Scuola Normale Superiore

Tesi di Perfezionamento in Matematica per la Finanza

Stochastic Optimal Control Problems for Pension Funds Management

Candidato: Salvatore Federico
Relatore: Prof. Fausto Gozzi

Anno Accademico: 2008-09

Ai miei nonni
A Gino
A Filomena

Acknowledgements

My special thanks are to my advisor, Prof. Fausto Gozzi: he introduced me to the mathematical field of Control Theory and it is due to him the most I know in this topic. But, mainly, I thank him for his kindness and patience during these years: we started as student and professor and we ended up as friends.

Special thanks go to Prof. Maurizio Pratelli, who introduced me to the Mathematics of Finance and gave his support in these years.

I thank also Prof. Giuseppe Da Prato for his teaching in Stochastic Analysis and for some useful suggestions on the infinite-dimensional problem treated in this thesis.

My thanks go also to Prof. Ben Goldys and to the Department of Statistics of the University of New South Wales for their invitation in Sydney in the period April-June 2008, where I wrote the contents of the first part of the third chapter. I thank Prof. Goldys for his kindness and for his guide and the Australian Research Council for the financial support.

I thank Prof. Bernt Øksendal and Prof. Giulia Di Nunno: I finished to write the thesis in Oslo in the months May-June 2009, when I was invited by Prof. Øksendal to visit the Centre of Mathematics for Applications of the University of Oslo. They were very warm with me and I worked really quietly there. I also thank the European Science Fundation - Advanced Mathematical Methods for Finance for the grant and the financial support for this visit.

I thank Prof. Andrzej Swiech for valuable comments on a suitable definition of visosity solution in infinite dimension for the problem of Chapter 2.

I thank Prof. Mogens Steffensen for his kind reading of the Introduction and for his valuable suggestions on the paragraph regarding the literature on Pension Funds that led to a real improvement of this part.

I thank Prof. Martino Grasselli for his nice care of me and for his very useful suggestions on the modeling features of the surplus appearing in the second chapter.

Finally I thank Dr. Elena Vigna and Dr. Marina Di Giacinto: the last chapter of the thesis is based on a joint work and it was really nice to work with them. In particular I thank Elena Vigna for having supplied to me a list of papers on the subject of Pension Funds.

Contents

Introduction

The present thesis is mainly devoted to present, study and develop the mathematical theory for a model of asset-liability management for pension funds.

Pension funds have become a very important subject of investigation for researchers in the last years. The reasons are quite clear: when life expectancy was relatively low, providing for old age was still not a major economic issue; now, since the average age of people and the expected lifetime have strongly increased in the last decades (and this trend is expected to continue), the way to organize a pension system providing for old age and mantaining economic efficiency and growth is a fundamental challenge for advances countries in the future.

Roughly speaking pension funds can be viewed as a form of forced savings, where external cashflows (contributions and benefits) enter into the dynamics of the fund wealth. The book [Davis; 1995] provides an overview of the economic issues related to the development of pension funds schemes to complement social security.

Basically we can distinguish two kinds of plans for pension funds:

- *Defined-benefit plans*, where the benefits are defined in advance and the contributions are adjusted in order to ensure that the fund remains in balance; in this case the financial risks are charged to the sponsor of the pension fund.

- *Defined-contribution plans*, where the contributions are defined in advance and the benefits depend on the return of the fund; in this case the financial risks are charged to the workers.

Although from a historical perspective defined-benefit plans have been more popular, in the last years most of the plans have been based on defined contributions. Since in the latter case the benefits are not fixed and the worker is directly exposed to the financial risk of the plan, a key issue in this context is the presence in the plan of the so called *minimum guarantee*. This is a lower bound for the benefits to pay to the workers in retirement, so it represents a downside protection against the investment-risk for the worker adhering to the the fund. Another important feature to consider dealing with pension funds is that they are usually constrained, by law, to keep their wealth above a certain level, which is the so called *solvency level*. Both these features (the minimum guarantee and the solvency level) lead to a restriction on the investment choices of the manager: basically a substantial part of the wealth will have to be invested in the riskless asset in order to guarantee that these two requirements are satisfied.

Finally, besides of the rules and of their structure, pension funds are characterized by their aims: within these rules they have the freedom to choose some variables and they should operate this choice in order to optimize some criterion. Such a criterion should take into account the point of view of the management of the fund and/or the point of view of the members of the fund.

We can summarize the description above, considering the management of a pension fund as a portfolio allocation problem with some specific features:

- the presence of external cashflows represented by the inflow of contributions $c(t)$, paid by the members who are adhering to the fund, and by the outflow of benefits $b(t)$, paid to the members in retirement;
- the presence of a minumum guarantee flow $g(t)$ for the benefits;
- the presence of a solvency level $l(t)$ for the wealth;
- the criterion (objective functional) to be optimized.

In this thesis we take the context of a standard Black-Scholes model for the market. Hence, denoting by $\theta(t)$ the proportion of wealth invested in the risky asset and by $X(t)$ the wealth of the fund, we are led to consider a dynamics of this kind for the wealth

$$dX(t) = \big[\big(r + \sigma\lambda\theta(t)\big)X(t)\big]dt + \sigma\theta(t)X(t)dB(t) + \Big[\ \underset{\substack{\big\downarrow \\ \text{contributions}}}{c(t)} \ - \ \underset{\substack{\big\downarrow \\ \text{benefits}}}{b(t)}\ \Big]\,dt,$$

with the constraints $X(t) \geq l(t)$ and $b(t) \geq g(t)$ and where r, σ, λ are the classical parameters of the Black-Scholes model, i.e. respectively the interest rate, the volatility of the risky asset and the risk premium. As usual in portfolio optimization problems, $\theta(\cdot)$ is the control variable to choose in order to optimize some objective.

Of course this is still a general description and the key issue is to model the terms $c(t), b(t), g(t), l(t)$. In general they should be considered as stochastic processes modelled on the financial and demographic features of the "world" and on the structure (and the aims) of the pension plan. The same consideration holds for choice of the criterion to optimize.

In Chapters 1 & 2 we deal with a pension fund operating over an infinite horizon and continuously open to entrance and exit of people. The model keeps a global perspective, in the sense that it considers the management of contributions and benefits of the whole community of members. The dynamics of the demographical processes in the model is still naive, as it considers the demographical variables as stationary (on the other hand we stress that a model keeping into account all these features in a nontrivial way is far from any analitical treatment, which is our aim). Indeed we consider the processes $c(\cdot)$ and $g(\cdot)$ as deterministic functions, while we set for the cumulative benefits process $b(\cdot)$ the structure

$$b(t) = g(t) + s\left(t, X(t+\xi)_{\xi\in[-T,0]}\right),$$

where $T > 0$ is a constant representing the time of adherence to the fund of the generic worker. The difference

$$b(t) - g(t) = s\left(t, X(t+\xi)_{\xi\in[-T,0]}\right)$$

represents a surplus contract defined in advance between the fund and the worker. It naturally depends on the past wealth of the fund during the period in which the worker in retirement has been adhering to the fund. When $s \equiv 0$ no surplus is paid to the members in retirement, who receive just the minimum guaranteee (this is the case treated in Chapter 1). On the other hand, in order to make the fund more appealing for the workers, it is natural (and this really happens) that the fund provides a surplus contract over the minimum guarantee. It is natural for pension funds to choose a kind of surplus contract which increasing on the growth of the fund in the past period $[t, t-T]$ and this is done in Chapter 2. The objective functional takes the point of view of the manager, but, when the surplus is nonzero (Chapter 2), to some extent the optimization of the manager's profit has as a direct consequence an improvement of the workers' benefits. Hence, the presence of the surplus term in the model seems to proceed towards the direction of a "well-planned" pension plan, in the sense that the pension plan is set in such a way that the interest of the manager meets the interest of the workers.

Chapter 4 is involved with a pension plan for a single pensioner in the so called decumulation phase. This means that the pension fund deals with the management of the pension of a single pensioner over a finite time horizon, i.e. for a certain number of years after the date of retirement of the worker. In this case no demographical feature is involved in the model, since it deals with an individual perspective (this is quite usual in some literature: see Subsection 0.2.2). Referring to the model described above, in Chapter 4 we will have $c(\cdot) \equiv 0$ and $b(\cdot) \equiv b_0 > 0$, while the objective functional, as usual in the individual perpective, keeps the point of view of the pensioner.

The optimal control problems arising in the thesis can be satisfactory treated in the case of Chapter 1 & 4, while this cannot be done in the case of Chapter 2, when the surplus term appears in the model. This is due to the presence of a delay term in the state equation, which makes the problem much more difficult to treat and very far from any known result in literature. In this case the analytical treatment of the problem is stopped at an unsatisfactory stage. For this reason we study the problem of Chapter 3, that can be considered as an easier reduction of the problem faced in Chapter 2. Although also this problem is far from the existing literature, in this case we are able to provide a satisfactory treatment of the problem and to give deeper answers to it.

0.1 A mathematical tool for portfolio management problems: Optimal Control Theory and Dynamic Programming

Since the celebrated papers [Merton; 1969] and [Merton; 1971], problems of optimal portfolio allocation have been naturally formulated (also) as stochastic control problems. As we said above, the problem of managing a pension fund can be viewed as an optimal portfolio allocation problem with external cashflows, which, due to the special rules of the fund, is subject to particular constraints. So, from the mathematical point of view, the problem of the optimal management of a pension fund can be quite naturally formulated as a stochastic control problem with constraints on the state (the wealth of the fund) and on the control (the investment strategy).

From a historical point of view, the mathematical theory of optimal control problems goes back to the 50s and has been developed basically along two lines:

1. *Dynamic Programming approach* (Bellman and the American school; we refer to the monograph [Bellman; 1957]);

2. *Maximum Principle approach* (Pontryagin and the Russian school; we refer to the monograph [Pontryagin, Boltyanski, Gamkerildze, Mishenko; 1962]).

In this thesis we follow the Dynamic Programming approach. We use the remainder of this subsection to give a brief informal description of it, discussing its main features and the possible difficulties. About the Maximum Principle approach we refer to [Yong, Zhou; 1999], Chapter 3. Here we only observe that, while the Dynamic Programming approach can be naturally extended from the deterministic case to the stochastic case, the same consideration does not hold for the Maximum Principle approach. Indeed the Maximum Principle requires the concept of *backward solution* to be stated, but, as we know, in the stochstic systems the time has a priviliged direction. This makes the extension to stochastic systems not trivial.

We give the description of the Dynamic Programming approach for both the deterministic and the stochastic case, but, for sake of simplicity, we restrict the description to the finite-horizon and unconstrained case. However, the same ideas can be adapted working case by case with the specific features of the problem. For a detailed description of this theory in both the deterministic and the stochastic case, we refer to two classic books dealing with this subject, i.e. [Fleming, Soner; 1993] and [Yong, Zhou; 1999]; the latter one contains in Chapter 5 also an analysis of the relationship existing between Dynamic Programming and Maximum Principle.

To start with, consider a one-dimensional controlled system

$$\begin{cases} x'(t) = b(t, x(t), u(t)), & t \in [0, T], \\ x(0) = x_0 \in \mathbb{R}, \end{cases} \tag{1}$$

where $x(\cdot)$ represents the state variable of the system taking values in $\mathbb{R}$ and $u(\cdot)$ the control variable of the system taking values in some compact control set $U \subset \mathbb{R}$. We consider the set of the admissible control functions

$$\mathcal{U}_{ad}^{det}[0, T] = \{u : [0, T] \to U \mid u(\cdot) \text{ is measurable}\}, \tag{2}$$

and suppose that b is sufficiently regualr to guarantee existence and uniqueness of solutions for the state equation (1) (e.g., we may suppose that b is Lipschitz continuous with respect to its arguments). The optimization problem consists in maximizing the functional

$$\mathcal{U}_{ad}^{det}[0, T] \ni u(\cdot) \longmapsto \int_0^T f(t, x(t), u(t))dt + h(x(T)), \tag{3}$$

where $x(\cdot) := x(\cdot; u(\cdot))$ is the unique solution to (1) and where f, h are given functions (we suppose by sake of simplicity also that f, g have sublinear growth in order to guarantee that the functional above is well-defined).

The stochastic counterpart of the deterministic system above would be a one-dimensional stochastic controlled diffusion

$$\begin{cases} dX(t) = b(t, X(t), u(t))dt + \sigma(t, X(t), u(t))dB(t), \qquad t \in [0, T], \\ X(0) = x_0 \in \mathbb{R}, \end{cases} \tag{4}$$

where $B(\cdot)$ is a standard brownian motion defined on some probability space $(\Omega, \mathcal{F}, \mathbb{P})$, $X(\cdot)$ represents the state variable of the system taking values in $\mathbb{R}$ and $u(\cdot)$ the control variable of the system taking values in some compact control set $U \subset \mathbb{R}$. In this case the set of admissible controls is a space of processes, e.g.

$$\begin{aligned} &\mathcal{U}_{ad}^{st}[0, T] = \{u : \Omega \times [0, T] \to U \mid \\ &\quad (u(t))_{t\in[0,T]} \text{ is progressively measurable with respect to } (\mathcal{F}_t^B)_{t\in[0,T]}\}. \end{aligned} \tag{5}$$

As well as in the deterministic case, we suppose that b, σ are sufficiently regualr to guarantee existence and uniqueness of solutions for the state equation (4) (e.g., we may suppose that b, σ are Lipschitz continuous with respect to their arguments). In this case the optimization problem consists in maximizing over the set of admissible controls the functional

$$\mathcal{U}_{ad}^{st}[0, T] \ni u(\cdot) \longmapsto \mathbb{E}\left[\int_0^T f(t, X(t), u(t))dt + h(X(T))\right]. \tag{6}$$

where $X(\cdot) := X(\cdot; u(\cdot))$ is the unique solution to (4) and where again, in order to guarantee that the functional above is well-defined, we suppose that f, h are given functions having sublinear growth.

***

The **first step** in Dynamic Programming consists in defining the latter problems for generic initial data $(s, x) \in [0, T] \times \mathbb{R}$, i.e. replacing the state equations (1)-(4), the set of the admissible control functions-processes (2)-(5) and the functional (3)-(6) respectively with

- in the deterministic case

$$\begin{cases} x'(t) = b(t, x(t), u(t)), \qquad t \in [s, T], \\ x(s) = y \in \mathbb{R}, \end{cases} \tag{7}$$

$$\mathcal{U}_{ad}^{det}[s, T] = \{u(\cdot) : [s, T] \to U \text{ measurable}\}, \tag{8}$$

$$\mathcal{U}_{ad}^{det}[s, T] \ni u(\cdot) \longmapsto J^{det}(s, y; u(\cdot)) = \int_s^T f(t, x(t), u(t))dt + h(x(T)), \tag{9}$$

where $x(\cdot) := x(\cdot; s, y, u(\cdot))$ is the solution to (7).

- in the stochastic case

$$\begin{cases} dX(t) = b(t, X(t), u(t))dt + \sigma(t, X(t), u(t))dB(t), \qquad t \in [s, T], \\ X(s) = y \in \mathbb{R}, \end{cases} \tag{10}$$

$$\mathcal{U}_{ad}^{st}[s,T] = \big\{u(\cdot) : \Omega \times [s,T] \to U \\ \text{progressively measurable with respect to } \mathcal{F}^B\big\}, \tag{11}$$

$$\mathcal{U}_{ad}^{st}[s,T] \ni u(\cdot) \longmapsto J^{st}(s, y; u(\cdot)) \\ = \mathbb{E}\left[\int_s^T f(t, X(t), u(t))dt + h(X(T))\right], \tag{12}$$

where $X(\cdot) := X(\cdot; s, y, u(\cdot))$ is the solution to (10).

***

The **second step** simply consists in defining the optimum function for this class of problems, i.e. the so-called *value function*:

- in the deterministic case

$$V(s, y) = \sup_{u(\cdot) \in \mathcal{U}_{ad}^{det}[s,T]} J^{det}(s, y; u(\cdot));$$

- in the stochastic case

$$V(s, y) = \sup_{u(\cdot) \in \mathcal{U}_{ad}^{st}[s,T]} J^{st}(s, y; u(\cdot)).$$

***

The problem now is to study the above function. Indeed, the main goal is to characterize the value function in order to use it to find optimal strategies for the problem as we will see below. Therefore, the **third step** consists in stating an equation solved by this function. The crucial point for stating this equation is the so-called *Dynamic Programming Principle*; in words this principle, as stated in [Bellman; 1957], is the following:

"An optimal policy has the property that whatever the initial state and the initial decision are, the remaining decisions must constitute an optimal policy with regard to the state resulting from the first decision."

Elaborated in mathematical language for deterministic systems, this principle means that, if $0 \le s \le \hat{s} \le T$, $\bar{u}(\cdot) \in \mathcal{U}_{ad}^{det}[s,T]$ and $\bar{x}(\cdot)$ is the solution to (7) under the control $\bar{u}(\cdot)$,

$$\begin{aligned}&\bar{u}(\cdot)|_{[s,T]} \text{ optimal on } [s,T] \text{ with the initial } (s,y)\\ &\qquad\qquad \Rightarrow \bar{u}(\cdot)|_{[\hat{s},T]} \text{ optimal on } [\hat{s},T] \text{ with the initial } (\hat{s},\bar{x}(\hat{s})).\end{aligned}$$

This holds also for stochastic systems, i.e. if $\bar{u}(\cdot) \in \mathcal{U}_{ad}^{st}[s,T]$ and $\bar{X}(\cdot)$ is the solution to (7) under the control $\bar{u}(\cdot)$, then

$$\begin{aligned}&\bar{u}(\cdot)|_{[s,T]} \text{ optimal on } [s,T] \text{ with the initial } (s,y)\\ &\qquad\qquad \Rightarrow \bar{u}(\cdot)|_{[\hat{s},T]} \text{ optimal on } [\hat{s},T] \text{ with the initial } (\hat{s},\bar{X}(\hat{s})).\end{aligned}$$

In terms of value function this statement can be seen as a consequence of the fact that the value function solves a functional equation.

- In the deterministic case

$$V(s,y) = \sup_{u(\cdot)\in\mathcal{U}^{det}[s,T]} \left[\int_s^{\hat{s}} f(t,x(t),u(t))dt + V(\hat{s},x(\hat{s}))\right], \quad 0 \le s \le \hat{s} \le T. \tag{13}$$

- In the stochastic case

$$V(s,y) = \sup_{u(\cdot)\in\mathcal{U}^{st}[s,T]} \mathbb{E}\left[\int_s^{\hat{s}} f(t,X(t),u(t))dt + V(\hat{s},X(\hat{s}))\right], \; 0 \le s \le \hat{s} \le T. \tag{14}$$

Of course (13)-(14) need to be proved. While in the deterministic framework the proof of (13) does not give trouble, in the stochastic framework the proof of (14) could present some problems. Indeed, when the value function is known to be continuous, the proof is quite standard (see, e.g., the classical references [Fleming, Soner; 1993] or [Yong, Zhou; 1999]). But if we do not know whether the value function is continuous or not, then a more subtle argument, requiring a measurable selection result, is needed to prove (14) (see e.g. [Soner; 2004] or [Soner, Touzi; 2002]).

***

Once we have proved (13)-(14), we wish to use them to study the value function. Unfortunately (13)-(14) result very difficult to treat. Therefore the idea of the **fourth step** is to write them in a differential form, getting the so called Hamilton-Jacobi-Bellman (HJB) equation. This equation is obtained passing formally to the limit, for $\hat{s} \downarrow s$, (13) or (14) and imposing the natural boundary condition given by the control problem. At this stage a quite substantial difference between the deterministc case and the stochastic case arises. Indeed, while the HJB equation for deterministic optimal control problems is a first order PDE, the HJB equation for stochastic optimal control problems is a second order PDE.

- In the detreministic case, passing to the limit (13) leads to the equation

$$\begin{cases} -v_s(s,y) = \mathcal{H}(s,y,v_y(s,y)), & (s,y) \in [0,T) \times \mathbb{R}, \\ v(T,y) = h(y), & x \in \mathbb{R}, \end{cases} \tag{15}$$

  where

$$\mathcal{H}(s,y,p) = \sup_{u \in U} \{ b(s,y,u)p + f(s,y,u) \}. \tag{16}$$

- In the stochstic case, passing to the limit (14) leads to the equation

$$\begin{cases} -v_s(s,y) = \mathcal{H}(s,y,v_y(s,y),v_{yy}(s,y)), & (s,y) \in [0,T) \times \mathbb{R}, \\ v(T,y) = h(y), & y \in \mathbb{R}, \end{cases} \tag{17}$$

  where

$$\mathcal{H}(s,y,p,P) = \sup_{u \in U} \left\{ \frac{1}{2}\sigma(s,x,u)^2 P + b(s,y,u)p + f(s,y,u) \right\}. \tag{18}$$

The function $\mathcal{H}$ in (15) and (17) is called Hamiltonian. We notice that in both cases the Hamiltonian is a concave function of its arguments, as it is the supremum of linear functions. (In the case of minimization problems the Hamiltonian would be convex.) This is a characteristic feature of HJB equation: the nonlinear part of the equation is concave (or convex).

We stress that the passage from (13)-(14) to (15)-(17) is a delicate point, because the argument to get the HJB equation from (13) and (14) is only formal (this is why we have replaced V with the formal function v). Indeed, what is true is the following.

- In the deterministic case, if the value function is $C^{1,1}([0,T] \times \mathbb{R})$, then it solves the HJB equation (15) (under further reasonable assumptions; see [Yong, Zhou; 1999], Chapter 4, Proposition 2.2).

- In the stochastic case, if the value function is $C^{1,2}([0,T] \times \mathbb{R})$, then it solves the HJB equation (17) (under further reasonable assumptions; see [Yong, Zhou; 1999], Chapter 4, Proposition 3.5).

Nevertheless it is not true in general that the value function has this kind of regularity and there are many nonpathological examples of nonsmooth value functions (see Subsection 0.1, or also [Bardi, Capuzzo-Dolcetta; 1997], Chapter 1, Example 3.1, and [Yong, Zhou; 1999], Chapter 4, Example 2.3). Moreover, even if the value function is actually smooth, usually it is not possible, working just with its definition, to prove regularity results for it going beyond the continuity.

***

The **fifth step** consists in studying the HJB equation in order to characterize the value function as its unique solution (in some sense). Now the question is how to approach this equation. Basically there are three main approaches to this problem (that may overlap each other):

(1) finding explicit solutions;

(2) using the classical theory of PDEs;

(3) using the theory of viscosity solutions of PDEs.

Before to proceed in this description, we want to stress that, roughly speaking, a good theory for PDEs is a theory which provides uniqueness of solutions and good possibilities to prove regularity results for these solutions.

(1) Finding an explicit solution to the HJB equation would be of course the best thing we can obtain. On the other hand the possibility of finding an explicit solution relies intrinsecally in the definition of the problem, because it is strongly related to the functional parameters of the problem (i.e. the functions b, σ, f, h). It is clear that we may expect to find an explicit solution only in very few cases and this is very uncomfortable.

(2) We distingush the first order case and the second order case, which are quite different. Moreover in the description, for sake of simplicity, we refer to PDEs of HJB type, i.e. to PDEs arising from control problems.

(i) **First order case.** For this case, in particular with regard to HJB euations, we refer to the classic book [Evans; 1998], Chapters 3 and 10. Here we observe that the usual concept of weak solutions (Lipschitz continuous function solving in classical sense the equation almost everywhere) does not guarantee in this case the desired uniqueness (see [Evans; 1998], Example at page 129); some extra-conditions on the data are needed. On the other hand, we cannot expect the existence of solutions in classical sense in general (see [Evans; 1998], Chapter 3, Section 2).

(ii) **Second order case.** In this case we have to distingush two subcases: nondegenerate case and degenerate case.

- *Nondegenerate case.* In the context of stochastic control problems an HJB equation is called *nondegenerate* if the the diffusion coefficient in the state equation is far from 0 uniformly with respect to its arguments (in more dimensions, the eigenvalues of the positive semidefinite diffusion matrix $\sigma\sigma^*$ are far from 0 uniformly with respect to the arguments).

  The classical theory of second order elliptic and parabolic nondegenerate PDEs has been well developed in the *linear* and *semilinear* case - in the context of stochastic control problems this corresponds to have respectively no control and control only in the drift in the state equation. For this theory we refer to the classic books [Gilbarg, Trudinger; 1983] for the elliptic case, [Ladyzenskaja, 1968] for the parabolic case and [Evans; 1998] for both them.

  A classical theory for *fully nonlinear* PDEs - in the context of stochastic control problems this corresponds to have the control in the diffusion coefficient - such that the nonlinear term is concave with respect to the second derivative (but, as we said, this is a natural property of HJB-type equations), has been developed later indipendently by Krylov and Evans (see [Evans; 1982], [Krylov; 1983] and the book [Krylov; 1987]).

- *Degenerate case.* If the equation is degenerate the classical theory basically presents the same problems of the first-order case, i.e. it does not guarantee uniqueness of weak solutions and we cannot expect in general the existence of classical solutions (see, e.g., Section 6.6 of [Gilbarg, Trudinger; 1983] for an example even in the linear case). This constitutes a limit for studying some HJB arising in the applications. Indeed, in particular control problems arising in finance unavoidably lead to have both the problematic features in the HJB equation: the presence of the control in the diffusion term of the state equation leads to a fully nonlinearity in the equation, while the fact that the diffusion term can vanish leads to a degeneracy in the equation.

(3) On the basis of the considerations done above, it is clear that the classical theory of PDEs cannot be considered completely satisfactory to treat in a wide generality HJB equations. This fact has represented for a long while an uncomfortable obstacle in the dynamic programming theory. Indeed, we would like to see the value function, even when it is not smooth, as unique solution, in some other weaker sense, of the associated HJB equation; but in which sense? The solution to this problem came in the early 80s, when Crandall and Lions introduced the concept of viscosity solutions for PDEs. We refer, for a description associated to control problems, to [Bardi, Capuzzo-Dolcetta; 1997] in the deterministic case and to the already cited books [Fleming, Soner; 1993] and [Yong, Zhou; 1999] in both the deterministic and the stochastic case. The concept of viscosity solution requires only continuity to be defined (actually even less) and seems to be the right one to approach in a wide generality this kind of equations. Indeed the theory of viscosity solutions fits very well the case of fully nonlinear as well as degenerate equations and provides in many cases a characterization of the value function as unique solution to the HJB equation in this sense. However, we have also to say that from a theoretical point of view the characterization of the value function as unique viscosity solution to the HJB equation is not easy to use (especially in the stochastic case) to give a solution to the problem in the sense of finding optimal controls in a form that can be used in the applications. Then, in order to overcome this obstacle and proceed towards the next steps to get computable optimal controls for the problem, the challenge is to prove regularity results for viscosity solutions. For this kind of results we give the following references.

- In the first order case we have such kind of results for both the finite-dimensional and the infinite-dimensional case.
  * For the finite dimensional framework, we refer to the books [Bardi, Capuzzo-Dolcetta; 1997] and [Cannarsa, Sinestrari; 2005] and the references therein.
  * For the infinite-dimensional framework, we refer to the paper [Federico, Goldys, Gozzi; 2009a] and to Chapter 3 of this thesis.
- In the second order case for degenerate HJB equations, there is not a general theory for this topic, but there are some papers which prove such results working, case by case, with the specific structure of the HJB equation. We mention
  * [Choulli, Taksar, Zhou; 2003], [Di Giacinto, Federico, Gozzi; 2009], [Morimoto, 2008]: in these papers basically arguments of convex analysis are used;

* [Zariphopoulou; 1994]: here the author uses a different technique, consisting in an approximation procedure of the equation.

Finally we should mention (although this does not involve the case of HJB equations) that the theory of viscosity solutions provides a new and very powerful perspective for proving regularity results in the context of fully nonlinear nondegenerate PDEs that are not covered by the concavity (convexity) assumptions of Krylov and Evans on the nonlinear term. The fundamental work [Caffarelli; 1989], where some interior a priori estimates for solutions of some classes of PDEs are proved, opened a way in this sense. The book [Cabré, Caffarelli; 1995a] provides a survey on this theory up to 1995. For this theory we refer also to the following list of papers:

- [Cabré, Caffarelli; 1995b],
- [Cabré, Caffarelli; 2003],
- [Caffarelli, Crandall, Kocan, Swiech; 1996],
- [Caffarelli, Yuan; 2000],
- [Escauriaza; 1993],
- [Swiech; 1997],
- [Wang I; 1992], [Wang II; 1992], [Wang III; 1992].

***

The **sixth step** uses the solution to the HJB equation to provide necessary and sufficient conditions to be a control $\bar{u}_{s,y}(\cdot)$ optimal for the problem starting from (s, y), which means
- in the deterministic case that $\bar{u}_{s,y}(\cdot) \in \mathcal{U}^{det}_{ad}[s,T]$ is such that

$$V(s,y) = J^{det}(s,y;\bar{u}_{s,y}(\cdot));$$

- in the stochastic case that $\bar{u}_{s,y}(\cdot) \in \mathcal{U}^{st}_{ad}[s,T]$ is such that

$$V(s,y) = J^{st}(s,y;\bar{u}_{s,y}(\cdot)).$$

These conditions are stated in a so called *classical verification theorem*.

- *Deterministic case.*

Theorem 0.1.1. *Let $v \in C^{1,1}([0,T] \times \mathbb{R})$ be a solution to* (15).

- *(Sufficient condition for optimality).*
  Let $(s, y) \in [0,T] \times \mathbb{R}$ and let $\bar{u}(\cdot) \in \mathcal{U}^{det}[s,T]$ and $\bar{x}(\cdot)$ the associated state trajectory. If for almost every $t \in [s,T]$

$$\mathcal{H}(t,\bar{x}(t),v_y(t,\bar{x}(t))) = b(t,\bar{x}(t),\bar{u}(t))v_y(t,\bar{x}(t)) + f(t,\bar{x}(t),\bar{u}(t)), \tag{19}$$

  then $v(s,y) = V(s,y)$ and $\bar{u}(\cdot)$ is optimal starting from (s,y).

– *(Necessary condition for optimality).*
Let $(s,y) \in [0,T] \times \mathbb{R}$ and let $\bar{u}(\cdot) \in \mathcal{U}^{det}[s,T]$ an optimal startegy for the initial (s,y) and $\bar{x}(\cdot)$ the associated state trajectory. If we know from the beginning that $v(s,y) = V(s,y)$, then $\bar{u}(\cdot)$ must maximize $\mathcal{H}$, i.e. must satify (19).

- *Stochastic case.*

Theorem 0.1.2. *Let $v \in C^{1,2}([0,T] \times \mathbb{R})$ be a solution to* (17).

– *(Sufficinent condition for optimality).*
Let $(s,y) \in [0,T] \times \mathbb{R}$ and let $\bar{u}(\cdot) \in \mathcal{U}^{st}[s,T]$ and $\bar{X}(\cdot)$ the associated state trajectory. If for almost every $(t,\omega) \in [s,T] \times \Omega$

$$\begin{aligned}
\mathcal{H}(t, \bar{X}(t), v_y(t, \bar{X}(t)), v_{yy}(t, \bar{X}(t))) \\
&= \frac{1}{2}\sigma^2(t, \bar{X}(t), \bar{u}(t)) v_{yy}(t, \bar{X}(t), \bar{u}(t)) \\
&\quad + b(t, \bar{X}(t), \bar{u}(t)) v_y(t, \bar{X}(t)) + f(t, \bar{X}(t), \bar{u}(t)), \quad (20)
\end{aligned}$$

then $v(s,y) = V(s,y)$ and $\bar{u}(\cdot)$ is optimal starting from (s,y).

– *(Necessary condition for optimality).*
Let $(s,y) \in [0,T] \times \mathbb{R}$ and let $\bar{u}(\cdot) \in \mathcal{U}^{st}[s,T]$ an optimal startegy for the initial (s,y) and $\bar{X}(\cdot)$ the associated state trajectory. If we know from the beginning that $v(s,y) = V(s,y)$, then $\bar{u}(\cdot)$ must maximize $\mathcal{H}$, i.e. must satify (20).

***

The **seventh step** concludes the program giving optimal control strategies. Indeed, the sixth step provides a way to construct optimal strategies starting from the solution to the HJB equation. The key issue is the study of the so called *closed loop equation*.

- *Deterministic case.* Suppose that

– For every $(t,z,p) \in [0,T] \times \mathbb{R}^2$, the map

$$U \ni u \longmapsto (t,z,u)p + f(t,z,u)$$

admits a unique maximizer $G^{det}_{max}(t,z,p)$.

– We know that the HJB equation has a solution $v \in C^{1,1}([0,T] \times \mathbb{R})$. Then we are able to define the so called *feedback map*

$$\begin{array}{cccc} G^{det}: & [0,T] \times \mathbb{R} & \longrightarrow & U, \\ & (t,z) & \longmapsto & G^{det}_{max}(t,z,v_y(t,z)). \end{array}$$

Then we have to study the state equation replicing the control $u(\cdot)$ with $G^{det}(t, x(t))$, i.e.

$$\begin{cases} x'(t) = b\left(t, x(t), G^{det}(t, x(t))\right), & t \in [s, T], \\ x(s) = y \in \mathbb{R}. \end{cases} \tag{21}$$

Then the last obstacle is to prove existence and uniqueness of a solution to (21). Indeed, straightly by Theorem 0.1.1, such a solution $x^*(\cdot)$ would provide an optimal strategy starting from (s, y), i.e.

$$u^*_{s,y}(t) = G^{det}(t, x^*(t)), \quad t \in [s, T]. \tag{22}$$

- *Stochastic case.*

  Suppose that

  - For every $(t, z, p, P) \in [0, T] \times \mathbb{R}^3$,

    $$U \ni u \longmapsto \frac{1}{2}\sigma(t, z, u)^2 P + b(t, z, u)p + f(t, z, u)$$

    admits a unique maximizer $G^{st}_{max}(t, z, p, P)$.
  - We know that the HJB equation has a solution $v \in C^{1,2}([0, T] \times \mathbb{R})$. Then we are able to define the so called *feedback map*

    $$\begin{array}{rccc} G^{st}: & [0, T] \times \mathbb{R} & \longrightarrow & U, \\ & (t, z) & \longmapsto & G^{st}_{max}(t, z, v_y(t, z), v_{yy}(t, z)). \end{array}$$

  Also in this case we have to study the state equation replicing the control $u(\cdot)$ with $G^{st}(t, X(t))$, i.e.

  $$\begin{cases} dX(t) = b(t, X(t), G^{st}(t, X(t)))dt + \sigma(t, X(t), G^{st}(t, X(t)))dB(t), \ t \in [s, T], \\ X(s) = y, \end{cases} \tag{23}$$

  Again it remains only to prove existence and uniqueness of a solution to (21). Then, straightly by Theorem 0.1.2, such a solution $X^*(\cdot)$ would provide an optimal strategy starting from (s, y), i.e.

  $$u^*_{s,y}(t) = G^{st}(t, X^*(t)), \quad t \in [s, T]. \tag{24}$$

***

We summarize the steps of the program described above.

