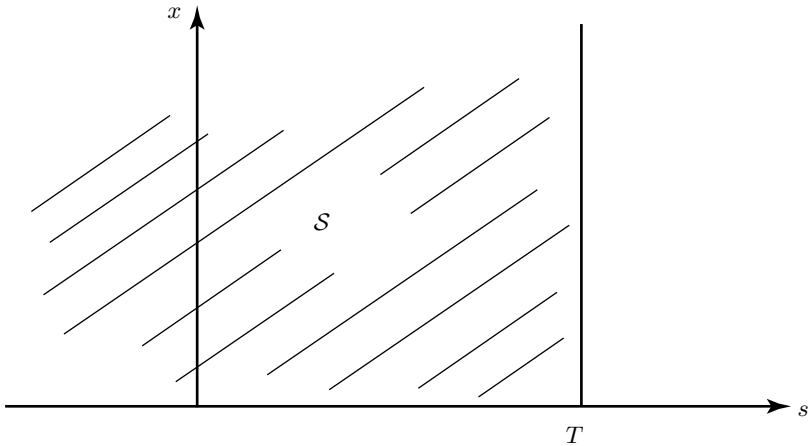


**Fig. 10.3.** The domain $\mathcal{S}$

We try a function $\phi$ of the form

$$\phi(s,x) = h(s) + k(s)x$$

for suitable functions $h(s), k(s)$. Then (i)–(vi) get the form

**(i)'** $h'(s) + k'(s)x + (\mu - \rho x - u)k(s) + e^{-\delta s}\frac{u^\gamma}{\gamma}$
$$+ \int_{\mathbb{R}} \{h(s) + k(s)(x + \gamma z) - h(s) - k(s)x - k(s)\gamma z\}\nu(dz) \leq 0$$

i.e.

$$e^{-\delta s}\frac{u^\gamma}{\gamma} + h'(s) + k'(s)x + (\mu - \rho x - u)k(s) \leq 0 \qquad \text{for all} \quad s < T,\ u \geq 0$$

**(ii)'** $h(T) = 0, \; k(T) = \lambda$

**(v)'** $h'(s) + k'(s)x + (\mu - \rho x - \hat{u})k(s) + e^{-\delta s}\frac{\hat{u}^\gamma}{\gamma} = 0$

**(vi)'** $\{h(\tau) + k(\tau)X(\tau)\}_{\tau \leq \tau_{\mathcal{S}}}$ is uniformly integrable.

From (i)' and (v)' we get

$$-k(s) + e^{-\delta s}\,\hat{u}^{\gamma-1} = 0$$

or
$$\hat{u} = \hat{u}(s) = \left(e^{\delta s}k(s)\right)^{\frac{1}{\gamma-1}}.$$

Combined with (v)' this gives

1) $k'(s) - \rho\,k(s) = 0$   so   $k(s) = \lambda\,e^{\rho(s-T)}$
2) $h'(s) = (\hat{u}(s) - \mu)k(s) - e^{-\delta s}\,\frac{\hat{u}^{\gamma}(s)}{\gamma}$,   $h(T) = 0$

Note that

$$h'(s) = \left(e^{\delta s}k(s)\right)^{\frac{1}{\gamma-1}}k(s) - \mu\,k(s) - e^{-\delta s}\,\frac{\left(e^{\delta s}k(x)\right)^{\frac{\gamma}{\gamma-1}}}{\gamma}$$

$$= e^{\frac{\delta s}{\gamma-1}}\,k(s)^{\frac{\gamma}{\gamma-1}} - \mu\,k(s) - e^{-\delta s(1-\frac{\gamma}{\gamma-1})}\cdot\frac{1}{\gamma}\cdot k(s)^{\frac{\gamma}{\gamma-1}}$$

$$= e^{\frac{\delta s}{\gamma-1}}\,k(s)^{\frac{\gamma}{\gamma-1}}\left[1 - \frac{1}{\gamma}\right] - \mu\,k(s) < 0\,.$$

Hence, since $h(T) = 0$, we have $h(s) > 0$ for $s < T$. Therefore

$$\phi(s,x) = h(s) + k(s)x \geq 0\,.$$

Clearly $\phi$ satisfies (i), (ii), (iv) and (v). It remains to check (vi), i.e. that

$$\left\{h(\tau) + k(\tau)X(\tau)\right\}_{\tau\leq T}$$

is uniformly integrable, and to check (iii).

For these properties to hold some conditions on $\nu$ must be imposed. We omit the details.

We conclude that if these conditions hold then

$$\hat{u}(s) = \lambda^{\frac{1}{\gamma-1}}\exp\left\{\frac{(\delta+\rho)s - \rho\,T}{\gamma-1}\right\};    \qquad s \leq T \qquad\qquad (10.3.1)$$

is the optimal control.

**Exercise 3.2.**

Define

$$J(u) = E\left[\int_0^{T_0} e^{-\delta t}\,\frac{u^{\gamma}(t)}{\gamma}dt + \lambda\,X(T_0)\right]$$

where

$$dX(t) = (\mu - \rho\,X(t) - u(t))dt + \sigma\,B(t) + \gamma\int_{\mathbb{R}} z\,\widetilde{N}(dt,dz); \qquad 0 \leq t \leq T_0\,.$$

The Hamiltonian is

$$H(t,x,u,p,q,r) = e^{-\delta t}\,\frac{u^{\gamma}}{\gamma} + (\mu - \rho\,x - u)p + \sigma\,q + \int_{\mathbb{R}} \gamma\,zr(t,z)\nu(dz).$$

The adjoint equation is

$$\begin{cases} d\hat{p}(t) = \rho\,\hat{p}(t)dt + \sigma\,\hat{q}(t)dB(t) + \displaystyle\int_{\mathbb{R}} \hat{r}(t,z)\widetilde{N}(dt,dz)\,; & t < T_0 \\ \hat{p}(T_0) = \lambda \end{cases}$$

Since $\lambda$ and $\rho$ are deterministic, we guess that $\hat{q} = \hat{r} = 0$ and this gives

$$\hat{p}(t) = \lambda\,e^{\rho(t-T_0)}.$$

Hence

$$H(t, \hat{X}(t), u, \hat{p}(t), \hat{q}(t), \hat{r}(t)) = e^{-\delta t}\tfrac{u^{\gamma}}{\gamma} + (\mu - \rho\,\hat{X}(t) - u)\hat{p}(t),$$

which is maximal when

$$u = \hat{u}(t) = \left(e^{\delta t}\,\hat{p}(t)\right)^{\frac{1}{\gamma-1}} = \lambda^{\frac{1}{\gamma-1}}\exp\left\{\frac{(\delta + \rho)t - \rho\,T_0}{\gamma - 1}\right\}. \tag{10.3.2}$$

## Exercise 3.3.

In this case we have

$$dX(t) = \begin{bmatrix} \displaystyle\int_{\mathbb{R}} u(t^-, \omega)z \\ \displaystyle\int_{\mathbb{R}} z^2 \end{bmatrix} \widetilde{N}(dt, dz) = \begin{bmatrix} \displaystyle\int_{\mathbb{R}} \gamma_1(t, X(t^-), u(t^-), z)\widetilde{N}(dt, dz) \\ \displaystyle\int_{\mathbb{R}} \gamma_2(t, X(t^-), u(t^-), z)\widetilde{N}(dt, dz) \end{bmatrix}$$

so the Hamiltonian is

$$H(t, x, u, p, q, r) = \int_{\mathbb{R}} \{u\,z\,r_1(t,z) + z^2 r_2(t,z)\}\nu(dz)$$

and the adjoint equations are $(g(x_1, x_2) = -(x_1 - x_2)^2)$

$$\begin{cases} dp_1(t) = \displaystyle\int_{\mathbb{R}} r_1(t^-, z)\widetilde{N}(dt, dz)\,; & t < T \\ p_1(T) = -2(X_1(T) - X_2(T)) \end{cases}$$

$$\begin{cases} dp_2(t) = \displaystyle\int_{\mathbb{R}} r_2(t^-, z)\widetilde{N}(dt, dz) \\ p_2(T) = 2(X_1(T) - X_2(T)). \end{cases}$$

Now $X_1(T) - X_2(T) = \displaystyle\int_0^T\int_{\mathbb{R}} \{u(t^-) - z\}\,\widetilde{N}(dt, dz)$. So if $\hat{u}$ is a given candidate for an optimal control we get

$$\hat{r}_1(t, z) = -2(\hat{u}(t) - z)z$$
$$\hat{u}_2(t, z) = 2(\hat{u}(t) - z)z.$$

This gives

$$H(t,x,u,\hat{p},\hat{q},\hat{r}) = \int_{\mathbb{R}} \{u\, z(-2(\hat{u}(t) - z)z) + z^2 2(\hat{u}(t) - z)z\}\nu(dz)$$

$$= -2u \int_{\mathbb{R}} \{\hat{u}(t)z^2 - z^3\}\nu(dz) + 2 \int_{\mathbb{R}} \{\hat{u}(t)z^3 - z^4\}\nu(dz).$$

This is a linear expression in $u$, so we guess that the coefficient of $u$ is 0, i.e. that

$$\hat{u}(t) = \frac{\int_{\mathbb{R}} z^3 \nu(dz)}{\int_{\mathbb{R}} z^2 \nu(dz)} \qquad \text{for all } (t,\omega) \in [0,T] \times \Omega. \qquad (10.3.3)$$

With this choice of $\hat{u}(t)$ all the conditions of the stochastic maximum principle are satisfied and we conclude that $\hat{u}$ is optimal.

Note that this implies that

$$\inf_u E\left[\left(F - \int_0^T u(t)dS_1(t)\right)^2\right] = E\left[\left(\int_0^T \int_{\mathbb{R}} \{z^2 - \hat{u}(t)z\}\widetilde{N}(dt,dz)\right)^2\right]$$

$$= \int_0^T \int_{\mathbb{R}} E[(z^2 - \hat{u}(t)z)^2]\nu(dz)dt$$

$$= T \int_{\mathbb{R}} \left[z^2 - \frac{\int_{\mathbb{R}} z^3 \nu(dz)}{\int_{\mathbb{R}} z^2 \nu(dz)} z\right]^2 \nu(dz).$$

We see that this is 0 if and only if

$$\int_{\mathbb{R}} z^3 \nu(dz) = z \int_{\mathbb{R}} z^2 \nu(dz) \qquad \text{for a.a. } z(\nu) \qquad (10.3.4)$$

i.e. iff $\nu$ is supported on one point $\{z_0\}$. *Only then is the market complete*! See [BDLØP] for more information.

## Exercise 3.4

We try to find $a, b$ such that the function

$$\varphi(s,x) = e^{-\rho s}\psi(x) := e^{-\rho s}(ax^2 + b)$$

satisfies the conditions of (the minimum version of) Theorem 3.1. In this case the generator is

$$A_\varphi^v(s,x) = e^{-\rho s} A_0^v \psi(x),$$

where

$$A_0^v \psi(x) = -\rho\psi(x) + v\psi'(x) + \frac{1}{2}\sigma^2 \psi''(x)$$

$$+ \int_{\mathbb{R}} \{\psi(x+z) - \psi(x) - z\psi'(x)\}\, \nu(dz).$$

Hence condition (i) of Theorem 3.1 becomes

$$A_0^v \psi(x) + x^2 + \theta v^2 =$$

$$- \rho(ax^2 + b) + v2ax + \frac{1}{2}\sigma^2 2a + a \int_{\mathbb{R}} z^2 \nu(dz) + x^2 + \theta v^2$$

$$= \theta v^2 + 2axv + x^2(1 - \rho a) + a\left(\sigma^2 + \int_{\mathbb{R}} z^2 \nu(dz)\right) - \rho b =: h(v).$$

The function $h$ is minimal when

$$v = u^*(x) = -\frac{ax}{\theta}. \qquad (10.3.5)$$

With this value of $v$ condition (v) becomes

$$x^2 \left[1 - \rho a - \frac{a^2}{\theta}\right] + a\left(\sigma^2 + \int_{\mathbb{R}} z^2 \nu(dz)\right) - \rho b = 0.$$

Hence we choose $a > 0$ and $b$ such that

$$a^2 + \rho\theta a - \theta = 0 \qquad (10.3.6)$$

and

$$b = \frac{a}{\rho}\left(\sigma^2 + \int_{\mathbb{R}} z^2 \nu(dz)\right). \qquad (10.3.7)$$

With these values of $a$ and $b$ we can easily check that

$$\varphi(s, x) := e^{-\rho s}(ax^2 + b)$$

satisfies all the conditions of Theorem 3.1. The corresponding optimal control is given by (10.3.5).

## Exercise 3.5

**b)** The Hamiltonian for this problem is

$$H(t, x, u, p, q, r) = x^2 + \theta u^2 + up + \sigma q + \int_{\mathbb{R}} r(t^-, z)\tilde{N}(dt, dz).$$

The adjoint equation is

$$\begin{cases} dp(t) = -2X(t)dt + q(t)dB(t) + \int_{\mathbb{R}} r(t^-, z)\tilde{N}(dt, dz) \; ; \; t < T \\ p(T) = 2\lambda X(T). \end{cases} \qquad (10.3.8)$$

By imposing the first and second order conditions, we see that $H(t, x, u, p, q, r)$ is minimal for

$$u = u(t) = \hat{u}(t) = -\frac{1}{2\theta}p(t). \qquad (10.3.9)$$

In order to find a solution of (10.3.8), we consider $p(t) = h(t)X(t)$, where $h : \mathbb{R} \to \mathbb{R}$ is a deterministic function such that

$$h(T) = 2\lambda.$$

Note that $u(t) = -\dfrac{h(t)X(t)}{2\theta}$ and

$$dX(t) = -\frac{h(t)X(t)}{2\theta}dt + \sigma dB(t) + \int_{\mathbb{R}} z\tilde{N}(dt, dz) \; ; \; X(0) = x.$$

Moreover, (10.3.8) turns into

$$dp(t) = h(t)dX(t) + X(t)h'(t)dt$$
$$= X(t)\Big[-\frac{h(t)^2}{2\theta} + h'(t)\Big]dt + h(t)\sigma dB(t) + h(t)\int_{\mathbb{R}} z\tilde{N}(dt, dz).$$

Hence $h(t)$ is the solution of

$$\begin{cases} h'(t) = \frac{h(t)^2}{2\theta} - 2 \; ; \; t < T \\ h(T) = 2\lambda. \end{cases} \tag{10.3.10}$$

The general solution of (10.3.10) is

$$h(t) = 2\sqrt{\theta}\,\frac{1 + \beta e^{\frac{2t}{\sqrt{\theta}}}}{1 - \beta e^{\frac{2T}{\sqrt{\theta}}}} \tag{10.3.11}$$

with $\beta = \frac{\lambda - \sqrt{\theta}}{\lambda + \sqrt{\theta}} e^{-\frac{2T}{\sqrt{\theta}}}$. By using the stochastic maximum principle, we can conclude that

$$u^*(t) = -\frac{h(t)}{2\theta}X(t)$$

is the optimal control, $p(t) = h(t)X(t)$ and $q(t) = \sigma h(t)$, $r(t^-, z) = h(t)z$, where $h(t)$ is given by (10.3.11).

## Exercise 3.6.

If we try a function of the form

$$\varphi(s, x) = e^{-\delta s}\psi(x)$$

then equations (i) and (v) for Theorem 3.1 combine to give the equation

$$\sup_{c \geq 0} \Big\{ \ln c - \delta\psi(x) + (\mu x - c)\psi'(x) + \frac{1}{2}\sigma^2 x^2 \psi''(x)$$
$$+ \int_{\mathbb{R}} \{\psi(x + x\theta z) - \psi(x) - x\theta z\psi'(x)\}\, \nu(dz) \Big\} = 0.$$

The function

$$h(c) := \ln c - c\psi'(x) \; ; \; c > 0$$

is maximal when

$$c = \hat{c}(x) = \frac{1}{\psi'(x)}.$$

If we set

$$\psi(x) = a \ln x + b$$

where $a, b$ are constants, $a > 0$, then this gives

$$\hat{c}(x) = \frac{x}{a},$$

and hence the above equation becomes

$$\ln x - \ln a - \delta(a \ln x + b) + \mu x \cdot \frac{a}{x} - 1 + \frac{1}{2}\sigma^2 x^2 \left(-\frac{a}{x^2}\right)$$

$$+ a \int_{\mathbb{R}} \left\{\ln(x + x\theta z) - \ln x - x\theta z \cdot \frac{1}{x}\right\} \nu(dz) = 0$$

or

$$(1 - \delta a) \ln x - \ln a - \delta b + \mu a - 1 - \frac{1}{2}\sigma^2 a$$

$$+ a \int_{\mathbb{R}} \{\ln(1 + \theta z) - \theta z\} \nu(dz) = 0, \text{ for all } x > 0.$$

This is possible if and only if

$$a = \frac{1}{\delta}$$

and

$$b = \delta^{-2} \left[\delta \ln \delta - \delta + \mu - \frac{1}{2}\sigma^2 + \int_{\mathbb{R}} \{\ln(1 + \theta z) - \theta z\} \nu(dz)\right].$$

One can now verify that if $\delta > \mu$ then with these values of $a$ and $b$ the function

$$\varphi(s, x) = e^{-\delta t}(a \ln x + b)$$

satisfies all the conditions of Theorem 3.1. We conclude that

$$\Phi(s, x) = e^{-\delta t}(a \ln x + b)$$

and that

$$c^*(x) = \hat{c} = \frac{x}{a}$$

(in feedback form) is an optimal consumption rate.

## 10.4 Exercises of Chapter 4

### Exercise 4.1.

**a)** The HJB equation, i.e. (vi) and (ix) of Theorem 4.2, for this problem gets the form

$$0 = \sup_{u \geq 0} \left\{ e^{-\delta s} \frac{u^\gamma}{\gamma} + \frac{\partial \phi}{\partial s} + (\mu x - u)\frac{\partial \phi}{\partial x} + \frac{\sigma^2 x^2}{2}\frac{\partial^2 \phi}{\partial x^2} \right.$$
$$\left. \int_{\mathbb{R}} \left\{\phi(s, x + \theta x z) - \phi(s, x) - \theta x z \frac{\partial \phi}{\partial x}(s, x)\right\} d\nu(z)\right\} \tag{10.4.1}$$

for $x > 0$.

We impose the first order conditions to find the supremum, which is obtained for

$$u = u^*(s, x) = \left(e^{\delta s} \frac{\partial \phi}{\partial x}\right)^{\frac{1}{\gamma-1}}. \tag{10.4.2}$$

We guess that $\phi(s, x) = Ke^{-\delta s}x^{\gamma}$ with $K > 0$ to be determined. Then

$$u^*(s, x) = (K\gamma)^{\frac{1}{\gamma-1}} x \tag{10.4.3}$$

and (10.4.1) turns into

$$\frac{1}{\gamma}(K\gamma)^{\frac{\gamma}{\gamma-1}} - K\delta + \left(\mu - (K\gamma)^{\frac{1}{\gamma-1}}\right) K\gamma + \frac{1}{2}\sigma^2 K\gamma(\gamma-1)$$

$$+ K \int_{\mathbb{R}} \{(1 + \theta z)^{\gamma} - 1 - \gamma\theta z\} \, \nu(dz) = 0$$

or

$$\gamma^{\frac{\gamma}{\gamma-1}} K^{\frac{1}{\gamma-1}} - \delta + \mu\gamma - \gamma^{\frac{\gamma}{\gamma-1}} K^{\frac{1}{\gamma-1}} + \frac{1}{2}\sigma^2\gamma(\gamma-1)$$

$$+ \int_{\mathbb{R}} \{(1 + \theta z)^{\gamma} - 1 - \gamma\theta z\} \, \nu(dz) = 0.$$

Hence

$$K = \frac{1}{\gamma}\left[\frac{1}{1-\gamma}\left(\delta - \mu\gamma + \frac{\sigma^2}{2}\gamma(1-\gamma) - \int_{\mathbb{R}} \{(1 + \theta z)^{\gamma} - 1 - \gamma\theta z\} \nu(dz)\right)\right]^{\gamma-1} \tag{10.4.4}$$

provided that

$$\delta - \mu\gamma + \frac{\sigma^2}{2}\gamma(1-\gamma) - \int_{\mathbb{R}} \{(1 + \theta z)^{\gamma} - 1 - \gamma\theta z\} \, \nu(dz) > 0.$$

With this choice of $K$ the conditions of Theorem 4.2 are satisfied and we can conclude that $\phi = \Phi$ is the value function.

**b)**
(i) First assume $\lambda \geq K$. Choose $\phi(s, x) = \lambda e^{-\delta s}x^{\gamma}$. By the same computations as in a), condition (vi) of Theorem 4.2 gets the form

$$\lambda \geq \frac{1}{\gamma}\left[\frac{1}{\gamma-1}\left(\delta - \mu\gamma + \frac{1}{2}\sigma^2\gamma(1-\gamma) - \int_{\mathbb{R}} \{(1 + \theta z)^{\gamma} - 1 - \gamma\theta z\} \nu(dz)\right)\right]^{\gamma-1}. \tag{10.4.5}$$

Since $\lambda \geq K$, the inequality (10.4.5) holds by (10.4.4).
By Theorem 4.2a), it follows that

$$\phi(s, x) = \lambda e^{-\delta s}x^{\gamma} \geq \Phi(s, x)$$

where $\Phi$ is the value function for our problem. On the other hand, $\phi(s, x)$ is obtained by the (admissible) control of stopping immediately ($\tau = 0$). Hence we also have

$$\phi(s, x) \leq \Phi(s, x).$$

We conclude that

$$\Phi(s, x) = \lambda e^{-\delta s}x^{\gamma}$$

in this case and $\tau^* = 0$ is optimal. Note that $D = \emptyset$.

(ii) Assume now $\lambda < K$. Choose $\phi(s,x) = Ke^{-\delta s}x^\gamma$. Then for all $(s,x) \in \mathbb{R} \times (0,\infty)$ we have

$$\phi(s,x) > \lambda e^{-\delta s}x^\gamma.$$

Hence we have $D = \mathbb{R} \times (0,\infty)$ and by Theorem 4.2a) we conclude that

$$\Phi(s,x) \le Ke^{-\delta s}x^\gamma.$$

On the other hand, we have seen in a) above that if we apply the control

$$u^*(s,x) = \left(K\gamma\right)^{\frac{1}{\gamma-1}}x$$

and never stop, then we achieve the performance $J^{(u^*)}(s,x) = Ke^{-\delta s}x^\gamma$. Hence

$$\Phi(s,x) = Ke^{-\delta s}x^\gamma$$

and it is optimal never to stop ($\tau^* = \infty$).


## 10.5 Exercises of Chapter 5

### Exercise 5.1.

In this case we put

$$dY(t) = \begin{bmatrix} dt \\ dX(t) \end{bmatrix} = \begin{bmatrix} 1 \\ \alpha \end{bmatrix} dt + \begin{bmatrix} 0 \\ \sigma \end{bmatrix} dB(t) + \begin{bmatrix} 0 \\ \beta \int\limits_{\mathbb{R}} z\, \tilde{N}(dt,dz) \end{bmatrix} + \begin{bmatrix} 0 \\ -(1+\lambda) \end{bmatrix} d\xi(t).$$

The generator if $\xi = 0$ is

$$A\phi = \frac{\partial \phi}{\partial s} + \alpha\frac{\partial \phi}{\partial x} + \tfrac{1}{2}\sigma^2\frac{\partial^2 \phi}{\partial x^2} + \int\limits_{\mathbb{R}} \left\{ \phi(s,x+\beta z) - \phi(s,x) - \beta z\frac{\partial \phi}{\partial x}(s,x) \right\}\nu(dz).$$

The non-intervention region $D$ is described by (see (5.2.5)

$$D = \left\{ (s,x); \sum_{i=1}^{k} \kappa_{ij}\frac{\partial \phi}{\partial y_i}(y) + \theta_j < 0 \quad \text{for all } j = 1,\dots,p \right\}$$

$$= \left\{ (s,x); -(1+\lambda)\frac{\partial \phi}{\partial x}(s,x) + e^{-\rho s} < 0 \right\}.$$

If we guess that $D$ has the form

$$D = \{(s,x); 0 < x < x^*\} \qquad \text{for some } x^* > 0$$

then by Theorem 5.2 we should have

$$A\phi(s,x) = 0 \qquad \text{for } 0 < x < x^*.$$

We try a solution $\phi$ of the form

$$\phi(s,x) = e^{-\rho s}\psi(x)$$

and get

$$A_0\psi(x) := -\rho\,\psi(x) + \alpha\,\psi'(x) + \tfrac{1}{2}\sigma^2\psi''(x) + \int_{\mathbb{R}} \{\psi(x+\beta\,z) - \psi(x) - \beta\,z\,\psi'(x)\}\nu(dz) = 0.$$

We now choose
$$\psi(x) = e^{r\,x} \qquad \text{for some constant } r \in \mathbb{R}$$

and get the equation

$$h(r) := -\rho + \alpha\,r + \tfrac{1}{2}\sigma^2 r^2 + \int_{\mathbb{R}} \{e^{r\,\beta\,z} - 1 - r\,\beta\,z\}\nu(dz) = 0.$$

Since $h(0) < 0$ and $\lim_{r\to\infty} h(r) = \lim_{r\to-\infty} h(r) = \infty$, we see that the equation $h(r) = 0$ has two solutions $r_1, r_2$ such that

$$r_2 < 0 < r_1\,.$$

Outside $D$ we require that

$$-(1+\lambda)\psi'(x) + 1 = 0$$

or
$$\psi(x) = \frac{x}{1+\lambda} + C_3\,, \qquad C_3 \text{ constant.}$$

Hence we put

$$\psi(x) = \begin{cases} C_1 e^{r_1 x} + C_2 e^{r_2 x} & ; \quad 0 < x < x^* \\ \frac{x}{1+\lambda} + C_3 & ; \quad x^* \le x \end{cases} \tag{10.5.1}$$

where $C_1, C_2$ are constants.

To determine $C_1, C_2, C_3$ and $x^*$ we have the four equations:

$$\psi(0) = 0 \;\Rightarrow\; C_1 + C_2 = 0. \tag{10.5.2}$$

Put $C_2 = -C_1$

$\psi$ continuous at $x = x^* \Rightarrow$
$$C_1(e^{r_1 x^*} - e^{r_2 x^*}) = \frac{x^*}{1+\lambda} + C_3 \tag{10.5.3}$$

$\psi \in C^1$ at $x = x^* \Rightarrow$
$$C_1(r_1 e^{r_1 x^*} - r_2 e^{r_2 x^*}) = \frac{1}{1+\lambda} \tag{10.5.4}$$

$\psi \in C^2$ at $x = x^* \Rightarrow$
$$C_1(r_1^2 e^{r_1 x^*} - r_2^2 e^{r_2 x^*}) = 0. \tag{10.5.5}$$

From (10.5.4) and (10.5.5) we deduce that

$$x^* = \frac{2(\ln|r_2| - \ln r_1)}{r_1 - r_2}\,. \tag{10.5.6}$$

Then by (10.5.4) we get the value for $C_1$, and hence the value of $C_3$ by (10.5.3).

With these values of $C_1, C_2, C_3$ and $x^*$ we must verify that $\phi(s,x) = e^{-\rho s}\psi(x)$ satisfies all the requirements of Theorem 5.2:

**(i)**: We have constructed $\phi$ such that $A\phi + f = 0$ in $D$. Outside $D$, i.e. for $x \geq x^*$, we have

$$e^{\rho s}(A\phi(s,x) + f(s,x) = A_0\psi(x) = -\rho\left(\frac{x}{1+\lambda} + C_3\right) + \alpha \cdot \frac{1}{1+\lambda}$$

$$= -\frac{\rho}{1+\lambda}x + \frac{\alpha}{1+\lambda} - \rho C_3 , \qquad \text{which is decreasing in } x.$$

So we need only to check that this holds for $x = x^*$, i.e. that

$$A_0\psi(x^*) \leq 0.$$

But this follows from the fact that $A_0\psi(x) = 0$ for all $x < x^*$ and $\psi \in C^2$.

**(ii)**: By construction we have

$$-(1+\lambda)\psi'(x) + 1 = 0 \qquad \text{for } x \geq x^*.$$

For $x < x^*$ the condition (ii) gets the form

$$F(x) := -(1+\lambda)C_1(r_1^{r_1 x} - r_2^{r_2 x}) + 1 \leq 0 .$$

We know that $F(x^*) = 0$ by (10.5.4) and

$$F'(x) = -(1+\lambda)C_1(r_1^2 e^{r_1 x} - r_2^2 e^{r_2 x}).$$

So $F'(x^*) = 0$ by (10.5.5) and hence (since $C_1 > 0$)

$$F'(x) > F'(x^*) = 0 \qquad \text{for } x < x^*.$$

Hence

$$F(x) < 0 \qquad \text{for } 0 < x < x^*.$$

The conditions (iii), (iv) and (v) are left to the reader to verify.

**(vi)**: This holds by construction of $\phi$.

**(vii–(x))**: These conditions claim the existence of an increasing process $\hat{\xi}$ such that $Y^{\hat{\xi}}(t)$ stays in $\bar{D}$ for all times $t$, $\hat{\xi}(t)$ is strictly increasing only when $Y(t) \notin D$, and if $Y(t) \notin \bar{D}$ then $\hat{\xi}(t)$ brings $Y(t)$ down to a point on $\partial D$. Such a singular control is called a *local time* at $\partial D$ of the process $Y(t)$ reflected downwards at $\partial D$. The existence and uniqueness of such a local time is proved in [CEM].

**(xi)**: This is left to the reader.

We conclude that the optimal dividend policy $\xi^*(t)$ is to take out exactly the amount of money needed to keep $X(t)$ on or below the value $x^*$. If $X(t) < x^*$ we take out nothing. If $X(t) > x^*$ we take out $X(t) - x^*$.

## Exercise 5.2

It suffices to prove that the function

$$\Phi_0(s, x_1, x_2) := Ke^{-\delta s}(x_1 + x_2)^\gamma$$

satisfies conditions (i)-(iv) of Theorem 5.2. In this case we have (see Section 5.3)

$$A^{(v)}\Phi_0(y) = A^{(c)}\Phi_0(y) = \frac{\partial\Phi_0}{\partial s} + (rx_1 - c)\frac{\partial\Phi_0}{\partial x_1} + \alpha x_2\frac{\partial\Phi_0}{\partial x_2} + \frac{1}{2}\beta^2 x_2^2\frac{\partial^2\Phi_0}{\partial x_2^2}$$

$$+ \int_{\mathbb{R}}\left\{\Phi_0(s, x_1, x_2 + x_2 z) - \Phi_0(s, x_1, x_2) - x_2 z\frac{\partial\Phi_0}{\partial x_2}(s, x_1, x_2)\right\}\nu(dz)$$

and $f(s, x_1, x_2, c) = e^{-\delta s}\dfrac{c^\gamma}{\gamma}$, so condition (i) becomes

(i)'     $A^c\Phi_0(s, x_1, x_2) + e^{-\delta s}\dfrac{c^\gamma}{\gamma} \leq 0$ for all $c \geq 0$.