1. Defining the problem for varying initial data.

2. Defining the value function, i.e. the function representing the optimal values of the problem with respect to the initial data.

3. Stating and proving the Dynamic Programming Principle, i.e. a functional equation solved by the value function.

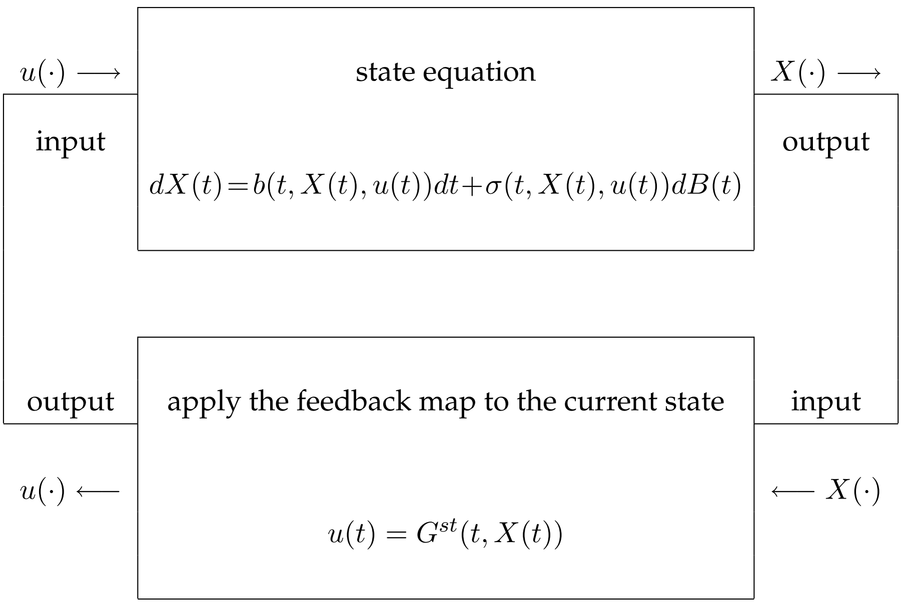
4. Passing to the limit the Dynamic Proggramming Principle in order to get the HJB equation, i.e. a PDE which is formally solved by the value function.

5. Studying (in some suitable sense) the HJB equation and its relationship with the value function.

6. Stating and proving Verification Theorems, yielding necessary and sufficient conditions of optimality.

7. Studying the closed loop equation arising from the feedack map in order to construct optimal feedback strategies for the problem according with the Verification Theorem proved.

The **final situation** of this program can be represented by the following picture (done for stochastic systems)



If everything works, the map G^{st} gives a "feedback" answer to our optimization problem: it gives the optimal current decision as function of the current state of the system. This is a very nice solution for the problem, as it is easily computable.

Viscosity solutions in a simple example

In order to exemplify the concept of viscosity solution and its use in optimal control problems, we give a simple example of a deterministic control problem whose HJB equation can be successfully approached with this concept. So, consider the following simple control problem. The state equation is

$$\begin{cases} y'(t) = u(t), \\ y(0) = x \in [-1,1], \end{cases}$$

where the measurable control function $u(\cdot)$ takes values in the control set $[-1,1]$. Denote by $y(\cdot; x, u(\cdot))$ the solution to this equation under the control $u(\cdot)$. The goal is to minimize the time at which the state variable $y(t)$ reaches the set $\{-1,1\}$, i.e. the functional

$$J(x; u(\cdot)) := \inf\{t \in [0,+\infty) \mid y(t; x, u(\cdot)) \in \{-1,1\}\}.$$

It is clear that the solution to this problem consists in keeping the strategy

$$\begin{cases} u(\cdot) \equiv 1, & \text{if } x \in (0,1], \\ u(\cdot) \equiv -1, & \text{if } x \in [-1,0), \\ \text{indifferently } u(\cdot) \equiv 1 \text{ or } u(\cdot) \equiv -1, & \text{if } x = 0. \end{cases}$$

The value function of this problem is clearly independent of time and explicitly computable at $t = 0$; as function of the only initial state x, it is the function

$$V(x) = \begin{cases} x + 1, & \text{for } x \in [-1,0], \\ -x + 1, & \text{for } x \in (0,1], \end{cases} \tag{25}$$

represented in Figure 1.1.

It is possible to state a dynamic programming principle also for this kind of control problems (the so called minimum time problems, which are different from the "standard" control problems described above) and to associate a HJB equation to V. In this case the HJB equation is

$$\begin{cases} |v'(x)| = 1, \\ v(-1) = v(1) = 0. \end{cases} \tag{26}$$

Clearly this equation does not admit any classical solution. We observe that the value function V solves this equation in classical sense at any point $x \in [-1,0) \cup (0,1]$. Therefore, the question is: in which sense does V solve the equation at $x = 0$? The "viscosity" answer to this question relies in this observation:

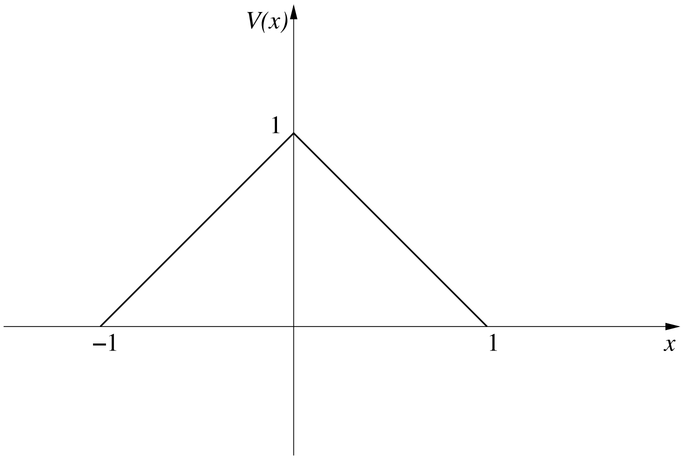


Figure 1: the value function V given in (25).

if we replace the function V with any smooth function $g \in C^1([-1,1];\mathbb{R})$ such that $g(0) = V(0)$ and $g(x) \geq V(x)$ in a neighborhood of 0 (see Figure 1.2), we have that g is a subolution of the equation at 0, i.e. $|g'(0)| \leq 1$. This leads to the concept of viscosity solution.

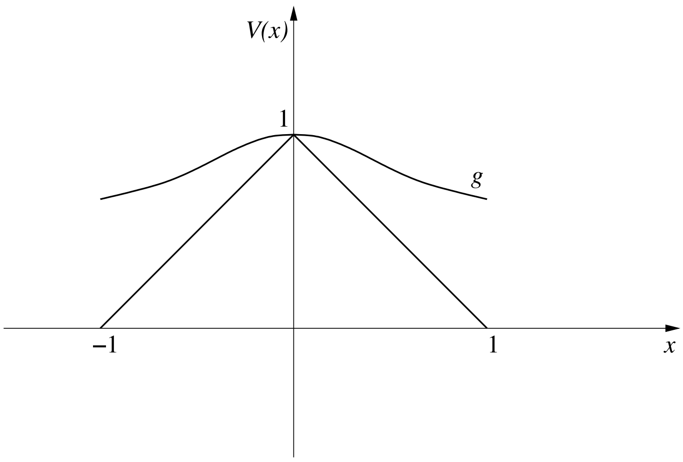


Figure 2: the subsolution viscosity property of V.

Roughly speaking a viscosity solution of a PDE is a continuous function such that

- it solves the equation in classical sense at the points where it is smooth;
- if at a point it is not smooth,
  - all the smooth functions touching it at that point and staying above

it in a neighborhood of that point must be subsolutions of the equation (*subsolution viscosity property*);

– all the smooth functions touching it at that point and staying under it in a neighborhood of that point must be supersolutions of the equation (*supersolution viscosity property*).

The value function of this simple toy problem solves the HJB equation (26) in this sense. Moreover it is the *unique* solution of the HJB equation in this sense, while the usual concept of generalized solution (i.e., a Lipschitz continuous function satisfying the equation almost everywhere) fails to provide uniqueness. Indeed, all the functions represented in Figure 1.3 are generalized solutions of the HJB equation (26). Instead, it is easy to see that at the local minimum points these functions would not satisfy the supersolution viscosity property. Therefore, in this case (and in many other ones), the concept of viscosity solution provides a characterization of the value function. This fact shows that the theory of generalized solutions is not satisfactory to treat control problems for getting a characterization of the value function by the HJB equation, whereas the theory of viscosity solutions provides that.

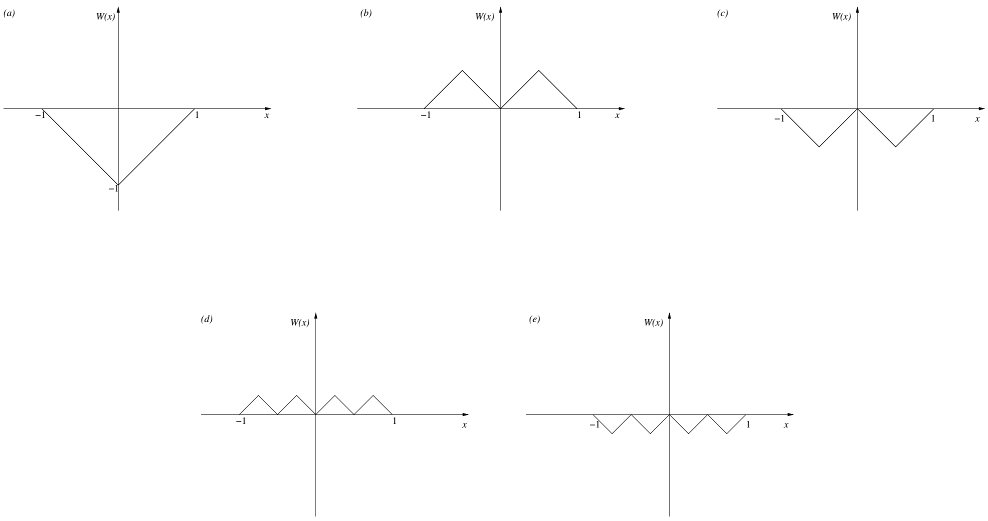


Figure 3: generalized solutions of HJB.

Finally we stress that the concept of viscosity solution is sign sensitive: if we replace the HJB equation above with the equation

$$\begin{cases} -|v'(x)| = -1, \\ v(-1) = v(1) = 0, \end{cases}$$

then the function V would be not a viscosity solution anymore, because the viscosity subsolution property would give $-|g'(0)| \leq -1$, i.e. $|g'(0)| \geq 1$, which is false. In its place the unique viscosity solution would become the function W of Figure 1.3-(a).

0.2 Literature on stochastic optimization for pension funds

The literature on stochastic optimization for pension funds is now quite rich. We try to give a quite compete list of the mathematical works on this subject, focusing the description on the ones which we consider closest to the issues of the present thesis. Basically we can divide them in two big classes:

- works keeping a collective perspective;
- works keeping an individual perpective.

From this point of view our model falls in the first class.

0.2.1 The collective perspective

The literature on this subject goes back to the 80s and 90s and, to a large extent, it culminated with the paper [Cairns; 2000]. As in our model, in this kind of literature the fund is open to entrance and exit of workers, so that at each time the fund collects contributions and pays benefits. In other words the inflow of contributions and the outflow of benefits happens at every time. However, we should say that, differently from our model, this literature was focused mainly on the analysis of defined benefits pension schemes. In such a framework, the control variables are the contribution rate and, sometimes, the investment on the risky market. Basically the fund choose the contribution rate (within some constraints) in order to manage its assets and liabilities during the time. We refer to the papers

- [Boulier, Michel, Wisnia; 1996],
- [Boulier, Trussant, Florens; 1995],
- [Cairns; 1996],
- [Cairns, Parker; 1997],
- [Cairns; 2000],
- [Haberman, Sung; 1994],
- [O'Brien; 1986],
- [Preisel, Jarner, Eliasen; 2008].

We give a quick description of the papers above, following the chronological order.

[O'Brien; 1986] is involved with the analysis of the stability of a continuous-time stochastic control system desribing a defined-benefit pension fund. Although the control formulation is very naive, the worth of this paper is that it is the first one introducing a continuous-time stochastic control formulation for pension funds. The approach of the paper is the following. Assuming an exponential growth for salary and population, a deterministic equation for the evolution of the fund is wrote down. The control variable is not just the contribution rate, but a variable controlling the contribution rate. Actually the admissible controls are restricted to constant functions, so that the problem is not properly a dynamic control problem. The control variable (a constant function, i.e. a constant) is chosen in order to make stable the value of the fund around a prescribed level.

The second step consists in adding some randomness to the model and to analyze the behaviour of the corresponding stochastic system under the control found in the deterministic model. Precisely, the variable corresponding to the market spot rate of the riskless asset and the variable corresponding to the sum of the growth rates of salary and population are now assumed gaussian random variables. A Lyapunov analysis of the corresponding stochastic system is performed.

[Haberman, Sung; 1994] is involved with a discrete-time stochastic model for a defined-benefit pension fund. The aim is to find the optimal streaming of contributions in order to minimize a quadratic functional measuring over the time the distance of contributions and wealth from prescribed targets. The problem is solved finding a backward recursive solution of the HJB equation.

[Boulier, Trussant, Florens; 1995] is involved with the study of a continuous-time stochastic model for a defined-benefit pension fund. The problem consists in minimizing the flow of contributions over the time. The control variables are the flow of contribution and the investment strategy on a standard Black-Scholes market. Moreover, the fund is subject to two constraints: it must be able to pay the (fixed) benefits at every time and it must keep its wealth positive. The problem is solved guessing and finding an explicit solution to the HJB equation. The paper [Boulier, Michel, Wisnia; 1996] follows the same line: the difference is on the objective functional.

Also [Cairns; 1996] is involved with a continuous-time stochastic model for a defined-benefit and defined-contribution pension fund. The paper analyzes the long-term bahaviour of the model under different possible investment startegies on a standard Black-Scholes market. A mean-variance analysis of the fund corresponding to these startegies is performed.

[Cairns, Parker; 1997] is involved with a discrete-time stochastic model for a defined-benefit pension fund. Under demographic stationarity assumptions, and assuming that the returns at each year constitute a i.d.d. sequence of random variables, a mean-variance analysis of the fund wealth is performed.

As we said, [Cairns; 2000] can be considered as a culminating paper on this kind of literature up to 2000. A quite general model for a defined-benefit pension fund is set and studied. Here the goal is to minimize an intertemporal functional depending on the current value of the fund and on the flow of contributions. The control variables are the contribution flow the investment strategy on a standard Black-Scholes market. A general analysis of the HJB equation is done and particulirized when the investment strategy is supposed to be fixed. This is done also for different kind of investment strategies and their effects are compared each other. Then the HJB equation is solved explicitely for different loss functions and the analysis of the corresponding optimal strategies is performed. Also conditions for the stationarity of the corresponding optimal fund are discussed.

The study is completed with a numerical analysis showing some empirical results and with some comments on the constrained cases.

In [Preisel, Jarner, Eliasen; 2008] is described and studied a model for a fund dealing with pension-life insurance products. Some dates are fixed and the rule of the fund consists in paying to its members, at these dates, a bonus related to the funding ratio (i.e. the ratio between the assets and the reserve) in the last period. This makes the model very similar to our model (see our definition of surplus in Chapter 2). The management is constrained to keep the funding ratio above 1 and can invest in a standard Black-Scholes market at every time. The investment strategy is chosen within any period in order to maximize an expected utility functional measuring the funding ratio at the end of the period. So, the optimal dynamics of the funding ratio is the result of the optimization done during every single period. It evolves as a discrete-time Markov's chain. The authors investigate the existence of a stationary distribution for such a process and then complete the study with some analytical approximations.

0.2.2 The individual perspective

The literature on this subject is more recent. In opposition to the collective perspective, it takes into account the management of contributions and benefits of a single representative partecipant. In this perspective, we can say that the pension fund considers the management of contributions and benefits of each member as a separate section that cannot communicate with the other ones corresponding to other members. The main consequence of this perpective is that, in these models, contributions and benefits are not paid at the same time.

We divide the works keeping this perspective in two classes:

- works on pension funds in the accumulation phase;
- works on pension funds in the decumulation phase.

We point out that in these works the words "accumulation phase" and "decumulation phase" are referred to the point of view of the member, i.e. the acumulation phase corresponds to his working lifetime during which he pays the contributions, while the decumulation phase corresponds to his pension lifetime during which he collects the benefits. In our model, which keeps the point of view of the management of the fund, it should be better to use the expression "accumulation phase" for the model studied in Section 2.2. However, in order to avoid confusion, we choose to call that "transitory phase".

Moreover, we treat separately also couple of papers on life and pension insurance contracts that, even if can be considered within this individual persepctive, due to their nature, cannot be considered at all as papers on pension funds in the accumulation or in the decumulation phase.

Papers on pension funds in the accumulation phase

We will focus on defined-contribution pension schemes. The literature on this subject is based on models where the pension fund collects the contributions of the partecipant during his working life and pays to him some benefits at retirement. The aim consists basically in reducing, by means of an appropriate investment strategy, the risk charged to the worker in this kind of pension schemes. This is done by defining and solving a finite-horizon (the time horizon of the worker) optimization problem with respect to some relevant quantities for the pensioner, tipically depending on the benefits collected by him at retirement or on the so-called replacement ratio, i.e. the ratio between the final wealth and the last salary. Here a list of papers dealing with this subject.

- [Battocchio, Menoncin; 2004],
- [Blake, Cairns, Dowd; 2001],
- [Booth; 1995],
- [Booth, Yakoubov; 2000],
- [Boulier, Huang, Taillard; 2001],
- [Cairns, Blake, Dowd; 2000],
- [Deelstra, Grasselli, Koehl; 2003],
- [Deelstra, Grasselli, Koehl; 2004],
- [Devolder, Bosch Princep, Dominguez Fabian; 2003],
- [Gao; 2008],

- [Haberman, Vigna; 2001],
- [Haberman, Vigna; 2002],
- [Khorasanee; 1995],
- [Khorasanee; 1998],
- [Knox; 1993],
- [Ludvik; 1994],
- [Xiao, Zhai, Qin; 2007].

We are going to describe

- [Battocchio, Menoncin; 2004],
- [Boulier, Huang, Taillard; 2001],
- [Cairns, Blake, Dowd; 2000],
- [Deelstra, Grasselli, Koehl; 2003],
- [Deelstra, Grasselli, Koehl; 2004],
- [Haberman, Vigna; 2001],
- [Haberman, Vigna; 2002],

dividing them in two subclasses: papers providing or not a minimum guarantee in the benefits.

- Papers without minimum guarantee

  Some papers on defined-contribution pension funds not providing a minimum guarantee at retirement are

  - [Battocchio, Menoncin; 2004],
  - [Cairns, Blake, Dowd; 2000],
  - [Haberman, Vigna; 2001],
  - [Haberman, Vigna; 2002].

  - *Discrete-time models*

    The papers [Haberman, Vigna; 2001] and [Haberman, Vigna; 2002] work in a discrete-time setting.

    In [Haberman, Vigna; 2001] the financial market is composed by two independent risky assets: a high-risky asset and a low-risky asset. Short selling of these assets are not allowed. The salary is supposed constant on time, i.e. salary risk is not considered. A final wealth target as well as intertemporal wealth targets depending on the financial parameters of the market are fixed and the optimization problem consists in approaching these targets (in the sense of a quadratic loss functional, i.e. values of the wealth far from the targets are penalized by a quadratic loss function). This kind of cost penalizes in the same way differences form the targets regardless of their sign. The problem is solved, by dynamic programming principle and by backward induction, guessing a quadratic structure for the solution.

The paper [Haberman, Vigna; 2002] represents an extension of [Haberman, Vigna; 2001]. It considers n assets instead of two assets and moreover it considers that these assets can be correlated with each other. Moreover in this case the cost functional is structured to penalize the differences from the targets in different way with respect to their sign, which is a more suitable assumption. The problem is solved with the same techniques of [Haberman, Vigna; 2001].

– *Continuous-time models*
[Battocchio, Menoncin; 2004], [Cairns, Blake, Dowd; 2000] work in a continuous-time setting.

In [Cairns, Blake, Dowd; 2000] the Vasicek model is considered for the interest rate process and the market is composed by the riskless asset, by a finite number of stock risky assets and by a continuous stream of zero-coupon bonds. The salary is risky, driven by the same sources of risk of the market and by an extra source of risk. The optimization is done maximizing the expected utilty from the replacement ratio. A qualitative study of the solution is done and, when the extra source of risk vanishes in the dynamics of the salary and the utility is a power function, the problem is solved finding explicit solutions.

In [Battocchio, Menoncin; 2004] the setting is very similar to the one of [Cairns, Blake, Dowd; 2000]: the Vasicek model is considered for the interest rate process and the market is composed by the riskless asset, by a stock risky asset and by a continuous stream of zero-coupon bonds. Here the salary risk is considered too, assuming that it depends on the same sources of risk of the market and on another source of risk. The difference here is that the authors consider also the inflation risk and suppose that the inflaction index is driven by the same sources of risk of the salary and that this index is tradeable: this makes this extra source of randomness hedgeable, so that the market results still complete. Finally the optimization is done maximizing the expected exponential utility from the terminal wealth and the problem is solved in closed form finding explicit solutions.

- Papers with minimum guarantee

Plans providing a minimum guarantee at retirement are introduced in the following papers:

- [Boulier, Huang, Taillard; 2001],
- [Deelstra, Grasselli, Koehl; 2003],
- [Deelstra, Grasselli, Koehl; 2004],
- [Sbaraglia, Papi, Briani, Bernaschi, Gozzi; 2003].

- *Discrete-time models*

  The paper [Sbaraglia, Papi, Briani, Bernaschi, Gozzi; 2003] provides a quite complex discrete-time model for an insurance contract. Some ideas and features of our model are taken from the model described therein. We stress that this model was set in collaboration with an italian insurance company (INA), so that it met the special requirements of this company. In this model the interest rate follows a stochastic dynamics and the financial market is composed by the riskless asset and by n risky assets; moreover transaction costs are considered. The policy-holder makes a single-sum deposit at the initial time. The management of the fund withdraws yearly from the fund a fraction of the positive part of the difference between the current value of the fund and the value of the fund at the previous year. The fund has to satisfy certain investment rules: if the fund's wealth is under a solvency level, the difference between this solvency level and the fund's wealth has to be invested in the riskless asset; this avoids improper behaviour of the manager. The fund pays to the policy-holder the maximum between its terminal value and a deterministic minimum guarantee promised at the initial time. The optimization problem consists in maximizing a performance index meseauring the net profit of the company, the average yield, the annual yield and the position with respect to the minimum guarantee pay-out. Such a criterion takes into account both the point of view of the management and of the policy-holder. The problem is approached by numerical simulations.
- *Continuous time models*

  The papers

  * [Boulier, Huang, Taillard; 2001],
  * [Deelstra, Grasselli, Koehl; 2003],
  * [Deelstra, Grasselli, Koehl; 2004]

  work with the minimum guarantee in a continuoustime setting.

  The paper [Boulier, Huang, Taillard; 2001] considers the Vasicek model for the interest rate and a market composed by the riskless asset, a risky asset and a continuous stream of zero-coupon bonds. The contributions flow is a deterministic process, so that salary risk is not considered. The fund guarantees a minimal annuity to the retired worker, i.e. pays to him a certain minimal flow of benefits from his retirement to his date of death. The date of death as well as the minimal flow of benefits are assumed deterministic. Using the bond market the minimal annuity simply becomes a minimum guarantee at the retirement date represented by a stochastic variable, which is a function of the prices of the bonds at the retirement date. Therefore the investment has to be done in order to ensure that at the retirement date the fund's wealth stays above this stochastic minimum guarantee. The optimization problem consists in maximizing the expected power utility from the terminal wealth. The problem is solved by martingale method and by means of backward stochastc differential equations.

The paper [Deelstra, Grasselli, Koehl; 2003] considers a stochastic dynamics for the interest rate covering as special case the Vasicek model and the Cox-Ingersoll-Ross model. The market is composed by the riskless asset, a risky asset and a zero-coupon bond with maturity T, where T is the terminal horizon for the investment. The fund starts with an initial endowement and collects during the time interval $[0, T]$ a contribution flow that it is a stochastic process. At the end of the period the fund has to pay a stochastic minimum guaratee to the worker in retirement plus a fraction of the surplus, i.e. the difference between the final wealth and the minimum guarantee. The remaining fraction of the surplus is taken by the manager, who optimizes the expected power utility of this fraction of surplus. Thus in this case the optimization takes the point of view of the manager (the utility function is chosen by the manager); nevertheless it is clear that such an optimization meets also the point of view of the worker: the manager will be induced to reach a high value of surplus, making also the interest of the worker. The problem is explicitely solved by martingale methods.

The paper [Deelstra, Grasselli, Koehl; 2004] considers a more general market model composed by a riskless asset and n risky assets; the interest rate, the drift vector of the risky assets and the volatility matrix of the risky assets are generic stochastic processes making the markete complete and arbitrage free. The worker pays a lump sum at the initial date and a stochastic contributions flow during his working life. The fund again ensures to the worker in retirement a stochastic minimum guarantee and pays to him also a fraction of the surplus, while the manager keeps the other part. The worker has to choose the minimum guaranteee contract among a set of possible contracts in order to maximize his expected power utility from the terminal benefits. The manager will manage the portfolio in order to maximize his own expected power utility from his own part of surplus. The problem is solved by martingale methods.

Papers on pension funds in the decumulation phase

In the previous subsection we have provided a list of the papers dealing with the management of pension funds in the so-called accumulation phase. However, another issue in the management of pension funds (arising as well in reality) is the analysis of the so-called decumulation phase. Indeed many pension schemes allow the member who retires not to convert the accumulated capital into an annuity immediately at retirement, but to defer the purchase of the annuity until a certain point of time after retirement. During this period, the member can withdraw periodically a certain amount of money from the fund within prescribed limits and the fund continues to invest in the risky market the pensioner's capital. The period of time can also be limited depending on the specific country's rules: usually freedom is given for a fixed number of years after retirement and at a certain age the annuity must be bought. Papers dealing with this subject are

- [Albrecht, Maurer; 2002],
- [Blake, Cairns, Dowd; 2003],
- [Charupat, Milevsky; 2002],
- [Gerrard, Hojgaard, Vigna; 2004],
- [Gerrard, Haberman, Vigna; 2004],
- [Gerrard, Haberman, Vigna; 2006],
- [Gerrard, Hojgaard, Vigna; 2008],
- [Kapur, Orszag; 1999],
- [Kohorasanee; 1996],
- [Milevsky, 2001],
- [Milevsky, Moore, Young; 2006],
- [Milevsky, Young; 2007],
- [Stabile; 2006],
- [Yaari; 1965].

We are going to describe

- [Gerrard, Hojgaard, Vigna; 2004],
- [Gerrard, Haberman, Vigna; 2004],
- [Gerrard, Haberman, Vigna; 2006],

In [Gerrard, Haberman, Vigna; 2004] the model considered for the market is the standard Black-Scholes model. The pensioner withdraws from the fund a constant amount of money for a fixed number of years after his retirement. The fund has to manage the investment between the riskless and the risky asset in order to reach a wealth target. Indeed, as in [Haberman, Vigna; 2001], a target function is fixed and the optimization problem consists in approaching this target (again in the sense of a quadratic loss functional, i.e. values of the wealth far from the targets are penalized by a quadratic loss function). The problem is solved finding explicit solutions.

The paper [Gerrard, Hojgaard, Vigna; 2004] represents in some sense an extension of [Gerrard, Haberman, Vigna; 2004]. In the first part the difference with respect to [Gerrard, Haberman, Vigna; 2004] is represented by the fact that here the pensioner is not constrained to a constant consumption, but he is allowed to choose a consumption strategy. In this case an intertemporal consumption target and a terminal wealth target are fixed and the optimization again consists in approaching these targets in the sense of a quadratic loss functional. The problem is solved finding explicit solutions. The second part of the paper extends further the model introducing a random time of death for the pensioner: in the case of death before the annuitization time, the optimization program must stop at such date. Also in this case the problem is solved finding explicit solutions.

[Gerrard, Haberman, Vigna; 2006] extends [Gerrard, Hojgaard, Vigna; 2004] from the point of view of the optimization adding an intertemporal target for the wealth and a bequest function in the case that death occurs before the annuitization time. The problem is solved finding explicit solutions.

Life and pension insurance contracts

The papers [Steffensen; 2004] and [Steffensen; 2006] deal with life-pension insurance contracts.

In [Steffensen; 2004], the income and outcome of external cashflows are modeled as a unique stream of (positive or negative) payments. The policy state of the life-pension insurance contract is modeled by a Markov chain. The control variables are the investment strategy in a Black-Scholes market and other variables affecting the stream of payments (roughly speaking, the company has to choose a portfolio-dividends strategy). The aim is to maximize a functional representing, to some extent, the stream of payments, keeping the point of view of the policy holder. The setting is quite general, in particular with regard to the choice of the objective functional. In this way the problem is suitable to fall into different classes of problems, on the basis of the choice of the weighting functions appearing in the objective functional. In particular different choices of these functions may lead either to a defined-contribution or to a defined-benefit framework. The problem is approached by the dynamic programming technique: the HJB equation is written (for special choices of weighting functions) and solved finding explicit solutions.

In [Steffensen; 2006] the “contribution” and the “benefit” stream are separated and the risky market is not present. The control variable is represented by the dividends. A utility process, whose dynamics is affected by the dynamics of dividends, is defined and the insurance company has to choose the strategy in order to maximize the expected total utility coming from this process. Also in this case, the problem is solved explicitly. Moreover, examples with constraints are discussed.

0.3 Plan of the thesis

The core of the thesis is represented by Chapters 1 & 2, where a model of pension fund, set up on the considerations done at the beginning of this introduction, is investigated. Chapter 3 treats a deterministic control problem; although such problem is not related to the financial topic of the thesis, we insert it because from a mathematical point of view it is related to the problem described in Chapter 2. Finally in Chapter 4 it is presented the study of a pension fund model in the decumulation phase, which is already studied in the literature; the novelty here consists in the fact that some financial constraints are added to the model.

***

In Chapter 1 we present a model of defined contribution pension fund providing a minimum guarantee. The main references for this chapter are the papers [Federico; 2008], published by the *Banach Center Publications*, and [Di Giacinto, Federico, Gozzi; 2009], accepted for publication by the journal *Finance and Stochastics*.

In the spirit of the description done at the beginning of this introduction, our aim is to propose and study a continuous-time stochastic model for a pension fund keeping a collective perspective. So, we imagine a defined-contributions pension fund with minimum guarantee, which is *continuously open to the entrance of new workers and to the exit of workers who have accrued the right to retirement*. In this persepective it keeps a collective perpective, because it does not consider the management of contributions and benefits of single workers, instead it works with the management of the cumulative flows of contributions and benefits of the community of people adhering to the fund.

We have to say that actually we do a strong assumption from the demographic point of view, i.e. we assume that the flow of people entering into the fund is constant on time and that each worker remains within the fund for a fixed time T. However, we stress that, even with our assumption of demographic stationarity, our setting cannot be considered simply as the sum of a series of individual problems. Indeed, due to the lag between the contribution time, i.e. the time during which the worker pays his contributions, and the retirement time, i.e. the date in which the worker collects his benefits, our fund can use part of the contributions paid by the workers adhering to the fund to pay the benefits to the workers who are retiring. From this point of view our model can be considered as a pay-as-you-go scheme. On the other hand our model provides for the workers retiring a minimum guarantee as benefits, which is given by the capitalization at a minimum guaranteed rate of the contributions paid by them. In this sense our model can be considered also as a funded pension scheme (see [Davis; 1995], Chapter 2, Section 4, for a general description of these different features for pension schemes). The model provides also a capital requirement, i.e. imposes to the manager to keep the wealth above a certain solvency level; this requirement is usual in pratice: it prevents an improper behaviour in the management in order to decrease the probability od defaults.

The optimization is viewed from the point of view of the manager, who is supposed to take benefits from a high current value of the fund (indeed usually the fee of the manager is related by dividends to the the absolute level of the fund's wealth). He is allowed to invest in a risky asset and in a riskless one, but borrowing and short selling are not allowed.

We separate the problem in two different phases: a first phase, over a finite horizon, in which the fund collects the contributions of people adhering to it and does not pay benefits because there are no retirements (transitory phase; based on [Federico; 2008])); a second phase, over an infinite horizon, in which the inflow of contributions and the outflow of benefits (stationary phase; based on [Di Giacinto, Federico, Gozzi; 2009]) are present.