This holds because we know by Example 3.2 that (see (3.1.21))

$$\sup_{c \geq 0}\left\{A^{(c)}\Phi(s, x_1, x_2) + e^{-\delta s}\frac{c^\gamma}{\gamma}\right\} = 0.$$

Since in this case $\theta = 0$ and

$$\kappa = \begin{bmatrix} -(1+\lambda) & 1-\mu \\ 1 & -1 \end{bmatrix}$$

we see that condition (ii) of Theorem 5.2 becomes

(ii)'     $- (1+\lambda)\dfrac{\partial\Phi_0}{\partial x_1} + \dfrac{\partial\Phi_0}{\partial x_2} \leq 0$

(ii)"     $(1-\mu)\dfrac{\partial\Phi_0}{\partial x_1} - \dfrac{\partial\Phi_0}{\partial x_2} \leq 0.$

Since

$$\frac{\partial\Phi_0}{\partial x_1} = \frac{\partial\Phi_0}{\partial x_2} = Ke^{-\delta s}\gamma(x_1 + x_2)^{\gamma-1}$$

we see that (ii)' and (ii)" hold trivially.

We leave the verification of conditions (iii)-(v) to the reader.

## 10.6 Exercises of Chapter 6

**Exercise 6.1.**

By using the same notation as in Chapter 6, we have here

$$Y^{(v)}(t) = \begin{bmatrix} s+t \\ X^{(v)}(t) \end{bmatrix} \; ; \; t \geq 0 \; ; \; Y^{(v)}(0^-) = \begin{bmatrix} s \\ x \end{bmatrix} = y \in \mathbb{R}^2$$

$$\Gamma(y, \zeta) = \Gamma(s, x, \zeta) = \begin{bmatrix} s \\ x+\zeta \end{bmatrix} \; ; \; (s, x, \zeta) \in \mathbb{R}^3$$

$$K(y, \zeta) = K(s, x, \zeta) = e^{-\rho s}(x + \lambda|\zeta|)$$

$$f(y) = f(s, x) = e^{-\rho s}x^2, \; g(y) = 0.$$

By symmetry we expect the continuation region to be of the form

$$D = \{(s, x) : -\bar{x} < x < \bar{x}\}$$

for some $\bar{x} > 0$, to be determined.

As soon as $X(t)$ reaches the unknown value $\bar{x}$ or $-\bar{x}$, there is an intervention and $X(t)$ is brought down (or up) to a certain value $\hat{x}$ (or $-\hat{x}$) where $-\bar{x} < -\hat{x} < 0 < \hat{x} < \bar{x}$. We determine $\bar{x}$ and $\hat{x}$ in the following computations.
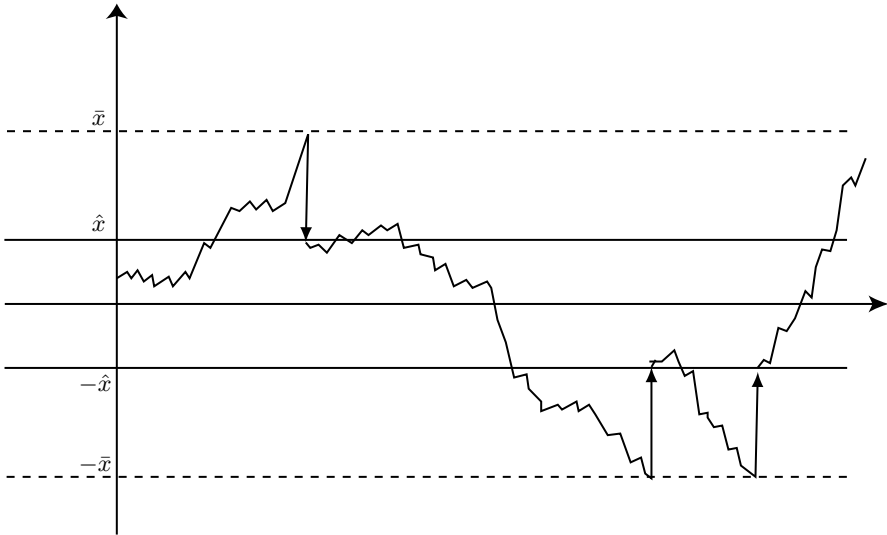


**Fig. 10.4.** The optimal strategy of Exercise 6.1

We guess that the value function is of the form

$$\phi(s, x) = e^{-\rho s} \psi(x).$$

In the continuation region $D$, we have by Theorem 6.2 (x)

$$A\phi + f = 0 \tag{10.6.1}$$

where $A$ is the generator of $Y$, i.e.

$$A\phi(s, x) = \frac{\partial \phi}{\partial s} + \frac{1}{2} \frac{\partial^2 \phi}{\partial x^2} + \int_{\mathbb{R}} \left\{ \phi(s, x + z) - \phi(s, x) - z \frac{\partial \phi}{\partial x}(s, x) \right\} \nu(dz).$$

In this case, equation (10.6.1) becomes

$$A_0\psi(x) + f(x) := -\rho\psi(x) + \frac{1}{2}\psi''(x) + \int_{\mathbb{R}} \{\psi(x + z) - \psi(x) - z\psi'(x)\}\nu(dz) + x^2 = 0.$$

We try a solution of the form

$$\psi(x) = C\cosh(\gamma x) + \frac{1}{\rho}x^2 + \frac{b}{\rho^2}$$

where $C$ is a constant (to be determined), $b = 1 + \int_{\mathbb{R}} z^2\nu(dz)$ and $\gamma > 0$ is the positive solution of the equation

$$F(\gamma) := -\rho + \frac{1}{2}\gamma^2 + \int_{\mathbb{R}}\{e^{\gamma z} - 1 - \gamma z\}\nu(dz) = 0.$$

Note that if we make no intervention at all, the value of $J^{(v)}(s,x)$ is

$$J^{(v)}(s,x) = e^{-\rho s}E^x\left[\int_0^\infty e^{-\rho t}\left(x + B(t) + \int_0^t\int_{\mathbb{R}} z\tilde{N}(ds,dz)\right)^2 dt\right]$$
$$= e^{-\rho s}\left(\frac{x^2}{\rho} + \frac{b}{\rho^2}\right).$$

Hence

$$0 \le \psi(x) \le \frac{x^2}{\rho} + \frac{b}{\rho^2}. \tag{10.6.2}$$

By (10.6.2) we obtain $C = -a$ where $a > 0$. We define

$$\psi_0(x) := \frac{1}{\rho}x^2 + \frac{b}{\rho^2} - a\cosh(\gamma x)$$

and put

$$\psi(x) = \psi_0(x) \; ; \; x \in D.$$

We recall that

$$D = \{(s,x) : \phi(s,x) < \mathcal{M}\phi(s,x)\} =$$
$$= \{x : \psi(x) < \mathcal{M}\psi(x)\}$$

and the intervention operator is in this case

$$\mathcal{M}\psi(x) = \inf\{\psi(x+\zeta) + c + \lambda|\zeta|; \zeta \in \mathbb{R}\}.$$

The first order condition for a minimum $\hat{\zeta} = \hat{\zeta}(x)$ of the function

$$G(\zeta) = \begin{cases} \psi(x+\zeta) + c + \lambda\zeta & \zeta > 0 \\ \psi(x+\zeta) + c - \lambda\zeta & \zeta < 0 \end{cases}$$

is the following

(i) $\zeta > 0$:   $\psi'(x+\zeta) + \lambda = 0 \Rightarrow \psi'(x+\zeta) = -\lambda$
(ii) $\zeta < 0$:   $\psi'(x+\zeta) - \lambda = 0 \Rightarrow \psi'(x+\zeta) = \lambda$.

Hence we look for points $\hat{x}, \bar{x}$ such that

$$-\bar{x} < -\hat{x} < 0 < \hat{x} < \bar{x}$$

and

$$\psi'(\hat{x}) = -\lambda$$
$$\psi'(-\hat{x}) = \lambda. \tag{10.6.3}$$

Note that since $\hat{x} < \bar{x}$, $\psi'(\hat{x}) = \psi_0'(\hat{x})$.

Arguing as in Example 6.5, we put

$$\psi(x) = \begin{cases} \psi_0(x) & ; \ -\bar{x} \leq x \leq \bar{x} \\ \psi_0(\hat{x}) + c + \lambda(x - \hat{x}) & ; \ x > \bar{x} \\ \psi_0(-\hat{x}) + c - \lambda(x + \hat{x}) & ; \ x < -\bar{x}. \end{cases} \tag{10.6.4}$$

We have to show that there exist $0 < \hat{x} < \bar{x}$ and a value of $a$ such that $\phi(s, x) := e^{-\rho s}\psi(x)$ satisfies all the requirements of (the minimum version of) Theorem 6.2. By symmetry we may assume $x > 0$ and $\zeta > 0$ in the following.

Continuity at $x = \bar{x}$ gives the equation

$$\psi_0(\hat{x}) + c + \lambda(\bar{x} - \hat{x}) = \psi_0(\bar{x}).$$

Differentiability at $x = \bar{x}$ gives the equation

$$\lambda = \psi_0'(\bar{x}).$$

Substituting for $\psi_0$ these equations give

$$\frac{\hat{x}^2}{\rho} - a\cosh(\gamma\hat{x}) - \lambda\hat{x} + c = \frac{\bar{x}^2}{\rho} - a\cosh(\gamma\bar{x}) - \gamma\bar{x} \tag{10.6.5}$$

and

$$\lambda = \frac{2\bar{x}}{\rho} - a\gamma\sinh(\gamma\bar{x}). \tag{10.6.6}$$

In addition we have required

$$\lambda = \psi_0'(\hat{x}) = \frac{2\hat{x}}{\rho} - a\gamma\sinh(\gamma\hat{x}). \tag{10.6.7}$$

As in Example 6.5 one can prove that for each $c > 0$ there exist $a = a^*(c) > 0$, $\hat{x} = \hat{x}(x) > 0$ and $\bar{x} = \bar{x}(c) > \hat{x}$ such that (10.6.4)-(10.6.6) hold. With these values of $a, \hat{x}$ and $\bar{x}$ it remains to verify that the conditions of Theorem 6.2 hold. We check some of them:

**(ii):** $\psi \leq \mathcal{M}\psi = \inf\{\psi(x - \zeta) + c + \lambda\zeta \ ; \ \zeta > 0\}$.

*First suppose $x \geq \bar{x}$.*
    If $x - \zeta \geq \bar{x}$ then

$$\psi(x - \zeta) + c + \lambda\zeta = \psi_0(\hat{x}) + c + \lambda(x - \zeta - \hat{x}) + c + \lambda\zeta = c + \psi(x) > \psi(x).$$

If $0 < x - \zeta < \bar{x}$ then

$$\psi(x - \zeta) + c + \lambda\zeta = \psi_0(x - \zeta) + c + \lambda\zeta,$$

which is minimal when

$$-\psi_0'(x - \zeta) + \lambda = 0$$

i.e. when

$$\zeta = \hat{\zeta} = x - \hat{x}.$$

This is the minimum point because
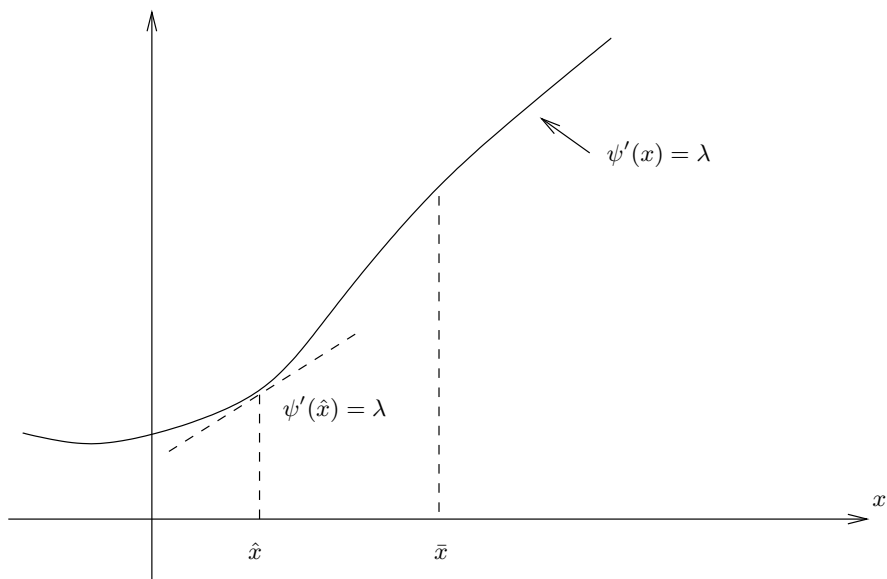
**Fig. 10.5.** The function $\psi(x)$ for $x > 0$

$$\psi_0''(\hat{x}) > 0.$$

See Figure 10.5.

This shows that

$$\mathcal{M}\psi(x) = \psi(x - \hat{\zeta}) + c + \lambda\hat{\zeta} = \psi(\hat{x}) + c + \lambda(x - \hat{x}) = \psi(x)$$

for $x > \hat{x}$.

*Next suppose* $0 < x < \bar{x}$.

Then

$$\mathcal{M}\psi(x) = \psi_0(\hat{x}) + c + \lambda(x - \hat{x}) > \psi(x)$$

if and only if

$$\psi(x) - \lambda x < \psi(\hat{x}) - \lambda\hat{x} + c.$$

Now the minimum of

$$H(x) := \psi(x) - \lambda x \quad \text{for } 0 < x < \bar{x}$$

is attained when

$$\psi'(x) = \lambda$$

i.e. when $x = \hat{x}$.

Therefore

$$\psi(x) - \lambda x \le \psi(\hat{x}) - \lambda\hat{x} < \psi(\hat{x}) - \lambda\hat{x} + c.$$

This shows that $\mathcal{M}\psi(x) > \psi(x)$ for all $0 < x < \bar{x}$.

Combined with the above we can conclude that

$$\mathcal{M}\psi(x) \geq \psi(x) \text{ for all } x > 0,$$

which proves (ii). Moreover,

$$\mathcal{M}\psi(x) > \psi(x) \text{ if and only if } 0 < x < \bar{x}.$$

Hence

$$D \cap (0, \infty) = (0, \bar{x}).$$

Finally we verify

**(vi):** $A\phi + f \geq 0$ for $x > \bar{x}$.

For $x > \bar{x}$, we have

$$A_0\psi(x) + f(x) = -\rho(\psi_0(\hat{x}) + c + \lambda(x - \hat{x})) + x^2.$$

This is nonnegative for all $x > \bar{x}$ iff it is nonnegative for $x = \bar{x}$, i.e. iff

$$-\rho\psi_0(\bar{x}) + \bar{x}^2 \geq 0. \tag{10.6.8}$$

By construction of $\psi_0$ we know that, for $x < \bar{x}$

$$-\rho\psi_0(x) + \frac{1}{2}\psi_0''(x) + \int_{\mathbb{R}}\{\psi_0(x+z) - \psi_0(x) - z\psi_0'(x)\}\nu(dz) + x^2 = 0.$$

Therefore (10.6.8) holds iff

$$\frac{1}{2}\psi_0''(\bar{x}) + \int_{\mathbb{R}}\{\psi_0(\bar{x}+z) - \psi_0(\bar{x}) - z\psi_0'(\bar{x})\}\nu(dz) \leq 0.$$

For this it suffices that

$$\int_{\mathbb{R}} z^2\nu(dz) \leq -\frac{\rho}{2}\psi_0''(\bar{x}). \tag{10.6.9}$$

**Conclusion.**

Suppose (10.6.9) holds. Then $\Phi(s, x) = e^{-\rho s}\psi(x)$, with $\psi(x)$ given by (10.6.4) and $a, \hat{x}, \bar{x}$ given by (10.6.5)-(10.6.7). The optimal impulse control is to do nothing while $|X(t)| < \bar{x}$, then move $X(t)$ down to $\hat{x}$ (respectively up to $-\hat{x}$) as soon as $X(t)$ reaches a value $\geq \bar{x}$ (respectively a value $\leq -\bar{x}$).

**Exercise 6.2.**

Here we put

$$Y^{(v)}(t) = \begin{bmatrix} s+t \\ X^{(v)}(t) \end{bmatrix}$$

$$Y^{(v)}(0^-) = \begin{bmatrix} s \\ x \end{bmatrix} = y$$

$$\Gamma(y, \zeta) = x - c - (1+\lambda)\zeta$$

$$K(y, \zeta) = e^{-\rho s}\zeta$$

$$f \equiv g \equiv 0$$

$$\mathcal{S} = \{(s, x) : x > 0\}.$$

We guess that the value function $\phi$ is of the form

$$\phi(s, x) = e^{-\rho s} \psi(x)$$

and consider the intervention operator

$$\mathcal{M}\psi(x) = \sup \left\{ \psi(x - c - (1 + \lambda)\zeta) + \zeta; 0 \leq \zeta \leq \frac{x - c}{1 + \lambda} \right\}. \tag{10.6.10}$$

Note that the condition on $\zeta$ is due to the fact that the impulse must be positive and $x - c - (1 + \lambda)\zeta$ must belong to $\mathcal{S}$. We distinguish between two cases:

1) $\mu > \rho$.
   In this case, suppose we wait until time $t_1$ and then take out

$$\zeta_1 = \frac{X(t_1) - c}{1 + \lambda}.$$

The corresponding value is

$$\begin{aligned} J^{(v_1)}(s, x) &= E^x \left[ \frac{e^{-\rho(t_1 + s)}}{1 + \lambda} (X(t_1) - c) \right] \\ &= E^x \left[ \frac{1}{1 + \lambda} \left( xe^{-\rho s} e^{(\mu - \rho)t_1} - ce^{-\rho(s + t_1)} \right) \right] \\ &\to \infty \text{ as } t_1 \to \infty. \end{aligned}$$

Therefore we obtain $\Phi(s, x) = +\infty$ in this case.

2) $\mu < \rho$.
   We look for a solution by using the results of Theorem 6.2. In this case condition (x) becomes

$$A_0 \psi(x) := -\rho \psi(x) + \mu x \psi'(x) + \tfrac{1}{2}\sigma^2 \psi''(x)$$

$$+ \int_{\mathbb{R}} \{\psi(x + \gamma x z) - \psi(x) - \gamma z \psi'(x)\} \nu(dz) = 0 \text{ in } D. \tag{10.6.11}$$

We try a solution of the form

$$\psi(x) = C_1 x^{\gamma_1} + C_2 x^{\gamma_2}$$

where $\gamma_1 > 1$, $\gamma_2 < 0$ are the solutions of the equation

$$F(\gamma) := -\rho + \mu\gamma + \frac{1}{2}\sigma^2 \gamma(\gamma - 1) + \int_{\mathbb{R}} \{(1 + \theta z)^\gamma - 1 - \theta z \gamma\} \nu'(dz) = 0.$$

We guess that the continuation region is of the form

$$D = \{(s, x) : 0 < x < \bar{x}\}$$

for some $\bar{x} > 0$ (to be determined).
   We see that $C_2 = 0$, because otherwise $\lim_{x \to 0} |\psi(x)| = \infty$.

   We guess that in this case it is optimal to wait till $X(t)$ reaches or exceeds a value $\bar{x} > c$ and then take out as much as possible, i.e. reduce $X(t)$ to 0. Taking the transaction costs into account this means that we should take out

$$\hat{\zeta}(x) = \frac{x - c}{1 + \lambda} \text{ for } x \geq \bar{x}.$$

We therefore propose that $\psi(x)$ has the form

$$\psi(x) = \begin{cases} C_1 x^{\gamma_1} \text{ for } 0 < x < \bar{x} \\ \frac{x-c}{1+\lambda} \text{ for } x \geq \bar{x}. \end{cases}$$

Continuity and differentiability of $\psi(x)$ at $x = \bar{x}$ give the equations

$$C_1 \bar{x}^{\gamma_1} = \frac{\bar{x} - c}{1 + \lambda}$$

and

$$C_1 \gamma_1 \bar{x}^{\gamma_1 - 1} = \frac{1}{1 + \lambda}.$$

Combining these we get

$$\bar{x} = \frac{\gamma_1 c}{\gamma_1 - 1} \text{ and } C_1 = \frac{\bar{x} - c}{1 + \lambda} \bar{x}^{-\gamma_1}.$$

With these values of $\bar{x}$ and $C_1$, we have to verify that $\psi$ satisfies all the requirements of Theorem 6.2. We check some of them:

**(ii):** $\psi \geq \mathcal{M}\psi$ on $\mathcal{S}$.

Here $\mathcal{M}\psi = \sup\{\psi(x - c - (1 + \lambda)\zeta) + \zeta \; ; \; 0 \leq \zeta \leq \frac{x-c}{1+\lambda}\}$.
If $x - c - (1 + \lambda)\zeta \geq \bar{x}$, then

$$\psi(x - c - (1 + \lambda)\zeta) + \zeta = \frac{x - 2c}{1 + \lambda} < \frac{x - c}{1 + \lambda} = \psi(x)$$

and if $x - c - (1 + \lambda)\zeta < \bar{x}$ then

$$h(\zeta) := \psi(x - c - (1 + \lambda)\zeta) + \zeta = C_1(x - c - (1 + \lambda)\zeta)^{\gamma_1} + \zeta.$$

Since

$$h'\left(\frac{x - c}{1 + \lambda}\right) = 1 \text{ and } h''(\zeta) > 0$$

we see that the maximum value of $h(\zeta) \; ; \; 0 \leq \zeta \leq \frac{x-c}{1+\lambda}$, is attained at $\zeta = \hat{\zeta}(x) = \frac{x-c}{1+\lambda}$.
Therefore

$$\mathcal{M}\psi(x) = \max\left(\frac{x - 2c}{1 + \lambda}, \frac{x - c}{1 + \lambda}\right) = \frac{x - c}{1 + \lambda} \text{ for all } x > c.$$

Hence $\mathcal{M}\psi(x) = \psi(x)$ for $x \geq \bar{x}$.
  For $0 < x < \bar{x}$ consider

$$k(x) := C_1 x^{\gamma_1} - \frac{x - c}{1 + \lambda}.$$

Since

$$k(\bar{x}) = k'(\bar{x}) = 0 \text{ and } k''(x) > 0 \text{ for all } x,$$

we conclude that

$$k(x) > 0 \text{ for } 0 < x < \bar{x}.$$

Hence

$$\psi(x) > \mathcal{M}\psi(x) \text{ for } 0 < x < \bar{x}.$$

**(vi):**  $A_0\psi(x) \leq 0$ for $x \in \mathcal{S}\backslash\bar{D}$ i.e. for $x > \bar{x}$.

For $x > \bar{x}$, we have

$$A_0\psi(x) = -\rho\frac{x-c}{1+\lambda} + \mu x \cdot \frac{1}{1+\lambda} = (1+\lambda)^{-1}[(\mu-\rho)x + \rho c].$$

Therefore we see that

$$A_0\psi(x) \geq 0 \text{ for all } x > \bar{x}$$
$$\Leftrightarrow (\mu - \rho)x + \rho c \leq 0 \text{ for all } x > \bar{x}$$
$$\Leftrightarrow (\mu - \rho)\bar{x} + \rho c \leq 0$$
$$\Leftrightarrow \bar{x} \geq \frac{\rho c}{\rho - \mu}$$
$$\Leftrightarrow \frac{\gamma_1 c}{\gamma_1 - 1} \geq \frac{\rho c}{\rho - \mu}$$
$$\Leftrightarrow \gamma_1 \leq \frac{\rho}{\mu}.$$

Since

$$F\left(\frac{\rho}{\mu}\right) \geq -\rho + \mu \cdot \frac{\rho}{\mu} + \frac{1}{2}\sigma^2\frac{\rho}{\mu}\left(\frac{\rho}{\mu} - 1\right) > 0$$

and $F(\gamma_1) = 0$, $\gamma_1 > 1$ we conclude that $\gamma_1 < \frac{\rho}{\mu}$ and hence (vi) holds.

### Exercise 6.3.

Here $f = g = 0$, $\Gamma(y,\zeta) = (s,0)$, $K(y,\zeta) = -c + (1-\lambda)x$ and $\mathcal{S} = \mathbb{R}^2$ ; $y = (s,x)$.
If there are no interventions, the process $Y(t)$ defined by

$$dY(t) = \begin{bmatrix} dt \\ dX(t) \end{bmatrix} = \begin{bmatrix} 1 \\ \mu \end{bmatrix} dt + \begin{bmatrix} 0 \\ \sigma \end{bmatrix} dB(t) + \begin{bmatrix} 0 \\ \int_{\mathbb{R}} \theta z \, \tilde{N}(dt, dz) \end{bmatrix}$$

has the generator

$$A\phi(y) = \frac{\partial\phi}{\partial s} + \mu\frac{\partial\phi}{\partial x} + \frac{1}{2}\sigma^2\frac{\partial^2\phi}{\partial x^2} + \int_{\mathbb{R}} \left\{\phi(s, x + \theta z) - \phi(s,x) - \theta z\frac{\partial\phi}{\partial x}(s,z)\right\}\nu(dz) \; ;$$

$y = (s,x)$.
    The intervention operator $\mathcal{M}$ is given by

$$\mathcal{M}\phi(y) = \sup\{\phi(\Gamma(y,\zeta)) + K(y,\zeta); \zeta \in Z \text{ and } \Gamma(y,\zeta) \in \mathcal{S}\} = \phi(s,0) + (1-\lambda)x - c \,.$$

If we try

$$\phi(s,x) = e^{-\rho s}\psi(x)$$

we get that

$$A\phi(s,x) = e^{-\rho s}A_0\psi(x),$$

where

$$A_0\psi(x) = -\rho\psi + \mu\psi'(x) + \frac{1}{2}\sigma^2\psi''(x) + \int_{\mathbb{R}} \{\psi(x + \theta z) - \psi(x) - \theta z\psi'(x)\}\nu(dz)$$

and
$$\mathcal{M}\phi(s,x) = e^{-\rho s}\mathcal{M}_0\psi(x),$$

where
$$\mathcal{M}_0\psi(x) = \psi(0) + (1-\lambda)x - c.$$

We guess that the continuation region $D$ has the form
$$D = \{(s,x); x < x^*\}$$

for some $x^* > 0$ to be determined. To find a solution $\psi_0$ of $A_0\psi_0 + f = A_0\psi_0 = 0$, we try
$$\psi_0(x) = e^{r\,x} \qquad (r \text{ constant})$$

and get
$$A_0\psi_0(x) = -\rho\, e^{r\,x} + \mu\, r\, e^{r\,x} + \tfrac{1}{2}\sigma^2 r^2 e^{r\,x}$$
$$+ \int_{\mathbb{R}} \{e^{r(x+\theta z)} - e^{r\,x} - r\,\theta\,z\,e^{r\,x}\}\nu(dz)$$
$$= e^{r\,x}h(r) = 0,$$

where
$$h(r) = -\rho + \mu\,r + \tfrac{1}{2}\sigma^2 r^2 + \int_{\mathbb{R}} \{e^{r\theta z} - 1 - r\,\theta\,z\}\nu(dz).$$

Choose $r_1 > 0$ such that
$$h(r_1) = 0 \qquad \text{(see the solution of Exercise 2.1)}.$$

Then we define
$$\psi(x) = \begin{cases} M\,e^{r_1 x}\,; & x < x^* \\ \psi(0) + (1-\lambda)x - c = M + (1-\lambda)x - c\,; & x \geq x^* \end{cases} \qquad (10.6.12)$$

for some constant $M = \psi(0) > 0$. If we require continuity and differentiability at $x = x^*$ we get the equations
$$M\,e^{r_1 x^*} = M + (1-\lambda)x^* - c \qquad (10.6.13)$$

and
$$M\,r_1 e^{r_1 x^*} = 1 - \lambda. \qquad (10.6.14)$$

This gives the following equations for $x^*$ and $M$:
$$k(x^*) := e^{-r_1 x^*} + r_1 x^* - 1 - \frac{r_1 c}{1-\lambda} = 0, \qquad M = \frac{1-\lambda}{r_1}e^{-r_1 x^*} > 0. \qquad (10.6.15)$$

Since $k(0) = -\frac{r_1 c}{1-\lambda} < 0$ and $\lim\limits_{x\to\infty} k(x) = \infty$, we see that there exists $x^* > 0$ s.t. $k(x^*) = 0$.

We must verify that with these values of $x^*$ and $M$ the conditions of Theorem 6.2 are satisfied. We consider some of them:

**(ii):**   $\psi(x) \geq \mathcal{M}_0\psi(x).$
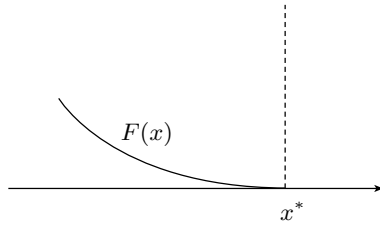
**Fig. 10.6.** The function $F$.

For $x \geq x^*$ we have $\psi(x) = \mathcal{M}_0\psi(x) = M + (1-\lambda)x - c$. For $x < x^*$ we have $\psi(x) = M\,e^{r_1 x}$ and $\mathcal{M}_0\psi(x) = M + (1-\lambda)x - c$. Define

$$F(x) = M\,e^{r_1 x} - (M + (1-\lambda)x - c); \qquad x \leq x^*.$$

See Figure 10.6. We have

$$F(x^*) = F'(x^*) = 0 \qquad \text{and} \qquad F''(x) = M\,r_1^2 e^{r_1 x} > 0.$$

Hence $F'(x) < 0$ and so $F(x) > 0$ for $x < x^*$. Therefore

$$\psi(x) \geq \mathcal{M}_0\psi(x) \qquad \text{for all } x.$$

**(vi):** $\quad A_0\psi \leq 0 \quad$ for $\quad x > x^*$:

For $x > x^*$ we have

$$A_0\psi(x) = -\rho[M + (1-\lambda)x - c] + \mu(1-\lambda) = -\rho(1-\lambda)x + \rho(c - M) + \mu(1-\lambda).$$

So

$$A_0\psi(x) \leq 0 \quad \text{for all } x > x^* \Leftrightarrow x \geq \frac{\mu}{\rho} + \frac{c - M}{1 - \lambda} \quad \text{for all } x > x^*$$

$$\Leftrightarrow x^* \geq \frac{\mu}{\rho} + \frac{c - M}{1 - \lambda}$$

$$\Leftrightarrow x^* \geq \frac{\mu}{\rho} + \frac{c}{1 - \lambda} - \frac{1}{r_1}e^{-r_1 x^*}$$

$$\Leftrightarrow e^{-r_1 x^*} + r_1 x^* - \frac{c}{1 - \lambda} \geq \frac{\mu}{\rho}$$

$$\Leftrightarrow 1 \geq \frac{\mu}{\rho} \Leftrightarrow \mu \leq \rho.$$

So we need to assume that $\mu \leq \rho$ for (vi) to hold.