From a mathematical point of view we are involved in two stochastic control problems: in the transitory phase the optimization is done over a finite horizon, in the stationary phase the optimization is done over in infinite horizon. However, they are both stochastic control problems with state and control constraints. We treat them by the Dynamic Programming approach, studying the associated HJB equations. In the transitory phase we prove that the value function of the problem is the unique viscosity solution of the associated HJB equation. In the stationary phase we go beyond: after having proved that the value functionis a viscosity solution of the associated HJB equation, we get a regularity result for it in the spirit of the fifth step of the program described above; then we study the closed loop equation and prove a verification theorem giving the optimal strategy in feedback form. However, in both cases we provide examples with explicit solution.

***

The main references for Chapter 2 are the paper [Federico; 2008], accepted for publication in *Finance and Stochastics*, and the working paper [Federico; WP]. This chapter is concerned with the study of the same problem of Chapter 1 when in the model a surplus term is added to the benefits, i.e. when it is supposed that the fund pays to its members in retirement something more over the minimum guarantee. The introduction of such a term is relevant from a financial point of view, because it makes the adherence to the fund more appealing to the workers. We suppose that this surplus is stated by a contract subscribed in advance between the fund and the workers. Roughly speaking such a contract provides that the retiring workers profit by part of the fund's return (referred to their contribution period), if it was sufficiently high. Since this contract depends on a return, it has to compare the fund's wealth at the current time with the fund's wealth at a past date. This leads to a delay term in the state equation, which makes the problem much more difficult to treat. We approach this delay problem by means of a representation in infinite-dimension, which seems to be the only one treatable in this case. Hence, the problem becomes an infinite-dimensional stochastic control problem with some specific features which makes it new with respect to the mathematical literature on this topic. In this case we close the study only at a viscosity stage, indicating future possible developements that our work leaves open for the problem.

***

Chapter 3 is out of the financial topic of this thesis; the main references for this chapter are the paper [Federico, Goldys, Gozzi; 2009a], submitted to the journal *SIAM - Journal on Control and Optimization*, and the subsequent paper [Federico, Goldys, Gozzi; 2009b], which is going to be submitted to the same journal. The problem described and studied therein arises basically from an attempt of finding a simplified version of the mathematical problem of Chapter 2, allowing a satisfactory answer to the problem. So, the problem faced in this chapter preserves the main mathematical feature of the problem studied in Chapter 2, i.e. the delay, and simplifies other features (in particular it is a deterministic problem), in order to allow us to go ahead with the analysis. In the perspective of this thesis, it should be viewed mainly as a toy model for a possible future approach to the problem of Chapter 2. Nevertheless we stress that it turns out to be interesting in itself from a theoretical point of view and also that it has intersting applications in Economics (in particular in growth models with "time-to-build") and in some financial problems.

As in Chapter 2 the delay problem is approached passing to the infinite-dimensional representation. In this case we prove not only that the value function is a viscosity solution to the associated HJB equation, but also that it has a regularity property that allows to define the feedback map. Then we study the closed loop equation and prove a nonstandard verification theorem showing that the feedback map defines optimal feedback strategies for the problem. In particular the part regarding the verification theorem contains the correction of a wrong result sometimes used in the literature in this kind of mathematical subject (see Subsection 3.3.3).

***

Chapter 4 deals with a different problem in the field of pension funds; the main reference is the working paper [Di Giacinto, Federico, Gozzi, Vigna; WP]. Here it is investigated is a problem already studied in literature, i.e. in the paper [Gerrard, Haberman, Vigna; 2004], described in the previous section. The model takes the point of view of a pensioner who delegates a manager to invest in the financial market, until a certain point of time after his retirement, the capital accumulated by contributions during his working life. The pensioner withdraws from this fund a fixed consumption rate and the optimization problem consists in reaching a fixed target at the end of this period. Here the novelty with respect to [Gerrard, Haberman, Vigna; 2004] is represented by the fact that we add constraints on the wealth and on the investment strategies, which is a crucial feature to make the model more realistic..

From a mathematical point of view the problem is very similar to the one arising in the transitory phase of the model of Chapter 1, since it is a stochastic control problem with finite horizon and with state and control constraints. In this case we work with explicit solution. In the first part we impose a constraint only on the strategies and the argument are quite standard. In the second part we impose the constraint also on the state. In this last case the explicit solution to the HJB equation we find is nontrivial and has a quite surprising and interesting similarity with the price of a European put option. In both cases we prove verification theorems for checking the optimality of the feedback strategies.

Remark 0.3.1. We want to clarify that this thesis has to be viewed in the spirit of Applied Mathematics, i.e. as a mathematical work on a mathematical subject inspired by motivations coming from the real world. Although the demographic assumptions (demographic stationarity) and the financial assumptions (constant interest rate) in the model described and studied in Chapters 1 & 2 make the problem far from reality, due to other features it is already very hard to approach it analitically. So, we are aware that generalizations are needed for the model described and studied in these chapters; nevertheless, also at this stage, it presents some special features making it very appealing and interesting from the mathematical point of view. Indeed the presence of the solvency level in the model leads to a stochastic control problem with state constraints, while the introduction in Chapter 2 of a surplus over the minimum guarantee leads to a stochastic delay problem approached by an infinite-dimensional representation. Both these mathematical features make nontrivial the study of the problem along the lines of the dynamic programming steps described above, opening interesting and difficult issues in the field of Optimal Control Theory. □

Chapter 1

A pension fund model with constraints

In this chapter we study a stochastic control problem for the optimal management of a defined contribution pension fund model with minimum guarantee and solvency constraint. The main references for this chapter are the papers [Federico; 2008] for Section 1.2 and [Di Giacinto, Federico, Gozzi; 2009] for Section 1.3.

We adopt the point of view of a fund manager who can invest in two assets (a risky one and a riskless one, in a standard Black and Scholes market) and maximizes an intertemporal utility function depending on the current level of fund wealth.

Our emphasis is posed on the constraints faced by the fund manager: the requirement of having a solvency level on the fund wealth, and the borrowing and short selling constraints on the allocation strategies.

The problem is similar to optimal portfolio selection problems but it has some special features due to the nature and the social target of the pension funds: the presence of contributions and benefits, the presence of constraints on the investment strategies, the presence of solvency constraints. This means that we require that the wealth of the running pension fund remains above a prescribed level, i.e. the so-called *solvency level*.

We focus the analysis on the role of the solvency constraint. We analyze the effect of such constraint on the admissible and on the optimal strategies: in particular we show that, for sufficiently high solvency level, the optimal portfolio strategies do not become trivial (i.e. the fund manager can still reinvest in the risky asset), even after that the solvency level has been reached.

We clarify that a model taking into account all the relevant features related to the optimal management of a real pension fund would be very difficult to treat and an analytical treatment would be substantially impossible at the present stage. So, to focus on the role of the solvency constraint keeping the problem treatable, we introduce some simplifying hypotheses on other features. We consider all demographic variables and the interest rate as constants and moreover, in the present chapter, we assume that no surplus is paid by the fund (see Section 1.1.3 for further details). The introduction of the surplus term is the object of Chapter 2.

From the mathematical point of view our problem is a stochastic optimal control problem with constraints on the control and on the state (deriving for the presence of investment and solvency constraints). Differently from some papers on optimal portfolio problems (see, e.g.,

- [Cadenillas, Sethi; 1997],
- [Choulli, Taksar, Zhou; 2003],
- [El Karoui, Jeanblanc, Lacoste; 2005],
- [Karatzas, Lehoczky, Sethi, Shreve; 1986],
- [Sethi, Taksar; 1992],
- [Sethi, Taksar, Presman; 1992],
- [Zariphopoulou; 1994]),

within our model the state boundary is not always an absorbing barrier: the optimal strategies can touch the boundary and come back in the interior keeping the same state dynamics. In [Duffie, Fleming, Soner, Zariphopoulou; 1997] and in [Sethi, Taksar; 1992] the state process can come back in the interior after touching the boundary too. In the first paper this happens thanks to the presence of a stochastic income in the special case of HARA utility functions (see also [Tebaldi, Schwartz; 2006] for a similar setting) while in the second one this is obtained taking different state dynamics when the boundary is reached, so using a completely different setting. This important modelling issue involves some nontrivial technical problems in the study of optimal strategies (see Subsectio 1.3.6).

We split the study of the problem in two different phases: a transitory phase and the stationary phase. In the trensitory phase no benefits are paid and the contributions collected give rise to a time-dependent entering cash-flow. In this case the dynamics of the wealth depends explicitally on the time variable. Moreover also the solvency level is a function of the time. Therefore the problem is strongly time-dependent and not easy to treat: we we will restrict our analysis just to some mathematical features of the problem and then we give an example with explicit solution. Instead in the stationary phase the external payment flows (contributions and benefits) are constant, so that the dynamics of the wealth is stationary with respect to the time variable. Moreover also the solvency level is assumed to be constant. This is an easier problem from the mathematical point of view, since we can get rid of the time variable. Therefore, in this case we are also able to find explicitally the optimal strategy and to give some qualitative comments about its behaviour with respect to the choice of the parameters of the model.

1.1 The model

In this section we give a detailed description of the model. Over an infinite continuous-time model we consider a financial market which is:

- competitive, i.e. we assume that the investor's behavior is optimizing: he optimizes his utility function on the whole time horizon;
- frictionless, i.e. all assets are perfectly divisible and there are no transaction costs or taxes;
- arbitrage free, i.e. there is no opportunity to gain without assuming risk with not null probability;
- default free, i.e. financial institutions issuing assets cannot default;
- continuously open, i.e. the investor can continuously trade in the market.

Moreover, we assume that:

- the investor is price taker, i.e. he cannot affect the probability distribution of the available assets: this hypothesis is usual in literature regarding financial management models of pension funds and it is realistic if the single agent does not invest a big amount of money; as a matter of fact, the volume of assets exchanged by pension funds is such that they could affect the price of assets (i.e. the investor may be price maker) but we do not deal with this fact here.
- the investor faces the following trading constraints: borrowing and short positions are not allowed;
- the investor maximizes the expected utility from the current fund wealth over an infinite horizon.

We impose that the pension fund wealth must be above a suitable positive function which we call *solvency level*.

It is supposed a demographic stationarity hypothesis, i.e. that the flow of people who enter into the fund starts at time $t = 0$ and is constant over time and that there is an exogenous constant $T > 0$ which is the time during which the members adhere to the pension fund. Therefore the exit flow of people is null in the interval $[0, T]$ and is constant after time T, balancing exactly the entrance flow.

1.1.1 The wealth dynamics

To set up the mathematical model we consider a complete probability space $(\Omega, \mathcal{F}, \mathbb{P})$ with a filtration $\{\mathcal{F}_t\}_{t\geq 0}$, where $t \geq 0$ is the time variable. The filtration $\{\mathcal{F}_t\}_{t\geq 0}$, describing the information structure, is generated by the trajectories of a one-dimensional standard Brownian motion $B(t)$, $t \geq 0$, defined on the same probability space and completed with the addition of the null measure sets of $\mathcal{F}$. Moreover we assume that $\mathcal{F}_\infty := \sigma\left(\bigcup_{t\geq 0} \mathcal{F}_t\right) = \mathcal{F}$. Sometimes we will use a starting point $s \geq 0$. In this case $\{\mathcal{F}_t^s\}_{t\geq s}$ will be the completion of the filtration generated by $B^s(t) = B(t) - B(s)$.

The financial market is composed of two kinds of assets: a riskless asset and a risky asset. The price of the riskless asset $S^0(t)$, $t \geq 0$, evolves according to the equation

$$\begin{cases} dS^0(t) = rS^0(t)dt, \\ S^0(0) = 1, \end{cases}$$

where $r \geq 0$ is the instantaneous spot rate of return. The price of the risky asset $S^1(t)$, $t \geq 0$, follows an Itô process and satisfies the stochastic differential equation

$$\begin{cases} dS^1(t) = \mu S^1(t)dt + \sigma S^1(t)dB(t), \\ S^1(0) = s_0^1, \end{cases}$$

where μ is the instantaneous rate of expected return and $\sigma > 0$ is the instantaneous rate of volatility. We assume that the market assigns a premium for the risky investmet, i.e. $\mu > r$. The drift μ can be expressed by the relation $\mu = r + \sigma\lambda$, where $\lambda > 0$ is the instantaneous risk premium of the market, i.e. the price that the market assigns to the randomness expressed by the standard Brownian motion $B(\cdot)$. The case $\lambda = 0$, i.e. $\mu = r$, is trivial in a (natural) context of risk aversion (as ours), since in this case the optimal investment is simply composed by the only riskless asset.

In our framework the interest rate is assumed to be constant. This assumption represents a restriction with respect to other works on the same subject, as

- [Battocchio, Menoncin; 2004],

- [Boulier, Huang, Taillard; 2001],
- [Cairns, Blake, Dowd; 2000],
- [Deelstra, Grasselli, Koehl; 2003],
- [Haberman, Vigna; 2001],
- [Haberman, Vigna; 2002],

where, on the other hand, the solvency constraint is not considered.

The state variable, represented by $X(t)$, $t \geq 0$, is the $\{\mathcal{F}_t\}_{t\geq 0}$-adapted process which gives the amount of the pension fund wealth at any time. We suppose that the pension fund starts its activity at the date $t = 0$ and that at this time it owns a starting amount of wealth $x \geq 0$.

The control variable, denoted by $\theta(t)$, $t \geq 0$, is the $\{\mathcal{F}_t\}_{t\geq 0}$-progressively measurable process which represents the proportion of fund wealth to invest in the risky asset. Therefore, the positivity of the wealth (due to the solvency constraints) and the borrowing and short selling constraints impose $\theta(t) \in [0,1]$ for every t. So the dynamics of wealth is expressed, formally, by the following state equation

$$\begin{cases} dX(t) = \dfrac{\theta(t)x(t)}{S^1(t)} dS^1(t) + \dfrac{[1-\theta(t)]\,X(t)}{S^0(t)}\, dS^0(t) + c(t)dt - b(t)dt, \quad t \geq 0, \\ X(0) = x_0 \geq 0, \end{cases} \tag{1.1}$$

where $\dfrac{\theta(t)X(t)}{S^1(t)}$ and $\dfrac{[1-\theta(t)]\,X(t)}{S^0(t)}$ are respectively the quantities in the portfolio of the risky asset and the riskless asset; the non-negative process $c(t)$, $t \geq 0$, indicates the flow of contributions and the non-negative process $b(t)$, $t \geq 0$, represents the flow of benefits.

The state equation (1.1) can be rewritten in the following way

$$\begin{cases} dX(t) = \big[\,(r + \sigma\lambda\theta(t))\,X(t) + c(t) - b(t)\big]dt + \sigma\theta(t)X(t)dB(t), \quad t \geq 0, \\ X(0) = x. \end{cases} \tag{1.2}$$

As we said before, we assume that a *solvency constraint* must be respected (see Subsection 1.1.4 for further explanations). More precisely the process $x(\cdot)$ describing the fund wealth is subject to the following constraint

$$X(t) \geq l(t), \quad \mathbb{P}\text{-a.s.}, \qquad \forall t \geq 0, \tag{1.3}$$

where the non-negative deterministic function $l(t)$, $t \geq 0$ represents the solvency level.

1.1.2 Contributions

In the population stationarity hypothesis, the flow of contributions $c(\cdot)$ can be considered exogenous. We assume that the workers who enter into the pension fund are a homogeneous class, i.e. a class of people having the same characteristics (same age at the entry date, same professional qualification, same level of skill, and so on). Moreover, as said, we suppose that their entrance flow is constant on time and that each participant adheres for a length of time represented by an exogenous constant $T > 0$. Due to these demographic assumptions, we assume that the aggragate contribution flow increases linearly on time in the interval $[0, T]$ and is equal to a constant $c > 0$ after time T. For instance we can imagine that each member who is adhering to the fund pays to the fund a contribution rate equal to αw, where $\alpha \in (0, 1)$ and where $w > 0$ is the (constant) wage rate (which has the dimension euros/time) earned by each member. Then, denoting by $\bar{c}$ the entrance flow of people into the fund, we can write the flow of aggregate contributions as

$$c(t) := \begin{cases} \alpha w \cdot \bar{c}t, & 0 \leq t \leq T, \\ \alpha w \cdot \bar{c}T, & t > T; \end{cases} \tag{1.4}$$

therefore, in this case, the aggregate contributions flow after time T is the constant $c = \alpha w \cdot \bar{c}T$.

The above hypothesis is a bit restrictive because the stochastic wage is an important and additional source of uncertainty for the fund manager. We observe that the introduction of an extra source of risk renders the market incomplete, as discussed and studied in [Cairns, Blake, Dowd; 2000] in absence of guarantee and in a continuous and finite time horizon. Nevertheless we stress again that we make this assumption in order to sempliy and focus on the effect of the solvency constraint.

1.1.3 Benefits

Due again to the demographic stationarity we assume that the flow of benefits starts at time T, when the first retirements occur, and that, after that date, it is given by a constant g representing the minimum guarantee flow. We assume that $g \geq c$, because g has to be, in some way, the capitalization of the contributions paid by the members who are retiring. For instance, we can imagine that the fund pays to the generic member in retirement as (lump sum) minimum guarantee the capitalization at a minimum guaranteed rate $\delta \in [0, r]$ of the contributions paid by him in the time interval during which he was adhering to the fund. In this case, coherently with (1.4), we can write the aggregate minimum guarantee flow, for $t \geq T$, as:

$$g = \bar{c} \int_{t-T}^{t} (\alpha w)\, e^{\delta(t-u)} du, \tag{1.5}$$

i.e.

$$g = \begin{cases} c, & \text{if } \delta = 0, \\ \bar{c} \cdot (\alpha w)\, \frac{e^{\delta T}-1}{\delta} & \text{if } \delta > 0; \end{cases} \tag{1.6}$$

in particular we have $g \geq c$. The previous inequality means in particular that, despite of the case $\delta = 0$, the current contributions do not permit to pay the current minimum guarantee. Nevertheless we will show that, in our setup for the benefits, under suitable assumptions on the solvency level, the fund manager can always pay the current benefits, mantaining the wealth level of the fund above the solvency level.

The inequality $\delta \leq r$ could be justified thinking to the fact that often the participants to the pension fund do not have time to enter to the financial market as the fund manager. Moreover we recall that in the actual market, but it is not the case of our framework which has neither transactional or informational costs, the fund manager can usually get higher interest rate than the fund members.

1.1.4 Solvency level

A solvency level may be imposed by law or by a supervisory authority to avoid improper behavior of the manager and to guarantee that the fund is able to pay at least part of the due benefits at each time $t \geq 0$. Without imposing this constraint the manager is allowed to use strategies that may bring him to mismatches with the social target of the pension fund. We assume that the solvency level $l(\cdot)$ imposed in (1.3) is a nondecreasing continuous function, which is constant after time T. More precisely we assume that the solvency level has the following structure:

- at the beginning the company should hold a given minimum startup level $l_0 := l(0) \geq 0$;
- for $t \in [0, T]$, the solvency level is the capitalization at a rate $\beta \leq r$ of the initial minimum wealth l_0 and of the aggregate contributions paid up to time t; therefore

$$l(t) = l_0\, e^{\beta t} + \int_0^t \alpha w \cdot \bar{c} s\, e^{\beta(t-s)} ds. \tag{1.7}$$

- after time T the solvency level is constant, i.e. $l(t) \equiv l := l(T)$ for $t \geq T$;

The rate β could be chosen, for example, by an authority with regard to the market's parameters.

1.1.5 Optimization in pension funds

To a large extent, the primary focus of a pension fund investment is the guarantee for the subscribers to obtain the promised benefits and the effective management of pension funds is severely restricted by regulatory authorities in order to enforce such a guarantee. For this reason despite the formal similarities, it is important to remark that the optimal allocation problem faced by a pension fund is radically different in its objectives from the problem faced by an investor having direct access to the market. While the investor is willing to optimize his welfare taking direct advantage from stock market opportunities, a pension fund subscription is usually a process of investment delegation forced by the social security laws.

It is well known that the process of investment delegation involves costs for the members and a potential divergence between the interests of the principal (the collectivity of subscribers) and the agent (the manager of the fund). Within our model forced delegation is costly. Indeed the members accept a guaranteed rate of return δ lower than the risk free rate r.

In order to incentive the manager to undertake risky investments and reduce this fixed cost, it is a common practice to introduce a variable component in the management fee proportional to the absolute level of fund's wealth (see, for example, [Starks; 1987] and [Goetzmann, Ingersoll, Ross; 2003]).

Hence we can say that basically the optimization criterion for the management of a pension fund can take into account two different points of view:

- The point of view of the members: the fund's manager is directly delegated by the members to invest in the risky market in order to perform their benefits.

- The point of view of the manager: the manager is led to invest in the risky asset in order to incentive his fee, which is proportional to the absolute level of fund's wealth. Observe that within this framework the participant to the pension fund has no direct benefit from risky investment, but only an indirect benefit. In fact, assuming the existence of a competitive market of pension funds' management (e.g., by insurance companies), if the manager is allowed to invest in the risky market the fixed delegation cost, i.e. the difference $r - \delta$, is expected to be reduced.

In our optimization problems we will use the second point of view. Nevertheless in Chapter 2 we will see how the introduction of the surplus term in the model will have some consequences also on the point of view of the members.

1.2 The transitory phase

The object of this section is the analysis of the transitory phase, corresponding to the time interval $[0, T]$. The main results are the proof of the continuity of the value function (that is not trivial here), the proof that the value function is a constrained viscosity solution of the HJB equation and its characterization by a uniqueness result in a special case. Moreover we close the section providing an example with explicit solution.

First of all we observe that the initial time $t = 0$ has been chosen as the first time of operation of the fund. However it also makes sense, in order to apply the dynamic programming techniques, to look to a pension fund that is already running after a given amount of time $s \in [0, T]$, in order to establish a decision policy from s on.

On the probability space of the Section 1.1 let $(\mathcal{F}_t^s)_{t\in[s,T]}$ be the completion of the filtration generated by the process $(B^s(t))_{t\in[s,T]} := (B(t) - B(s))_{t\in[s,T]}$; the control process $(\theta(t))_{t\in[s,T]}$ is a $(\mathcal{F}_t^s)$-progressively measurable process with values in $[0, 1]$.

Let us set an initial time $s \in [0, T]$ and a given amount of wealth x at time s. In the interval $[0, T]$ the state equation becomes, according to (1.2) and with the hypotheses just stated on the contribution term,

$$\begin{cases} dX(t) = [r + \sigma\lambda\theta(t)]\, X(t)\, dt + kt\, dt + \sigma\theta(t)X(t)dB^s(t), \quad t \in [s, T], \\ X\,(s) = x. \end{cases} \tag{1.8}$$

Theorem 1.2.1. *For any $(\mathcal{F}_t^s)_{t\geq s}$-progressively measurable $[0, 1]$-valued process $\theta(\cdot)$*

- *equation (1.8) admits on the filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t^s)_{t\in[s,T]}, \mathbb{P})$, a unique strong solution;*
- *this solution belongs to the space $C_{\mathcal{P}}([s, T]; L^p(\Omega, \mathbb{P}))$ of the p-continuous progressively measurable processes for any $p \in [1, +\infty)$.*

Proof. See Theorem 6.16, Chapter 1, of [Yong, Zhou; 1999] or Section 5.6.C of [Karatzas, Shreve; 1991]. □

We denote the unique strong solution to (1.8) by $X\,(t; s, x, \theta(\cdot))$.

1.2.1 The optimization problem

In this transitory phase we study a finite horizon optimization problem in the interval $[0, T]$ related to an objective functional with this form:

$$\mathbb{E}\left[\int_0^T e^{-\rho t} U\left(t, X\left(t\right)\right) dt + f(X(T))\right]. \tag{1.9}$$

Here $\rho > 0$ is the individual discount factor of the manager and U is his utility function. So, according with the considerations of Subsection 1.1.5, the first term in the optimization criterion (1.9) takes into account the manager's point of view. Instead the main role of the exit/bequest function f in this context consists in allowing to link the problem in this transitory phase with the problem in the stationary phase-Thus we assume that in the transitory phase the fund's manager takes care about the future of the fund after time T only by means of the exit/bequest function f (see Remark 1.3.9).

The problem lives in the set

$$\mathcal{C} := \{(s, x) \in \mathbb{R}^2 \mid s \in [0, T],\ x \geq l(s)\}.$$

We assume that the current utility function U and the exit/bequest utility function f satisfy the following assumptions:

Hypothesis 1.2.2. The current utility function U is such that

(i) $U : \mathcal{C} \to \mathbb{R}$ has the structure $U(s, x) = u(x - l(s))$, where $u : [0, +\infty) \to \mathbb{R}$;

(ii) $u \in C([0, +\infty); \mathbb{R})$;

(iii) u is increasing and concave. □

Hypothesis 1.2.3. The exit/bequest utility function f is such that

(i) $f \in C([\, l(T), +\infty); \mathbb{R})$;

(ii) f is increasing and concave. □

Remark 1.2.4. We note that

- the utility functions U and f are defined where the wealth process $X(\cdot)$ must live;
- all the functions of the form $u(x) = \frac{(x-x_0)^\gamma}{\gamma}$, for $x_0 \leq 0, \gamma \in (0, 1)$, always give rise to functions U satisfying Hypothesis 1.2.2.
- all the functions of the form $f(x) = \frac{(x-x_0)^\gamma}{\gamma}$, for $x_0 \leq 0, \gamma \in (0, 1)$, satisfy Hypothesis 1.2.3.

□

For $(s,x) \in \mathcal{C}$, the problem consists in maximizing, over the set of the admissible strategies, i.e. the strategies keeping the state variable above $l(\cdot)$ (see (1.10) for a formal definition), the functional (1.9).

1.2.2 The set of admissible strategies

In this transitory phase the set of the admissible strategies, for $(s,x) \in \mathcal{C}$, is given by

$$\Theta_{ad}(s,x) := \Big\{ \theta : [s,T] \times \Omega \longrightarrow [0,1] \text{ prog. meas. with respect to } \{\mathcal{F}_t^s\}_{t\in[s,T]} \mid X\left(t; s, x, \theta(\cdot)\right) \geq l\left(t\right),\ t \in [s,T] \Big\}. \quad (1.10)$$

We show that the set $\Theta_{ad}(s,x)$, for $(s,x) \in \mathcal{C}$, is not empty:

Proposition 1.2.5. *Let $(s,x) \in \mathcal{C}$ and let $X(t) := X(t; s, x, 0)$; then*

$$X(t) - l(t) \geq (x - l(s))\, e^{r(t-s)}, \qquad t \in [s,T]; \quad (1.11)$$

In particular, for each $(s,x) \in \mathcal{C}$, the null strategy belongs to $\Theta_{ad}(s,x)$, so that $\Theta_{ad}(s,x)$ is not empty.

Proof. Let $(s,x) \in \mathcal{C}$ and let $X(t) := X(t; s, x, 0)$ be the state trajectory associated with the null startegy; the dynamics of $X(\cdot)$ is given by

$$\begin{cases} dX(t) = rX(t)\, dt + kt\, dt, \\ X(s) = x; \end{cases}$$

the "dynamics" of the solvency level $l(\cdot)$ is given by

$$\begin{cases} dl(t) = \beta l(t)\, dt + kt\, dt, \\ l(s) = l(s). \end{cases}$$

The claim follows taking in account that $\beta \leq r$. □

We define the lateral boundary as the set

$$\partial^* \mathcal{C} := \{(s,x) \in \mathcal{C} \mid s \in [0,T],\ x = l(s)\}. \quad (1.12)$$

We have the following behaviour of the lateral boundary with respect to the parameter β:

Proposition 1.2.6. *Let $(s,x) \in \partial^*\mathcal{C}$.*
(i) *If $\beta < r$, $s > T$, then $X(t;s,x,0) > l(t)$ for all $t \in (s,T]$. We stress this property saying that $\partial^*\mathcal{C}$ is reflecting.*
(ii) *If $\beta = r$, then $\Theta_{ad}(s,x) = \{0\}$ and $X(t;s,x,0) = l(t)$ for all $t \in [s,T]$. We stress this property saying that $\partial^*\mathcal{C}$ is absorbing.*

Proof. (i) This straightly follows looking at the proof of Proposition 1.2.5.

(ii) Let $\beta = r$, $s \in [0,T)$, $\theta(\cdot) \in \Theta_{ad}(s,l(s))$ and set $X(t) := X(t;s,l(s),\theta(\cdot))$; the dynamics of $X(t)$ is given by

$$\begin{cases} dX(t) = rX(t)\,dt + kt\,dt + \sigma\theta(t)X(t)d\tilde{B}^s(t), \\ X(s) = l(s), \end{cases} \tag{1.13}$$

where, thanks to Girsanov's Theorem A.1.1, the process $\tilde{B}^s(\cdot) := B^s(\cdot) + \lambda(\cdot - s)$ is a Brownian motion under the probability $\tilde{\mathbb{P}} = \exp\left(-\lambda B^s(T) - \frac{\lambda^2}{2}(T-s)\right) \cdot \mathbb{P}$ in the interval $[s,T]$. Since $X \in C([s,T];L^p(\Omega,\mathbb{P}))$ for any $p \geq 1$, by Hölder's inequality it holds also

$$\tilde{\mathbb{E}}\left[\int_s^T |X(t)|^2 dt\right] < +\infty,$$

so that

$$\tilde{\mathbb{E}}\left[\int_s^t X(r) d\tilde{B}^s(r)\right] = 0, \quad \forall t \in [s,T].$$

Thus we can pass (1.13) to the expectations getting

$$\begin{cases} d\tilde{\mathbb{E}}[X(t)] = r\tilde{\mathbb{E}}[X(t)]\,dt + kt\,dt, \\ \tilde{\mathbb{E}}[X(s)] = l(s). \end{cases} \tag{1.14}$$

By (1.14) we have $\tilde{\mathbb{E}}[X(t;s,x,\theta(\cdot))] = l(t)$ for $t \in [s,T]$; moreover by assumption $X(t) \geq l(t)$ for $t \in [s,T]$, so that we get that $X(t) = l(t)$ almost surely for any $t \in [s,T]$. This implies that $\theta(\cdot) \equiv 0$, so that we can conclude that the only admissible strategy starting from $(s,l(s))$ is the null one and that the corresponding state trajectory remains on the boundary. □

1.2.3 The value function

For $(s,x) \in \mathcal{C}$, $\theta(\cdot) \in \Theta_{ad}(s,x)$, we define

$$J(s,x;\theta(\cdot)) := \mathbb{E}\left[\int_0^T e^{-\rho t} U\left(t, X\left(t;s,x,\theta)\cdot)\right)\right) dt + f(X(T;s,x,,\theta(\cdot)))\right].$$

The stochastic control problem consists in studying, for $(s,x) \in \mathcal{C}$, the value function

$$V(s,x) := \sup_{\theta(\cdot) \in \Theta_{ad}(s,x)} J(s,x;\theta(\cdot)), \tag{1.15}$$

and, when possible, in finding an optimal control strategy for the problem in the sense of the following definition.

Definition 1.2.7. Let $(s,x) \in \mathcal{C}$.