*Conclusion.*

Let

$$\psi(s, x) = e^{-\rho s}\psi(x)$$

where $\psi$ is given by (10.6.12). Assume that

$$\mu \leq \rho \,.$$
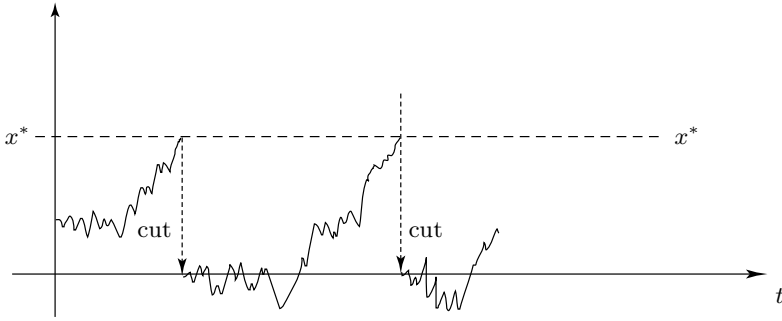
**Fig. 10.7.** The optimal forest management of Exercise 6.3

Then
$$\phi(s,x) = \sup_v J^{(v)}(s,x)$$

and the optimal strategy is to cut the forest every time the biomass reaches the value $x^*$ (see Figure 10.7).

## 10.7 Exercises of Chapter 7

### Exercise 7.1

As in Exercise 6.3, we have
$$
\begin{aligned}
&f = g = 0 \\
&\Gamma(y,\zeta) = (s,0) \ ; \ y = (s,x) \\
&K(y,\zeta) = (1-\lambda)x - c \\
&\mathcal{S} = [0,\infty) \times \mathbb{R}
\end{aligned}
$$

If there is no intervention, then $\phi_0 \equiv 0$ and
$$\mathcal{M}\phi_0 = \sup\{(1-\lambda)\zeta - c; \zeta \le x\} = (1-\lambda)x - c.$$

Hence
$$\phi_1(y) = \sup_{\tau \le \tau_\mathcal{S}} E^y[\mathcal{M}\phi_0(Y(\tau))] = \sup_{\tau \le \tau_\mathcal{S}} E^y\big[e^{-\rho(s+\tau)}((1-\lambda)X(\tau) - c)\big]. \quad (10.7.1)$$

This is an optimal stopping problem that can be solved by exploiting the three basic variational inequalities. We assume that the continuation region
$$D_1 = \{\phi_1 > \mathcal{M}\phi_0\}$$

is of the form
$$D_1 = \{(s,x) \ ; \ x < x_1\} \text{ for some } x_1 > 0$$

and that the value function has the form $\phi_1(s,x) = e^{-\rho s}\psi_1(x)$ for some function $\psi_1$. On $D_1$, $\psi_1$ is the solution of

$$-\rho\psi_1(x)+\mu\psi_1'(x)+\tfrac{1}{2}\sigma^2\psi_1''(x)+\int_{\mathbb{R}}\{\psi_1(x+\theta z)-\psi_1(x)-\theta z\psi_1'(x)\}\nu(dz)=0. \quad (10.7.2)$$

A solution of (10.7.2) is
$$\psi_1(x)=Ae^{\gamma_1 x}+Be^{\gamma_2 x}$$

where $\gamma_2<0$ and $\gamma_1>1$, $A$ and $B$ arbitrary constants to be determined.
We choose $B=0$ and put $A_1=A>0$. We get

$$\psi_1(x)=\begin{cases}A_1e^{\gamma_1 x} & x<x_1\\(1-\lambda)x-c & x\geq x_1.\end{cases}$$

We impose the continuity and differentiability conditions of $\psi_1$ at $x=x_1$.

(i) Continuity: $A_1e^{\gamma_1 x_1}=(1-\lambda)x_1-c$.
(ii) Differentiability: $A_1\gamma_1e^{\gamma_1 x_1}=1-\lambda$.

We get $A_1=\frac{(1-\lambda)}{\gamma_1}e^{-\gamma_1 x_1}$ and $x_1=\frac{1}{\gamma_1}+\frac{c}{1-\lambda}$.
As a second step, we evaluate

$$\phi_2(y)=\sup_\tau E^y[\mathcal{M}\phi_1(Y(\tau))].$$

We suppose $\phi_2(s,x)=e^{-\rho s}\psi_2(x)$ and consider

$$\mathcal{M}\psi_1(x)=\sup\{\psi_1(0)+(1-\lambda)\zeta-c;\zeta\leq x\}=\psi_1(0)+(1-\lambda)x-c=(1-\lambda)x+A_1-c.$$

Hence
$$\phi_2(y)=\sup_{\tau\leq\tau_S}E^y\left[e^{-\rho(s+\tau)}((1-\lambda)X(\tau)-(c-A_1))\right]. \quad (10.7.3)$$

By the same argument as before, we get $\Phi_2(s,x)=e^{-\rho s}\psi_2(x)$, where

$$\psi_2(x)=\begin{cases}A_2e^{\gamma_1 x} & x<x_2\\(1-\lambda)x+A_1-c & x\geq x_2\end{cases}$$

where $x_2=\frac{1}{\gamma_1}+\frac{c-A_1}{1-\lambda}$ and $A_2=\frac{1-\lambda}{\gamma_1}e^{-\gamma_1 x_2}$. Note that $x_2<x_1$ and $A_2>A_1$.
Since $\mathcal{M}\phi_0$ and $\mathcal{M}\phi_1$ have linear growth, the conditions of Theorem 7.2 are satisfied. Hence $\phi_1$ and $\phi_2$ are the solutions for our impulse control problems when respectively one intervention and two interventions are allowed. The impulses are given by $\zeta_1=\zeta_2=x$ and $\tau_1=\inf\{t:X(t)\geq x_2\}$ and $\tau_2=\inf\{t>\tau_1:X(t)\geq x_1\}$.

## Exercise 7.2

Here we have (see the notation of Chapter 6)

$$f=g\equiv 0$$
$$K(x,\zeta)=\zeta$$
$$\Gamma(x,\zeta)=x-(1+\lambda)\zeta-c$$
$$\mathcal{S}=\{(s,x)\;;\;x>0\}.$$

We put $y=(s,x)$ and suppose $\phi_0(s,x)=e^{-\rho s}\psi_0(x)$. Since $f=g=0$ we have

$$\phi_0(y) = 0$$

and $\mathcal{M}\psi_0(y) = \sup\{\zeta : 0 \le \zeta \le \frac{x-c}{1+\lambda}\} = (\frac{x-c}{1+\lambda})^+$. As a second step, we consider

$$\phi_1(s,x) = \sup_{\tau \le \tau_S} E^x[\mathcal{M}\phi_0(X(\tau))] = \sup_{\tau \le \tau_S} E^x\left[e^{-\rho(\tau+s)}\frac{(X(\tau+s)-c)^+}{1+\lambda}\right]. \quad (10.7.4)$$

We distinguish between 3 cases

**(a)** $\mu > \rho$
Then

$$\phi_1(s,x) \ge \frac{xe^{(\mu-\rho)(t+s)} - ce^{-\rho(t+s)}}{1+\lambda}.$$

Hence if $t \to +\infty$

$$\phi_1(s,x) \to +\infty.$$

We obtain $\mathcal{M}\phi_1(s,x) = +\infty$ and clearly $\phi_n = +\infty$ for all $n$. In this case, the optimal stopping time does not exist.

**(b)** $\mu < \rho$
In this case we try to put $\phi_1(s,x) = e^{-\rho s}\psi_1(x)$ and solve the optimal stopping problem (10.7.4) by using Theorem 2.2.

We guess that the continuation region is of the form $D = \{0 < x < x_1^*\}$ and solve

$$L_0\psi_1(x) := -\rho\psi_1(x) + \mu x\psi_1'(x) + \frac{\sigma^2 x^2}{2}\psi_1''(x)$$

$$+ \int_{\mathbb{R}}\{\psi_1(x+\theta xz) - \psi_1(x) - \theta xz\psi_1'(x)\}\nu(dx) = 0. \quad (10.7.5)$$

A solution of equation (10.7.5) is

$$\psi_1(x) = c_1 x^{\gamma_1} + c_2 x^{\gamma_2}$$

where

$$\gamma_2 < 0 \text{ and } \gamma_1 > 1, \text{ and } c_1, c_2 \text{ are arbitrary constants.}$$

Since $\gamma_2 < 0$, we put $c_2 = 0$. We obtain

$$\psi_1(x) = \begin{cases} c_1 x^{\gamma_1} & 0 < x < x_1^* \\ \frac{x-c}{1+\lambda} & x \ge x_1^*. \end{cases}$$

By imposing the condition of continuity and differentiability, we can compute $c_1$ and $x_1^*$. The result is

1. $x_1^* = \frac{\gamma_1 c}{\gamma_1 - 1}$

2. $c_1 = \frac{1}{\gamma_1(1+\lambda)}\left(\frac{\gamma_1 c}{\gamma_1 - 1}\right)^{1-\gamma_1}$

Note that $\gamma_1 > 1$ and $x_1^* > c$. We check some of the conditions of Theorem 2.2:

**(ii)** $\psi_1(x) \ge \mathcal{M}\psi_0(x)$ for all $x$:

We know that $\phi_1(x) = \mathcal{M}\phi_0(x)$ for $x > x_1^*$. Consider $h_1(x) := \psi_1(x) - \mathcal{M}\psi_0(x)$. We have

$$h_1(x_1^*) = 0$$
$$h_1'(x_1^*) = 0$$
$$h_1''(x_1^*) = c_1\gamma_1(\gamma_1 - 1)(x_1^*)^{\gamma_1 - 2} > 0.$$

Hence $x_1^*$ is a minimum for $h_1$ and $\psi_1(x) \geq \mathcal{M}\psi_0(x)$ for every $0 < x < x_1^*$.

**(vi)** $\quad L_0\psi_1 \leq 0 \quad$ for all $\quad x > 0$:

Clearly $L_0\psi_1 = 0$ for $0 < x < x_1^*$.
If $x > x_1^*$ then $L_0\psi_1(x) = ((\mu - \rho)x + c\rho)\frac{1}{1+\lambda} \leq 0$ iff $x \geq \frac{c\rho}{\rho - \mu}$.
Define

$$F(\gamma) := x^{-\gamma}L_0(x^\gamma) = -\rho + \mu\gamma + \frac{1}{2}\gamma(\gamma - 1) + \int_{\mathbb{R}} \{(1 + \theta z)^\gamma - 1 - \gamma\theta z\}\, \nu(dz).$$

Then we know that $F(\gamma_2) = F(\gamma_1) = 0$ and $F'(\gamma) > 0$ for $\gamma \geq \gamma_1$. Since $F\left(\frac{\rho}{\mu}\right) > 0$ we have that $\frac{\rho}{\mu} > \gamma_1$, which implies that $x_1^* = \dfrac{\gamma_1 c}{\gamma_1 - 1} > \dfrac{\rho c}{\rho - \mu}$. Hence $L_0\psi_1(x) \leq 0$ for all $x \geq x_1^*$.

We conclude that $\phi_1(s, x) = e^{-\rho s}\psi_1(x)$ actually solves the optimal stopping problem (10.7.4).

Next we consider

$$\mathcal{M}\psi_1(x) = \sup\left\{\psi_1(x - (1 + \lambda)\zeta - c) + \zeta; 0 \leq \zeta \leq \frac{x - c}{1 + \lambda}\right\} = \frac{(x - c)^+}{1 + \lambda}$$

and repeat the same procedure to find $\psi_2$.

By induction, we obtain $\mathcal{M}\psi_n = \mathcal{M}\psi_{n-1} = \mathcal{M}\psi_1 = \mathcal{M}\psi_0$. Consequently, we also have

$$\Phi = \Phi_1$$

and $\Phi(s, x) = \Phi_n(s, x)$ for every $n$. Moreover, we achieve the optimal result with just one intervention.

**(c)** $\mu = \rho$
This case is left to the reader.

## 10.8 Exercises of Chapter 8

### Exercise 8.1

In this case we have

$$\Gamma_1(\zeta, x_1, x_2) = x_1 - \zeta - c - \lambda|\zeta|$$
$$\Gamma_2(\zeta, x_1, x_2) = x_2 + \zeta$$
$$K(\zeta, x_1, x_2) = c + \lambda|\zeta|$$
$$g = 0$$
$$f(s, x_1, x_2, u) = \frac{e^{-\delta s}}{\gamma}u^\gamma.$$

The generator is given by

$$L^u = \frac{\partial}{\partial t} + (rx_1 - u)\frac{\partial}{\partial x_1} + \alpha x_2 \frac{\partial}{\partial x_2} + \frac{\sigma^2}{2}x_2^2 \frac{\partial^2}{\partial x_2^2}.$$

Let $\phi(s, x_1, x_2)$ be the value function of the optimal consumption problem

$$\sup_{w \in \mathcal{W}} E^y\left[\int_0^\infty e^{-\delta(s+t)}\frac{u(t)^\gamma}{\gamma}dt\right] \; ; \; y = (s, x_1, x_2)$$

with $c, \lambda > 0$ and $\phi_0(s, x_1, x_2)$ the corresponding value function in the case when there are no transaction costs, i.e.

$$c = \lambda = 0.$$

In order to prove that

$$\Phi(s, x_1, x_2) \le Ke^{-\delta s}(x_1 + x_2)^\gamma = \Phi_0(s, x_1, x_2)$$

we check the hypotheses of Theorem 8.1a):

   **(vi)**    $L^u \phi_0 + f \le 0$

Since $\phi_0$ is the value function in the absence of transaction costs, we have

$$\sup_{u \ge 0} \{L^u \phi_0 + f\} = 0 \qquad \text{in } \mathbb{R}^3.$$

Note that

$$\mathcal{M}\phi_0(s, x_1, x_2) = \sup_{\zeta \in \mathbb{R}\setminus\{0\}} \phi_0(s, x_1 - \zeta - \lambda|\zeta| - c, x_2 + \zeta)$$

$$= \sup_{\zeta \in \mathbb{R}\setminus\{0\}} Ke^{-\delta s}(x_1 + x_2 - c - \lambda|\zeta|)^\gamma = Ke^{-\delta s}(x_1 + x_2 - c)^\gamma.$$

Therefore

$$D = \{\phi_0 > \mathcal{M}\phi_0\} = \mathbb{R}^3.$$

Hence we can conclude that

$$e^{-\delta s}(x_1 + x_2)^\gamma \ge \Phi(s, x_1, x_2).$$

## Exercice 8.2a)

The HJBQVI's for this problem can be formulated in one equation as follows:

$$\min\left(\inf_{u \in \mathbb{R}}\left\{\frac{\partial \varphi}{\partial t} + u\frac{\partial \varphi}{\partial x} + \frac{1}{2}\sigma^2\frac{\partial^2 \varphi}{\partial x^2} + x^2 + u^2\right\}, \varphi - \mathcal{M}\varphi\right) = 0,$$

where

$$\mathcal{M}\varphi(t, x) = \inf_{\zeta \in \mathbb{R}}\{\varphi(t, x + \zeta) + c\}.$$

Since $\varphi$ is a candidate for the value function $\Phi$ it is reasonable to guess that, for each $t$, $\varphi(t, z)$ is minimal for $z = 0$. Hence

$$\mathcal{M}\varphi(t, x) = c, \text{ attained for } \zeta = \zeta^*(x) = -x.$$

Motivated by Exercise 4.3 we try a function $\varphi$ of the form

$$\varphi(t, x) = \begin{cases} a(t)x^2 + b(t) & ; \ |x| < x^*(t) \\ c & ; \ |x| \geq x^*(t) \end{cases}$$

where $a(t)$, $b(t)$ and $x^*(t)$ are functions to be determined. With this $\varphi$ the above HJBQVI becomes, for $|x| < x^*(t)$,

$$\inf_{u \in \mathbb{R}} \left\{ \frac{\partial \varphi}{\partial t} + u \frac{\partial \varphi}{\partial x} + \frac{1}{2} \sigma^2 \frac{\partial^2 \varphi}{\partial x^2} + x^2 + u^2 \right\}$$

$$= \inf_{u \in \mathbb{R}} \{ a'(t)x^2 + b'(t) + ua(t)2x + \sigma^2 a(t) + x^2 + u^2 \}$$

$$= x^2 \left[ a'(t) - \frac{1}{4} a^2(t) + 1 \right] + b'(t) + \sigma^2 a^2(t) = 0,$$

and this minimum is attained at

$$u = u^*(t, x) = -\frac{1}{2} \frac{\partial \varphi}{\partial x} = -a(t)x.$$

Together with the boundary values

$$\varphi(T, x) = 0$$

we therefore get that $(a(t), b(t))$ must be the unique solution of the two equations

$$a'(t) - \frac{1}{4} a^2(t) + 1 = 0 \ ; \ a(T) = 0$$

$$b'(t) + \sigma^2 a^2(t) = 0 \ ; \ b(T) = 0.$$

Hence

$$a(t) = \frac{2 \left( 1 - e^{-(T-t)} \right)}{1 + e^{-(T-t)}} > 0 \ ; \ 0 \leq t \leq T$$

and

$$b(t) = \sigma^2 \int_t^T a^2(s) ds \ ; \ 0 \leq t \leq T.$$

Finally we determine $x^*(t)$ by requiring $\varphi$ to be continuous at $x = x^*(t)$:

$$a(t)(x^*(t))^2 + b(t) = c$$

which gives

$$x^*(t) = a^{-1}(t)[c - b(t)].$$

If $b(0) < c$ then with this choice of $a(t)$, $b(t)$, $x^*(t)$ all the conditions of $\varphi$ in Theorem 8.1 can be verified and we conclude that

$$\varphi(s, x) = \Phi(s, x) = \begin{cases} a(s)x^2 + b(s) & ; \ |x| < x^*(s) \\ c & ; \ |x| \geq x^*(s) \end{cases}$$

with optimal combined control $w^* = (u^*, v^*)$ described by
(i) If $|x| < x^*(s)$ use $u^*(s, x) = -a(s)x$ and no impulse control.
(ii) If $|x| \geq x^*(s)$ use the impulse

$$\zeta^*(x) = -x$$

to bring the system down to 0.

See Figure 10.8
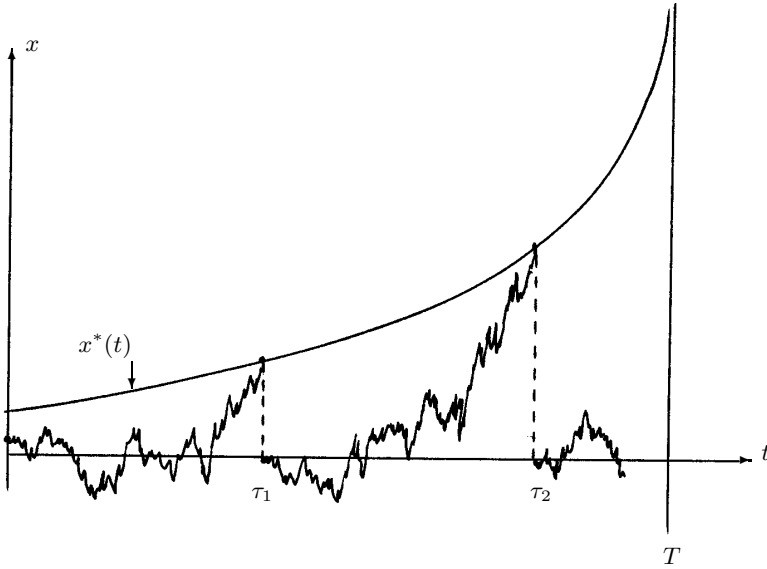
**Fig. 10.8.** The optimal combined control of Exercise 8.2

## 10.9 Exercises of Chapter 9

### Exercise 9.1

Because of the symmetry of $h$, we assume that the continuation region is of the form

$$D = \{\phi > h\} = \{(s, x) : -x^* < x < x^*\}$$

with $x^* > 0$.

We assume that the value function $\phi(s, x) = e^{-\rho s}\psi(x)$. On $D$, $\phi$ is the solution of

$$L\phi(s, x) = 0 \tag{10.9.1}$$

where $L = \frac{\partial}{\partial t} + \frac{1}{2}\frac{\partial^2}{\partial x^2}$. We obtain

$$L_0\psi(x) := -\rho\psi(x) + \frac{1}{2}\psi''(x) = 0. \tag{10.9.2}$$

The general solution of equation (10.9.2) is

$$\psi(x) = c_1 e^{\sqrt{2\rho}\,x} + c_2 e^{-\sqrt{2\rho}\,x}.$$

We must have $\psi(x) = \psi(-x)$, hence $c_1 = c_2$. We put $c_1 = \frac{1}{2}c$ and assume $c > 0$. We impose continuity and differentiability conditions at $x = x^*$:

(i)   Continuity at $x = x^*$

$$\frac{1}{2}c\left(e^{\sqrt{2\rho}\,x^*} + e^{-\sqrt{2\rho}\,x^*}\right) = Kx^*$$
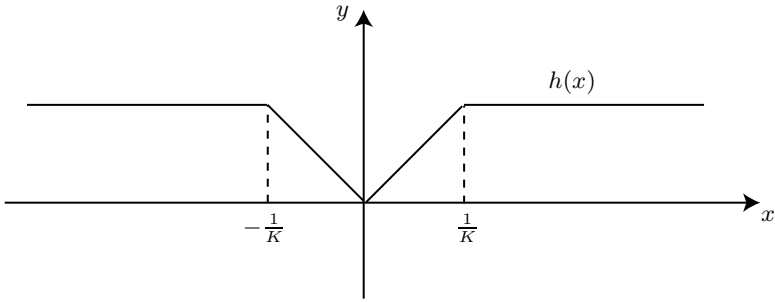
**Fig. 10.9.** The function $h(x)$

(ii)   Differentiability at $x = x^*$

$$\frac{1}{2}c\sqrt{2\rho}\left(e^{x^*\sqrt{2\rho}}\frac{1}{2} - e^{-\sqrt{2\rho}\,x^*}\right) = K.$$

Then $x^*$ is the solution of

$$\frac{1}{x^*\sqrt{2\rho}} = \frac{e^{x^*\sqrt{2\rho}} - e^{-\sqrt{2\rho}\,x^*}}{e^{x^*\sqrt{2\rho}} + e^{-x^*\sqrt{2\rho}}} = \mathrm{tgh}(x^*\sqrt{2\rho})$$

and

$$c = \frac{K}{\sqrt{2\rho}}\frac{1}{+\sinh(x^*\sqrt{2\rho})}.$$

We must check if $x^* < \frac{1}{K}$. If we put $z^* = x^*\sqrt{2\rho}$, then $z^*$ is the solution of

$$\frac{1}{z^*} = \mathrm{tgh}(z^*).$$

We distinguish between two cases:
   **Case 1.**   For $K < \dfrac{1}{x^*} = \dfrac{\sqrt{2\rho}}{z^*}$ we have

$$\psi(x) = \begin{cases} 1 & |x| > \frac{1}{K} \\ K|x| & x^* < |x| < \frac{1}{K} \\ c\cosh(x\sqrt{2\rho}) & |x| < x^* \end{cases}$$

Since $\psi$ is not $C^2$ at $x = x^*$ we prove that $\psi$ is a viscosity solution for our optimal stopping problem.

(i)   We first prove that $\psi$ is a viscosity subsolution.
   Let $u$ belong to $C^2(\mathbb{R})$ and $u(x) \geq \psi(x)$ for all $x \in \mathbb{R}$ and let $y_0 \in \mathbb{R}$ be such that

$$u(y_0) = \psi(y_0).$$

Then $\psi$ is a viscosity subsolution if and only if

$$\max(L_0 u(y_0), h(y_0) - u(y_0)) \geq 0 \text{ for all such } u, y_0. \tag{10.9.3}$$
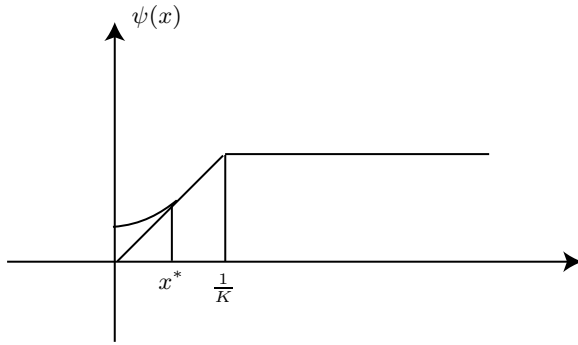
**Fig. 10.10.** The function $\psi(x)$ in Case 1 for $x \geq 0$

We need to check (10.9.3) only for $y_0 = x^*$. We have

$$u(x^*) = \psi(x^*) = h(x^*)$$

i.e. $(h - u)(x^*) = 0$. Hence $\max(L_0 u(x^*), h(x^*) - u(x^*)) \geq (h - u)(x^*) = 0$.

(ii)   We prove that $\psi$ is a viscosity supersolution.
    Let $v$ belong to $C^2(\mathbb{R})$ and $v(x) \leq \psi(x)$ for every $x \in \mathbb{R}$ and let $y_0 \in \mathbb{R}$ be such that $v(y_0) = \psi(y_0)$. Then $\psi$ is a viscosity supersolution if and only if

$$\max(L_0 v(y_0), h(y_0) - v(y_0)) \leq 0 \text{ for all such } v, y_0.$$

We check it only for $x = x^*$. Then

$$h(x^*) = \psi(x^*) = v(x^*).$$

Since $v \leq \psi$, $x = x^*$ is a maximum point for $H = h - \psi$. We have

$$H(x^*) = 0$$
$$H'(x^*) = 0$$
$$H''(x^*) = h''(x^*) - \psi''(x^*_-) \leq 0.$$

Hence $L_0 h(x^*) \leq L_0 \psi(x^*) \leq 0$. Since $\psi$ is both a viscosity supersolution and subsolution, $\psi$ is a viscosity solution.

**Case 2.**   We consider now the case when $K \geq \frac{\sqrt{2\rho}}{z^*}$.
    In this case, the continuation region is given by

$$D = \left\{ -\frac{1}{K} < x < \frac{1}{K} \right\}$$

i.e. $x^* = \frac{1}{K}$.
    We have

$$\psi(x) = \begin{cases} 1 & |x| \geq \frac{1}{K} \\ c \cosh(z\sqrt{2\rho}) & |x| < \frac{1}{K} \end{cases}$$

$$\psi(x)$$
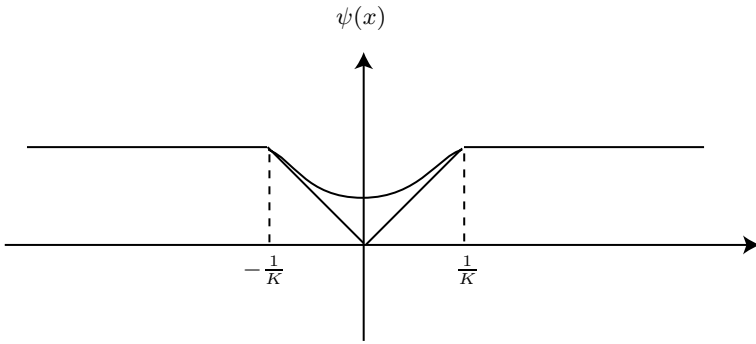


**Fig. 10.11.** The function $\psi(x)$ in Case 2

$\psi$ is not $C^1$ at $|x| = \frac{1}{K}$. We prove that it is a viscosity solution.

(i)   $\psi$ is a viscosity subsolution.

Let $u$ belong to $C^2(\mathbb{R})$ and $u(x) \geq \psi(x)$ for every $x \in \mathbb{R}$. Let $y_0$ be such that $u(y_0) = \psi(y_0)$. Then $\psi$ is a viscosity subsolution if and only if

$$\max(L_0 u(y_0), h(y_0) - u(y_0)) \geq 0. \tag{10.9.4}$$

For $y_0 = x^*$, $h(x^*) = \psi(x^*) = u(x^*)$. Hence (10.9.4) is trivially satisfied.

(ii)   $\psi$ is a viscosity supersolution.

Actually there does not exist any $v \in C^2(\mathbb{R})$ such that $v(\frac{1}{K}) = \psi(\frac{1}{K})$ and $v \leq \psi$. Heuristically, this happens because $\psi$ has an angle at $x = \frac{1}{K}$:

Suppose that there exists $v \in C^2(\mathbb{R})$ such that $v \leq \psi$, $v(\frac{1}{K}) = \psi(\frac{1}{K})$. We consider $H := \psi - v$. We have $H(\frac{1}{K}) = 0$ and $H \geq 0$. Hence we must have

$$H'\left(\left(\frac{1}{K}\right)^-\right) := \lim_{x \to (\frac{1}{K})^-} H'(x) \leq 0 \text{ and } H'\left(\left(\frac{1}{K}\right)^+\right) := \lim_{x \to (\frac{1}{K})^+} H'(x) \geq 0.$$

Therefore

$$\psi'\left(\left(\frac{1}{K}\right)^-\right) - v'\left(\left(\frac{1}{K}\right)^-\right) \leq \psi'\left(\left(\frac{1}{K}\right)^+\right) - v'\left(\left(\frac{1}{K}\right)^+\right),$$

which implies that $\psi'\left(\left(\frac{1}{K}\right)^-\right) \leq \psi'\left(\left(\frac{1}{K}\right)^+\right)$ since $v \in C^2(\mathbb{R})$.

Since $\psi'$ does not satisfy this inequality, we conclude that there does not exist any $v \in C^2(\mathbb{R})$ such that $\psi \geq v$ and $\psi(\frac{1}{K}) = v(\frac{1}{K})$. ($v$ cannot be even $C^1$!).

We can conclude that $\psi$ is a viscosity solution for our problem.

### Exercise 9.2

We intend to apply (the minimum version of) Theorem 9.8 and note that in this case we have, with $y = (s, x) \in \mathbb{R}^2$,

$$L^u \varphi(y) = \frac{\partial \varphi}{\partial s} + u \frac{\partial \varphi}{\partial x} + \frac{1}{2} \frac{\partial^2 \varphi}{\partial x^2} + \int_{\mathbb{R}} \left\{ \varphi(s, x+z) - \varphi(s, x) - z \frac{\partial \varphi}{\partial x}(s, x) \right\} \nu(dz)$$

$$\Gamma(y, \zeta) = y + (0, \zeta), \ K(y, \zeta) = ce^{-\rho s}, \ f(y, u) = e^{-\rho s}(x^2 + \theta u^2),$$

$$g = 0 \text{ and}$$

$$M\varphi(y) = \inf \left\{ \varphi(s, x+\zeta) + ce^{-\rho s} \ ; \ \zeta \in \mathbb{R} \backslash \{0\} \right\} = \varphi(s, 0) + ce^{-\rho s}.$$

We first prove that

$$\Phi(x) < \Phi_1(x) = e^{-\rho s}(ax^2 + b)$$

(see the solution of Exercise 3.4). Since clearly

$$\Phi(x) \le \Phi_1(x)$$

it suffices, by Theorem 9.8, to prove that $\Phi_1(x)$ does not satisfy equation (9.3.6) in the viscosity sense. In particular, (9.3.6) implies that

$$e^{-\rho s}(ax^2 + b) = \Phi_1(s, x) \le M\Phi_1(s, x) = e^{-\rho s}(b + c) \ ; \ x \in \mathbb{R}.$$

Since $a > 0$ this is impossible.