(i) A control $\theta^*(\cdot) \in \Theta_{ad}(s,x)$ is called optimal for the couple (s,x) if

$$J(s,x;\theta^*(\cdot)) = V(s,x).$$

(ii) Let $\varepsilon > 0$; a control $\theta^\varepsilon(\cdot) \in \Theta_{ad}(s,x)$ is called ε-optimal for the couple (s,x) if

$$J(s,x;\theta^\varepsilon(\cdot)) \geq V(s,x) - \varepsilon.$$

□

Proposition 1.2.8. *Let us suppose that Hypotheses 1.2.2 and 1.2.3 hold true. Then there exists a constant $C > 0$ such that $V(s,x) \leq C(1+x)$ for all $(s,x) \in \mathcal{C}$.*

Proof. Let $(s,x) \in \mathcal{C}$. By Hypotheses 1.2.2-(iii) and 1.2.3 there exists $C > 0$ such that $U(t,y) \leq C(1+y)$ for any $t \in [s,T]$, $y \geq l(t)$, and $f(y) \leq C(1+y)$ for any $y \geq l(T)$. Let $\theta(\cdot) \in \Theta_{ad}(s,x)$; then, setting $X(t) := X(t;s,x,\theta(\cdot))$, we have

$$\mathbb{E}\left[\int_s^T e^{-\rho t}[U(t,X(t))]dt + f(X(T))\right] \leq C\,\mathbb{E}\left[\int_s^T e^{-\rho t}(1+X(t))dt + (1+X(T))\right]. \tag{1.16}$$

Taking into account that $X \in C([s,T];L^2(\Omega))$, we have

$$\mathbb{E}\left[\int_s^t \theta(r)X(r)dB^s(r)\right] = 0, \qquad \forall t \in [s,T].$$

Therefore we can pass to the expectations in the state equation getting

$$\begin{cases} d\mathbb{E}[X(t)] = r\mathbb{E}[X(t)]dt + kt\,dt + \sigma\lambda\mathbb{E}[\theta(t)X(t)]dt \leq (r+\sigma\lambda)\mathbb{E}[X(t)]dt + kT\,dt, \\ \mathbb{E}[X(s)] = x; \end{cases}$$

thus, for some $C > 0$,

$$\mathbb{E}[X(t)] \leq \left(x + \frac{kT}{r+\sigma\lambda}\right)e^{(r+\sigma\lambda)(t-s)} - \frac{kT}{r+\sigma\lambda} \leq C(1+x). \tag{1.17}$$

The estimate (1.17) does not depend on the control. Thus the claim follows putting (1.17) into (1.16) and taking the supremum over $\theta(\cdot) \in \Theta_{ad}(s,x)$. □

Proposition 1.2.9. *Let $(s,x) \in \mathcal{C}$; then*

$$V(s,x) \geq \frac{u(x-l(s))}{\rho}\left(e^{-\rho s} - e^{-\rho T}\right) + f\big(l(T) + x - l(s)\big).$$

Proof. By (1.11) and by monotonicity of u, f, we can deduce that

$$\begin{aligned} J(s,x;0) &\geq \int_s^T e^{-\rho t} u(x-l(s))dt + f\big(l(T)+x-l(s)\big) \\ &= \frac{u(x-l(s))}{\rho}\left(e^{-\rho s}-e^{-\rho T}\right) + f\big(l(T)+x-l(s)\big); \end{aligned}$$

so the claim follows. □

Proposition 1.2.10. *Let $s \in [0,T]$; the function $x \mapsto V(s,x)$ is concave on $[l(s),+\infty)$.*

Proof. Fix $x, x' \geq l(s)$; for $\gamma \in [0,1]$, set $x_\gamma := \gamma x + (1-\gamma)x'$; of course $x_\gamma \geq l(s)$. We have to prove that

$$V(s,x_\gamma) \geq \gamma V(s,x) + (1-\gamma)V(s,x'). \tag{1.18}$$

Take $\theta(\cdot) \in \Theta_{ad}(s,x)$ and $\theta'(\cdot) \in \Theta_{ad}(s,x')$ ε-optimal for x, x' respectively and $X(\cdot), X'(\cdot)$ the corresponding trajectories; then

$$\begin{aligned} \gamma V(s,x) + (1-\gamma)V(s,x') &\leq \gamma\big[J(s,x;\theta(\cdot))+\varepsilon\big] + (1-\gamma)\big[J(s,x';\theta'(\cdot))+\varepsilon\big] \\ &= \varepsilon + \gamma J(s,x;\theta(\cdot)) + (1-\gamma)J(s,x';\theta'(\cdot)) \\ &= \varepsilon + \gamma\mathbb{E}\left[\int_s^T e^{-\rho t}U\big(t,X(t)\big)dt + f(X(T))\right] \\ &\quad +(1-\gamma)\mathbb{E}\left[\int_s^T e^{-\rho t}U\big(t,X'(t)\big)dt + f(X'(T))\right] \\ &= \mathbb{E}\left[\int_s^T e^{-\rho t}\left[\gamma U\big(t,X(t)\big) + (1-\gamma)U\big(t,X'(t)\big)\right]dt\right] \\ &\quad +\mathbb{E}\left[\gamma f(X(T)) + (1-\gamma)f(X'(T))\right] + \varepsilon. \end{aligned}$$

The concavity of u, f implies that

$$\gamma U\big(t,X(t)\big) + (1-\gamma)U\big(t,X'(t)\big) \leq U\big(t,\gamma X(t)+(1-\gamma)X'(t)\big), \qquad \forall t \in [s,T],$$

$$\gamma f\big(X(t)\big) + (1-\gamma)f\big(X'(t)\big) \leq f\big(\gamma X(t)+(1-\gamma)X'(t)\big), \qquad \forall t \in [s,T].$$

Consequently, if we set $X_\gamma(\cdot) := \gamma X(\cdot) + (1-\gamma)X'(\cdot)$, we get

$$\gamma V(s,x) + (1-\gamma)V(s,x') \leq \varepsilon + \mathbb{E}\left[\int_s^T e^{-\rho t}U\big(X_\gamma(t)\big)dt + f(X_\gamma(T))\right].$$

If there exists $\theta_\gamma(\cdot) \in \Theta(s,x_\gamma)$ such that $X_\gamma(\cdot) \leq X(\cdot;s,x_\gamma,\theta_\gamma(\cdot))$, then we would have

$$\varepsilon + \mathbb{E}\left[\int_s^T e^{-\rho t}U\big(X_\gamma(t)\big)dt + f(X_\gamma(T))\right] \leq \varepsilon + J(s,x_\gamma;\theta_\gamma(\cdot)) \leq \varepsilon + V(s,x_\gamma),$$

i.e.

$$\gamma V(s,x) + (1-\gamma)V(s,x') \leq \varepsilon + V(s,x_\gamma)$$

and therefore, by the arbitrariness of ε, the claim (1.18) would be proved. We will show that

$$\theta_\gamma(t) := a(t)\theta(t) + d(t)\theta'(t),$$

where

$$a(\cdot) = \gamma \frac{X(\cdot)}{X_\gamma(\cdot)}, \quad d(\cdot) = (1-\gamma)\frac{X'(\cdot)}{X_\gamma(\cdot)},$$

is good. The admissibility of $\theta_\gamma(\cdot)$ is clear since:

(i) for any $t \in [s,T]$ we have $\theta\,(t)\,, \theta'\,(t) \in [0,1]$, and $a\,(t) + d\,(t) = 1$, so that by convexity of $[0,1]$ we get $\theta_\gamma\,(t) \in [0,1]$;

(ii) by construction $X_\gamma(t) \geq l(t)$ for any $t \in [s,T]$.

Note that actually we will prove that $X_\gamma(\cdot) = X(\cdot\,; s, x_\gamma, \theta_\gamma(\cdot))$. The equation satisfied by $X_\gamma(\cdot)$ in the interval $[s,T]$ is

$$\begin{aligned}
dX_\gamma(t) &= \gamma\, dX(t) + (1-\gamma)\, dX'(t) \\
&= \gamma\Big[\big[(r+\sigma\lambda\theta(t))X(t) + kt\big]dt + \sigma\theta(t)X(t)dB^s(t)\Big] \\
&\quad +(1-\gamma)\Big[\big[(r+\sigma\lambda\theta'(t))X'(t) + kt\big]dt + \sigma\theta'(t)X'(t)dB^s(t)\Big] \\
&= \Big[rX_\gamma(t) + \sigma\lambda\big(\gamma\theta(t)X(t) + (1-\gamma)\theta'(t)X'(t)\big) + kt\Big]dt \\
&\quad +\sigma\big[\gamma\theta(t)X(t) + (1-\gamma)\theta'(t)X'(t)\big]dB^s(t) \\
&= [rX_\gamma(t) + kt]\,dt + \sigma\lambda\left[\gamma\theta(t)\frac{X(t)}{X_\gamma(t)} + (1-\gamma)\theta'(t)\frac{X'(t)}{X_\gamma(t)}\right]X_\gamma(t)\,dt \\
&\quad +\sigma\left[\gamma\theta(t)\frac{X(t)}{X_\gamma(t)} + (1-\gamma)\theta'(t)\frac{X'(t)}{X_\gamma(t)}\right]X_\gamma(t)dB^s(t) \\
&= \Big[\big(r+\sigma\lambda\theta_\gamma(t)\big)X_\gamma(t) + kt\Big]dt + \sigma\theta_\gamma(t)X_\gamma(t)dB^s(t)
\end{aligned}$$

and the claim follows. □

Proposition 1.2.11. *Let u or f be strictly increasing and let $s \in [0,T]$. Then $x \mapsto V(s,x)$ is strictly increasing on $[l(s), +\infty)$.*

Proof. Let $l(s) \leq x \leq x'$; Writing the equation for $Y(\cdot) := X(\cdot; s, x', \theta(\cdot)) - X(\cdot; s, x, \theta(\cdot))$, we can see that $Y(\cdot)$ solves a linear SDE with nonnegative initial datum. Therefore $Y(\cdot) \geq 0$, $\mathbb{P}$-a.s., i.e. $X(t; s, x, \theta(\cdot)) \leq X(t; s, x', \theta(\cdot))$ for all $\theta(\cdot) \in \Theta_{ad}(s,x)$ and in particular $\Theta_{ad}(s,x) \subset \Theta_{ad}(s,x')$. Moreover by monotonicity of u, f we get $J(s,x;\theta(\cdot)) \leq J(s,x',\theta(\cdot))$ for all $\theta(\cdot) \in \Theta_{ad}(s,x)$, so that $V(s,\cdot)$ is increasing. Now we prove that this function is strictly increasing. We can note that, if a concave and increasing function is not strictly inceasing, then such a function must be definetively constant on a right half line $[\bar{x}, +\infty)$; we show that this is not our case.

If we have $\lim_{x\to\infty} u(x) = +\infty$ or $\lim_{x\to\infty} f(x) = +\infty$, then, by Proposition 2.2.12, we must have also $\lim_{x\to+\infty} V(s,x) = +\infty$ and the claim follows. Instead let us suppose that we have both $\lim_{x\to+\infty} u(x) = \bar{u} < +\infty$ and $\lim_{x\to\infty} f(x) = \bar{f} < +\infty$ and suppose by contradiction that $V(s,\cdot)$ is constant on $[\bar{x}, +\infty)$ for some $\bar{x} \geq l(s)$. Again by Proposition 2.2.12 we must have

$$V(s,\bar{x}) = \lim_{x\to+\infty} V(s,x) \geq \frac{\bar{u}}{\rho}\, e^{-\rho s}(1 - e^{-\rho(T-s)}) + \bar{f};$$

on the other hand, taking into account (1.17), the concavity and the monotonicity of u, f, we can write, for any $\theta(\cdot) \in \Theta_{ad}(s,\bar{x})$, setting $X(t) := X(t; s, \bar{x}, \theta(\cdot))$,

$$\begin{aligned} J(s,\bar{x};\theta(\cdot)) &= \int_s^T e^{-\rho t}\mathbb{E}[u(X(t) - l(t))]dt + \mathbb{E}\left[f\big(X(T)\big)\right] \\ &\leq \int_s^T e^{-\rho t} u(\mathbb{E}[X(t)] - l(t))dt + f\big(\mathbb{E}\left[X(T)\right]\big) \\ &\leq \int_s^T e^{-\rho t} u(C(1+\bar{x}))dt + f\big(C(1+\bar{x})\big) \\ &= \frac{u(C(1+\bar{x}))}{\rho}\, e^{-\rho s}(1 - e^{-\rho(T-s)}) + f\big(C(1+\bar{x})\big), \end{aligned}$$

i.e., since u or f is strictly increasing,

$$\begin{aligned} V(s,x) &\leq \frac{u(C(1+\bar{x}))}{\rho}\, e^{-\rho s}(1 - e^{-\rho(T-s)}) + f\big(C(1+\bar{x})\big) \\ &< \frac{\bar{u}}{\rho}\, e^{-\rho s}(1 - e^{-\rho(T-s)}) + \bar{f}; \end{aligned}$$

thus a contradiction arises and the claim is proved. $\square$

1.2.4 Continuity of the value function

In this section we will prove that the value function is continuous on

$$\mathcal{C} = \{(s,x) \in \mathbb{R}^2 \mid x \geq l(s)\}.$$

We will prove this result by some lemmata.

Lemma 1.2.12. *Let $s \in [0,T]$, $\varepsilon > 0$; the function $[l(s)+\varepsilon, +\infty) \to [0,+\infty),\ x \mapsto V(s,x)$ is Lipschitz continuous.*

Proof. The claim follows by the fact that the function $[l(s), +\infty) \rightarrow \mathbb{R}$, $x \mapsto V(s,x)$ is concave and increasing. This implies that this function is continuous in the interior part of its domain $(l(s), +\infty)$ and Lipschitz continuous on $[l(s)+\varepsilon, +\infty)$ for any $\varepsilon > 0$. □

Let us define, for $a \geq 0$, the curves

$$\mathcal{L}_a := \{(s, l(s)+a) \mid s \in [0,T]\};$$

we analyze the behaviour of the value function along these curves.

Lemma 1.2.13. *Let $u(0) \geq 0$, $a \geq 0$; then the value function is nonincreasing along the curve $\mathcal{L}_a$, i.e. the function $[0,T] \rightarrow \mathbb{R}$, $s \mapsto V(s, l(s)+a)$ is nonincreasing.*

Proof. Let $s \in [0,T]$, $s' \in (s,T]$ and let x, x' be such that $(s,x), (s',x') \in \mathcal{L}_a$, for some $a \geq 0$; let us consider $X(t) := X(t; s, x, 0)$; we get

$$V(s,x) \geq \int_s^{s'} e^{-\rho t} U(t, X(t)) dt + V(s', X(s')) \geq V(s', x')$$

where the first inequality follows by the dynamic programming principle (see Theorem 1.2.20 and Remark 1.2.21), while the second one follows taking into account (1.11), which gives that $X(s') \geq x'$, the fact that the utility function u is positive and Proposition 1.2.11. □

Remark 1.2.14. Let $0 \leq s \leq s' \leq T$; from [Yong, Zhou; 1999], Chapter 1, Theorem 2.10, we see that we can map in a natural way a strategy starting at time s into a strategy starting at time s'. Indeed, consider the measurable space $(C[s,T], \mathcal{B}(C[s,T]))$, endowed with the filtration $(\mathcal{B}_t(C[s,T]))_{t\in[s,T]}$ defined in the following way: $(\mathcal{B}_t(C[s,T]))$ is the σ-algebra on $C[s,T]$ induced by the projection

$$\begin{array}{rccc} \pi: & C[s,T] & \longrightarrow & (C[s,t], \mathcal{B}(C[s,t])) \\ & \zeta(\cdot) & \longmapsto & \zeta(\cdot)|_{[s,t]}, \end{array}$$

i.e. the smallest σ-algebra which makes π measurable; intuitively a measurable application with respect to $\mathcal{B}_t(C[s,T])$ is an application which does not distinguish between two functions of $C[s,T]$ which coincide on $[s,t]$. If $(\theta_s(t))_{t\in[s,T]}$ is a strategy starting from s, there exists a process ψ on $(C[s,T], \mathcal{B}(C[s,T]))$, adapted with respect to $(\mathcal{B}(C[s,T]))_{t\in[s,T]}$, such that

$$\theta_s(t) = \psi(t, B^s(\cdot)), \quad t \in [s,T];$$

then we can consider the strategy

$$\theta_{s'}(t) = \psi(t - s' + s, B^{s'}(\cdot)), \quad t \in [s',T],$$

starting from s'; we denote by $\Gamma_{s,s'}$ the map $\theta_s \mapsto \theta_{s'}$. $\square$

The following lemma is the crucial key to prove the continuity of the value function:

Lemma 1.2.15. *Let $u(0) \geq 0$; then the value function is continuous along the curves $\mathcal{L}_a$ for any $a \geq 0$, i.e. the function $[0,T] \to \mathbb{R}$, $s \mapsto V(s, l(s) + a)$ is continuous.*

Proof. Fix $a \geq 0$ and $s \in [0,T)$, let $s' \in (s,T]$ and let x, x' be such that $(s,x), (s',x') \in \mathcal{L}_a$, i.e $x' - x = l(s') - l(s)$. Take a control $\theta_s(\cdot) \in \Theta_{ad}(s,x)$, set $\varepsilon := s' - s$ and consider, for $t \in [s,T]$, the process $X_s(t) := X(t; s, x, \theta_s(\cdot))$ and, for $t \in [s',T]$, the process $Y_{s'}(t)$ given by

$$\begin{cases} dY_{s'}(t) = (r + \sigma\lambda\theta_{s'}(t))Y_{s'}(t)dt + k \cdot (t - \varepsilon)dt + \sigma\theta_{s'}(t)Y_{s'}(t)dB^{s'}(t), \\ Y_{s'}(s') = x, \end{cases}$$

where $\theta_{s'}(\cdot) = \Gamma_{s,s'}(\theta_s(\cdot))$. Denote by the symbol $\stackrel{\mathcal{L}}{=}$ the equality in law of two random variables. We strightly have $X_s(t - \varepsilon) \stackrel{\mathcal{L}}{=} Y_{s'}(t)$ and therefore, by the assumption $\theta(\cdot) \in \Theta_{ad}(s,x)$, we get $Y_{s'}(t) \geq l(t - \varepsilon)$ almost surely for every $t \in [s',T]$. Define the "semi-feedback" strategy $\tilde{\theta}_{s'}(\cdot)$ starting from s' by

$$\tilde{\theta}_{s'}(t) = \theta_{s'}(t)\frac{Y_{s'}(t)}{X_{s'}(t)},$$

where $X_{s'}(\cdot)$ denotes the solution to the state equation starting from x' at time s' under the strategy $\tilde{\theta}_{s'}(\cdot)$; we want to show that $\tilde{\theta}_{s'}(\cdot)$ takes values in $[0,1]$ and that $\tilde{\theta}_{s'}(\cdot) \in \Theta_{ad}(s',x')$. The dynamics of $X_{s'}(\cdot) - Y_{s'}(\cdot)$ is given by

$$\begin{cases} d(X_{s'}(t) - Y_{s'}(t)) = r(X_{s'}(t) - Y_{s'}(t))\,dt + \varepsilon k\,dt \\ X_{s'}(s') - Y_{s'}(s') = l(s') - l(s); \end{cases} \tag{1.19}$$

the dynamics of $l(t) - l(t - \varepsilon)$ is given by

$$\begin{cases} d(l(t) - l(t - \varepsilon)) = \beta(l(t) - l(t - \varepsilon))\,dt + \varepsilon k\,dt \\ l(s') - l(s' - \varepsilon) = l(s') - l(s) \geq 0; \end{cases} \tag{1.20}$$

comparing (1.19), (1.20) and taking into account that $\beta \leq r$, we get

$$X_{s'}(t) - Y_{s'}(t) \geq l(t) - l(t - \varepsilon). \tag{1.21}$$

As a byproduct of (1.21) we get that $\tilde{\theta}_{s'}(\cdot)$ takes values in the set $[0,1]$ and, since $Y_{s'}(t) \geq l(t - \varepsilon)$, that $\theta_{s'}(\cdot) \in \Theta(s',x')$.

Let δ be the modulus of uniform continuity of u and let δ' be the modulus of uniform continuity of l. We have proved that $X_{s'}(t) \geq Y_{s'}(t) \stackrel{\mathcal{L}}{=} X_s(t-\varepsilon)$, for $t \in [s', T]$; thus, taking also into account that U is increasing with respect to the second argument, we have

$$\begin{aligned}
\mathbb{E}\left[\int_{s'}^{T} e^{-\rho t} U(t, X_{s'}(t)) dt\right] &\geq \mathbb{E}\left[\int_{s'}^{T} e^{-\rho t} U(t, X_s(t-\varepsilon)) dt\right] \\
&= \mathbb{E}\left[\int_{s}^{T-\varepsilon} e^{-\rho(t+\varepsilon)} U(t+\varepsilon, X_s(t)) dt\right] \\
&= e^{-\rho\varepsilon}\mathbb{E}\left[\int_{s}^{T-\varepsilon} e^{-\rho t} u(X_s(t) - l(t+\varepsilon)) dt\right] \\
&\geq e^{-\rho\varepsilon}\mathbb{E}\left[\int_{s}^{T-\varepsilon} e^{-\rho t} u(X_s(t) - l(t) - \delta'(\varepsilon)) dt\right] \\
&\geq e^{-\rho\varepsilon}\mathbb{E}\left[\int_{s}^{T-\varepsilon} e^{-\rho t} u(X_s(t) - l(t)) dt\right] - C\delta(\delta'(\varepsilon)) \\
&= e^{-\rho\varepsilon}\mathbb{E}\left[\int_{s}^{T} e^{-\rho t} u(X_s(t) - l(t)) dt\right] - C\delta(\delta'(\varepsilon)) \\
&\quad - e^{-\rho\varepsilon}\mathbb{E}\left[\int_{T-\varepsilon}^{T} e^{-\rho t} U(t, X_s(t)) dt\right],
\end{aligned} \tag{1.22}$$

for a suitable constant $C > 0$. Note that

$$\delta(\delta'(\varepsilon)) \to \text{ when } \varepsilon \to 0. \tag{1.23}$$

By mean-square continuity of X_s and by uniform continuity of u

$$\mathbb{E}\left[\int_{T-\varepsilon}^{T} e^{-\rho t} U(t, X_s(t)) dt\right] \longrightarrow 0, \qquad \text{as } \varepsilon \to 0. \tag{1.24}$$

Since $X_{s'}(t) \geq Y_{s'}(t) \stackrel{\mathcal{L}}{=} X_s(t-\varepsilon)$, for $t \in [s', T]$, we get

$$\mathbb{E}\left[f(X_{s'}(T))\right] \geq \mathbb{E}\left[f(X_s(T-\varepsilon))\right]; \tag{1.25}$$

moreover, since X_s is mean-square continuous and f is uniformly continuous, by Lemma A.1.2 we get

$$\mathbb{E}\left[|f(X_s(T-\varepsilon)) - f(X_s(T))|^2\right] \longrightarrow 0, \tag{1.26}$$

when $\varepsilon \to 0$, so that, combining (1.25) and (1.26),

$$\mathbb{E}\left[f(X_{s'}(T))\right] \geq \mathbb{E}\left[f(X_s(T))\right] - \eta(\varepsilon), \qquad \text{with } \eta(\varepsilon) \to 0 \text{ when } \varepsilon \to 0. \tag{1.27}$$

Fix $s \in [0, T]$; combining (1.22), (1.23), (1.24) and (1.27), we get that, for any $\varepsilon > 0$ and any control $\theta_s(\cdot) \in \Theta_{ad}(s, x)$, there exists a control $\tilde{\theta}_{s'}(\cdot) \in \Theta_{ad}(s', x')$ such that

$$J(s', x'; \tilde{\theta}_{s'}(\cdot)) \geq e^{-\rho\varepsilon} J(s, x; \theta_s(\cdot)) - \omega(\varepsilon), \tag{1.28}$$

and $\omega(\varepsilon) \to 0$ when $\varepsilon \to 0$. Therefore passing to the supremum over $\theta_s(\cdot) \in \Theta_{ad}(s,x)$ the right hand-side of (1.28) we get

$$V(s',x') \geq e^{-\rho\varepsilon} V(s,x) - \omega(\varepsilon), \quad \text{with } \omega(\varepsilon) \to 0 \text{ when } \varepsilon \to 0. \tag{1.29}$$

This shows that the value function is lower semicontinuous from the right along $\mathcal{L}_a$ and upper semicontinuous from the left along $\mathcal{L}_a$. Since it is nonincreasing by Lemma 1.2.13, it must be continuous along $\mathcal{L}_a$. □

Let us define, for $\varepsilon > 0$, the sets

$$S_\varepsilon := \{(s,x) \in \mathbb{R}^2 \mid x \geq l(s) + \varepsilon\}.$$

We show that the value function is continuous on these sets.

Lemma 1.2.16. *Let $u(0) \geq 0$; then the value function is continuous on the sets S_ε, for any $\varepsilon > 0$.*

Proof. Let $s \in [0,T]$, $\varepsilon > 0$ and consider the function $[l(s)+\varepsilon) \to \mathbb{R}$, $x \mapsto V(s,x)$; by Lemma 1.2.12 this function is Lipschitz continuous: we want to estimate its Lipschitz constant. The function $[l(s), +\infty) \to \mathbb{R}$, $x \mapsto V(s,x)$ is concave (so that the incremental ratios are nonincreasing) and increasing (so that the incremental ratios are positive); thus, if we set

$$M_{s,\varepsilon} := \frac{V(s, l(s)+\varepsilon) - V(s, l(s))}{\varepsilon},$$

we get that $M_{s,\varepsilon}$ is good as Lipshitz constant for the function $[l(s)+\varepsilon, +\infty) \to \mathbb{R}$, $x \mapsto V(s,x)$. By Lemma 1.2.15 there exists

$$M_\varepsilon := \max_{s\in[0,T]} M_{s,\varepsilon} = \max_{s\in[0,T]} \frac{V(s, l(s)+\varepsilon) - V(s, l(s))}{\varepsilon};$$

thus the functions $[l(s)+\varepsilon, +\infty) \to [0,+\infty)$, $x \mapsto V(s,x)$, $s \in [0,T]$, are Lipschitz continuous with respect the same Lipschitz constant M_ε.

This uniform Lipschitz continuity of $V(s,\cdot)$ together with Lemma 1.2.15 yield the claim.

□

Lemma 1.2.17. *Let $u(0) \geq 0$, $s \in [0,T]$; the function $[l(s), +\infty) \to [0+\infty)$, $x \mapsto V(s,x)$ is continuous at $l(s)$.*

Proof. Of course the function is lower semicontinuous at $l(s)$, since it is increasing; we will prove that it is also upper semicontinuous at $l(s)$. We have to distinguish the two cases when the boundary is absorbing or not, i.e. when $\beta < r$ or $\beta = r$ (see Proposition 1.2.6).

Case 1: $\beta < r$. Fix $s \in (0, T]$ and take $s' \in [0, s)$; consider $X_{s'}(t) := X(t; s', l(s'), 0)$; then the function $t \mapsto V(t, X_{s'}(t))$ is nonincreasing by dynamic programming principle (see Theorem 1.2.20), because u is positive. Moreover, looking at the proof of Proposition 1.2.5, we see that $X_{s'}(s) > l(s)$, due to the assumption $\beta < r$, and that $X_{s'}(s) \downarrow l(s)$, when $s' \uparrow s$, due to the continuous dependence on the initial datum s' of the state equation. Consider also the value function along $\mathcal{L}_0$ in the time interval $[s', s]$; by Lemma 1.2.15 it is continuous, so that

$$\limsup_{x \downarrow l(s)} V(s, x) = \limsup_{s' \uparrow s} V(s, X_{s'}(s)) \leq \limsup_{s' \uparrow s} V(s', l(s')) = V(s, l(s)),$$

where the inequality holds since $t \mapsto V(t, X_{s'}(t))$ is nonincreasing and the last equality holds since the value function is continuous along $\mathcal{L}_0$; therefore the claim is proved for $s \in (0, T]$.
In the case $s = 0$, we can argue as well as before by extending also for $s < 0$ the problem with $k = 0$, setting the solvency level $l(s) \equiv l_0$ and defining the value function in obvious way.

Case 2: $\beta = r$. In this case we proceed directly with estimates on the state equation. Let $x > l(s)$, $\varepsilon > 0$ and let D be the density of $\mathbb{P}$ with respect to the probability measure $\tilde{\mathbb{P}}$ given by the Girsanov transformation (see Theorem A.1.1), which belongs to $L^p(\Omega, \tilde{\mathbb{P}})$, for any $p \in [1, +\infty)$. For any $\theta(\cdot) \in \Theta_{ad}(s, x),\ t \in [s, T]$, we have, by Hölder and Markov inequalities

$$\begin{aligned}
\mathbb{P}\{X(t; s, x, \theta(\cdot)) - l(t) > \varepsilon\} &= \mathbb{E}\left[I_{\{X(t;s,x,\theta(\cdot)) - l(t) > \varepsilon\}}\right] \\
&= \tilde{\mathbb{E}}\left[I_{\{X(t;s,x,\theta(\cdot)) - l(t) > \varepsilon\}}\, D\right] \\
&\leq \left(\tilde{\mathbb{E}}[D^2]\right)^{1/2} \left(\tilde{\mathbb{E}}\left[I_{\{X(t;s,x,\theta(\cdot)) - l(t) > \varepsilon\}}\right]\right)^{1/2} \\
&= \left(\tilde{\mathbb{E}}[D^2]\right)^{1/2} \left(\tilde{\mathbb{P}}\{X(t; s, x, \theta(\cdot)) - l(t) > \varepsilon\}\right)^{1/2} \\
&\leq \frac{\left(\tilde{\mathbb{E}}[D^2]\right)^{1/2}}{\varepsilon^{1/2}} \left(\tilde{\mathbb{E}}\left[X(t; s, x, \theta(\cdot)) - l(t)\right]\right)^{1/2}.
\end{aligned}$$

Let us estimate $\tilde{\mathbb{E}}\left[X(t; s, x, \theta(\cdot)) - l(t)\right]$; we have, arguing as in the proof of Proposition 1.2.5,

$$\begin{cases} d\tilde{\mathbb{E}}[X(t)] = rX(t)dt + ktdt, \\ \tilde{\mathbb{E}}[X(s)] = x, \end{cases}$$

and

$$\begin{cases} dl(t) = rl(t)dt + ktdt, \\ l(s) = l(s), \end{cases}$$

so that

$$\tilde{\mathbb{E}}\left[X(t;s,x,\theta(\cdot)) - l(t)\right] = (x - l(s))\, e^{r(t-s)}.$$

Take a sequence $(\varepsilon_n, \delta_n)$ such that $\varepsilon_n > 0$, $\delta_n > 0$. We have shown that, for any $n \in \mathbb{N}$, we can find x_n such that, for every $n \in \mathbb{N}$,

- $l(s) < x_n < l(s) + 1/n$;
- $\mathbb{P}\{X(t;s,x_n,\theta(\cdot)) - l(t) > \varepsilon_n\} < \delta_n$, for every $\theta(\cdot) \in \Theta_{ad}(s,x_n), t \in [s,T]$.

Moreover we can estimate $\mathbb{E}\left[(X(t;s,x,\theta(\cdot)) - l(t))^2\right]$ uniformly with respect to $t, x, \theta(\cdot)$ varying in the sets $[s,T]$, $[l(s), l(s)+1]$, $\Theta_{ad}(s,x)$; indeed, by the Dynkin's formula applied to the function $\psi(t,x) = (x - l(t))^2$, we have

$$\mathbb{E}\left[(X(t;s,x,\theta(\cdot)) - l(t))^2\right] = (x - l(s))^2 + \mathbb{E}\Bigg[\int_s^t \Big(-2(X(r) - l(r))l'(r)$$
$$+ 2(X(r) - l(r))\left((r + \sigma\lambda\theta(r))X(r) + kr\right) + \sigma^2\theta(r)^2 X(r)^2\Big) dr\Bigg]$$

and the right-handside is dominated by a constant C (not dependent on $t \in [s,T]$, $x \in [l(s), l(s)+1]$, $\theta(\cdot) \in \Theta_{ad}(s,x)$) by mean-square continuity of $X(\cdot)$ and since $\theta(\cdot)$ takes values in $[0,1]$. Thus we can split the expectation and write, again by Hölder's inequality,

$$\begin{aligned}
&\mathbb{E}[X(t;s,x_n,\theta(\cdot)) - l(t)] \\
&\qquad \le \varepsilon_n + \mathbb{E}\big[I_{\{X(t;s,x_n,\theta(\cdot)) - l(t) > \varepsilon_n\}}(X(t;s,x_n,\theta(\cdot)) - l(t))\big] \\
&\qquad\qquad \le \varepsilon_n + C\delta_n^{1/2},
\end{aligned}$$

for some constant C not dependent on n. Hence, for such a point $x_n > l(s)$, we have, by concavity and monotonicity of u, f and by Jensen's inequality,

$$V(s,x_n) \le \int_s^T u(\varepsilon_n + C\delta_n^{1/2})dt + f(\varepsilon_n + C\delta_n^{1/2} + l(T)).$$

If we take $(\varepsilon_n, \delta_n)$ such that $(\varepsilon_n, \delta_n) \to 0$ when $n \to \infty$, the right hand-side in the previous inequality tends to $V(s, l(s))$ and so the claim is proved. $\square$

Lemma 1.2.18. *Let $u(0) \ge 0$; then the value function is continuous on $\mathcal{C}$.*

Proof. It remains only to prove the continuity at the boundary, as in the interior part it was proved in Lemma 1.2.16. By Lemma 1.2.17 we know that

$$V(s, l(s) + \varepsilon) \downarrow V(s, l(s)), \quad \text{for any } s \in [0,T],$$

(when $\varepsilon \downarrow 0$) and moreover, by Lemma 1.2.15, we know that $s \mapsto V(s, l(s))$ is continuous. Therefore by Dini's lemma, $V(\cdot, l(\cdot) + \varepsilon) \to V(\cdot, l(\cdot))$ uniformly when $\varepsilon \downarrow 0$. This convergence, togheter with the continuity of $s \mapsto V(s, l(s))$, is enough to prove the claim. □

In order to conclude we have to remove the assumption $u(0) \geq 0$:

Proposition 1.2.19. *The value function is continuous on $\mathcal{C}$.*

Proof. Let $u(0) = -c < 0$; consider the function $u^c(\cdot) := u(\cdot) + c$ and let V^c be the value function associated with this utility function; by Lemma 1.2.18 V^c is continuous on $\mathcal{C}$. Moreover $V(s, x) = V^c(s, x) - \frac{c}{\rho}(e^{-\rho s} - e^{-\rho T})$ and so the claim follows. □

1.2.5 Dynamic programming

We study the optimization problem following a dynamic programming approach. The core of the dynamic programming is the so-called dynamic programming principle, which in our context can be stated as follows.