Next we prove that

$$\Phi(x) < \Phi_2(x)$$

where $\Phi_2(x)$ is the solution of Exercise 6.1. It has the form $\Phi_2(s, x) = e^{-\rho s} \psi_2(x)$, with

$$\psi_2(x) = \begin{cases} \psi_0(x) \ ; \ |x| \le \bar{x} \\ \psi_0(\hat{x}) + c \ ; \ |x| > \bar{x} \end{cases}$$

where

$$\psi_0(x) = \frac{1}{\rho} x^2 + \frac{b}{\rho^2} - a \cosh(\gamma x)$$

is a solution of

$$-\rho \psi_0(x) + \frac{1}{2} \psi_0''(x) + \int_{\mathbb{R}} \left\{ \psi_0(x+z) - \psi_0(x) - z\psi_0'(x) \right\} \nu(dz) + x^2 = 0.$$

Since clearly $\Phi(s, x) \le \Phi_2(s, x)$, it suffices to prove that $\Phi_2(s, x)$ does not satisfy (9.3.6) in viscosity sense. In particular, (9.3.6) implies that

$$L^u \Phi_2(s, x) + e^{-\rho s}(x^2 + \theta u^2) \ge 0 \text{ for all } u \in \mathbb{R}.$$

For $|x| < \bar{x}$ this reads

$$u\psi_0'(x) + \theta u^2 \ge 0 \text{ for all } u \in \mathbb{R}, \ |x| < \bar{x}. \tag{10.9.5}$$

The function

$$h(u) := \left( \frac{2x}{\rho} - a\gamma \sinh(\gamma x) \right) u + \theta u^2 \ ; \ u \in \mathbb{R}$$

is minimal when

$$u = \hat{u} = \frac{1}{2\theta} \left( a\gamma \sinh(\gamma x) - \frac{2x}{\rho} \right)$$

with corresponding minimum value

$$h(\hat{u}) = -\frac{1}{4\theta} \left( a\gamma \sinh(\gamma x) - \frac{2x}{\rho} \right)^2 .$$

Hence (10.9.5) cannot possibly hold and we conclude that $\Phi_2$ cannot be a viscosity solution of (9.3.6). Hence $\Phi \neq \Phi_2$ and hence $\Phi(x) < \Phi_2(x)$ for some $x$.

# References

[A]        D. Applebaum: Lévy Processes and Stochastic Calculus. Cambridge University Press 2003.

[Aa]       K. Aase: Optimum portfolio diversification in a general continuous-time model. Stoch. Proc. and their Appl. 18 (1984), 81–98.

[Am]       A.L. Amadori: Nonlinear integro-differential operators arising in option pricing: a viscosity solutions approach. J. Differential and Integral Eq. 13 (2003), 787–811.

[AKL]      A.L. Amadori, K.H. Karlsen and C. La Chioma: Non-linear degenerate integro-partial differential evolution equations related to geometric Lévy processes and applications to backward stochastic differential equations. Stochastics and Stoch. Rep. 76 (2004), 147–177.

[AMS]      M. Akian, J.L. Menaldi and A. Sulem: On an investment-consumption model with transaction costs. SIAM J. Control and Optim. 34 (1996), 329-364.

[AT]       O. Alvarez and A. Tourin: Viscosity solutions of nonlinear integro-differential equations. Ann. Inst. Henri Poincaré 13 (1996), 293–317.

[B]        G. Barles: Solutions de Viscosité des Équations de Hamilton-Jacobi. Math. & Appl. 17. Springer-Verlag 1994.

[B-N]      O. Barndorff-Nielsen: Processes of normal inverse Gaussian type. Finance & Stochastics 1 (1998), 41–68.

[Be]       J. Bertoin: Lévy Processes. Cambridge Univ. Press 1996.

[Ben]      A. Bensoussan: Stochastic maximum principle for distributed parameter systems. Journal of Franklin Institute 315 (1983), 387–406.

[Bi]       J.-M. Bismut: Conjugate convex functions in optimal stochastic control. J. Math. Anal. Appl. 44 (1973), 384–404.

[BCa]      M. Bardi and I. Capuzzo-Dolcetta: Optimal Control and Viscosity Solutions of Hamilton-Jacobi-Bellman Equations. Birkhäuser 1997.

[BCe]      B. Bassan and C. Ceci: Optimal stopping problems with discontinuous reward: regularity of the value function and viscosity solutions. Stochastics and Stoch. Rep. 72 (2002), 55–77.

[BDLØP]    F.E. Benth, G. Di Nunno, A. Løkka, B. Øksendal and F. Proske : Explicit representation of the minimal variance portfolio in markets driven by Lévy processes. Math. Finance 13 (2003), 55–72.

[BK]     F.E. Benth and K.H. Karlsen: Portfolio Optimization in Lévy Markets. World Scientific (to appear).

[BKR1]   F.E. Benth, K. Karlsen and K. Reikvam: Optimal portfolio management rules in a non-Gaussian market with durability and intertemporal substitution. Finance & Stochastics 4 (2001), 447–467.

[BKR2]   F.E. Benth, K. Karlsen and K. Reikvam: Optimal portfolio selection with consumption and nonlinear integro-differential equations with gradient constraint: a viscosity solution approach. Finance & Stochastics 5 (2001), 275–303.

[BL]     A. Bensoussan and J.L. Lions: Impulse Control and Quasi-Variational Inequalities. Gauthiers-Villars, Paris, 1984.

[BT]     D.P. Bertsekas and J.N. Tsitsiklis: An analysis of stochastic shortest path problems. Math. Oper. Res. 16 (1991), 580–595.

[BØ1]    K.A. Brekke and B. Øksendal: Optimal switching in an economic activity under uncertainty. SIAM J. Control Opt. 32 (1994), 1021–1036.

[BØ2]    K.A. Brekke and B. Øksendal: A verification theorem for combined stochastic control and impulse control. In L. Decreusefond et al. (eds.): Stochastic Analysis and Related Topics 6, Birkhäuser, Basel (1998), 211–220.

[BS]     G. Barles and P.E. Souganidis: Convergence of approximation schemes for fully nonlinear second-order equations. Asymptotic analysis 4 (1991), 271–283.

[C]      T. Chan: Pricing contingent claims on stocks driven by Lévy processes. Annals of Appl. Prob. 9 (1999), 504–528.

[CB]     C. Ceci and B. Bassan: Mixed optimal stopping and stochastic control problems with semi-continuous final reward for diffusion processes. Stochastics and Stoch. Rep. 76 (2004), 323–347.

[CEM]    M. Chaleyat-Maurel, N. El Karoui and B. Marchal: Réflexion discontinue et systèmes stochastiques. Annals of Prob. 8 (1980), 1049–1067.

[CIL]    M.G. Crandall, H. Ishii and P.-L. Lions: User's guide to viscosity solutions of second order partial differential equations. Bulletin AMS 27 (1992), 1–67.

[CMS]    J. Ph. Chancelier, M. Messaoud and A. Sulem: A policy iteration algorithm for fixed point problems with nonexpansive operators. Research Report 2004-264, Cermics/ENPC, 2004.

[CØS]    J.-Ph. Chancelier, B. Øksendal and A. Sulem: Combined stochastic control and optimal stopping, and application to numerical approximation of combined stochastic and impulse control. in A.N. Shiryaev (editor): Stochastic Financial Mathematics, Steklov Math. Inst., Moscow 237, 2002, 149 –173.

[CT]     R. Cont and P. Tankov: Financial Modelling with Jump Processes. Chapman & Hall/CRC Press 2003.

[DN]     M.H.A. Davis and A. Norman: Portfolio selection with transaction costs. Math. Oper. Res. 15 (1990), 676–713.

[DS]     F. Delbaen and W. Schachermayer: A general version of the fundamental theorem of asset pricing. Math. Ann. 300 (1994), 463–520.

[DZ]     K. Duckworth and M. Zervos: A model for investment decisions with switching costs. Annals of Appl. Prob. 11 (2001), 239–260.

[E]      H.M. Eikseth: Optimization of dividends with transaction costs. Manuscript 2001.

[Eb]      E. Eberlein: Application of generalized hyperbolic Lévy motion to fi-
          nance. In O.E. Barndorff-Nielsen (editor): Lévy Processes. Birkhäuser
          2001, 319–336.
[EK]      E. Eberlein and U. Keller: Hyperbolic distributions in finance. Bernouilli
          1 (1995), 281–299.
[F]       N.C. Framstad: Combined Stochastic Control for Jump Diffusions with
          Applications to Economics. Cand. Scient Thesis, University of Oslo 1997.
[FØS1]    N.C. Framstad, B. Øksendal and A. Sulem: Optimal consumption and
          portfolio in a jump diffusion market. In A. Shiryaev and A. Sulem (eds.):
          Mathematical Finance. INRIA, Paris 1998, 8–20.
[FØS2]    N.C. Framstad, B. Øksendal and A. Sulem: Optimal consumption and
          portfolio in a jump diffusion market with proportional transaction costs.
          J. Math. Economics 35 (2001), 233–257.
[FØS3]    N.C. Framstad, B. Øksendal and A. Sulem: Sufficient stochastic maximum
          principle for optimal control of jump diffusions and applications to finance.
          J. Optim. Theory and Applications 121 (2004), 77–98.
[FS]      W. Fleming and M. Soner: Controlled Markov Processes and Viscosity
          Solutions. Springer-Verlag 1993.
[GS]      I.I. Gihman and A.V. Skorohod: Controlled Stochastic Processes.
          Springer-Verlag 1979.
[H]       U. Haussmann: A Stochastic Maximum Principle for Optimal Control of
          Diffusions. Longman Scientific & Technical 1986.
[HST]     M.J. Harrison, T. Selke and A. Taylor: Impulse control of a Brownian
          motion. Math. Oper. Res. 8 (1983), 454–466.
[I]       K. Itô: Spectral type of the shift transformation of differential processes
          with stationary increments. Trans. Am. Math. Soc. 81 (1956), 253–263.
[Is1]     H. Ishii: Viscosity solutions of nonlinear second order elliptic PDEs as-
          sociated with impulse control problems. Funkciala Ekvacioj. 36 (1993),
          132–141.
[Is2]     H. Ishii: On the equivalence of two notions of weak solutions, viscosity so-
          lutions and distribution solutions. Funkciala Ekvacioj. 38 (1995), 101–120.
[Ish]     Y. Ishikawa: Optimal control problem associated with jump processes.
          Appl. Math. & Optim. 50 (2004), 21–65.
[J]       M. Jeanblanc-Picqué: Impulse control method and exchange rate. Math.
          Finance 3 (1993), 161–177.
[JK]      E.R. Jakobsen and K.H. Karlsen: A maximum principle for semicontinu-
          ous functions applicable to integro-partial differential equations. Preprint,
          Univ. of Oslo, 18/2003.
[JS]      J. Jacod and A. Shiryaev: Limit Theorems for Stochastic Processes.
          Springer-Verlag 1987.
[J-PS]    M. Jeanblanc-Picqué and A.N. Shiryaev: Optimization of the flow of div-
          idends. Russian Math. Survey 50 (1995), 257–277.
[K]       N.V. Krylov: Controlled Diffusion Processes. Springer-Verlag 1980.
[Ka]      O.Kallenberg: Foundations of Modern Probability. Second Edition.
          Springer-Verlag 2002.
[Ko1]     R. Korn: Portfolio optimization with strictly positive transaction costs
          and impulse control. Finance & Stochast. 2 (1998), 85–114.
[Ko2]     R. Korn: Optimal Portfolios: Stochastic Models for Optimal Investment
          and Risk Management in Continuous Time. World Scientific 1997.

[Ku]     H.J. Kushner: Necessary conditions for continuous parameter stochastic optimization problems. SIAM J. Control. 10 (1972), 550–565.

[KD]     H.J. Kushner and P. Dupuis: Numerical Methods for Stochastic Control Problems in Continuous Time. Springer-Verlag 1992.

[KS]     I. Karatzas and S. Shreve: Brownian Motion and Stochastic Calculus. Second Edition. Springer-Verlag 1991.

[L]      A. Løkka: Martingale representation and functionals of Lévy processes. Stoch. Anal. and Appl. 22 (2004), 867–892.

[LS]     S. Levental and A.V. Skorohod: A necessary and sufficient condition for absence of arbitrage with tame portfolios. Ann. Appl. Prob. 5 (1995), 906–925.

[LZ]     R.R. Lumley and M. Zervos: A model for investment in the natural resource industry with switching costs. Math. Oper. Res. 26 (2001), 637–653.

[LST]    B. Lapeyre, A. Sulem and D. Talay: Simulation of Financial Models: Mathematical Foundations and Applications. Cambridge University Press (to appear).

[M]      R. Merton: Optimal consumption and portfolio rules in a continuous time model. J. Economic Theory 3 (1971), 373–413.

[Ma]     C. Makasu: On some optimal stopping and stochastic control problems with jump diffusions. Ph.D. Thesis, University of Zimbabwe 2002.

[Me]     J.-L. Menaldi: Optimal impulse control problems for degenerate diffusions with jumps. Acta Appl. Math. 8 (1987), 165–198.

[MS]     M. Mnif and A. Sulem: Optimal risk control and divident pay-outs under excess of loss reinsurance. Research Report, RR-5010, November 2003, Inria-Rocquencourt.

[MØ]     G. Mundaca and B. Øksendal: Optimal stochastic intervention control with application to the exchange rate. J. Math. Economics 29(1998), 225–243.

[Ø1]     B. Øksendal: Stochastic Differential Equations. 6th Edition. Springer-Verlag 2003.

[Ø2]     B. Øksendal: Stochastic control problems where small intervention costs have big effects. Appl. Math. Optim. 40 (1999), 355–375.

[Ø3]     B. Øksendal: An Introduction to Malliavin Calculus with Applications to Economics. NHH Lecture Notes 1996.

[ØR]     B.Øksendal and K. Reikvam: Viscosity solutions of optimal stopping problems. Stochastics and Stoch. Rep. 62 (1998), 285–301.

[ØS]     B. Øksendal and A. Sulem: Optimal consumption and portfolio with both fixed and proportional transaction costs. SIAM J. Control and Optim, 40 (2002), 1765–1790.

[ØS2]    B. Øksendal and A. Sulem: Applied Stochastic Control of Jump Diffusions. Springer Verlag Universitext, to appear.

[ØUZ]    B. Øksendal, J. Ubøe and T. Zhang: Nonrobustness of some impulse control problems with respect to intervention costs. Stoch. Anal. and Appl. 20 (2002), 999–1026.

[P]      P. Protter: Stochastic Integration and Differential Equations. Second Edition. Springer-Verlag 2003.

[Ph]     H. Pham: Optimal stopping of controlled jump diffusion processes: a viscosity solution approach. J. Math. Systems, Estimation and Control 8 (1998), 1–27.

[Pu]      M.L. Puterman: Markov Decision Processes: Discrete Stochastics Dy-
          namic Programming. Probability and Mathematical Statistics: applied
          probability and statistics section. Wiley 1994.
[PBGM]    L.S. Pontryagin, V.G. Boltyanskii, R.V. Gamkrelidze and E.F.
          Mishchenko: The Mathematical Theory of Optimal Processes. J. Wiley
          & Sons 1962.
[S]       K. Sato: Lévy Processes and Infinitely Divisible Distributions. Cambridge
          Univ. Press 1999.
[Sc]      W. Schoutens: Lévy Processes in Finance. Wiley 2003.
[S1]      A. Sulem: A solvable one-dimensional model of a diffusion inventory sys-
          tem. Math. Oper. Res. 11 (1986), 125–133.
[S2]      A. Sulem: Explicit solution of a two-dimensional deterministic inventory
          problem. Math. Oper. Res. 11 (1986), 134–146.
[SeSy]    A. Seierstad and K. Sydsæter: Optimal Control Theory with Economic
          Applications. North-Holland 1987.
[SS]      S.E. Shreve and H.M. Soner: Optimal investment and consumption with
          transaction costs. Ann. Appl. Probab. 4 (1994), 609–692.
[V]       H. Varner: Some Impulse Control Problems with Applications to Eco-
          nomics. Cand. Scient. Thesis, University of Oslo, 1997.
[W]       Y. Willassen: The stochastic rotation problem: A generalization of Faust-
          mann's formula to stochastic forest growth. J. Economic Dynamics and
          Control 22 (1998), 573–596.
[YZ]      J. Yong and X.Y. Zhou: Stochastic Controls. Springer-Verlag 1999.

# Notation and Symbols

| | |
|---|---|
| $\mathbb{R}^n$ | $n$-dimensional Euclidean space. |
| $\mathbb{R}^+$ | the non-negative real numbers. |
| $\mathbb{R}^{n \times m}$ | the $n \times m$ matrices (real entries). |
| $\mathbb{Z}$ | the integers |
| $\mathbb{N}$ | the natural numbers |
| $\mathbb{R}^n \simeq \mathbb{R}^{n \times 1}$ | i.e. vectors in $\mathbb{R}^n$ are regarded as $n \times 1$-matrices. |
| $I_n$ | the $n \times n$ identity matrix. |
| $A^T$ | the transposed of the matrix $A$. |
| $\mathcal{P}(\mathbb{R}^k)$ | set of functions $f : \mathbb{R}^k \to \mathbb{R}$ of at must polynomial growth, i.e. there exists constants $C, m$ such that: $|f(y)| \le C(1 + |y|^m)$ for all $y \in \mathbb{R}^k$. |
| $C(U, V)$ | the continuous functions from $U$ into $V$. |
| $C(U)$ | the same as $C(U, \mathbb{R})$. |
| $C_0(U)$ | the functions in $C(U)$ with compact support. |
| $C^k = C^k(U)$ | the functions in $C(U, \mathbb{R})$ with continuous derivatives up to order $k$. |
| $C_0^k = C_0^k(U)$ | the functions in $C^k(U)$ with compact support in $U$. |
| $C^{k+\alpha}$ | the functions in $C^k$ whose $k$'th derivatives are Lipschitz continuous with exponent $\alpha$. |
| $C^{1,2}(\mathbb{R} \times \mathbb{R}^n)$ | the functions $f(t, x); \mathbb{R} \times \mathbb{R}^n \to \mathbb{R}$ which are $C^1$ w.r.t. $t \in \mathbb{R}$ and $C^2$ w.r.t. $x \in \mathbb{R}^n$. |
| $C_b(U)$ | the bounded continuous functions on $U$. |
| $|x|^2 = x^2$ | $\displaystyle\sum_{n=1}^{n} x_i^2$ if $x = (x_1, \ldots, x_n)$. |
| $x \cdot y$ | the dot product $\displaystyle\sum_{n=1}^{n} x_i y_i$ if $x = (x_1, \ldots, x_n)$, $y = (y_1, \ldots, y_n)$. |
| $x^+$ | $\max(x, 0)$ if $x \in \mathbb{R}$. |
| $x^-$ | $\max(-x, 0)$ if $x \in \mathbb{R}$. |
| $\operatorname{sign} x$ | $\begin{cases} 1 & \text{if } x \ge 0 \\ -1 & \text{if } x > 0. \end{cases}$ |
| $\sinh(x)$ | hyperbolic sine of $x$ $\left(= \frac{e^x - e^{-x}}{2}\right)$ |

| | |
|---|---|
| $\cosh(x)$ | hyperbolic cosine of $x$ $(= \frac{e^x + e^{-x}}{2})$ |
| $\operatorname{tgh}(x)$ | $\frac{\sinh(x)}{\cosh(x)}$ |
| $s \wedge t$ | the minimum of $s$ and $t$    $(= \min(s,t))$. |
| $s \vee t$ | the maximum of $s$ and $t$    $(= \max(s,t))$. |
| $\delta_x$ | the unit point mass at $x$. |
| $\operatorname{Argmax}_{u \in U} f(u)$ | $\{u^* \in U; f(u^*) \geq f(u), \forall u \in U\}$ |
| $:=$ | equal to by definition. |
| $\overline{\lim}, \underline{\lim}$ | the same as $\liminf$, $\limsup$. |
| $\operatorname{supp} f$ | the support of the function $f$. |
| $\nabla f$ | the same as $Df = \left[\frac{\partial f}{\partial x_i}\right]_{i=1}^n$. |
| $\partial G$ | the boundary of the set $G$. |
| $\bar{G}$ | the closure of the set $G$. |
| $G^0$ | the interior of the set $G$. |
| $\chi_G$ | the indicator function of the set $G$; |
| | $\chi_G(x) = 1$ if $x \in G$,   $\chi_G(x) = 0$ if $x \notin G$. |
| $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \geq 0}, P)$ | filtered probability space. |
| $\Delta \eta_t$ | the jump of $\eta_t$ defined by $\Delta \eta_t = \eta_t - \eta_t^-$. |
| $P$ | the probability law of $\eta_t$. |
| $N(t, U)$ | see (1.1.2). |
| $\nu(U)$ | $E[N(1,U)]$ see (1.1.3). |
| $\tilde{N}(dt, dz)$ | see (1.1.7). |
| $B(t)$ | Brownian motion. |
| $P \ll Q$ | the measure $P$ is absolutely continuous w.r.t. |
| | the measure $Q$. |
| $P \sim Q$ | $P$ is equivalent to $Q$, i.e. $P \ll Q$ and $Q \ll P$. |
| $E_Q$ | the expectation w.r.t. the measure $Q$. |
| $E$ | the expectation w.r.t. a measure which is clear |
| | from the context (usually $P$). |
| $E[Y] = E^\mu[Y] = \int Y \, d\mu$ | the expectation of the random variable $Y$ |
| | w.r.t. the measure $\mu$. |
| $[X, Y]$ | quadratic covariation of $X$ and $Y$: see Definition 1.27. |
| $\mathcal{T}$ | set of all stopping times $\leq \tau_{\mathcal{S}}$ see (2.1.1). |
| $\tau_G$ | the first exit time from the set $G$ of a process $X_t$ : |
| | $\tau_G = \inf\{t > 0; X_t \notin G\}$. |
| $\Delta_N Y(t)$ | the jump of $Y$ caused by the jump of $N$, see (5.2.2). |
| $\check{Y}(t^-)$ | $Y(t^-) + \Delta_N Y(t)$ (see (6.1.7)). |
| $\Delta_\xi Y(t)$ | the jump of $Y$ caused by the singular control $\xi$. |
| $\Delta_\xi \phi$ | see (5.2.3). |
| $\xi^c(t)$ | continuous part of $\xi(t)$, i.e. the process obtained by |
| | removing the jumps of $\xi(t)$. |
| $\pi/K$ | the restriction of the measure $\pi$ to the set $K$. |
| $A = A_Y$ | the generator of jump diffusion $Y$. |
| $\mathcal{M}$ | intervention operator: see Definition 6.1. |
| VI | variational inequality. |
| QVI | quasi-variational inequality. |
| HJB | Hamilton-Jacobi-Bellman equation. |
| HJBVI | Hamilton-Jacobi-Bellman variational inequality. |
| HJBQVI | Hamilton-Jacobi-Bellman quasi-variational inequality. |

| | |
|---|---|
| SDE | Stochastic differential equation |
| cadlag | right continuous with left limits. |
| caglad | left continuous with right limits. |
| i.i.d. | independent identically distributed. |
| iff | if and only if. |
| a.a., a.e., a.s. | almost all, almost everywhere, almost surely. |
| w.r.t. | with respect to. |
| s.t. | such that. |

# Index

# Universitext

*DoCarmo, M. P.:* Differential Forms and Applications

*Duistermaat, J. J.; Kolk, J. A. C.:* Lie Groups

*Edwards, R. E.:* A Formal Background to Higher Mathematics Ia, and Ib

*Edwards, R. E.:* A Formal Background to Higher Mathematics IIa, and IIb

*Emery, M.:* Stochastic Calculus in Manifolds

*Endler, O.:* Valuation Theory

*Erez, B.:* Galois Modules in Arithmetic

*Everest, G.; Ward, T.:* Heights of Polynomials and Entropy in Algebraic Dynamics

*Farenick, D. R.:* Algebras of Linear Transformations

*Foulds, L. R.:* Graph Theory Applications

*Frauenthal, J. C.:* Mathematical Modeling in Epidemiology

*Friedman, R.:* Algebraic Surfaces and Holomorphic Vector Bundles

*Fuks, D. B.; Rokhlin, V. A.:* Beginner's Course in Topology

*Fuhrmann, P. A.:* A Polynomial Approach to Linear Algebra

*Gallot, S.; Hulin, D.; Lafontaine, J.:* Riemannian Geometry

*Gardiner, C. F.:* A First Course in Group Theory

*Gårding, L.; Tambour, T.:* Algebra for Computer Science

*Godbillon, C.:* Dynamical Systems on Surfaces

*Godement, R.:* Analysis I

*Goldblatt, R.:* Orthogonality and Spacetime Geometry

*Gouvêa, F. Q.:* $p$-Adic Numbers

*Gustafson, K. E.; Rao, D. K. M.:* Numerical Range. The Field of Values of Linear Operators and Matrices

*Gustafson, S. J.; Sigal, I. M.:* Mathematical Concepts of Quantum Mechanics

*Hahn, A. J.:* Quadratic Algebras, Clifford Algebras, and Arithmetic Witt Groups

*Hájek, P.; Havránek, T.:* Mechanizing Hypothesis Formation

*Heinonen, J.:* Lectures on Analysis on Metric Spaces

*Hlawka, E.; Schoißengeier, J.; Taschner, R.:* Geometric and Analytic Number Theory

*Holmgren, R. A.:* A First Course in Discrete Dynamical Systems

*Howe, R., Tan, E. Ch.:* Non-Abelian Harmonic Analysis

*Howes, N. R.:* Modern Analysis and Topology

*Hsieh, P.-F.; Sibuya, Y. (Eds.):* Basic Theory of Ordinary Differential Equations

*Humi, M., Miller, W.:* Second Course in Ordinary Differential Equations for Scientists and Engineers

*Hurwitz, A.; Kritikos, N.:* Lectures on Number Theory

*Huybrechts, D.:* Complex Geometry: An Introduction

*Isaev, A.:* Introduction to Mathematical Methods in Bioinformatics

*Iversen, B.:* Cohomology of Sheaves

*Jacod, J.; Protter, P.:* Probability Essentials

*Jennings, G. A.:* Modern Geometry with Applications

*Jones, A.; Morris, S. A.; Pearson, K. R.:* Abstract Algebra and Famous Inpossibilities

*Jost, J.:* Compact Riemann Surfaces

*Jost, J.:* Postmodern Analysis

*Jost, J.:* Riemannian Geometry and Geometric Analysis

*Kac, V.; Cheung, P.:* Quantum Calculus

*Kannan, R.; Krueger, C. K.:* Advanced Analysis on the Real Line

*Kelly, P.; Matthews, G.:* The Non-Euclidean Hyperbolic Plane

*Kempf, G.:* Complex Abelian Varieties and Theta Functions

*Kitchens, B. P.:* Symbolic Dynamics

*Kloeden, P.; Ombach, J.; Cyganowski, S.:* From Elementary Probability to Stochastic Differential Equations with MAPLE

*Kloeden, P. E.; Platen; E.; Schurz, H.:* Numerical Solution of SDE Through Computer Experiments

*Kostrikin, A. I.:* Introduction to Algebra

*Krasnoselskii, M. A.; Pokrovskii, A. V.:* Systems with Hysteresis

*Luecking, D. H., Rubel, L. A.:* Complex Analysis. A Functional Analysis Approach

*Ma, Zhi-Ming; Roeckner, M.:* Introduction to the Theory of (non-symmetric) Dirichlet Forms

*Mac Lane, S.; Moerdijk, I.:* Sheaves in Geometry and Logic

*Marcus, D. A.:* Number Fields

*Martinez, A.:* An Introduction to Semiclassical and Microlocal Analysis

*Matoušek, J.:* Using the Borsuk-Ulam Theorem

*Matsuki, K.:* Introduction to the Mori Program

*Mc Carthy, P. J.:* Introduction to Arithmetical Functions

*Meyer, R. M.:* Essential Mathematics for Applied Field

*Meyer-Nieberg, P.:* Banach Lattices

*Mikosch, T.:* Non-Life Insurance Mathematics

*Mines, R.; Richman, F.; Ruitenburg, W.:* A Course in Constructive Algebra

*Moise, E. E.:* Introductory Problem Courses in Analysis and Topology

*Montesinos-Amilibia, J. M.:* Classical Tessellations and Three Manifolds

*Morris, P.:* Introduction to Game Theory

*Nikulin, V. V.; Shafarevich, I. R.:* Geometries and Groups

*Oden, J. J.; Reddy, J. N.:* Variational Methods in Theoretical Mechanics

*Øksendal, B.:* Stochastic Differential Equations

*Øksendal, B.; Sulem, A.:* Applied Stochastic Control of Jump Diffusions

*Poizat, B.:* A Course in Model Theory

*Polster, B.:* A Geometrical Picture Book

*Porter, J. R.; Woods, R. G.:* Extensions and Absolutes of Hausdorff Spaces

*Radjavi, H.; Rosenthal, P.:* Simultaneous Triangularization

*Ramsay, A.; Richtmeyer, R. D.:* Introduction to Hyperbolic Geometry

*Rees, E. G.:* Notes on Geometry

*Reisel, R. B.:* Elementary Theory of Metric Spaces

*Rey, W. J. J.:* Introduction to Robust and Quasi-Robust Statistical Methods

*Ribenboim, P.:* Classical Theory of Algebraic Numbers

*Rickart, C. E.:* Natural Function Algebras

*Rotman, J. J.:* Galois Theory

*Rubel, L. A.:* Entire and Meromorphic Functions

*Rybakowski, K. P.:* The Homotopy Index and Partial Differential Equations

*Sagan, H.:* Space-Filling Curves

*Samelson, H.:* Notes on Lie Algebras

*Schiff, J. L.:* Normal Families

*Sengupta, J. K.:* Optimal Decisions under Uncertainty

*Séroul, R.:* Programming for Mathematicians

*Seydel, R.:* Tools for Computational Finance

*Shafarevich, I. R.:* Discourses on Algebra

*Shapiro, J. H.:* Composition Operators and Classical Function Theory

*Simonnet, M.:* Measures and Probabilities

*Smith, K. E.; Kahanpää, L.; Kekäläinen, P.; Traves, W.:* An Invitation to Algebraic Geometry

*Smith, K. T.:* Power Series from a Computational Point of View

*Smoryński, C.:* Logical Number Theory I. An Introduction

*Stichtenoth, H.:* Algebraic Function Fields and Codes

*Stillwell, J.:* Geometry of Surfaces

*Stroock, D. W.:* An Introduction to the Theory of Large Deviations

*Sunder, V. S.:* An Invitation to von Neumann Algebras

*Tamme, G.:* Introduction to Étale Cohomology

*Tondeur, P.:* Foliations on Riemannian Manifolds

*Verhulst, F.:* Nonlinear Differential Equations and Dynamical Systems

*Wong, M. W.:* Weyl Transforms

*Xambó-Descamps, S.:* Block Error-Correcting Codes

*Zaanen, A.C.:* Continuity, Integration and Fourier Theory

*Zhang, F.:* Matrix Theory

*Zong, C.:* Sphere Packings

*Zong, C.:* Strange Phenomena in Convex and Discrete Geometry

*Zorich, V. A.:* Mathematical Analysis I

*Zorich, V. A.:* Mathematical Analysis II