Theorem 1.2.20. *The value function V satisfies the dynamic programming equation, i.e. for every $s \in [0, T]$, $x \in [l(s), +\infty)$ and for any family of stopping times $\left(\tau^{\theta(\cdot)}\right)_{\theta(\cdot) \in \Theta_{ad}(s,x)}$ taking values in $[s, T]$, the following functional equation holds true:*

$$V(s, x) = \sup_{\theta(\cdot) \in \Theta_{ad}(s,x)} \mathbb{E}\left[\int_s^{\tau^{\theta(\cdot)}} e^{-\rho t} U(t, X(t; s, x, \theta(\cdot)))\, dt + V\left(\tau^{\theta(\cdot)}, X\left(\tau^{\theta(\cdot)}; s, x, \theta(\cdot)\right)\right)\right]. \tag{1.30}$$

Proof. Actually we give here only a heuristic proof.[1] For simplicity of notation we suppress the possible dependence of τ on $\theta(\cdot)$, i.e. we will write

[1]The proof is only heuristic because of the definition of the strategy $\bar{\theta}$; this is a pointwise definition and it gives trouble about measurability, because of the uncountability of the space Ω.

simply τ to intend $\tau^{\theta(\cdot)}$. Of course we have

$$\begin{aligned}
V(s,x) &= \sup_{\theta(\cdot)\in\Theta_{ad}(s,x)} \mathbb{E}\left[\int_s^T e^{-\rho t}U(t,X(t;s,x,\theta(\cdot)))dt\right]\\
&= \sup_{\theta(\cdot)\in\Theta_{ad}(s,x)} \mathbb{E}\Bigg[\int_s^\tau e^{-\rho t}U(t,X(t;s,x,\theta(\cdot)))dt\\
&\qquad + \int_\tau^T e^{-\rho t}U(t,X(t;s,x,\theta(\cdot)))dt\Bigg]\\
&= \sup_{\theta(\cdot)\in\Theta_{ad}(s,x)} \mathbb{E}\Bigg[\int_s^\tau e^{-\rho t}U(t,X(t;s,x,\theta(\cdot)))dt\\
&\qquad + \mathbb{E}\left[\int_\tau^T e^{-\rho t}U(t,X(t;s,x,\theta(\cdot)))dt\middle|\mathcal{F}^s_\tau\right]\Bigg]\\
&= \sup_{\theta(\cdot)\in\Theta_{ad}(s,x)} \mathbb{E}\left[\int_s^\tau e^{-\rho t}U(X(t;s,x,\theta(\cdot)))dt + J\left(\tau,X\left(\tau;s,x,\theta(\cdot)\right)\right)\right]\\
&\le \sup_{\theta(\cdot)\in\Theta_{ad}(x)} \mathbb{E}\left[\int_s^\tau e^{-\rho t}U(X(t;s,x,\theta(\cdot)))dt + V\left(\tau,X\left(\tau;s,x,\theta(\cdot)\right)\right)\right].
\end{aligned}$$

Conversely, for fixed $\varepsilon > 0$, for any (s',y) such that $s' \in [s,T]$, $y \ge l(s')$, let $\theta^\varepsilon_{s',y}(\cdot)$ a control ε-optimal for the pair (s',y), i.e. $J(s',y;\theta^\varepsilon_{s',y}(\cdot)) \ge V(s',y) - \varepsilon$. Let $\theta(\cdot) \in \Theta_{ad}(s,x)$ and define the control

$$\bar\theta(t) = \begin{cases} \theta(t), & \text{if } t \in [s,\tau],\\ \theta^\varepsilon_{\tau,X(\tau;s,x,\theta(\cdot))}(t), & \text{if } t \in [\tau,T]. \end{cases}$$

Of course we have $\bar\theta(\cdot) \in \Theta_{ad}(s,x)$, so that

$$\begin{aligned}
V(s,x) \ge J(s,x;\bar\theta(\cdot)) &= \mathbb{E}\Bigg[\int_s^\tau e^{-\rho t}U(t,X(t;s,x,\theta(\cdot)))dt\\
&\quad + \int_\tau^T e^{-\rho t}U\left(t,X\left(t;\tau,X(\tau;s,x,\theta(\cdot)),\theta^\varepsilon_{\tau,X(\tau;s,x,\theta(\cdot))}\right)\right)dt\Bigg]\\
&= \mathbb{E}\Bigg[\int_s^\tau e^{-\rho t}U(t,X(t;s,x,\theta(\cdot))dt\\
&\quad + \mathbb{E}\left[\int_\tau^T e^{-\rho t}U\left(t,X\left(t;\tau,X(\tau;s,x,\theta(\cdot)),\theta^\varepsilon_{\tau,X(\tau;s,x,\theta(\cdot))}\right)\right)dt\;\middle|\;\mathcal{F}^s_\tau\right]\Bigg]\\
&= \mathbb{E}\left[\int_s^\tau e^{-\rho t}U(t,X(t;T,x,\theta(\cdot)))dt + J\left(\tau,X(\tau;s,x,\theta(\cdot));\theta^\varepsilon_{\tau,X(\tau;s,x;\theta(\cdot))}(\cdot)\right)\right]\\
&\ge \mathbb{E}\left[\int_s^\tau e^{-\rho t}U(t,X(t;s,x,\theta(\cdot)))dt + V(\tau,X(\tau;s,x,\theta(\cdot)))\right] - \varepsilon.
\end{aligned}$$

By taking the supremum over all $\theta(\cdot) \in \Theta_{ad}(s,x)$ and by arbitrariness of ε, we get the desired inequality and so the claim. $\square$

Remark 1.2.21. We did not give a satisfactory proof of Theorem 1.2.20, but we want to comment about it: in [Yong, Zhou; 1999], Chapter 4, Theorem 3.3, it is contained a proof of this statement when the value function is continuous: therein the state is unconstrained, but the argument can be easily adapted to our case. A general proof of this statement, where the continuity of the value function is not required, is contained in [Soner; 2004]: it requires a measurable selection result.

However, we want to point out that we have proved the continuity of our value function using in Lemma 1.2.13 and in Lemma 1.2.17 only the inequality

$$V(s,x) \geq \int_s^{s'} e^{-\rho t} U(t, X(t;s,x,0))dt + V(s', X(s';s,x,0)),$$
$$0 \leq s \leq s' \leq T,\ x \geq l(s),$$

which can be proved without any measurable selection argument, because in this case we are on a deterministic trajectory. Therefore we can use the argument of [Yong, Zhou; 1999] in order to prove the dynamic programming principle without loss of generality. □

We want to write the HJB equation associated to our problem. To this aim we introduce the following Hamiltonian function

$$\mathcal{H}(s,x,p,Q) := \sup_{\theta \in [0,1]} \mathcal{H}_{cv}(s,x,p,Q;\theta), \quad s \in [0,T], \quad x \in [l(s), +\infty),\ p, Q \in \mathbb{R},$$

where

$$\mathcal{H}_{cv}(s,x,p,Q;\theta) := e^{-\rho s} U(s,x) + [(r + \sigma\lambda\theta)x + ks]p + \frac{1}{2}\sigma^2\theta^2 x^2 Q. \quad (1.31)$$

Formally the HJB equation on the domain $\mathcal{C}$ associated with our problem is

$$\begin{cases} -v_s(s,x) - \mathcal{H}(s,x,v_x(s,x),v_{xx}(s,x)) = 0, & (s,x) \in \overset{\circ}{\mathcal{C}}, \\ v(T,x) = f(x), & x \in [l(T), +\infty); \end{cases} \quad (1.32)$$

setting

$$\mathcal{H}^0_{cv}(x,p,Q;\theta) := \sigma\lambda\theta x p + \frac{1}{2}\sigma^2\theta^2 x^2 Q,$$

we can write

$$\mathcal{H}(s,x,p,Q) = e^{-\rho s} U(s,x) + (rx + ks)p + \sup_{\theta \in [0,1]} \mathcal{H}^0_{cv}(x,p,Q;\theta).$$

To calculate the Hamiltonian we can observe that the function

$$\mathcal{H}^0_{cv}(x,p,Q;\theta) = \sigma\lambda\theta x p + \frac{1}{2}\sigma^2\theta^2 x^2 Q,$$

when $p \geq 0$, $Q \leq 0$, $p^2 + Q^2 > 0$, has a unique maximum point over $\theta \in [0,1]$ given by

$$\theta^* = \left(-\frac{\lambda p}{\sigma x Q}\right) \wedge 1$$

(where we mean that, for $Q = 0$, it is $\theta^* = 1$) and

$$\mathcal{H}^0\left(x,p,Q\right) := \sup_{\theta\in[0,1]} \mathcal{H}^0_{cv}\left(x,p,Q;\theta\right) = \begin{cases} -\frac{\lambda^2 p^2}{2Q}, & \text{if} \quad \theta^* < 1, \\ \sigma\lambda x p + \frac{1}{2}\sigma^2 x^2 Q, & \text{if} \quad \theta^* = 1. \end{cases}$$

When $p = Q = 0$ each $\theta \in [0,1]$ is a maximum point of $\mathcal{H}^0_{cv}$ and $\mathcal{H}^0\left(x,0,0\right) = 0$.

1.2.6 The HJB equation: viscosity solutions

We cannot hope to have explicit solutons for (1.32) in general. Moreover, since the diffusion coefficient can vanish, the equation is degenerate and therefore the theory by Krylov & Evans on parabolic PDEs cannot be applied to get existence of regular solutions (see Section 0.1). So we use the viscosity approach to the equation.

Let us consider (1.32) on $\mathcal{C}$. In (1.12) we have introduced the lateral boundary

$$\partial^*\mathcal{C} := \{(s,x) \in \mathcal{C} \mid s \in [0,T),\ x = l(s)\};$$

let us introduce also the set

$$\text{Int}^*(\mathcal{C}) := \text{Int}(\mathcal{C}) \cup \{\{0\} \times (l_0, +\infty)\}.$$

Next we give the definition of viscosity solution to (1.32) (for a survey on viscosity solutions of second order PDEs, see [Crandall, Ishii, Lions; 1992]).

Definition 1.2.22. **(i)** A continuous function $v : \mathcal{C} \to \mathbb{R}$ is called a viscosity subsolution of the HJB equation (1.32) on $\text{Int}^*(\mathcal{C}) \cup \partial^*\mathcal{C}$ if

$$v(T,x) \leq f(x), \quad x \in [l(T), +\infty),$$

and if, for any couple $\psi \in C^2(\mathcal{C};\mathbb{R})$ and $(s_M, x_M) \in \text{Int}^*(\mathcal{C}) \cup \partial^*\mathcal{C}$ such that (s_M, x_M) is a local maximum for $v - \psi$ on $(\mathcal{C})$, we have

$$-\psi_s\left(s_M, x_M\right) - \mathcal{H}\left(s_M, x_M, \psi_x(s_M,x_M), \psi_{xx}(s_M,x_M)\right) \leq 0.$$

(ii) A continuous function $v : \mathcal{C} \to \mathbb{R}$ is called a viscosity supersolution of the HJB equation (1.32) on $\text{Int}^*(\mathcal{C})$ if

$$v(T,x) \geq f(x), \quad x \in [l(T), +\infty),$$

and if, for any couple $\psi \in C^2(\mathcal{C};\mathbb{R})$ and $(s_M, x_M) \in \text{Int}^*(\mathcal{C})$ such that (s_M, x_M) is a local minimum for $v - \psi$ on $\text{Int}^*(\mathcal{C})$, we have

$$-\psi_s(s_M, x_M) - \mathcal{H}(s_M, x_M, \psi_x(s_M, x_M), \psi_{xx}(s_M, x_M)) \leq 0.$$

(iii) A continuous function $v : \mathcal{C} \to \mathbb{R}$ is called a constrained viscosity solution to the HJB equation (1.32) on $\mathcal{C}$ if it is a viscosity subsolution on $\text{Int}^*(\mathcal{C}) \cup \partial^*\mathcal{C}$ and a viscosity supersolution on $\text{Int}^*(\mathcal{C})$. □

Now we can state and prove the following result.

Theorem 1.2.23. *The value function V is a constrained viscosity solution to the HJB equation* (1.32) *on $\mathcal{C}$.*

Proof. (i) Here we prove that V is a viscosity supersolution on $\text{Int}^*(\mathcal{C})$. First of all notice that $V(T, x) = f(x)$, for $x \geq l(T)$, so that the terminal boundary condition is satisfied. Now let $\psi \in C^2(\mathcal{C};\mathbb{R})$ and let $(s_m, x_m) \in \text{Int}^*(\mathcal{C})$ be such that (s_m, x_m) is a local minimum point for $V - \psi$. For proving the supersolution property on $Int * (\mathcal{C})$ we can assume without loss of generality that

$$V(s_m, x_m) = \psi(s_m, x_m), \qquad V(s,x) \geq \psi(s,x), \qquad \forall (s,x) \in \mathcal{C}. \tag{1.33}$$

Let $\theta \in [0,1]$ and set $X(t) := X(t; s_m, x_m, \theta)$. Let us define

$$\tau^\theta = \inf\{t \in [s_m, T] \mid (t, X(t)) \notin \text{Int}^*(\mathcal{C})\},$$

with the convention $\inf \emptyset = T$; of course τ^θ is a stopping time and, by continuity of trajectories, we have $\tau^\theta > s_m$ almost surely. By (1.33) we get, for any $t \in [s_m, \tau^\theta]$,

$$V(t, X(t)) - V(s_m, x_m) \geq \psi(t, X(t)) - \psi(s_m, x_m).$$

Let $h \in (s_m, T]$ and set $\tau_h^\theta := \tau^\theta \wedge h$; by the dynamic programming principle (1.30) we get, for any $\theta \in [0,1]$,

$$\begin{aligned} 0 &\geq \mathbb{E}\left[\int_{s_m}^{\tau_h^\theta} e^{-\rho t} U(t, X(t))\, dt + V(\tau_h^\theta, X(\tau_h^\theta)) - V(s_m, x_m)\right] \\ &\geq \mathbb{E}\left[\int_{s_m}^{\tau_h^\theta} e^{-\rho t} U(t, X(t))\, dt + \psi(\tau_h^\theta, X(\tau_h^\theta)) - \psi(s_m, x_m)\right]. \end{aligned} \tag{1.34}$$

Applying the Dynkin's formula to the function $\psi(t,x)$ with the process $X(t)$, we get

$$\begin{aligned} \mathbb{E}\left[\psi(\tau_h^\theta, X(\tau_h^\theta)) - \psi(s_m, x_m)\right] = \mathbb{E}\Bigg[\int_{s_m}^{\tau_h^\theta} \Big[\psi_s(t, X(t)) \\ + [(r + \sigma\lambda\theta)X(t) + kt]\,\psi_x(t, X(t)) + \frac{1}{2}\sigma^2\theta^2 X(t)^2 \psi_{xx}(t, X(t))\Big] dt\Bigg] \end{aligned}$$

and thus by (1.34) we have

$$
\begin{aligned}
0 \quad \geq \quad & \mathbb{E}\left[\int_{s_m}^{\tau_h^\theta}\left[e^{-\rho t}U\left(t,X\left(t\right)\right)dt+\psi_s(t,X(t))+\left[(r+\sigma\lambda\theta)X(t)+kt\right]\psi_x(t,X(t))\right.\right.\\
& \left.\left.+\frac{1}{2}\sigma^2\theta^2X(t)^2\psi_{xx}(t,X(t))\right]dt\right];
\end{aligned}
$$

Thus, for any $\theta\in[0,1]$, we get

$$
0\geq\mathbb{E}\left[\int_{s_m}^{\tau_h^\theta}\left[\psi_s(t,X(t))+\mathcal{H}_{cv}(t,X(t),\psi_x(t,X(t)),\psi_{xx}(t,X(t));\theta)\right]dt\right];
$$

thus we can write, for $\theta\in[0,1]$,

$$
\begin{aligned}
0\geq\mathbb{E}\Bigg[&\frac{1}{h-s_m}\int_{s_m}^{h}I_{[s_m,\tau^\theta]}(t)\big[\psi_s(t,X(t))\\
&+\mathcal{H}_{cv}(t,X(t),\psi_x(t,X(t)),\psi_{xx}(t,X(t));\theta)\big]dt\Bigg];
\end{aligned}
$$

now, by the continuity properties of ψ and $\mathcal{H}_{cv}$, passing to the limit for $h\to s_m$, we get by dominated convergence

$$
-\psi_s(s_m,x_m)-\mathcal{H}_{cv}\left(s_m,x_m,\psi'\left(s_m,x_m\right),\psi''\left(s_m,x_m\right);\theta\right)\geq 0.
$$

By the arbitrariness of θ we have proved that V is a supersolution on $\mathrm{Int}^*(\mathcal{C})$.

(ii) Here we prove that V is a viscosity subrsolution on $\mathrm{Int}^*(\mathcal{C})\cup\partial^*\mathcal{C}$. Notice again that $V(T,x)=f(x)$, for $x\geq l(T)$, so that the terminal boundary condition is satisfied. Let $\psi\in C^2(\mathcal{C};\mathbb{R})$ and $(s_M,x_M)\in\mathrm{Int}^*(\mathcal{C})\cup\partial^*\mathcal{C}$ such that (s_M,x_M) is a local maximum point for $V-\psi$. For proving the subsolution property we can assume, without loss of generality, that

$$
V(s_M,x_M)=\psi(s_M,x_M),\qquad V(s,x)\leq\psi(s,x),\qquad\forall(s,x)\in\mathcal{C}. \tag{1.35}
$$

We must prove that

$$
-\psi_s(s_M,x_M)-\mathcal{H}\left(s_M,x_M,\psi_x(s_M,x_M),\psi_{xx}(s_M,x_M)\right)\leq 0.
$$

Let us suppose by contradiction that this relation is false. Then there exists $\nu>0$ such that

$$
0<\nu<-\psi_s(s_M,x_M)-\mathcal{H}\left(s_M,x_M,\psi_x(s_M,x_M),\psi_{xx}(s_M,x_M)\right)
$$

The functions $U,\psi,\mathcal{H}$ are continuous. Therefore there exists $\varepsilon>0$ such that, if $(t,x)\in B:=B((s_M,x_M),\varepsilon)\cap\mathcal{C}$, we have, for any $\theta\in[0,1]$,

$$
\begin{aligned}
0<\frac{\nu}{2} \quad &< \quad -\psi_s(s,x)-\mathcal{H}\left(s,x,\psi_x(s,x),\psi_{xx}(s,x)\right)\\
&\leq \quad -\psi_s(s,x)-\mathcal{H}_{cv}\left(s,x,\psi_x(s,x),\psi_{xx}(s,x);\theta\right) \qquad (1.36)
\end{aligned}
$$

Take any control strategy $\theta(\cdot) \in \Theta_{ad}(s_M, x_M)$ and let $X(t) := X(t; s_M, x_M, \theta(\cdot))$. Define the stopping time

$$\tau^\theta := \inf \{t \in [s_M, T] \mid (t, X(t)) \notin B\},$$

with the convention $\inf \emptyset = T$; of course, by continuity of trajectories, we have $\tau^\theta > s_M$ almost surely. Now we can apply (1.36) to $X(t)$, for $t \in [s_M, \tau^\theta]$, getting

$$0 < \frac{\nu}{2} < -\psi_s(t, X(t)) - \mathcal{H}_{cv}(t, X(t), \psi_x(t, X(t)), \psi_{xx}(t, X(t)); \theta(t)); \quad (1.37)$$

integrating (1.37) on $[s_M, \tau^\theta]$ and taking the expectations we get

$$\begin{aligned} 0 &< \frac{\nu}{2}\mathbb{E}\left[\tau^\theta - s_M\right] \\ &\leq -\mathbb{E}\left[\int_{s_M}^{\tau^\theta} \psi_s(t, X(t)) + \mathcal{H}_{cv}(t, X(t), \psi_x(t, X(t)), \psi_{xx}(t, X(t)); \theta(t))\, dt\right]; \end{aligned}$$

we claim that there exists a constant $\delta > 0$, independent of the control $\theta(\cdot) \in \Theta_{ad}(s_M, x_M)$, such that $\frac{\nu}{2}\mathbb{E}[\tau^\theta - s_M] \geq \delta$; we will prove this fact in Lemma 1.2.24. Thus, assuming that, we can write for every $\theta(\cdot) \in \Theta_{ad}(s_M, x_M)$

$$\delta \leq -\mathbb{E}\left[\int_{s_M}^{\tau^\theta} (\psi_s(t, X(t)) + \mathcal{H}_{cv}(t, X(t), \psi_x(t, X(t)), \psi_{xx}(t, X(t)); \theta(t)))\ dt\right].$$

Appliyng the Dynkin formula to X on $[s_M, \tau^\theta]$ we get

$$\psi(s_M, x_M) - \mathbb{E}\left[\psi(\tau^\theta, X(\tau^\theta))\right] \geq \delta + \mathbb{E}\left[\int_{s_M}^{\tau^\theta} e^{-\rho t} U(t, X(t))\, dt\right];$$

from (1.35) we get

$$V(s_M, x_M) - \mathbb{E}\left[V(\tau^\theta, X(\tau^\theta))\right] \geq \delta + \mathbb{E}\left[\int_{s_M}^{\tau^\theta} e^{-\rho t} U(t, X(t))\, dt\right];$$

on the other hand, if we choose a $\delta/2$ optimal control $\theta(\cdot) \in \Theta_{ad}(s_M, x_M)$, we get

$$V(s_M, x_M) - \delta/2 \leq \mathbb{E}\left[\int_{s_M}^{\tau^\theta} e^{-\rho t} U(t, X(t))\, dt + V(\tau^\theta, X(\tau^\theta))\right].$$

So a contradiction arises and we have the claim. □

Lemma 1.2.24. *For any $\theta(\cdot) \in \Theta_{ad}(s_M, x_M)$ let τ^θ be the stopping time defined in the part (ii) of the proof of Theorem 1.2.23. There exists a constant $\delta > 0$ independent of $\theta(\cdot) \in \Theta_{ad}(s_M, x_M)$ such that*

$$\mathbb{E}\left[\tau^\theta - s_M\right] \geq \delta.$$

Proof. For $\theta(\cdot) \in \Theta_{ad}(s_M, x_M)$, let $X(t) := X(t; s_M, x_M, \theta(\cdot))$ and apply the Dynkin formula to the process $X(\cdot)$ with $\varphi(t,x) = (t - s_M)^2 + (x - x_M)^2$ on $[s_M, \tau^\theta]$; we get

$$\mathbb{E}\left[\left(\tau^\theta - s_M\right)^2 + \left(X(\tau^\theta) - x_M\right)^2\right]$$
$$= \mathbb{E}\left[\int_{s_M}^{\tau^\theta} \left[2(t - s_M) + 2(X(t) - x_M)\left[(r + \sigma\lambda\theta(t))X(t) + kt\right] + \sigma^2\theta(t)^2 X(t)^2\right] dt\right].$$

So, considering that $\theta(t) \in [0,1]$ and that for every $t \in [s_M, \tau^\theta]$ we have $|X(t)| \le x_M + \varepsilon$, we can find $K > 0$ such that

$$\begin{aligned} (T - s_M)^2 \wedge \varepsilon^2 &\le \mathbb{P}\{\tau^\theta = T\}(T - s_M)^2 + \mathbb{P}\{\tau^\theta < T\}\varepsilon^2 \\ &\le \mathbb{E}\left[\int_{s_M}^{\tau^\theta} K\, dt\right] = K\,\mathbb{E}\left[\tau^\theta - s_M\right]; \end{aligned}$$

this estimate does not depend on $\theta(\cdot)$ and therefore the claim is proved. □

Remark 1.2.25. In the definition 1.2.22 of constrained viscosity solution we could replace the request that $V - \psi$ has a local maximum (resp. minimum) at (s_M, x_M) (resp. (s_m, x_m)) with the request that it has a right (with respect to the time variable) local maximum (resp. local minimum) at (s_M, x_M) (resp. (s_m, x_m)), i.e., for some $\varepsilon > 0$,

$$V(s_M, x_M) - \psi(s_M, x_M) \ge V(s,x) - \psi(s,x),$$
$$\text{for } (s,x) \in \left\{[s_M, s_M + \varepsilon] \times [x_M - \varepsilon, x_M + \varepsilon]\right\} \cap \mathcal{C}$$

(resp. the analogous for the minimum). Then we could prove exactly as in the proof of Theorem 1.2.23 that V is a constrained viscosity solution also in this stronger sense[2].

1.2.7 The HJB equation: comparison and uniqueness

The definition 1.2.22 of constrained viscosity solution which we have given is the natural version in the parabolic case of a quite standard definition of constrained viscosity solution for HJB elliptic equations arising in optimal control problems with infinite time horizon and state constraints. In particular the condition of viscosity subsolution on $\partial^*\mathcal{C}$ plays the role of a boundary condition. This boundary condition was introduced by Soner in [Soner; 1986]

[2] The test functions ψ on which testing the HJB equation are in this case more than in the case of the definitions 1.2.22.

in the deterministic case. In the stochastic case it was used by Katsoulakis in [Katsoulakis; 1994], Zariphopoulou in [Zariphopoulou; 1994] and Ishii & Loreti in [Ishii, Loreti; 2002]. For the study of viscosity solutions of second order fully nonlinear equations with boundary conditions see [Ishii, Lions; 1990]. In particular in [Soner; 1986] and [Ishii, Loreti; 2002] this boundary condition turns out to be strong enough to guarantee, under a cone-like condition for the state equation at the boundary (see assumption (A4) in [Ishii, Loreti; 2002]), the uniqueness for the solution to the HJB equation. The natural version of this cone-like condition in the parabolic context holds true in our case when $\beta < r$. However, in the cited references the optimal control problem is time-homogeneous and over an infinite time-horizon, so that the associated PDE problem is elliptic. Our problem is instead strongly time-dependent, because both the state equation and the state constraint depend on time, and this leads to a parabolic PDE problem, so that we cannot use directly such results. Although we believe that they can be adapted to our parabolic case, we do not analyze the uniqueness topic in the case $\beta < r$, since we consider this secondary for our aim. Instead we treat the case $\beta = r$: in this case the fact that the boundary is absorbing yields a Dirichlet-type boundary condition and we can use the techniques and some results contained [Fleming, Soner; 1993].

So let $\beta = r$; as we have shown in Remark 1.2.6, in this case the boundary is absorbing and the only admissible strategy for the initial point $(t, l(t))$, $t \in [0, T)$, is the null one; thus in this case the value function is explicitely computable on the lateral boundary $\partial^* \mathcal{C}$, i.e. $V(t, l(t)) = g(t)$, where $g : [0, T] \to \mathbb{R}$ is a known function.

Notice that, if u, f are bounded, then V is obviously bounded. Since this assumption simplifies the study, we will assume it. Taking into account Proposition 1.2.9, it is straightforward to show that in this case V is uniformly continuous.

Theorem 1.2.26. *Let $\beta = r$, let u be Lipschitz continuous and let u, f be bounded. Then the value function V is the unique bounded and uniformly continuous viscosity solution to (1.32) on Int$^*(\mathcal{C})$ which satisfies the boundary condition $v(s, l(s)) = g(s)$, $s \in [0, T]$.*

Proof. We will give two proofs of the Theorem. We call main proof the first one, because it is suitable to be generalized to the case of sublinear coefficients for the state equation for the control problem. We call alternative proof the second one, which is strongly related to the linear structure of the state equation of the control problem. We refer to Remark 1.2.27 for a more detailed discussion.

Main proof.

Step 1. First we transform the equation in order to eliminate the term corresponding to the drift of the state equation and make nicer (not time-dependent) the constraint. So let v be a bounded and uniformly continuous viscosity solution to (1.32) on $\mathrm{Int}^*(\mathcal{C})$ and define

$$w : [0,T] \times [l_0, +\infty) \to \mathbb{R}, \quad w(s,x) := v\left(s, h(s,x)\right),$$

where

$$h(s,x) := xe^{rs} + \int_0^s kt\, e^{r(s-t)} dt.$$

It is straightforward to prove, taking into account that $l'(s) = rl(s) + ks$, that v is a bounded and uniformly continuous viscosity solution to (1.32) on $\mathrm{Int}^*(\mathcal{C})$ with lateral boundary condition $v(s,l(s)) = g(s)$, $s \in [0,T]$, and terminal boundary condition $v(T,x) = f(x)$, $x \geq l(T)$, if and only if w is a bounded viscosity solution on $[0,T) \times (l_0, +\infty)$, with lateral boundary condition $w(s,l_0) = g(s)$, $s \in [0,T]$, and terminal boundary condition $w(T,x) = f\left(h(T,x)\right)$, $x \geq l_0$, of the equation

$$-u_s(s,x) - \tilde{\mathcal{H}}(s,x,u_x(s,x),u_{xx}(s,x)) = 0, \tag{1.38}$$

where

$$\tilde{\mathcal{H}}(s,x,p,Q) := e^{-\rho s} U\left(s, h(s,x)\right) + \mathcal{H}^0(s,x,p,Q),$$
$$s \in [0,T), x \in (l_0, +\infty),\ p,\ Q \in \mathbb{R}.$$

Therefore uniqueness of bounded and uniformly continuous viscosity solutions for (1.32) on $\mathrm{Int}^*(\mathcal{C})$ holds with lateral boundary condition $v(s,l(s)) = g(s)$, $s \in [0,T]$, and terminal boundary condition $v(T,x) = f(x)$, $x \geq l(T)$, if and only if uniqueness of bounded and uniformly continuous viscosity solutions for (1.38) on $[0,T) \times (l_0, +\infty)$ holds with lateral boundary condition $w(s,l_0) = g(s)$, $s \in [0,T]$, and terminal boundary condition $w(T,x) = f\left(h(T,x)\right)$, $x \geq l_0$.

Step 2. Set

$$O := [0,T) \times (l_0, +\infty), \quad \partial^* O := \big\{(s,l_0) \mid s \in [0,T)\big\} \ \cup\ \big\{(T,x) \mid x \in [l_0, +\infty)\big\}.$$

We want to prove that, if w, $\tilde{w}$ are respectively a viscosity subsolution and a viscosity supersolution to (1.38) on O such that

$$\sup_{\partial^* O} (w - \tilde{w}) \leq 0, \tag{1.39}$$

then

$$\sup_{\bar{O}} (w - \tilde{w}) \leq 0,$$

that is enough for proving the claim. So we suppose on the countrary that

$$\sup_{\bar{O}} (w - \tilde{w}) \geq 6\delta > 0 \tag{1.40}$$

and prove that this leads to a contradiction. Thanks to (1.40) and to the continuity of w, $\tilde{w}$, we can find $(\bar{t}, \bar{x}) \in (0,T) \times (l_0, +\infty)$ such that

$$w(\bar{s}, \bar{x}) - \tilde{w}(\bar{s}, \bar{x}) \geq 5\delta. \tag{1.41}$$

Step 3. Due to the assumption on u, we get that the function

$$(s,x) \mapsto e^{-\rho s} U(s, h(s,x))$$

is Lipschitz continuous with respect to x uniformly in $s \in [0,T]$. Therefore it is straightforward to check that a result like Lemma V.7.1 in [Fleming, Soner; 1993] holds true in our case (with the unbounded domain), due to the assumption on u and to the linear structure of the drift coefficient and the diffusion coefficient in our equation. Indeed we could show, following the proof of Lemma V.7.1 in [Fleming, Soner; 1993], that there exists $C > 0$ such that

$$\begin{aligned} \Big| \tilde{\mathcal{H}}\Big(s, y, \alpha(x-y) - \frac{\varepsilon_0}{x+y+1}, B\Big) - \tilde{\mathcal{H}}\Big(s, x, \alpha(x-y) + \frac{\varepsilon_0}{x+y+1}, A\Big) \Big| \\ \leq C \left[\alpha |x-y|^2 + |x-y| + \varepsilon_0 + \varepsilon_0^2/\alpha \right], \end{aligned} \tag{1.42}$$

for every $(s,x,y) \in (0,T) \times (l_0, +\infty)^2$, $\alpha > 0$, $\varepsilon_0 > 0$, and $A, B \in \mathbb{R}$ such that

$$\begin{pmatrix} A & 0 \\ 0 & -B \end{pmatrix} \leq 3\alpha \begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix} + \frac{\varepsilon_0 + \varepsilon_0^2/\alpha}{(x+y+1)^2} \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix},$$

in the sense of the usual partial order in the space of simmetrix 2×2 matrices $\mathcal{S}(\mathbb{R}^2)$.

Step 4. Let us define, for $\rho_0 > 0$,

$$w^{\rho_0}(s,x) := w(s,x) - \frac{\rho_0}{s}, \qquad (s,x) \in (0,T] \times [l_0, +\infty).$$

Since

$$\frac{d}{ds}\left(-\frac{\rho_0}{s}\right) = \rho_0 / s^2 > 0,$$

we get that w^{ρ_0} is a viscosity solution to (1.38) on $(0,T) \times (l_0, +\infty)$.

Step 5. For $\alpha, \varepsilon_0, \rho_0, \beta_0 > 0$, let us consider the function

$$\Phi(s,x,y) := w^{\rho_0}(s,x) - \tilde{w}(s,y) - \frac{\alpha}{2}|x-y|^2 - \varepsilon \log(x+y+1) + \beta_0 (s-T),$$

defined for $(s, x, y) \in (0, T] \times [l_0, +\infty)^2$. Since $w, \tilde{w}$ are bounded, for any $\alpha > 0$, $\beta_0 > 0$, $\varepsilon_0 > 0$, $\rho_0 > 0$ there exists a maximum point $(\hat{s}_\alpha, \hat{x}_\alpha, \hat{y}_\alpha) \in (0, T] \times [l_0, +\infty)^2$ for Φ on $(0, T] \times [l_0, +\infty)^2$. We claim that there exist $\alpha^*, \beta_0, \varepsilon^*, \rho_0 > 0$ such that

$$(\hat{s}_\alpha, \hat{x}_\alpha, \hat{y}_\alpha) \in (0, T) \times (l_0, +\infty) \times (l_0, +\infty), \ \ \forall \alpha \geq \alpha^*, \ \forall \varepsilon_0 \leq \varepsilon^*. \tag{1.43}$$

Take

$$\beta_0 = \delta/T, \ \ \varepsilon^* = \delta/2\bar{x}, \ \ \rho_0 = \delta/\bar{s},$$

and let $(\bar{s}, \bar{x}) \in (0, T) \times (l_0, +\infty)$ be the point verifying (1.41); then we have, for any $\alpha > 0$,

$$\Phi(\bar{s}, \bar{x}, \bar{x}) \geq 2\delta. \tag{1.44}$$

By uniform continuity of $\tilde{w}, w$, there exists $\eta > 0$ such that, for every $s \in [0, T]$, $x, y \in [l_0, +\infty)$,

$$|x - y|^2 < \eta \implies |\tilde{w}(s, x) - \tilde{w}(s, y)| < \delta, \ \ |w(s, x) - w(s, y)| < \delta.$$

Let $K = \|\tilde{w}\|_\infty + \|w\|_\infty$, $\alpha_0 = \frac{2K}{\eta}$. Let $\alpha \geq \alpha_0$; then, for any $(s, x, y) \in (0, T] \times [l_0, +\infty)^2$ such that $|x - y|^2 \geq \eta$, we have

$$\Phi(s, x, y) = w^{\rho_0}(s, x) - \tilde{w}(s, y) - \frac{\alpha}{2}|x - y|^2 - (|x| + |y|) + \beta(s - T) \leq 0;$$

therefore, due to (1.44), we must have $|\hat{x}_\alpha - \hat{y}_\alpha| < \eta$. Now, if $\hat{s}_\alpha = T$, taking into account (1.39) we have

$$\begin{aligned} \Phi(\hat{s}_\alpha, \hat{x}_\alpha, \hat{y}_\alpha) &\leq w(T, \hat{x}_\alpha) - \tilde{w}(T, \hat{y}_\alpha) \\ &\leq w(T, \hat{x}_\alpha) - \tilde{w}(T, \hat{x}_\alpha) + \delta \leq \delta, \end{aligned}$$

which contradicts (1.44). Then suppose that $\hat{s}_\alpha \in (0, T)$ and that $\hat{x}_\alpha = l_0$. In the same way we have

$$\begin{aligned} \Phi(\hat{s}_\alpha, l_0, \hat{y}_\alpha) &\leq w(T, \hat{x}_\alpha) - \tilde{w}(T, \hat{y}_\alpha) \\ &\leq w(\hat{s}_\alpha, l_0) - \tilde{w}(\hat{s}_\alpha, l_0) + \delta \leq \delta, \end{aligned}$$

which again contradicts (1.44). We get the same contradiction if we suppose $\hat{y}_\alpha = l_0$, so that we can conclude that (1.43) holds for suitable $\alpha^*, \beta_0, \varepsilon^*, \rho_0 > 0$.

Step 6. Let $\alpha^*, \beta_0, \varepsilon^*, \rho_0 > 0$ as above and take

$$\varepsilon_0 = \min\left\{\varepsilon^*, \frac{\beta_0}{2C}\right\},$$

where C is the constant appearing in (1.42). Let $\alpha \geq \alpha_0$, and consider the function

$$\varphi(s, x, y) := \frac{\alpha}{2}|x - y|^2 + \varepsilon_0(x + y) - \beta_0(s - T), \quad (s, x, y) \in (0, T) \times (l_0, +\infty)^2.$$

Since

$$w^{\rho_0}(\hat{s}_\alpha,\hat{x}_\alpha)-\tilde{w}(\hat{s}_\alpha,\hat{y}_\alpha)-\varphi(\hat{s}_\alpha,\hat{x}_\alpha,\hat{y}_\alpha)=\max_{[0,T]\times[l_0,+\infty)^2}\Big(w^{\rho_0}(s,x)-\tilde{w}(s,y)-\varphi(s,x,y)\Big),$$

with $(\hat{s}_\alpha,\hat{x}_\alpha,\hat{y}_\alpha)\in(0,T)\times(l_0,+\infty)^2$, by Crandall-Hishii maximum principle (see Theorem V.6.1 in [Fleming, Soner; 1993]), there exist $a,b\in\mathbb{R}$, $A,B\in\mathbb{R}$ such that

$$\begin{cases} b-a=\beta_0,\\ (a,\alpha(\hat{x}_\alpha-\hat{y}_\alpha)+\varepsilon_0,A)\in cD^{+(1,2)}w^{\rho_0}(\hat{s}_\alpha,\hat{x}_\alpha),\\ (b,\alpha(\hat{x}_\alpha-\hat{y}_\alpha)-\varepsilon_0,B)\in cD^{-(1,2)}\tilde{w}(\hat{s}_\alpha,\hat{y}_\alpha),\\ \begin{pmatrix} A & 0\\ 0 & -B\end{pmatrix}\leq 3\alpha\begin{pmatrix} 1 & -1\\ -1 & 1\end{pmatrix}+\dfrac{\varepsilon_0+\varepsilon_0^2/\alpha}{(x+y+1)^2}\begin{pmatrix} 1 & 1\\ 1 & 1\end{pmatrix}, \end{cases}$$

where the sets $cD^{+(1,2)}w^{\rho_0}(\hat{s}_\alpha,\hat{x}_\alpha)$, $cD^{-(1,2)}\tilde{w}(\hat{s}_\alpha,\hat{x}_\alpha)$ are defined as in Definition V.4.2 in [Fleming, Soner; 1993]. Viscosity properties of w^{ρ_0} and $\tilde{w}$ yield (see Section V.4 in [Fleming, Soner; 1993])

$$-a-\tilde{\mathcal{H}}\left(\hat{s}_\alpha,\hat{x}^\alpha,\alpha(\hat{x}_\alpha-\hat{y}_\alpha)-\frac{\varepsilon_0}{\hat{x}_\alpha+\hat{y}_\alpha+1},A\right)\leq 0$$

and

$$-b-\tilde{\mathcal{H}}\left(\hat{s}_\alpha,\hat{y}^\alpha,\alpha(\hat{x}_\alpha-\hat{y}_\alpha)-\frac{\varepsilon_0}{\hat{x}_\alpha+\hat{y}_\alpha+1},B\right)\geq 0.$$

Subtracting and taking into account (1.42) we get

$$\begin{aligned}\beta_0\leq\tilde{\mathcal{H}}&\left(\hat{s}_\alpha,\hat{y}_\alpha,\alpha(\hat{x}_\alpha-\hat{y}_\alpha)-\frac{\varepsilon_0}{\hat{x}_\alpha+\hat{y}_\alpha+1},B\right)\\ &-\tilde{\mathcal{H}}\left(\hat{s}_\alpha,\hat{x}_\alpha,\alpha(\hat{x}_\alpha-\hat{y}_\alpha)-\frac{\varepsilon_0}{\hat{x}_\alpha+\hat{y}_\alpha+1},A\right)\\ &\leq C\left[\alpha|\hat{x}_\alpha-\hat{y}_\alpha|^2+|\hat{x}_\alpha-\hat{y}_\alpha|+\varepsilon_0+\varepsilon_0^2/\alpha\right].\end{aligned}\tag{1.45}$$

Step 7. In the same setting of the previous step, set

$$h(r):=\sup\{|\tilde{w}(s,x)-\tilde{w}(s,y)|\mid s\in[0,T],\ x,y\in[l_0,+\infty),\ |x-y|^2\leq r\};$$

of course h is increasing and, by uniform continuity of $\tilde{w}$, $\lim_{r\downarrow 0}h(r)=0$; moreover, since $\tilde{w}$ is bounded, also h is bounded by a constant C_h. Then for any $x,y\in[l_0,+\infty)$ we have

$$|\tilde{w}(s,x)-\tilde{w}(s,y)|\leq h(|x-y|^2).$$

Suppose $\hat{x}_\alpha \leq \hat{y}_\alpha$; since $(\hat{s}_\alpha, \hat{x}_\alpha, \hat{y}_\alpha)$ maximizes Φ over $(0,T] \times [l_0, +\infty)^2$, we have

$$\begin{aligned}\Phi(\hat{s}_\alpha, \hat{x}_\alpha, \hat{y}_\alpha) &= w^{\rho_0}(\hat{s}_\alpha, \hat{x}_\alpha) - \tilde{w}(\hat{s}_\alpha, \hat{y}_\alpha) - \frac{\alpha}{2}|\hat{x}_\alpha - \hat{y}_\alpha|^2 \\ &\quad -\varepsilon_0 \log(\hat{x}_\alpha + \hat{y}_\alpha + 1) + \beta_0(\hat{t}_\alpha - T) \\ &\geq w^{\rho_0}(\hat{s}_\alpha, \hat{x}_\alpha) - \tilde{w}(\hat{s}_\alpha, \hat{x}_\alpha) - \varepsilon_0 \log(2\hat{x}_\alpha + 1) + \beta_0(\hat{s}_\alpha - T),\end{aligned}$$

so that

$$\alpha|\hat{x}_\alpha - \hat{y}_\alpha|^2 \leq 2|\tilde{w}(\hat{s}_\alpha, \hat{x}_\alpha) - \tilde{w}(\hat{s}_\alpha, \hat{y}_\alpha)| \leq 2h(|\hat{x}_\alpha - \hat{y}_\alpha|^2).$$

Similarly, if $\hat{x}_\alpha \leq \hat{y}_\alpha$, we have

$$\begin{aligned}\Phi(\hat{t}_\alpha, \hat{x}_\alpha, \hat{y}_\alpha) &= w^{\rho_0}(\hat{s}_\alpha, \hat{x}_\alpha) - \tilde{w}(\hat{s}_\alpha, \hat{y}_\alpha) - \frac{\alpha}{2}|\hat{x}_\alpha - \hat{y}_\alpha|^2 \\ &\quad -\varepsilon_0 \log(\hat{x}_\alpha + \hat{y}_\alpha + 1) + \beta_0(\hat{t}_\alpha - T) \\ &\geq w^{\rho_0}(\hat{s}_\alpha, \hat{y}_\alpha) - \tilde{w}(\hat{s}_\alpha, \hat{y}_\alpha) \\ &\quad -\varepsilon_0 \log(2\hat{y}_\alpha + 1) + \beta_0(\hat{s}_\alpha - T),\end{aligned}$$

getting again (and so for any $\hat{x}_\alpha$, $\hat{y}_\alpha$)

$$\alpha|\hat{x}_\alpha - \hat{y}_\alpha|^2 \leq 2|\tilde{w}(\hat{s}_\alpha, \hat{x}_\alpha) - \tilde{w}(\hat{s}_\alpha, \hat{y}_\alpha)| \leq 2h(|\hat{x}_\alpha - \hat{y}_\alpha|^2). \tag{1.46}$$

This implies

$$\alpha|\hat{x}_\alpha - \hat{y}_\alpha|^2 \leq C_h,$$

which togheter with (1.46) yields

$$\alpha|\hat{x}_\alpha - \hat{y}_\alpha|^2 + |\hat{x}_\alpha - \hat{y}_\alpha| \leq 2h(C_h/\alpha) + (C_h/\alpha)^{1/2}. \tag{1.47}$$

We can put (1.47) in (1.45) and pass to the limit for $\alpha \to +\infty$ getting

$$\beta_0 \leq C\varepsilon_0 \leq \beta_0/2.$$

Since $\beta_0 > 0$, this yields a contradiction and the proof is complete. □

Alternative proof.

Case 1. Let us suppose $l_0 > 0$. After the transformation of step 1 of the main proof, we operate a further transormation on the equation setting, for $s \in [0,T]$, $x \in [\log l_0, +\infty)$,

$$z(s,x) := w(s, e^x).$$

Then it is straightforward to show that w is a viscosity solution to (1.38) on $[0,T) \times (l_0, +\infty)$, with lateral boundary condition $w(s, l_0) = g(s)$, $s \in [0,T]$, and terminal boundary condition $w(T,x) = f(h(T,x))$, $x \geq l_0$, if and only if

z is a viscosity solution on $[0,T)\times(\log l_0,+\infty)$, with lateral boundary condition $z(s,\log l_0)=g(s)$, $s\in[0,T]$, and terminal boundary condition $z(T,x)=f\left(h(T,e^x)\right)$, $x\in[\log l_0,+\infty)$, of

$$-u_s(s,x)-\mathcal{H}_{log}(s,x,u_x(s,x),u_{xx}(s,x))=0, \tag{1.48}$$

where

$$\mathcal{H}_{log}(s,x,p,Q):=\sup_{\theta\in[0,1]}\left[e^{-\rho t}U\left(s,h(s,e^x)\right)+\frac{1}{2}\sigma^2\theta^2Q+\left(\sigma\lambda\theta-\frac{1}{2}\sigma^2\theta^2\right)p\right].$$

Thus uniqueness of bounded and continuous viscosity solutions holds for (1.38) on $[0,T)\times(l_0,+\infty)$, with lateral boundary condition $w(s,l_0)=g(s)$ and terminal boundary condition $w(T,x)=f(h(T,x))$, if and only uniqueness of bounded and continuous viscosity solutions holds for (1.48) on $[0,T)\times(\log l_0,+\infty)$, with lateral boundary condition $z(s,\log l_0)=g(s)$, $s\in[0,T]$, and terminal boundary condition $z(T,x)=f(h(T,e^x))$, $x\geq\log l_0$. The uniqueness for the latter problem can be proved with a slight modification of the proof of Theorem V.9.1 in [Fleming, Soner; 1993].

Case 2. Let us suppose $l_0=0$. In this case with the same transformation we get that w is a viscosity solution to (1.38) on $[0,T)\times(l_0,+\infty)$, with lateral boundary condition $w(s,l_0)=g(s)$, $s\in[0,T]$, and terminal boundary condition $w(T,x)=f(h(t,x))$, $x\geq l_0$, if and only if z is a viscosity solution on $[0,T)\times\mathbb{R}$, with terminal boundary condition $z(T,x)=f\left(h(T,e^x)\right)$, $x\in\mathbb{R}$, of

$$-u_s(s,x)-\mathcal{H}_{log}(s,x,u_x(s,x),u_{xx}(s,x))=0.$$

Uniqueness for the latter problem is a straight consequence of Theorem V.9.1 in [Fleming, Soner; 1993]. □

Remark 1.2.27. The main proof of the latter thorem basically follows the line of the proof of Theorem V.8.1 of [Fleming, Soner; 1993], but we stress that it was needed to adapt the argument to our case. This proof works also if the drift and the diffusion coefficient are replaced by generic coefficients having sublinear growth with respect to x.

About the alternative proof we stress that it is not possible to follow directly the line of the proof of Theorem V.9.1 of [Fleming, Soner; 1993]. Indeed that proof requires boundedness for the drift and the diffusion coefficients in the state equation and can be generalized at most to the case of strictly sublinear coefficient with respect to x, i.e. growing as x^α, $\alpha\in[0,1)$, which is not our case. The transformation done in the alternative proof basically consists to take the logarithm of the state variable, so that basically relies in the fact that

$$dX(t) = a\theta(t)X(t)dt + b\theta(t)X(t)dB(t) \implies \\ d\log(X(t)) = \frac{dX(t)}{X(t)} = \left(a\theta(t) - \frac{b^2}{2}\theta(t)^2\right)dt + b\theta(t)dB(t).$$

Hence we can define and work with the new state variable $Y(t) = \log(X(t))$ following the dynamics

$$dY(t) = d\log(X(t)) = \left(a\theta(t) - \frac{b^2}{2}\theta(t)^2\right)dt + b\theta(t)dB(t).$$

This simplification is strongly related to the linear structure of the state equation, so that the same technique cannot be replied if we consider generic coefficients having sublinear growth with respect to x.

Finally we stress that the lateral boundary condition disappears in the alternative proof when we transform the equation in the case $l_0 = 0$, that is the lateral boundary condition is redundant in this case. This is coherent with the fact that the state constraint is automatically satisfied in this case. $\square$

1.2.8 An example with explicit solution

In this subsection we show how the problem can be solved in closed form when some constraints on the parameters and a special form of u and f are considered. Let $\gamma \in (0,1)$; here we assume that

$$\begin{cases} (i)\ u(y) = \dfrac{y^\gamma}{\gamma}, \quad y \geq 0, \\ (ii)\ f(x) = \kappa\,\dfrac{(x-l(T))^\gamma}{\gamma}, \quad x \geq l(T), \ \ \kappa \geq 0, \\ (iii)\ \beta = r, \\ (iv)\ \lambda \leq \sigma(1-\gamma). \end{cases} \tag{1.49}$$

Following [Merton; 1969] and [Merton; 1971], we look for a solution to equation (1.32) of the following form

$$v(s,x) = C(s)e^{-\rho t}\frac{(x-l(s))^\gamma}{\gamma}, \ \ (s,x) \in \mathcal{C}.$$

We have, for $(t,x) \in \mathcal{C} \setminus \partial^* \mathcal{C}$ (by the symbols $v_s(0,x)$, $v_s(T,x)$ we respectively mean $v_s(0^+,x)$, $v_s(T^-,x)$,

$$\begin{cases} v_s(s,x) = -C(s)\rho e^{-\rho s}\left[\dfrac{(x-l(s))^\gamma}{\gamma} + l'(s)(x-l(s))^{\gamma-1}\right] + C'(s)e^{-\rho s}\dfrac{(x-l(s))^\gamma}{\gamma}, \\ \\ v_x(s,x) = C(s)e^{-\rho s}(x-l(s))^{\gamma-1}, \\ \\ v_{xx}(s,x) = C(s)e^{-\rho s}(\gamma-1)(x-l(s))^{\gamma-2}. \end{cases}$$

Note that, for $(s,x) \in \mathcal{C} \setminus \partial^*\mathcal{C}$,

$$-\frac{\lambda v_x(s,x)}{\sigma x v_{xx}(s,x)} = \frac{\lambda}{\sigma(1-\gamma)} \cdot \frac{(x-l(s))}{x} \le 1, \tag{1.50}$$

so that, by (1.49)-(iv), for $(s,x) \in \mathcal{C} \setminus \partial^*\mathcal{C}$,

$$\mathcal{H}^0(x, v_x(s,x), v_{xx}(s,x)) = -\frac{\lambda^2 v_x(s,x)^2}{2\, v_{xx}(s,x)} = Ce^{-\rho s}\frac{\lambda^2}{2(1-\gamma)} \cdot (x-l(s))^\gamma.$$

Putting the expressions for the derivatives of v into (1.32), we get, taking into account that $l'(s) = rl(s) + ks$,

$$\left[\frac{1}{\gamma}C'(s) + C(s)\left[\frac{\rho}{\gamma} - r - \frac{\lambda^2}{2(1-\gamma)}\right] - \frac{1}{\gamma}\right](x-l(s))^\gamma = 0.$$

Therefore, if $C(s)$ is the unique solution to the ordinary differential equation

$$\begin{cases} C'(s) + \left[\rho - \gamma r - \frac{\lambda^2\gamma}{2(1-\gamma)}\right] C(s) = 0, \\ C(T) = \kappa e^{\rho T}, \end{cases} \tag{1.51}$$

then

$$v(s,x) = C(s)e^{-\rho s}\frac{(x-l(s))^\gamma}{\gamma}, \quad (s,x) \in \mathcal{C}, \tag{1.52}$$

is a solution (in classical sense) to (1.32) on $\mathcal{C} \setminus \partial^*\mathcal{C}$. Moreover such v satisfies the lateral boundary condition

$$v(s, l(s)) = 0, \quad s \in [0,T]$$

and the terminal boundary condition

$$v(T,x) = f(x), \quad x \ge l(T).$$

Note that condition 1.49-(iv) guarantees that the maximum point in the Hamiltonian is smaller than 1, so the no borrowing constraint is never active: this allows to keep $\tilde{\mathcal{H}}^0$ in the form which is suitable to find the explicit solution. Indeed, when $\lambda > \sigma(1-\gamma)$ it is not difficult to see that $\tilde{V}(x) < v(x)$ for any $x > l$ using the fact that v is the value function of a problem with larger control set whose optimal trajectory is not admissible for our problem.

Lemma 1.2.28. *Let v be defined by* (1.52)*.*

(i) Let $(s,x) \in \mathcal{C} \setminus \partial^*\mathcal{C}$, let $\theta(\cdot) \in \Theta_{ad}(s,x)$ be such that $X(t; s, x, \theta(\cdot)) > l(t)$ almost surely for every $t \in [s,T]$ and set $X(\cdot) := X(\cdot; s, x, \theta(\cdot))$. Then the following fundamental identity holds:

$$v(s,x) = J(s,x;\theta(\cdot)) + \mathbb{E}\left[\int_s^T \Big(\mathcal{H}_{cv}\big(t, X(t), v_x(t,X(t)), v_{xx}(t,X(t)); \theta(t)\big) - \mathcal{H}\big(t, X(t), v_x(t,X(t)), v_{xx}(t,X(t))\big)\Big)\, dt\right]. \tag{1.53}$$

(ii) Let $(s,x) \in \mathcal{C}$ and $\theta(\cdot) \in \Theta_{ad}(s,x)$. Then

$$v(s,x) \geq J(s,x;\theta(\cdot)). \tag{1.54}$$

Proof. (i) Let $(s,x) \in \mathcal{C} \setminus \partial^*\mathcal{C}$ and let $\theta(\cdot) \in \Theta_{ad}(s,x)$ have the property that $X(t)$ never touches the lateral boundary as stated by the hypothesis of the first claim. Then we can apply the Dynkin formula to $X(\cdot)$ on the interval $[s,T]$ with the function v and taking into account the expression of $\mathcal{H}_{cv}$ getting

$$\begin{aligned}&\mathbb{E}\left[v(T,X(T)) - v(s,x)\right] = \\ &\mathbb{E}\left[\int_s^T \Big(v_s(t,X(t)) + \mathcal{H}_{cv}\big(t,X(t),v_x(t,X(t)),v_{xx}(t,X(t));\theta(t)\big) - e^{-\rho t}U(t,X(t))\Big)\,dt\right].\end{aligned}$$

Since v solves in classical sense the HJB equation (1.32) on $\mathcal{C} \setminus \partial^*\mathcal{C}$, we can write

$$\begin{aligned}E\left[v(T,X(T)) - v(s,x)\right] = E\Bigg[\int_s^T &\Big(\mathcal{H}_{cv}\big(t,X(t),v_x(t,X(t)),v_{xx}(t,X(t));\theta(t)\big) \\ &- e^{-\rho t}U(t,X(t)) - \mathcal{H}\big(t,X(t),v_x(t,X(t)),v_{xx}(t,X(t))\big)\Big)\,dt\Bigg].\end{aligned}$$

Taking into account that $v(T,x) = f(x)$ for $x \geq l(T)$ the equality above can be rewritten to get

$$\begin{aligned}v(s,x) = \mathbb{E}\Bigg[&\int_s^T e^{-\rho t}U(t,X(t))dt + f(X(T)) \\ &+ \int_s^T \Big(\mathcal{H}\big(t,X(t),v_x(t,X(t)),v_{xx}(t,X(t))\big) \\ &\qquad - \mathcal{H}_{cv}\big(t,X(t),v_x(t,X(t)),v_{xx}(t,X(t));\theta(t)\big)\Big)\,dt\Bigg],\end{aligned}$$

i.e. the desired identity (1.53).

(ii) First let $(s,x) \in \mathcal{C} \setminus \partial^*\mathcal{C}$ and $\theta(\cdot) \in \Theta_{ad}(s,x)$. Set again $X(\cdot) := X(\cdot;s,x,\theta(\cdot))$ and define the stopping time

$$\tau := \inf\{t \geq s \mid X(t) = l(t)\},$$

with the convention $\inf\emptyset = T$. Since $(s,x) \in \mathcal{C} \setminus \partial^*\mathcal{C}$, by continuity of trajectories of $X(\cdot)$ we have $\tau > s$. Therefore for any fixed $\varepsilon > 0$ we have $X(t) \in \mathcal{C} \setminus \partial^*\mathcal{C}$ for every $t \in [s, s \vee (\tau - \varepsilon)]$. Therefore we can argue as above in the interval $[s, s \vee (\tau - \varepsilon)]$ getting

$$v(s,x) = \mathbb{E}\left[\int_s^{s\vee(\tau-\varepsilon)} e^{-\rho t} U(t, X(t))dt + e^{-\rho(s\vee(\tau-\varepsilon))} v(X(s \vee (\tau - \varepsilon)) \right]$$
$$+ \mathbb{E}\left[\int_s^{s\vee(\tau-\varepsilon)} \Big(\mathcal{H}\big(t, X(t), v_x(t, X(t)), v_{xx}(t, X(t))\big) \right.$$
$$\left. - \mathcal{H}_{cv}\big(t, X(t), v_x(t, X(t)), v_{xx}(t, X(t)); \theta(t)\big) \Big) \, dt \right].$$

The first term of the right hand-side in the previous equality for $\varepsilon \to 0$ tends to

$$E\left[\int_s^{\tau} e^{-\rho t} U(t, X(t))dt + e^{-\rho\tau} v(\tau, X(\tau)) \right].$$

Therefore passing to the liminf (or limsup) the previous equality, since the second term of the right hand-side is always positive, we get

$$v(s,x) \geq E\left[\int_s^{\tau} e^{-\rho t} U(t, X(t))dt + e^{-\rho\tau} v\left(\tau, X(\tau)\right) \right]. \tag{1.55}$$

Due to Proposition 1.2.6-(ii), we have $\theta(t) = 0$ and $X(t) = l(t)$ for $t \in [\tau, T]$. Therefore, if $\tau > T$,

$$v(\tau, X(\tau)) = 0 = \int_{\tau}^{T} e^{-\rho t} U(t, X(t))dt + f(X(T)).$$

If $\tau = T$ we have as well

$$v(\tau, X(\tau)) = v(T, X(T)) = f(X(T)).$$

In definitive we can rewrite (1.55) as

$$v(s,x) \geq E\left[\int_s^{T} e^{-\rho t} U(t, X(t))dt + f(X(T)) \right],$$

getting (1.54).

Now let $(s,x) \in \partial^*\mathcal{C}$. By Proposition 1.2.6-(ii) we know that $\Theta_{ad}(s,x) = \{0\}$ and $X(t; s, x, 0) = l(t)$ for $t \in [s, T]$, so that $J(s, x; 0) = 0$. In particular, since $v(s,x) = v(s, l(s)) = 0$ we have (1.54) with the equality. □

We define the feedback map

$$G(t,x) := \begin{cases} -\dfrac{\lambda v_x(t,x)}{\sigma x v_{xx}(t,x)} = \dfrac{\lambda}{\sigma(1-\gamma)} \cdot \dfrac{x - l(t)}{x}, & \text{if } (t,x) \in \mathcal{C} \setminus \partial^*\mathcal{C}, \\ 0, & \text{if } (t,x) \in \partial^*\mathcal{C}, \end{cases} \tag{1.56}$$

and note that, thanks to (1.50), for every $(t, x) \in \mathcal{C} \setminus \partial^* \mathcal{C}$,

$$\mathcal{H}(x, v_x(t, x), v_{xx}(t, x)) = \mathcal{H}_{cv}(x, v_x(t, x), v_{xx}(t, x); G(t, x)). \tag{1.57}$$

Given $(s, x) \in \mathcal{C}$, the so-called closed-loop equation associated with this map is the stochastic differential equation

$$\begin{cases} dX(t) = \big[\left[r + \sigma\lambda G(t, X(t)) \right] X(t) + kt \big] dt + \sigma G(t, X(t)) X(t) dB^s(t), \\ X(s) = x, \end{cases}$$

i.e.

$$\begin{cases} dX(t) = \left[rX(t) + \dfrac{\lambda^2}{1-\gamma}(X(t) - l(t)) + kt \right] dt + \dfrac{\lambda}{1-\gamma}(X(t) - l(t)) dB^s(t), \\ X(s) = x. \end{cases} \tag{1.58}$$

Lemma 1.2.29. *For any $(s, x) \in \mathcal{C}$ there exists a unique process $X_G(\cdot; s, x)$ solution to the closed-loop equation* (1.58)*. Moreover $(t, X_G(t; s, x)) \in \mathcal{C}$ for every $t \in [s, T]$ and*

- *if $x > l(s)$, then $X_G(t; s, x) > l(t)$ almost surely for every $t \in [s, T]$;*
- *if $x = l(s)$, then $X_G(t; s, x) = l(t)$ almost surely for every $t \in [s, T]$; in particular $X_G(\cdot; s, x)$ is deterministic.*

Proof. Existence and uniqueness of solutions follow by the standard theory, since the coefficients are Lipschitz continuous.

Let $x > l(s)$, set $X_G(\cdot) := X_G(\cdot; s, x)$ and define the process $Y(t) := X_G(t) - l(t)$, $t \in [s, T]$. Then, taking into account that $dl(t) = [rl(t) + kt]dt$, we see that $Y(\cdot)$ solves

$$\begin{cases} dY(t) = \left[r + \dfrac{\lambda^2}{1-\gamma} \right] Y(t)\, dt + \dfrac{\lambda}{1-\gamma}\, Y(t) dB^s(t), \\ Y(s) = x - l(s) > 0, \end{cases}$$

i.e. $Y(\cdot)$ is a geometric Brownian motion with strictly positive initial point. Therefore $Y(t) > 0$ almost surely for $t \in [s, T]$, i.e. $X(t) > l(t)$ almost surely for $t \in [s, T]$.

Let $x = l(s)$. Arguing as above we see that in this case it has to be $Y(t) = 0$ almost surely for $t \in [s, T]$, i.e. $X(t) = l(t)$ almost surely for $t \in [s, T]$. □

Theorem 1.2.30 (Verification)**.** *Let $(s, x) \in \mathcal{C}$. Then $V(s, x) = v(s, x)$, where v is the function defined in* (1.52)*, and the feedback strategy*

$$\theta_G^{s,x}(t) := \frac{\lambda}{\sigma(1-\gamma)} \cdot \frac{X_G(t; s, x) - l(t)}{X_G(t; s, x)}, \quad t \in [s, T], \tag{1.59}$$

is the unique optimal strategy optimal starting from (s, x).

Proof. Let $(s, x) \in \mathcal{C} \setminus \partial^* \mathcal{C}$. By Lemma 1.2.28-(ii) we know that, for any $\theta(\cdot) \in \Theta_{ad}(s, x)$, we have

$$v(s, x) \geq J(s, x; \theta(\cdot));$$

this shows that $V(s, x) \leq v(s, x)$. On the other hand, take feedback strategy $\theta_G^{s,x}(\cdot)$ defined in (1.59); we know by Lemma 1.2.29 that $X(t; s, x, \theta_G^{s,x}(\cdot)) > l(t)$ for every $t \in [s, T]$; therefore we can write the fundamental identity (1.53) for $X(t; s, x, \theta_G^{s,x}(\cdot))$; by (1.57) we get

$$v(s, x) = J(s, x; \theta_G^{s,x}(\cdot)) \leq V(s, x).$$

Therefore we see that $\theta_G^{s,x}(t)$ is optimal. On the other hand, due to the fact that $v = V$, we see from (1.53) that any other optimal strategy $\theta(\cdot) \in \Theta_{ad}(s, x)$ must satisfy $\theta(\cdot) = G(\cdot; X(\cdot; s, x, \theta(\cdot))$, i.e. must be in closed loop form. Due to the uniqueness of solutions of the closed loop equation, we get uniqueness of optimal strategies.

Now let $(s, x) \in \partial^* \mathcal{C}$. By Proposition 1.2.6-(ii) in this case the only admissible strategy is $\theta(\cdot) = 0$ and moreover we have

$$V(s, x) = J(s, x; 0) = v(s, x).$$

By Lemma 1.2.29 and by (1.59) we have $\theta_G^{s,x}(\cdot) \equiv 0$, which gives the claim also in this case. $\square$

1.3 The stationary phase

The object of this section is the analysis of the stationary phase, corresponding to the time interval $[T, +\infty)$. We will solve completely the problem showing that the value function is a classical solution of the HJB equation (passing through the viscosity approach) and a verification theorem yielding an optimal feedback strategy. Moreover we will provide an example with explicit solution.

The initial time for the optimization problem is here $t = T$. Again it also makes sense, in order to apply the dynamic programming techniques, to look to a pension fund that is already running after a given amount of time $s \geq T$, in order to establish a decision policy from s on.

On the probability space of Section 1.1 let $(\mathcal{F}_t^s)_{t \geq s}$ be the completion of the filtration generated by the process $(B^s(t))_{t \geq s} := (B(t) - B(s))_{t \geq s}$; the control process $(\theta(t))_{t \geq s}$ is a $(\mathcal{F}_t^s)$-progressively measurable process with values in $[0, 1]$.

Let us set an initial time $s \geq T$ and a given amount of wealth x at time s. In the interval $[s,+\infty)$ the state equation becomes, according to (1.2) and with the hypotheses just stated on the contribution term,

$$\begin{cases} dX(t) = [r + \sigma\lambda\theta(t)]\, X(t)\, dt - q\, dt + \sigma\theta(t)X(t)dB^s(t), \quad t \geq s, \\ X\,(s) = x, \end{cases} \tag{1.60}$$

where $q := g - c$.

Theorem 1.3.1. *For any $(\mathcal{F}_t^s)_{t\geq s}$-progressively measurable $[0,1]$-valued process $\theta(\cdot)$*

- *equation* (1.60) *admits on the filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t^s)_{t\geq s}, \mathbb{P})$, a unique strong solution;*
- *this solution belongs to the space $C_{\mathcal{P}}([s,+\infty); L^p(\Omega,\mathbb{P}))$ of the p-mean continuous progressively measurable processes for any $p \in [1,+\infty)$.*

Proof. See Theorem 6.16, Chapter 1, of [Yong, Zhou; 1999] or Section 5.6.C of [Karatzas, Shreve; 1991]. □

We denote the unique strong solution to (1.60) by $X\,(t; s, x, \theta(\cdot))$.

1.3.1 The optimization problem

In this stationary phase we study an infinite horizon optimization problem in the interval $[T,+\infty)$ related to an objective functional with this form:

$$\mathbb{E}\left[\int_T^{+\infty} e^{-\rho(t-T)}\tilde{U}\,(X\,(t))\, dt\right]. \tag{1.61}$$

Here $\rho > 0$ is the manager's individual discount factor and $\tilde{U}$ is the manager's utility function. So, according with the considerations of Subsection 1.1.5, this criterion takes into account the manager's point of view. For $x \geq l = l(T)$, the problem is to maximize, over the set of the admissible strategies, i.e. the strategies keeping the state variable above l (see (1.63) for a formal definition), the functional (1.61).

We assume that the utility function $\tilde{U}$ satisfies the following assumptions:

Hypothesis 1.3.2. .

(i) $\tilde{U} : [l,+\infty) \to \mathbb{R}\cup\{-\infty\}$ belongs to the class $C^2\,((l,+\infty)\,;\mathbb{R})$ is increasing and $\tilde{U}' > 0$, $\tilde{U}'' < 0$.

(ii) For given $C > 0$ and $\beta \in [0, 1)$ we have $\tilde{U}(x) \le C\left(1 + x^{\beta}\right)$, where

$$\rho > \beta r + \frac{\lambda^2}{2} \cdot \frac{\beta}{1 - \beta}. \tag{1.62}$$

□

Note that

- The utility function is defined where the wealth process $X(\cdot)$ must live.

- We do not assume that Inada like conditions, i.e.

$$\lim_{x \to l^+} \tilde{U}'(x) = +\infty, \quad \lim_{x \to +\infty} \tilde{U}'(x) = 0,$$

  hold for the utility function $\tilde{U}$. We will do it only in the special case of Subsection 1.3.7.

- All utility functions of the form $\tilde{U}(x) = \frac{(x - x_0)^{\gamma}}{\gamma}$, for $x_0 \le l$ and $\gamma \in (-\infty, 0) \cup (0, 1)$ always satisfy Hypothesis 1.3.2-(i). In the case when $x_0 = l$ they also satisfy the Inada like conditions mentioned above. This fact will be used later in Subsection 1.3.7 to give examples.

- Hypothesis 1.3.2-(ii) guarantees the finiteness of the value function, as is proved in Proposition 1.3.11. A more general condition could be given on the line of what is done in [Karatzas, Lehoczky, Sethi, Shreve; 1986], Section 2. As we will see in Subsection 1.3.7 in the case of $\tilde{U}(x) = \frac{(x-l)^{\gamma}}{\gamma}$, condition (1.62) may be sharp and even cases with $\rho \le 0$ may be treated (when $\gamma < 0$).

1.3.2 The set of admissible strategies

In this stationary phase the set of the admissible strategies, for $s \ge T$, $x \ge l$, is given by

$$\begin{aligned}\tilde{\Theta}_{ad}(s, x) := \Big\{ &\theta : [s, +\infty) \times \Omega \to [0, 1] \\ &\text{progressively measurable w.r.t. } (\mathcal{F}_t^s)_{t \ge s} \mid X(t; s, x, \theta(\cdot)) \ge l,\ t \ge s \Big\}. \end{aligned} \tag{1.63}$$

Remark 1.3.3. The set $\tilde{\Theta}_{ad}(s, x)$ is independent of $s \ge T$ in the following sense. By Theorem 2.10, Chapter 1, of [Yong, Zhou; 1999], there is a one-to-one correspondence between the strategies starting from T and the strategies starting from $s > T$; to make more clear this point, consider the measurable space $(C([T, +\infty); \mathbb{R}), \mathcal{B}(C([T, +\infty); \mathbb{R})))$, with the filtration $(\mathcal{B}_t(C([T, +\infty); \mathbb{R})))_{t \ge T}$

defined in the following way: $\mathcal{B}_t(C([T,+\infty);\mathbb{R}))$ is the σ-algebra on the space $C([T,+\infty);\mathbb{R})$ induced by the projection

$$\begin{array}{rccc}\pi: & C([T,+\infty);\mathbb{R}) & \longrightarrow & (C([T,t];\mathbb{R}),\mathcal{B}(C([T,t];\mathbb{R}))) \\ & \zeta(\cdot) & \longmapsto & \zeta(\cdot)|_{[T,t]},\end{array}$$

i.e. the smallest σ-algebra which makes π measurable; intuitively a measurable application with respect to $\mathcal{B}_t(C([T,+\infty);\mathbb{R}))$ is an application which does not distinguish between two functions of $C[T,+\infty)$ which coincide on $[T,t]$. If $(\theta_T(t))_{t\geq T}$ is a strategy starting from T, then there exists an adapted process ψ on $[T,+\infty)\times C([T,+\infty);\mathbb{R})$ such that

$$\theta_T(t)(\cdot)=\psi(t,B^T(\cdot)),\quad t\geq T;$$

then the shifted strategy

$$\theta_s(t)(\cdot)=\psi(t-s+T,B^s(\cdot)),\quad t\geq s,$$

starts from s. Since the state equation is homogeneous on time, it is easy to check that we have $\theta_T\in\Theta_{ad}(T,x)$ if and only if $\theta_s\in\Theta_{ad}(s,x)$. □

Now we give a lemma on the nonemptiness of the set of admissible strategies.

Lemma 1.3.4. *Given any $s\geq T$, $x\geq l$ the set of admissible strategies $\tilde{\Theta}_{ad}(s,x)$ is nonempty if and only if the control $\theta(\cdot)\equiv 0$ is admissible. This happens if and only if*

$$x\geq\frac{q}{r}. \tag{1.64}$$

In particular the set of admissible strategies $\tilde{\Theta}_{ad}(s,x)$ is nonempty for every $s\geq T$, $x\geq l$ if and only if

$$l\geq\frac{q}{r}. \tag{1.65}$$

Proof. Thanks to Remark 1.3.3 we can take without loss of generality $s=T$. Let $x\geq l$, it is clear that, if $\theta(\cdot)\equiv 0$ is admissible at (T,x), then $\Theta(T,x)$ is nonempty. We prove the opposite. Assume that $\tilde{\Theta}_{ad}(T,x)$ is nonempty; let $\theta(\cdot)$ be an admissible strategy and set $X(t):=X(t;T,x,\theta(\cdot))$. By Girsanov's Theorem A.1.1, under the probability $\tilde{\mathbb{P}}=\exp\left(-\lambda B^T(t)-\frac{1}{2}\lambda^2(t-T)\right)\cdot\mathbb{P}$ (depending on t, defined on $\mathcal{F}_t^T$ and equivalent to $\mathbb{P}$), the process $\tilde{B}^T(\cdot)=\lambda(\cdot-T)+B^T(\cdot)$ is a Brownian motion on $[T,t]$ and we have

$$X(t)=x+\int_T^t rX(\tau)\,d\tau-\int_T^t q\,d\tau+\int_T^t\sigma\theta(t)\,X(\tau)\,d\tilde{B}^T(\tau),\quad t\geq T. \tag{1.66}$$

Since $X \in C([s,T];L^p(\Omega,\mathbb{P}))$ for any $p \geq 1$, it holds also

$$\tilde{\mathbb{E}}\left[\int_T^t |X(\tau)|^2 d\tau\right] < +\infty, \quad \forall t \in [T,+\infty),$$

so that

$$\tilde{\mathbb{E}}\left[\int_T^t X(\tau) d\tau\right] = 0, \ \ \forall t \in [T,+\infty).$$

Thus we can pass (1.66) to the expectations getting

$$\tilde{\mathbb{E}}\left[X\left(t\right)\right] = x + \int_T^t r\tilde{\mathbb{E}}\left[X\left(\tau\right)\right] d\tau - \int_T^t q d\tau. \tag{1.67}$$

This implies that the deterministic function $g\left(\cdot\right) = \tilde{\mathbb{E}}\left[X\left(\cdot\right)\right]$ satisfies on $[T,t]$ the same ordinary differential equation as $X\left(\cdot;T,x,0\right)$. Since by assumption $X\left(\tau\right) \geq l$ almost surely with respect to $\mathbb{P}$ for any $\tau \in [T,t]$, it has to be also $X\left(t\right) \geq l$ almost surely with respect to $\tilde{\mathbb{P}}$ for any $\tau \in [T,t]$; thus taking the expectations $\tilde{\mathbb{E}}$ under $\tilde{\mathbb{P}}$ we get $g(\cdot) = \tilde{\mathbb{E}}\left[X\left(\cdot\right)\right] \geq l$ on $[T,t]$. Therefore $X(\cdot\,;T,x,0) = g(\cdot) \geq l$ on $[T,t]$. The first claim is proved by the arbitrariness of t.

Let us prove the second claim. For $\theta(\cdot) \equiv 0$ the state equation (1.60) becomes the following deterministic equation

$$\begin{cases} dX(t) = \left(rX(t) - q\right)dt, \quad t \geq T, \\ X(T) = x \geq l. \end{cases} \tag{1.68}$$

From this equation it is then easy to see that $X(t;T,x,0) \geq l$, for any $t \geq T$, if and only if

$$rx - q \geq 0 \Longleftrightarrow x \geq \frac{q}{r}.$$

This gives the second statement. □

Remark 1.3.5. Lemma 1.3.4 substantially states that the good solvency level must be such that the return rl from it is greater than q, i.e. the balance between the contribution and the benefit rate. In other words the solvency level l must be above the present value $\frac{q}{r}$ of the perpetual annuity, which is obtained discounting at the instantaneous risk free rate r the balance between benefit and contribution rate q, i.e. the present value of the total outcomes, over the whole time horizon. This may remind what happens, in a different setting, in [Cadenillas, Sethi; 1997] and [Sethi, Taksar, Presman; 1992], where models with subsistence consumption are considered. In the case of stochastic interest rates and demographic risk this would be a stochastic constraint.

We observe that in the expression (1.7) for the solvency level all the quantities are given by the market except for l_0 and β which may be chosen by a supervisory authority. In other words the authority fix the liquidity l_0 needed to start a pension fund and the capitalization rate $\beta \leq r$ of the startup level l_0 and of the aggregate contribution flow that the fund must guarantee in the accumulation period. This choice should be always such that $l := l(T)$ satisfy (1.65) and may vary depending on the goals of the authority itself. For example high l_0 and β will force the fund manager to keep more prudential behaviours in order to avoid default. $\square$

From now on we will always assume that $\tilde{\Theta}_{ad}(T, x)$ is nonempty over all $x \in [l, +\infty)$, i.e. that (1.65) holds true. We will often divide the two cases $rl = q$ and $rl > q$, since, as shown in Proposition 1.3.6, they behave differently.

Proposition 1.3.6. *Let $s \geq T$.*

- *If $rl - q > 0$, then $X(t; s, x, 0) > l$ for all $t > T$. We stress this property saying that the boundary $\{l\} \times [T, +\infty)$ is reflecting.*
- *If $rl - q = 0$, then $\tilde{\Theta}_{ad}(s, l) = \{0\}$ and $X(t; T, l, 0) = l$ for all $t \geq T$. We stress this property saying that the boundary $\{l\} \times [T, +\infty)$ is absorbing.*

Proof. By 1.67, for any $\theta(\cdot) \in \tilde{\Theta}_{ad}(s, l)$ it must be

$$\tilde{\mathbb{E}}[X(t; s, x, \theta(\cdot))] = l, \quad \forall t \geq s,$$

so that we can conclude that $\tilde{\Theta}_{ad}(s, l)$ is made only by the null strategy. The other statements follow taking $\theta(\cdot) \equiv 0$ in the state equation. $\square$

1.3.3 The value function

For $s \geq T$, $x \geq l$, $\theta(\cdot) \in \tilde{\Theta}_{ad}(s, x)$, we define

$$\tilde{J}(s, x; \theta(\cdot)) := \mathbb{E}\left[\int_s^{+\infty} e^{-\rho(t-T)} \tilde{U}\left(X\left(t; s, x, \theta(\cdot)\right)\right) dt\right]. \tag{1.69}$$

We show that (1.69) is well-defined for every $s \geq T$, $x \geq l$, $\theta(\cdot) \in \tilde{\Theta}_{ad}(s, x)$:

Proposition 1.3.7. *Assume Hypothesis 1.3.2. The functional $\tilde{J}$ defined in (1.69) is well-defined, i.e. for any $s \geq T$, $x \geq l$, $\theta(\cdot) \in \tilde{\Theta}_{ad}(s, x)$ we have*

$$\mathbb{E}\left[\int_s^{+\infty} e^{-\rho(t-T)} \tilde{U}^{+}(X(t; s, x, \theta(\cdot))) dt\right] < +\infty.$$

Proof. Let $s \geq T$, $x \geq l$, $\theta(\cdot) \in \tilde{\Theta}(s, x)$ and set $X(t) = X(t; T, x, \theta(\cdot))$; by Hypothesis 1.3.2-(ii) we have

$$\mathbb{E}\left[\int_s^{+\infty} e^{-\rho(t-T)} \tilde{U}^{+}\left(X(t)\right) dt\right] \leq C \int_s^{+\infty} e^{-\rho(t-T)} \left(1 + \mathbb{E}\left[X(t)^{\beta}\right]\right) dt. \tag{1.70}$$

Let $t \geq s$; by Girsanov's Theorem A.1.1 under the probability

$$\tilde{\mathbb{P}} = \exp\left(-\lambda B^s(t) - \frac{1}{2}\lambda^2(t-s)\right) \cdot \mathbb{P}$$

(depending on t, defined on $\mathcal{F}_t^s$ and equivalent to $\mathbb{P}$), the process

$$\tilde{B}^s(\cdot) = \lambda(\cdot - s) + B^s(\cdot)$$

is a Brownian motion on $[s, t]$. So, we have

$$X(t) = x + \int_s^t rX(\tau)\,d\tau - \int_s^t q\,d\tau + \int_s^t \sigma\theta(t)\,X(\tau)\,d\tilde{B}^s(\tau).$$

As in the proof of Lemma 1.3.4 we can pass to the expectations and get

$$\tilde{\mathbb{E}}[X(t)] = x + \int_s^t r\tilde{\mathbb{E}}[X(\tau)]d\tau - \int_s^t q\,d\tau.$$

This shows that $\tilde{\mathbb{E}}[X(t)] \leq e^{r(t-s)}x$ for every $t \geq T$. Now

$$\mathbb{E}\left[X(t)^\beta\right] = \tilde{\mathbb{E}}\left[X(t)^\beta \exp\left(\lambda\tilde{B}^s(t) - \frac{1}{2}\lambda^2(t-s)\right)\right],$$

and by Hölder's inequality

$$\begin{aligned}
&\tilde{\mathbb{E}}\left[X(t)^\beta \exp\left(\lambda\tilde{B}^s(t) - \frac{1}{2}\lambda^2(t-s)\right)\right] \\
&\qquad \leq \left(\tilde{\mathbb{E}}\left[X(t)\right]\right)^\beta \cdot \left(\tilde{\mathbb{E}}\left[\exp\left(\lambda\tilde{B}^s(t) - \frac{1}{2}\lambda^2(t-s)\right)\frac{1}{1-\beta}\right]\right)^{1-\beta} \\
&\qquad\qquad = x^\beta e^{\left(\beta r + \frac{\beta}{1-\beta}\cdot\frac{\lambda^2}{2}\right)(t-s)} \leq x^\beta e^{\left(\beta r + \frac{\beta}{1-\beta}\cdot\frac{\lambda^2}{2}\right)(t-T)}.
\end{aligned}$$

The above inequalities, together with (1.62), imply that, for some $C' > 0$ independent of $s, x, \theta(\cdot)$,

$$\begin{aligned}
&\mathbb{E}\left[\int_s^{+\infty} e^{-\rho(t-T)}\tilde{U}^+(X(t))\,dt\right] \\
&\qquad \leq C\int_s^{+\infty} e^{-\rho(t-T)}\left(1 + x^\beta e^{\left(\beta r + \frac{\beta}{1-\beta}\cdot\frac{\lambda^2}{2}\right)(t-T)}\right)dt \\
&\qquad\qquad \leq C' e^{-\rho(s-T)}\left(1 + x^\beta\right),
\end{aligned}$$

which gives the claim. □

Due to the previous result from now on we suppose that Hypothesis 1.3.2 holds true. We can associate to the problem the value function defined by

$$\tilde{V}(s,x) := \sup_{\theta(\cdot)\in\tilde{\Theta}_{ad}(s,x)} \tilde{J}(s,x;\theta(\cdot)), \quad s \geq T, x \geq l. \tag{1.71}$$

The stochastic control problem consists in studying this function and, when possible, in finding an optimal control strategy in the sense of the following definition.

Definition 1.3.8. Let $s \geq T, x \geq l$.

(i) A control $\theta^*(\cdot) \in \Theta_{ad}(s,x)$ is called optimal for the couple (s,x) if

$$\tilde{J}(s,x;\theta^*(\cdot)) = \tilde{V}(s,x).$$

(ii) For fixed $\varepsilon > 0$, a control $\theta^\varepsilon(\cdot) \in \Theta_{ad}(s,x)$ is called ε-optimal for the couple (s,x) if

$$\tilde{J}(s,x;\theta^\varepsilon(\cdot)) \geq \tilde{V}(s,x) - \varepsilon.$$

□

Remark 1.3.9. The problem of Section 1.2 and the problem of the present section can be linked to form a unique optimization problem in the whole interval $[0,+\infty)$. Indeed we can imagine to want to optimize a functional of this form over the whole interval $[0,+\infty)$

$$\int_0^{+\infty} e^{-\rho t} U(t, X(t)) dt$$

Thanks to the dynamic programming principle this can be done by choosing as exit function f in (1.61) the value function of the optimization problem in the stationary phase. □

Since we are in the stationary case, using the properties of the set of admissible strategies (see Remark 1.3.3), we can prove the following.

Proposition 1.3.10. *We have the following dependence on time for the value function:*

$$\tilde{V}(s,x) = e^{-\rho(s-T)} \tilde{V}(T,x), \quad s \geq T, \ x \geq l. \tag{1.72}$$

Proof. We only sketch the proof, as the argument are quite standard. Since both the state equation and the set of the admissible strategies are autonomous for $s \geq T$ (see Remark 1.3.3), performing the change of variable $t' = t - (s-T)$ we get, by equality in law,

$$\begin{aligned}
\tilde{V}(s,x) &= \sup_{\theta(\cdot)\in\tilde{\Theta}_{ad}(s,x)} \mathbb{E}\left[\int_s^{+\infty} e^{-\rho(t-T)} \tilde{U}\left(X\left(t;s,x,\theta(\cdot)\right)\right) dt\right] \\
&= e^{-\rho(s-T)} \cdot \sup_{\theta(\cdot)\in\tilde{\Theta}_{ad}(T,x)} \mathbb{E}\left[\int_T^{+\infty} e^{-\rho(t'-T)} \tilde{U}\left(X\left(t';T,x,\theta(\cdot)\right)\right) dt'\right] \\
&= e^{-\rho(s-T)} \tilde{V}(T,x),
\end{aligned}$$

i.e. the claim. $\square$

Thanks to Proposition (1.72) we are reduced to study the function

$$x \mapsto \tilde{V}(T,x), \quad x \geq l.$$

With a slight abuse of notation we set

$$\tilde{\Theta}_{ad}(x) := \tilde{\Theta}_{ad}(T,x), \;\; \tilde{J}(x;\theta(\cdot)) := \tilde{J}(T,x;\theta(\cdot)), \;\; \tilde{V}(x) := \tilde{V}(T,x).$$

Similarly we will write $B(\cdot)$ for $B^T(\cdot)$. We give now a result about the finiteness of the value function.

Proposition 1.3.11. *We have $\tilde{V}(\cdot) > -\infty$ on $(l,+\infty)$. Moreover*

(i) when $rl = q$, we have $\tilde{V}(l) > -\infty$ if and only if $\tilde{U}(l) > -\infty$;

(ii) when $rl > q$, we have $\tilde{V}(l) > -\infty$ if and only if $\tilde{U}$ is integrable in a right neighborhood of l.

Finally there exists $K > 0$ such that $\tilde{V}(x) \leq K\left(1 + x^{\beta}\right)$ for every $x \geq l$, where β is given by Hypothesis 1.3.2-(ii).

Proof. *Estimates from below for $x > l$.* First of all we show that $\tilde{V}(\cdot) > -\infty$ on $(l,+\infty)$. Indeed, since the null strategy is always admissible we have, for every $x \geq l$,

$$\tilde{V}(x) \geq \tilde{J}(x;0) = \int_T^{+\infty} e^{-\rho(t-T)}\tilde{U}(X(t;T,x,0))\,dt.$$

But, recalling that $X(t;T,x,0)$ satisfies (1.68), we have

$$X(t;T,x,0) = e^{r(t-T)}\left[x - \frac{q}{r}\right] + \frac{q}{r}.$$

Since (1.65) holds, then $x - \frac{q}{r} > 0$, so $X(t;T,x,0) \geq x$ for every $t \geq T$ and

$$\tilde{V}(x) \geq \tilde{J}(x;0) \geq \int_T^{+\infty} e^{-\rho(t-T)}\tilde{U}(x)\,dt = \frac{\tilde{U}(x)}{\rho},$$

which gives the claim.

Estimates from below for $x = l$, case (i). The above argument also says that $\tilde{V}(l) > -\infty$ when $\tilde{U}(l) > -\infty$. Moreover, when $rl = q$ the only admissible strategy at $x = l$ is the null one that keeps the state in l at every time (Proposition 1.3.6); so when $\tilde{U}(l) = -\infty$ also $\tilde{V}(l) = -\infty$.

Estimates from below for $x = l$, case (ii). Assume now that $rl > q$, so

$$\tilde{V}(l) \geq \tilde{J}(l;0) = \int_T^{+\infty} e^{-\rho(t-T)}\tilde{U}\left(e^{r(t-T)}\left[l - \frac{q}{r}\right] + \frac{q}{r}\right)dt. \tag{1.73}$$

Setting $z = e^{r(t-T)}\left[l - \frac{q}{r}\right] + \frac{q}{r}$ the above integral becomes equal to a given constant multiplied by

$$\int_l^{+\infty} \left(z - \frac{q}{r}\right)^{-\frac{\rho}{r}-1} \tilde{U}(z)\,dz$$

which is not $-\infty$. Indeed, taken any $z_0 > l$, the integrability of $\tilde{U}$ in a neighborhood of l says that $\int_l^{z_0} \left(z - \frac{q}{r}\right)^{-\frac{\rho}{r}-1} \tilde{U}(z)\,dz$ is finite, while for the term $\int_{z_0}^{+\infty} \left(z - \frac{q}{r}\right)^{-\frac{\rho}{r}-1} \tilde{U}(z)\,dz$ we have two cases. Either U remains negative over all $(l, +\infty)$ and in this case the integral is not $-\infty$ thanks to the term $\left(z - \frac{q}{r}\right)^{-\frac{\rho}{r}-1}$, or becomes positive after a certain point and in this case the integral is immediately greater than $-\infty$.

Take now $\tilde{U}$ which is not integrable in a right neighborhood of l. To prove the claim it is enough to show that, for every $\theta(\cdot) \in \tilde{\Theta}_{ad}(l)$, setting $X(t) := X(t; T, l, \theta(\cdot))$ we have

$$\mathbb{E}\left[\int_T^{T+1} e^{-\rho(t-T)}\tilde{U}(X(t))dt\right] = -\infty.$$

By the state equation we have

$$X(t) = l + \int_T^t rX(\tau)\,d\tau + \int_T^t \sigma\lambda\theta(\tau)X(\tau)\,d\tau - \int_T^t q\,d\tau + \int_T^t \sigma\theta(t)\,X(\tau)\,dB(\tau);$$

passing to the expectations and taking into account that $\theta(\cdot) \in [0, 1]$, $X(\cdot) \geq 0$, and the comparison criterion for ODE, we get the estimate

$$\mathbb{E}[X(t)] \leq le^{(r+\sigma\lambda)(t-T)}.$$

So we finally get by Jensen's inequality

$$\mathbb{E}\left[\int_T^{T+1} e^{-\rho(t-T)}\tilde{U}(X(t))dt\right] \leq \int_T^{T+1} e^{-\rho(t-T)}\tilde{U}\left(le^{(r+\sigma\lambda)(t-T)}\right)dt.$$

Applying a change of variable like the one done in formula (1.73) we get the claim.

Estimates from above. First, if $\lim_{z\to+\infty} \tilde{U}(z) =: \bar{U} < +\infty$ then, for every $x \geq l$,

$$\tilde{V}(x) \leq \int_T^{+\infty} e^{-\rho(t-T)}\bar{U}dt = \frac{\bar{U}}{\rho},$$

so in this case $\tilde{V}$ is finite and bounded. If $\bar{U} = +\infty$, then the claim follows as in the proof of Proposition 2.2.8, since therein the estimate does not depend on the control $\theta(\cdot) \in \tilde{\Theta}_{ad}(x)$. □

Proposition 1.3.12. *The value function $\tilde{V}$ is concave.*

Proof. Take two initial values x_1 and x_2 such that $x_1, x_2 \geq l$. Suppose, without loss of generality, that $x_1 < x_2$, $\tilde{V}(x_1) > -\infty$, and set $x_\gamma := \gamma x_1 + (1-\gamma)x_2$, $\gamma \in [0,1]$. We have to prove that

$$\tilde{V}(x_\gamma) \geq \gamma \tilde{V}(x_1) + (1-\gamma)\tilde{V}(x_2). \tag{1.74}$$

Let $\theta_1(\cdot) \in \tilde{\Theta}_{ad}(x_1)$, $\theta_2(\cdot) \in \tilde{\Theta}_{ad}(x_2)$ be ε-optimal for x_1 and x_2 respectively. Set $X_1(\cdot) := X(\cdot\,; T, x_1, \theta(\cdot))$ and $X_2(\cdot) := X(\cdot\,; T, x_2, \theta'(\cdot))$. We have

$$\begin{aligned}
\gamma\tilde{V}(x_1) + (1-\gamma)\tilde{V}(x_2) &< \gamma\left[\tilde{J}(x_1;\theta_1(\cdot)) + \varepsilon\right] + (1-\gamma)\left[\tilde{J}(x_2;\theta_2(\cdot)) + \varepsilon\right] \\
&= \varepsilon + \gamma\tilde{J}(x_1;\theta_1(\cdot)) + (1-\gamma)\tilde{J}(x_2;\theta_2(\cdot)) \\
&= \varepsilon + \gamma\mathbb{E}\left[\int_T^{+\infty} e^{-\rho(t-T)}\tilde{U}\left(X_1(t)\right)dt\right] \\
&\quad + (1-\gamma)\mathbb{E}\left[\int_T^{+\infty} e^{-\rho(t-T)}\tilde{U}\left(X_2(t)\right)dt\right] \\
&= \varepsilon + \mathbb{E}\left[\int_T^{+\infty} e^{-\rho(t-T)}\left[\gamma\tilde{U}\left(X_1(t)\right) + (1-\gamma)\tilde{U}\left(X_2(t)\right)\right]dt\right].
\end{aligned}$$

The concavity of $\tilde{U}$ implies that

$$\gamma\tilde{U}\left(X_1(t)\right) + (1-\gamma)\tilde{U}\left(X_2(t)\right) \leq \tilde{U}\left(\gamma X_1(t) + (1-\gamma)X_2(t)\right), \qquad \forall t \geq T.$$

Consequently, if we set $X_\gamma(\cdot) := \gamma X_1(\cdot) + (1-\gamma)X_2(\cdot)$, then we get

$$\gamma\tilde{V}(x_1) + (1-\gamma)\,\tilde{V}(x_2) < \varepsilon + \mathbb{E}\left[\int_T^{+\infty} e^{-\rho(t-T)}\tilde{U}\left(X_\gamma(t)\right)dt\right].$$

If there exists $\theta_\gamma(\cdot) \in \tilde{\Theta}_{ad}(x_\gamma)$ such that $X_\gamma(\cdot) = X\left(\cdot; T, x_\gamma, \theta_\gamma(\cdot)\right)$, then we would have

$$\varepsilon + \mathbb{E}\left[\int_T^{+\infty} e^{-\rho(t-T)}\tilde{U}\left(X_\gamma(t)\right)dt\right] = \varepsilon + \tilde{J}(x_\gamma;\theta_\gamma(\cdot)) \leq \varepsilon + \tilde{V}(x_\gamma),$$

i.e.

$$\gamma\tilde{V}(x_1) + (1-\gamma)\tilde{V}(x_2) < \varepsilon + \tilde{V}(x_\gamma)$$

and therefore, by the arbitrariness of ε, the claim (1.74) would be proved. To find such a $\theta_\gamma(\cdot)$, let us write the equation satisfied by $X_\gamma(\cdot)$. Recalling (1.60) we get

$$\begin{aligned}
dX_\gamma(t) &= \gamma dX_1(t) + (1-\gamma)dX_2(t) \\
&= \gamma\left[\Big(\left(r + \sigma\lambda\theta_1(t)\right)X_1(t) - q\Big)dt + \sigma\theta_1(t)X_1(t)dB(t)\right] \\
&\quad + (1-\gamma)\left[\Big(\left(r + \sigma\lambda\theta_2(t)\right)X_2(t) - q\Big)dt + \sigma\theta_2(t)X_2(t)dB(t)\right] \\
&= \left[rX_\gamma(t) + \left[\gamma\frac{X_1(t)}{X_\gamma(t)}\theta_1(t) + (1-\gamma)\frac{X_2(t)}{X_\gamma(t)}\theta_2(t)\right]\sigma\lambda X_\gamma(t) - q\right]dt \\
&\quad + \sigma\left[\gamma\frac{X_1(t)}{X_\gamma(t)}\theta_1(t) + (1-\gamma)\frac{X_2(t)}{X_\gamma(t)}\theta_2(t)\right]X_\gamma(t)dB(t).
\end{aligned}$$

Then defining the control

$$\theta_\gamma(t) = a\,(t)\,\theta_1(t) + d\,(t)\,\theta_2(t),$$

where

$$a\,(\cdot) = \gamma\frac{X_1(\cdot)}{X_\gamma(\cdot)}, \quad d\,(\cdot) = (1-\gamma)\frac{X_2(\cdot)}{X_\gamma(\cdot)},$$

we have

$$\begin{cases} dX_\gamma(t) = [[r + \sigma\lambda\theta_\gamma(t)]\,X_\gamma(t) - q]\,dt + \sigma\theta_\gamma(t)X_\gamma(t)dB(t), \quad t \geq T, \\ X_\gamma(T) = \gamma x_1 + (1-\gamma)x_2 = x_\gamma \geq l, \end{cases}$$

so we get $X_\gamma(\cdot) = X\,(\cdot; T, x_\gamma, \theta_\gamma(\cdot))$. The admissibility of $\theta_\gamma(\cdot)$ is clear since:

(i) for every $t \geq T$, we have $\theta_1\,(t)\,, \theta_2\,(t) \in [0,1]$ and $a\,(t) + d\,(t) = 1$, so by convexity of $[0,1]$ we get $\theta_\gamma\,(t) \in [0,1]$;

(ii) by construction $X_\gamma(t) \geq l$, almost surely with respect to $\mathbb{P}$ for any $t \geq T$.

The claim follows. □

Proposition 1.3.13. *The value function $\tilde{V}$ is strictly increasing.*

Proof. First we verify that the value function $\tilde{V}$ is increasing showing that

$$l \leq x \leq x' \implies \tilde{V}(x) \leq \tilde{V}(x').$$

Take any $\theta(\cdot) \in \tilde{\Theta}_{ad}(x)$. Writing the equation for $Y(\cdot) := X(\cdot; T, x', \theta(\cdot)) - X(\cdot; T, x, \theta(\cdot))$, we can see that $Y(\cdot)$ solves a linear SDE with nonnegative initial datum. Therefore $Y(\cdot) \geq 0$, $\mathbb{P}$-a.s., i.e.

$$X(\cdot\,; T, x, \theta(\cdot)) \leq X(\cdot\,; T, x', \theta(\cdot)), \quad \mathbb{P}\text{-a.s.}.$$

So it is also $\theta(\cdot) \in \tilde{\Theta}_{ad}(x')$, i.e. $\tilde{\Theta}_{ad}\,(x) \subset \Theta_{ad}\,(x')$. Moreover, by monotonicity of the utility function $\tilde{U}$ we have

$$\begin{aligned} x \leq x' \Longrightarrow \tilde{U}\,(X\,(\cdot; T, x, \theta(\cdot))) \leq \tilde{U}\,(X\,(\cdot; T, x_2, \theta(\cdot)))\,,\ \mathbb{P}\text{-a.s.} \\ \Longrightarrow \tilde{J}\,(x; \theta(\cdot)) \leq \tilde{J}\left(x'; \theta(\cdot)\right). \end{aligned}$$

Since $\tilde{\Theta}_{ad}\,(x) \subset \tilde{\Theta}_{ad}\,(x')$ the above implies $\tilde{V}\,(x) \leq \tilde{V}\,(x')$.

The strict monotonicity of the value function $\tilde{V}$ is a direct consequence of monotonicity and concavity (see, e.g., the proof in [Zariphopoulou; 1994], p. 63). Indeed, if $\tilde{V}$ is not strictly monotone then it must be constant on a half line $[\bar{x}, +\infty)$. We show that this cannot be true.
By (1.3.3) we have, for every $y \geq l$,

$$\tilde{V}\,(y) \geq \frac{\tilde{U}\,(y)}{\rho}.$$

So, if $\lim_{z\to+\infty}\tilde{U}(z)=+\infty$ then $\lim_{z\to+\infty}\tilde{V}(z)=+\infty$ and the claim follows.

Take now $\lim_{z\to+\infty}\tilde{U}(z)=:\tilde{U}_\infty<+\infty$. In this case we must have

$$\tilde{V}(\bar{x})=\lim_{y\to+\infty}\tilde{V}(y)\geq\frac{\tilde{U}_\infty}{\rho}.$$

On the other hand, for every $\theta(\cdot)\in\tilde{\Theta}_{ad}(\bar{x})$, we get

$$\begin{aligned}\tilde{J}(\bar{x};\theta(\cdot))&=\mathbb{E}\left[\int_T^{+\infty}e^{-\rho(t-T)}\tilde{U}(X(t;T,\bar{x},\theta(\cdot)))\,dt\right]\\&\leq\int_T^{+\infty}e^{-\rho(t-T)}\tilde{U}\left(e^{(r+\lambda\sigma)(t-T)}\bar{x}\right)dt.\end{aligned}$$

Fix $T_1>T$. Calling $\tilde{U}_{T_1}=\tilde{U}\left(e^{(r+\lambda\sigma)(T_1-T)}\bar{x}\right)$, we have $\tilde{U}_{T_1}=\tilde{U}_\infty-\varepsilon$, for some $\varepsilon>0$. We can write

$$\begin{aligned}\tilde{J}(\bar{x};\theta(\cdot))&\leq\int_T^{T_1}e^{-\rho(t-T)}\tilde{U}_{T_1}dt+\int_{T_1}^{+\infty}e^{-\rho(t-T)}\tilde{U}_\infty dt\\&=\frac{\tilde{U}_{T_1}}{\rho}\left[1-e^{-\rho(T_1-T)}\right]+\frac{\tilde{U}_\infty}{\rho}e^{-\rho(T_1-T)}\\&\leq\frac{\tilde{U}_\infty}{\rho}-\frac{1-e^{-\rho(T_1-T)}}{\rho}\varepsilon\\&\leq\tilde{V}(\bar{x})-\frac{1-e^{-\rho(T_1-T)}}{\rho}\varepsilon.\end{aligned}$$

This is a contradiction and so the claim follows. □

Proposition 1.3.14. *The value function $\tilde{V}$ is continuous in $(l,+\infty)$ and Lipschitz continuous in $[a,+\infty)$, for any $a>l$. Moreover:*

- *if $rl>q$ and $\tilde{V}(l)>-\infty$, then $\tilde{V}$ is uniformly continuous in $[l,+\infty)$;*
- *if $rl=q$ and $\tilde{U}(l)>-\infty$, then $\tilde{V}$ is uniformly continuous in $[l,+\infty)$.*

Proof. The Lipschitz continuity of $\tilde{V}$ in the interval $[a,+\infty)$, for any $a>l$ (and so also the continuity of $\tilde{V}$ in the interval $(l,+\infty)$), is a straightforward consequence of concavity and strict monotonicity.

(i) Let $rl>q$ and $\tilde{V}(l)>-\infty$, i.e. $\tilde{U}$ is integrable in a right neighborhood of l (see Proposition 1.3.11 - (ii)). We have to show that

$$\lim_{x\to l^+}[\tilde{V}(x)-\tilde{V}(l)]=0.$$

Since $rl>q$, the control strategy $\theta(\cdot)\equiv 0$ at the starting point l gives rise to a trajectory which is strictly increasing. Let $x>l$. Applying the control

$\theta(\cdot) \equiv 0$ to the state equation (1.60) with initial point $X(T) = l$, the corresponding trajectory is deterministic and it reaches the point x at time $\hat{t}$ such that $X\left(\hat{t}; T, l, 0\right) = x$, i.e.

$$e^{r\left(\hat{t}-T\right)}\left[l - \frac{q}{r}\right] + \frac{q}{r} = x \qquad \Longleftrightarrow \qquad \hat{t} = T + \frac{1}{r}\log\frac{rx-q}{rl-q}.$$

Now, by the dynamic programming principle (Proposition 1.3.16), we have

$$\tilde{V}(l) \geq \int_T^{\hat{t}} e^{-\rho(t-T)}\tilde{U}\left(X\left(t; T, l, 0\right)\right)dt + e^{-\rho\left(\hat{t}-T\right)}\tilde{V}\left(X\left(\hat{t}; T, l, 0\right)\right),$$

which gives

$$0 \leq \tilde{V}(x) - \tilde{V}(l) \leq -\int_T^{\hat{t}} e^{-\rho(t-T)}\tilde{U}\left(X\left(t; T, l, 0\right)\right)dt + \left(1 - e^{-\rho\left(\hat{t}-T\right)}\right)\tilde{V}(x)$$

(notice that the first inequality is a consequence of the monotonicity of the value function given in Proposition 1.3.13). Observing that

$$x \to l \Longrightarrow \hat{t} \to T$$

and using the integrability of $\tilde{U}$ we get the claim.

(ii) Let $rl = q$ and $\tilde{V}(l) > -\infty$, i.e. by Proposition 1.3.11-(i) $\tilde{U}(l) > -\infty$. in this case $\tilde{V}(l) = \frac{\tilde{U}(l)}{\rho}$ and of course, without loss of generality for our goal, we can suppose $\tilde{U}(l) \geq 0$. Let $t_0 \geq T$, take $l_n \downarrow l$ and define

$$\tilde{V}^{t_0}(l_n) := \sup_{\theta(\cdot)\in\tilde{\Theta}_{ad}(l_n)} \mathbb{E}\left[\int_T^{t_0} e^{-\rho(t-T)}\tilde{U}\left(X\left(t; T, l_n, \theta(\cdot)\right)\right)dt\right].$$

Arguing as in the proof of Lemma 1.2.17-(2), one can prove that for any fixed t_0 the convergence

$$\tilde{V}^{t_0}(l_n) \longrightarrow \frac{1 - e^{-\rho(t_0-T)}}{\rho}\tilde{U}(l), \quad n \to \infty$$

holds true. We prove now that for any fixed $\varepsilon > 0$ there exists $t_0^* \geq T$ independent of n such that, for any $t_0 \geq t_0^*$, it holds

$$\left|\tilde{V}^{t_0}(l_n) - \tilde{V}(l_n)\right| < \varepsilon/2. \tag{1.75}$$

Indeed, let $t_0 \geq T$; taking into account the inequality

$$\left|\sup_\xi f(\xi) - \sup_\xi g(\xi)\right| \leq \sup_\xi |f(\xi) - g(\xi)|$$

and arguing as in the proof of Proposition 1.3.7, we can find $C' > 0$ such that, for any $t_0 \geq T$,

$$\begin{aligned}\left|\tilde{V}^{t_0}(l_n) - \tilde{V}(l_n)\right| &\leq \sup_{\theta(\cdot)\in\tilde{\Theta}_{ad}(l_n)} \mathbb{E}\left[\int_{t_0}^{+\infty} e^{-\rho(t-T)}\tilde{U}\left(X\left(t\right)\right)dt\right] \\ &\leq e^{-\rho(t_0-T)}C'\left(1+l_n^{\beta}\right) \\ &\leq e^{-\rho(t_0-T)}C'\left(1+l_1^{\beta}\right),\end{aligned}$$

where $\beta \in [0,1)$ is given by Hypothesis 1.3.2-(ii). So we can find t_0^* such that (1.75) holds true. Moreover we can find $\bar{t}_0 \geq t_0^*$ such that

$$\left|\tilde{V}(l) - \frac{1-e^{-\rho(\bar{t}_0-T)}}{\rho}\tilde{U}(l)\right| < \varepsilon/2.$$

Therefore

$$\begin{aligned}&\limsup_{n\to\infty}\left|\tilde{V}(l_n) - \tilde{V}(l)\right| \\ &\quad\leq \limsup_{n\to\infty}\left[\left|\tilde{V}(l_n) - \tilde{V}^{\bar{t}_0}(l_n)\right| + \left|\tilde{V}^{\bar{t}_0}(l_n) - \frac{1-e^{-\rho(\bar{t}_0-T)}}{\rho}\right|\right. \\ &\qquad\qquad\left. + \left|\tilde{V}(l) - \frac{1-e^{-\rho(\bar{t}_0-T)}}{\rho}\tilde{U}(l)\right|\right] < \varepsilon,\end{aligned}$$

so that, by the arbitrariness of ε, the claim follows. □

Remark 1.3.15. From the proof of Proposition 1.3.14-(i) it follows that, when $\tilde{U}\left(l\right)$ is finite (and so also $\tilde{V}\left(l\right)$) and $rl > q$, we have, for $x > l$,

$$\frac{\tilde{V}\left(x\right)-\tilde{V}\left(l\right)}{x-l} \leq -\frac{(rl-q)^{\rho/r}}{x-l}\int_l^x (rz-q)^{-1-\rho/r}\,U(z)dz + \frac{1-\left(\frac{rx-q}{rl-q}\right)^{-\frac{\rho}{r}}}{x-l}\tilde{V}\left(x\right),$$

so, recalling that $\tilde{V}'\left(l^+\right)$ must exist by the concavity of $\tilde{V}$, it follows

$$\tilde{V}'\left(l^+\right) \leq \frac{\rho}{rl-q}\left[\frac{\tilde{|U(l)|}}{\rho} + \tilde{V}\left(l\right)\right].$$

This means, in particular that $\tilde{V}'\left(l^+\right)$ is finite. On the other hand when $rl = q$, $\tilde{U}'\left(l^+\right) = +\infty$ and $\tilde{U}(l) > -\infty$ (hence $\tilde{V}$ is finite and continuous at l), then $\tilde{V}'\left(l^+\right)$ is infinite. Indeed in this case $\tilde{V}\left(l\right) = \frac{\tilde{U}(l)}{\rho}$ while

$$\tilde{V}\left(x\right) \geq \tilde{J}\left(x;0\right) \geq \frac{\tilde{U}\left(x\right)}{\rho},$$

so $\tilde{V}'\left(l^+\right) \geq \frac{\tilde{U}'(l^+)}{\rho} = +\infty$. See Section 1.3.7 for an example. □

1.3.4 Dynamic programming

Also in this case we follow a dynamic programming approach. The dynamic programming principle, in this stationary context can be stated as follows.

Theorem 1.3.16. *The value function $\tilde{V}$ satisfies the dynamic programming equation, i.e. for every $x \geq l$ and for any family of stopping times $\left(\tau^{\theta(\cdot)}\right)_{\theta(\cdot)\in\Theta_{ad}(x)}$ taking values in $[T, +\infty)$, the following functional equation holds*

$$\tilde{V}(x) = \sup_{\theta(\cdot)\in\tilde{\Theta}_{ad}(x)} \mathbb{E}\left[\int_T^{\tau^{\theta(\cdot)}} e^{-\rho(t-T)}\tilde{U}\left(X\left(t;T,x,\theta(\cdot)\right)\right)dt \right.$$
$$\left. + e^{-\rho(\tau^{\theta(\cdot)}-T)}\tilde{V}\left(X\left(\tau^{\theta(\cdot)};T,x,\theta(\cdot)\right)\right)\right]. \quad (1.76)$$

Remark 1.3.17. We do not give the proof and refer to [Yong, Zhou; 1999], Chapter 4, Theorem 3.3, in the case of continuous value functions and to [Soner; 2004] in the general case, where a measurable selection argument is used. As in the transitory phase, we want to point out that we have proved the continuity of our value function in Proposition 1.3.14 only using

$$\tilde{V}(x) \geq \int_T^{\hat{t}} e^{-\rho(t-T)}\tilde{U}(X(t;T,x,0))dt + e^{\rho(\hat{t}-T)}\tilde{V}(X(\hat{t};T,x,0)), \quad \hat{t} \geq T,$$

which can be proved without any measurable selection argument, because in this case we are on a deterministic trajectory. Therefore we can use the argument of [Yong, Zhou; 1999] in order to prove the dynamic programming principle without loss of generality. $\square$

The HJB equation formally associated with $\tilde{V}$ is

$$\rho v(x) - \tilde{\mathcal{H}}\left(x, v'(x), v''(x)\right) = 0, \qquad x \in [l, +\infty), \quad (1.77)$$

where, for $x \geq l$ and $p, Q \in \mathbb{R}$,

$$\tilde{\mathcal{H}}(x,p,Q) = \sup_{\theta\in[0,1]} \tilde{\mathcal{H}}_{cv}(x,p,Q;\theta),$$

with

$$\tilde{\mathcal{H}}_{cv}(x,p,Q;\theta) = \tilde{U}(x) + \left([r+\sigma\lambda\theta]x - q\right)p + \frac{1}{2}\sigma^2x^2\theta^2Q.$$

Note that calling $\mathcal{L}^\theta$, $\theta \in [0,1]$, the parabolic operator defined by

$$\left[\mathcal{L}^\theta f\right](x) = \frac{1}{2}\sigma^2\theta^2x^2f''(x) + \left([r+\sigma\lambda\theta]x - q\right)f'(x), \quad x \geq l,\ f \in C^2\left([l,+\infty);\mathbb{R}\right), \quad (1.78)$$

we can write

$$\tilde{\mathcal{H}}_{cv}\left(x, v'(x), v''(x);\theta\right) = \tilde{U}(x) + \left[\mathcal{L}^\theta v\right](x), \qquad x \geq l.$$

To calculate the Hamiltonian we observe that the function to optimize over $\theta \in [0,1]$ is

$$\tilde{\mathcal{H}}_{cv}^{0}\left(x,p,Q;\theta\right) := \sigma\lambda\theta x p + \frac{1}{2}\sigma^2\theta^2 x^2 Q,$$

which is exactly the function $\mathcal{H}_{cv}^{0}$ defined in Subsection 1.2.5. In the same way

$$\tilde{\mathcal{H}}^{0}(x,p,Q) := \sup_{\theta\in[0,1]} \tilde{\mathcal{H}}_{cv}^{0}\left(x,p,Q;\theta\right) = \mathcal{H}^{0}(x,p,Q).$$

1.3.5 The HJB equation: viscosity solutions and regularity

Let us consider the HJB equation (1.77). This is a second order PDE which is degenerate elliptic. The concept of viscosity solution we use here is given by the following definition. We refer to the literature cited in Subsection 1.2.7 for more details on the definition of constrained viscosity solutions of elliptic equations.

Definition 1.3.18. **(i)** A continuous function $v : (l, +\infty) \to \mathbb{R}$ is called a viscosity subsolution (respectively supersolution) of the equation (1.77) in $(l, +\infty)$ if, for any $\psi \in C^2\left((l, +\infty); \mathbb{R}\right)$ and for any maximum point $x_M \in (l, +\infty)$ (respectively minimum point $x_m \in (l, +\infty)$) of $v - \psi$, we have

$$\rho v(x_M) - \mathcal{H}\left(x_M, \psi'(x_M), \psi''(x_M)\right) \leq 0$$

$$(\text{respectively } \rho v(x_m) - \mathcal{H}\left(x_m, \psi'(x_m), \psi''(x_m)\right) \geq 0).$$

(ii) A continuous function $v : (l, +\infty) \to \mathbb{R}$ is called a viscosity solution to equation (1.77) in $(l, +\infty)$ if it is both a viscosity subsolution and a viscosity supersolution in $(l, +\infty)$.
(iii) A continuous function $v : [l, +\infty) \to \mathbb{R}$ is called a viscosity subsolution to equation (1.77) on $[l, +\infty)$ if, for any $\psi \in C^2\left([l, +\infty)\right); \mathbb{R})$ and for any maximum point $x_M \in [l, +\infty)$ of $v - \psi$, it follows

$$\rho\psi(x_M) - \mathcal{H}\left(x_M, \psi'(x_M), \psi''(x_M)\right) \leq 0.$$

(iv) A continuous function $v : [l, +\infty) \to \mathbb{R}$ is called a constrained viscosity solution to equation (1.77) if it is viscosity subsolution on $[l, +\infty)$ and a viscosity supersolution in $(l, +\infty)$. □

We can state and prove the following result.

Theorem 1.3.19. *The value function $\tilde{V}$ defined in* (1.3.3) *is a viscosity solution to the HJB equation* (1.77) *in $(l, +\infty)$. If $\tilde{U}$ is finite in l then $\tilde{V}$ is a constrained viscosity solution to the HJB equation* (1.77) *on $[l, +\infty)$.*

Proof. We have to show that $\tilde{V}$ is:

(i) a viscosity supersolution in $(l+\infty)$;

(ii) a viscosity subsolution in $(l+\infty)$;

(iii) a viscosity subsolution in l when $\tilde{U}$ is finite in l.

(i) Let $x_m \in (l,+\infty)$, $\psi \in C^2((l,+\infty);\mathbb{R})$ be such that x_m is a local minimum point for the function $\tilde{V}-\psi$. Of course for our goal we can assume without loss of generality that

$$\tilde{V}(x_m) = \psi(x_m); \qquad \tilde{V}(x) \geq \psi(x),\ \forall x \in (l,+\infty). \tag{1.79}$$

Let $\varepsilon > 0$ be such that $x_m - \varepsilon > l$. For $\theta \in [0,1]$, let us set $X(t) := X(t;T,x_m,\theta)$. Consider the stopping time $\tau^\theta = \inf\{t \geq T \mid |X(t) - x_m| \geq \varepsilon\}$ and notice that $\tau^\theta > T$ almost surely. By (1.79) we get, for any $t \geq T$,

$$e^{-\rho(t-T)}\tilde{V}(X(t)) - \tilde{V}(x_m) \geq e^{-\rho(t-T)}\psi(X(t)) - \psi(x_m).$$

Let $h > T$. Setting $\tau_h^\theta := \tau^\theta \wedge h$, by the dynamic programming principle (1.76) we get, for any $\theta \in [0,1]$,

$$\begin{aligned} 0 &\geq \mathbb{E}\left[\int_T^{\tau_h^\theta} e^{-\rho(t-T)}\tilde{U}(X(t))\,dt + e^{-\rho\left(\tau_h^\theta - T\right)}\tilde{V}\left(X\left(\tau_h^\theta\right)\right) - \tilde{V}(x_m)\right] \\ &\geq \mathbb{E}\left[\int_T^{\tau_h^\theta} e^{-\rho(t-T)}\tilde{U}(X(t))\,dt + e^{-\rho\left(\tau_h^\theta - T\right)}\psi\left(X\left(\tau_h^\theta\right)\right) - \psi(x_m)\right]. \end{aligned} \tag{1.80}$$

Applying the Dynkin formula to the function $(t,x) \mapsto e^{-\rho(t-T)}\psi(x)$ with the process $X(t)$, we get ($\mathcal{L}^\theta$ is defined in (1.78))

$$\begin{aligned} &\mathbb{E}\left[e^{-\rho\left(\tau_h^\theta - T\right)}\psi\left(X\left(\tau_h^\theta\right)\right) - \psi(x_m)\right] \\ &\qquad\qquad = \mathbb{E}\left[\int_T^{\tau_h^\theta} e^{-\rho(t-T)}\left[-\rho\psi(X(t)) + \left(\mathcal{L}^\theta\psi\right)(X(t))\right]dt\right] \end{aligned}$$

and thus by (1.80) we have, for any $\theta \in [0,1]$,

$$0 \geq \mathbb{E}\left[\int_T^{\tau_h^\theta} e^{-\rho(t-T)}\left[-\rho\psi(X(t)) + \tilde{\mathcal{H}}_{cv}\left(X(t), \psi'(X(t)), \psi''(X(t));\theta\right)\right]dt\right].$$

Divide now by $\tau_h^\theta - T$ and let $h \to T$. By continuity of $\tilde{\mathcal{H}}_{cv}$ we easily get by dominated convergence

$$0 \geq -\rho\psi(x_m) + \tilde{\mathcal{H}}_{cv}\left(x_m, \psi'(x_m), \psi''(x_m);\theta\right), \qquad \theta \in [0,1].$$

Taking the supremum over $\theta \in [0,1]$ the claim follows.

(ii) Let $x_M \in (l,+\infty)$ and $\psi \in C^2((l,+\infty);\mathbb{R})$ be such that x_M is a local maximum point of $\tilde{V}-\psi$ in $(l,+\infty)$. Again for our goal we can assume without loss of generality that

$$\tilde{V}(x_M) = \psi(x_M); \qquad \tilde{V}(x) \le \psi(x),\ \forall x \in (l,+\infty)\,. \tag{1.81}$$

We must prove that

$$\rho\psi(x_M) - \tilde{\mathcal{H}}\left(x_M, \psi'(x_M), \psi''(x_M)\right) \le 0.$$

Let us suppose by contradiction that this relation is false. Then there exists $\nu > 0$ such that

$$0 < \nu < \rho\psi(x_M) - \tilde{\mathcal{H}}\left(x_M, \psi'(x_M), \psi''(x_M)\right).$$

By continuity of $\tilde{U}$, ψ, $\tilde{\mathcal{H}}$, there exists $\varepsilon \in (0, x_M - l)$ such that, for any $x \in (x_M - \varepsilon, x_M + \varepsilon)$, we have

$$\frac{\nu}{2} \le \rho\psi(x) - \tilde{\mathcal{H}}\left(x, \psi'(x), \psi''(x)\right). \tag{1.82}$$

Let $\theta(\cdot) \in \tilde{\Theta}_{ad}(x_M)$ and set $X(t) := X(t;T,x_M,\theta(\cdot))$. Define the stopping time

$$\tau^\theta := \inf\left\{t \ge T \mid |X(t) - x_M| \ge \varepsilon\right\} \wedge 2T.$$

Of course $\tau^\theta > T$ almost surely by continuity of trajectories of $X(\cdot)$. Now we take (1.82) for $x = X(t)$, multiply it by $e^{-\rho(t-T)}$, integrate it on $[T,\tau^\theta]$ and calculate its expected value obtaining

$$\begin{aligned}&\frac{\nu}{2}\,\mathbb{E}\left[\int_T^{\tau^\theta} e^{-\rho(t-T)}dt\right]\\ &\le \mathbb{E}\left[\int_T^{\tau^\theta} e^{-\rho(t-T)}\left(\rho\psi(X(t)) - \sup_{\theta\in[0,1]}\left[\left(\mathcal{L}^\theta\psi\right)(X(t))\right] - \tilde{U}(X(t))\right)dt\right],\end{aligned}$$

from which it follows

$$\begin{aligned}&\frac{\nu}{2}\,\mathbb{E}\left[\int_T^{\tau^\theta} e^{-\rho(t-T)}dt\right]\\ &\qquad\le \mathbb{E}\left[\int_T^{\tau^\theta} e^{-\rho(t-T)}\left(\rho\psi(X(t)) - \left[\mathcal{L}^{\theta(t)}\psi\right](X(t)) - \tilde{U}(X(t))\right)dt\right].\end{aligned} \tag{1.83}$$

Similarly to what done in (i), we apply the Dynkin formula to the function $(t,x) \mapsto e^{-\rho(t-T)}\psi(x)$ with the process $X(t)$. We get

$$\begin{aligned}&\mathbb{E}\left[e^{-\rho\left(\tau^\theta - T\right)}\psi(X(\tau^\theta)) - \psi(x_M)\right]\\ &\qquad\qquad = \mathbb{E}\left[\int_T^{\tau^\theta} e^{-\rho(t-T)}\left(\left[\mathcal{L}^{\theta(t)}\psi\right](X(t)) - \rho\psi(X(t))\right)dt\right].\end{aligned} \tag{1.84}$$

From (1.81), (1.83) and (1.84) it follows, rearranging the terms,

$$\tilde V(x_M) \geq \mathbb{E}\left[\int_T^{\tau^\theta} e^{-\rho(t-T)}\tilde U(X(t))dt + e^{-\rho\left(\tau^\theta - T\right)}\tilde V\left(X\left(\tau^\theta\right)\right)\right] + \frac{\nu}{2}\,\mathbb{E}\left[\int_T^{\tau^\theta} e^{-\rho(t-T)}dt\right]. \quad (1.85)$$

We claim that there exists a constant $\alpha > 0$ independent of $\theta(\cdot)$ such that

$$\mathbb{E}\left[\int_T^{\tau^\theta} e^{-\rho(t-T)}dt\right] \geq \alpha.$$

We will prove this fact in Lemma 1.3.20 below. Therefore by (1.85) we get

$$\tilde V(x_M) \geq \mathbb{E}\left[\int_T^{\tau^\theta} e^{-\rho(t-T)}\tilde U(X(t))dt + e^{-\rho\left(\tau^\theta - T\right)}\tilde V\left(X\left(\tau^\theta\right)\right)\right] + \frac{\nu}{2}\alpha$$

contradicting the dynamic programming principle

$$\tilde V(x_M) = \sup_{\theta(\cdot)\in\tilde\Theta_{ad}(x_M)} \mathbb{E}\left[\int_T^{\tau^\theta} e^{-\rho(t-T)}\tilde U(X(t))dt + e^{-\rho\left(\tau^\theta - T\right)}\tilde V\left(X\left(\tau^\theta\right)\right)\right].$$

(iii) Let $\tilde U$ be finite in l. If $rl = q$ then $\tilde V$ is continuous in l (see Proposition 1.3.14) and $\rho\tilde V(l) = \tilde U(l)$; so the subsolution inequality is immediate from the fact that $\tilde{\mathcal{H}}(x,p,Q)$ is always nonnegative for $x \geq l$, $p \geq 0$.

So, consider now the case $rl > q$. Take $\psi \in C^2([l,+\infty);\mathbb{R})$ such that l is a maximum point of $\tilde V - \psi$ in $[l,+\infty)$. Then we can argue exactly as in point (ii) to get the claim taking right neighborhoods of l instead of whole neighborhoods. □

Lemma 1.3.20. *For any $\theta(\cdot) \in \tilde\Theta_{ad}(x_M)$ let τ^θ be the stopping time defined in the part (ii) of the proof of Theorem 1.3.19. There exists a constant $\alpha > 0$ independent of $\theta(\cdot) \in \tilde\Theta_{ad}(x_M)$ such that*

$$\mathbb{E}\left[\tau^\theta - T\right] \geq \alpha.$$

Proof. For the controls such that $\mathbb{P}\{\tau^\theta < 2T\} < 1/2$, we have the estimate

$$\mathbb{E}\left[\int_T^{\tau^\theta} e^{-\rho(t-T)}dt\right] \geq \frac{1}{2}\left[\frac{1-e^{-\rho T}}{\rho}\right].$$

Therefore we can suppose without loss of generality that $\mathbb{P}\{\tau^\theta < 2T\} \geq 1/2$. For $\theta(\cdot) \in \tilde\Theta_{ad}(x_M)$, let $X(t) := X(t;T,x_M,\theta(\cdot))$ and apply the Dynkin formula to the process $X(\cdot)$ with the function $\varphi(t,x) = e^{-\rho(t-T)}(x-x_M)^2$ on $[T,\tau^\theta]$; we get

$$\mathbb{E}\left[e^{-\rho(\tau^\theta - T)}\left(X(\tau^\theta) - x_M\right)^2\right]$$
$$= \mathbb{E}\Bigg[\int_T^{\tau^\theta} e^{-\rho(t-T)}\big[-\rho(X(t)-x_M)^2$$
$$+ 2(X(t)-x_M)\left[(r+\sigma\lambda\theta(t))X(t) - q\right] + \sigma^2\theta(t)^2 X(t)^2\big]\,dt\Bigg].$$

So, considering that $\tau^\theta \leq 2T$, $\theta(t) \in [0,1]$ and that, for $t \in [T,\tau^\theta]$, we have $|X(t)| \leq x_M + \varepsilon$, we can find $K > 0$ such that

$$\frac{1}{2}\,\varepsilon^2 e^{-\rho T} \leq \mathbb{P}\{\tau^\theta < T\}\,\varepsilon^2 e^{-\rho T} \leq \mathbb{E}\left[\int_T^{\tau^\theta} K e^{-\rho(t-T)}\,dt\right];$$

this estimate does not depend on $\theta(\cdot)$ and therefore the claim is proved. □

Remark 1.3.21. The above proof recall some argument used in Theorem 1 of [Choulli, Taksar, Zhou; 2003]. However our problem does not fit exactly into the results contained in [Choulli, Taksar, Zhou; 2003] or in other papers. In Theorem 3.1 of [Zariphopoulou; 1994], pp. 65–69, a different proof of the existence results is given for an HJB equation similar to ours (featuring state constraints and unboudedness of the data). □

Remark 1.3.22. We are not proving here a comparison theorem. This should be possible, e.g., arguing as in [Zariphopoulou; 1994] Theorem 4.1, p. 69–74, even if our HJB equation is different (see also [Ishii, Loreti; 2002] for uniqueness results in presence of state constraints). We do not do it here for brevity since the comparison result is not essential for our applications. □

Now we work to prove the smoothness of $\tilde{V}$. For this purpose the following lemmata are needed.

Lemma 1.3.23. *Let g be a concave function on $\mathbb{R}$ such that $g(x) = g(x_0) + a(x - x_0)$ for $x \leq x_0$ and $g(x) = g(x_0) + b(x - x_0)$ for $x \geq x_0$, where $a > b$. Then for each $\varepsilon > 0$, $c \in [b,a]$, there exists a concave $C^2(\mathbb{R};\mathbb{R})$ function f such that $f \geq g$, $f(x_0) = g(x_0)$, $f'(x_0) = c$ and $f''(x_0) \leq -\varepsilon^{-1}$.*

Proof. See Lemma 2, p. 1958, of [Choulli, Taksar, Zhou; 2003]. □

Lemma 1.3.24. *Let I be a given interval in $\mathbb{R}$, $g \in C^0(I;\mathbb{R})$ and let x_0 be an interior point of I. Assume that there exists a sequence $(x_n)_{n\in\mathbb{N}}$ such that $x_n < x_0$, $\exists g'(x_n)$ for every $n \in \mathbb{N}$, and $g'(x_n) \to -\infty$ as $x_n \to x_0$. Then $D^+g(x_0) = \emptyset$, where $D^+g(x_0)$ is the superdifferential of g at x_0.*

Proof. If $p \in D^+g(x_0)$ then, for every x in a given neighborhood of x_0, we have

$$g(x) - g(x_0) \le p(x - x_0) + o(x - x_0),$$

so

$$\liminf_{x \to x_0^-} \frac{g(x) - g(x_0)}{x - x_0} \ge p. \tag{1.86}$$

On the other hand, for every $n \in \mathbb{N}$ we have

$$g(x_n) - g(x_0) = -\left[g'(x_n)(x_0 - x_n) + o(x_0 - x_n)\right],$$

so

$$\lim_{n \to +\infty} \frac{g(x_n) - g(x_0)}{x_n - x_0} = \lim_{n \to +\infty} g'(x_n) = -\infty$$

which contradicts (1.86). □

Theorem 1.3.25. *The value function $\tilde{V}$ belongs to the class $C^2((l, +\infty); \mathbb{R})$.*

Proof. We first prove that $\tilde{V}$ is differentiable. Since $\tilde{V}$ is concave, by Alexandrov's Theorem we know that for almost every $x \in (l, +\infty)$ there exists $\tilde{V}'(x)$ and $\tilde{V}''(x)$ (in the sense of the Taylor expansion). Let us suppose, by contradiction, that there exists $x_0 \in (l, +\infty)$ such that $\nexists \tilde{V}'(x_0)$. Then by concavity the right and left derivatives $\tilde{V}'(x_0^+)$ and $\tilde{V}'(x_0^-)$ exist and $\tilde{V}'(x_0^-) > \tilde{V}'(x_0^+)$. Moreover the subdifferential $D^-\tilde{V}(x_0)$ has to be empty and the superdifferential $D^+\tilde{V}(x_0)$ has to be the interval $\left[\tilde{V}'(x_0^+), \tilde{V}'(x_0^-)\right]$. Now, using Lemma 1.3.23 with

$$g(x) = \begin{cases} \tilde{V}(x_0) + \tilde{V}'(x_0^+)(x - x_0), & \text{when} \quad x \ge x_0, \\ \tilde{V}(x_0) + \tilde{V}'(x_0^-)(x - x_0), & \text{when} \quad x < x_0, \end{cases}$$

we get that for every ε there exists $f_\varepsilon \in C^2(\mathbb{R}; \mathbb{R})$ such that

- $f_\varepsilon(x_0) = \tilde{V}(x_0)$;
- $f_\varepsilon(x) \ge g(x) \ge \tilde{V}(x)$ for $x \in (l, +\infty)$;
- $f_\varepsilon'(x_0) = \tilde{V}'(x_0^+)$;
- $f_\varepsilon''(x_0) \le \varepsilon^{-1}$.

Since $\tilde{V}$ is a viscosity solution to the HJB equation (1.77)), then

$$\rho\tilde{V}(x_0) \le (rx_0 - q) f_\varepsilon'(x_0) + \tilde{U}(x_0) + \tilde{\mathcal{H}}^0\left(x_0, f_\varepsilon'(x_0), f_\varepsilon''(x_0)\right).$$

For ε sufficiently small the above implies

$$\rho\tilde{V}(x_0) < (rx_0 - q)\tilde{V}'(x_0^-) + \tilde{U}(x_0). \tag{1.87}$$

On the other hand, let $(x_n)_{n\in\mathbb{N}}$ be a sequence such that, for every $n \in \mathbb{N}$, $x_n < x_0$, $\exists \tilde{V}'(x_n)$, $\tilde{V}''(x_n)$ and $\tilde{V}'(x_n) \to \tilde{V}'\left(x_0^-\right)$, $\tilde{V}''(x_n) \to Q$ for some $Q \in [-\infty, 0]$ when $x_n \to x_0$. Then we have

$$\rho \tilde{V}(x_n) = (rx_n - q)\,\tilde{V}'(x_n) + \tilde{U}(x_n) + \tilde{\mathcal{H}}^0\left(x_n, \tilde{V}'(x_n), \tilde{V}''(x_n)\right).$$

Passing to the limit for $n \to +\infty$ we get, if $Q > -\infty$

$$\rho \tilde{V}(x_0) = (rx_0 - q)\,\tilde{V}'\left(x_0^-\right) + \tilde{U}(x_0) + \tilde{\mathcal{H}}^0\left(x_0, \tilde{V}'\left(x_0^-\right), Q\right), \tag{1.88}$$

if $Q = -\infty$

$$\rho \tilde{V}(x_0) = (rx_0 - q)\,\tilde{V}'\left(x_0^-\right) + \tilde{U}(x_0). \tag{1.89}$$

Both equalities (1.88) and (1.89) are not compatible with (1.87), so a contradiction arise and $\tilde{V}$ must be differentiable at x_0. Continuous differentiability of V follows from its concavity.

We now prove the twice differentiability. Again by the Alexandrov Theorem, there exists a set $\mathcal{A} \subset (l, +\infty)$ such that the Lebesgue measure of $\mathcal{A}^c = (l, +\infty) - \mathcal{A}$ is zero and $\tilde{V}$ is twice differentiable at every point of $\mathcal{A}$. Let $x_0 \in (l, +\infty)$. Take any sequence $\{x_n\}_{n\in\mathbb{N}} \subset \mathcal{A}$ such that $x_n \to x_0$. Then, by the continuous differentiability of $\tilde{V}$, we get that $\tilde{V}(x_n) \to \tilde{V}(x_0)$ and $\tilde{V}'(x_n) \to \tilde{V}'(x_0) > 0$ (note that $\tilde{V}'(x_0) > 0$ since $\tilde{V}$ is concave and strictly increasing).

First of all we observe that each element of the sequence $\tilde{V}''(x_n)$ belongs to $(-\infty, 0]$, so there exists at least a subsequence converging either to $-\infty$ or to a finite nonpositive limit. Now we prove that the limit exists and does not depend on the sequence. Let $(y_n)_{n\in\mathbb{N}}$ and $(z_n)_{n\in\mathbb{N}}$ two sequences in $\mathcal{A}$ such that $y_n \to x_0$, $z_n \to x_0$ and $\tilde{V}''(y_n) \to Q_1$, $\tilde{V}''(z_n) \to Q_2$ with $Q_1, Q_2 \in [-\infty, 0]$, $Q_1 \neq Q_2$. Therefore, by the HJB equation (1.77)), we have

$$\rho \tilde{V}(y_n) = (ry_n - q)\,\tilde{V}'(y_n) + \tilde{U}(y_n) + \tilde{\mathcal{H}}^0\left(y_n, \tilde{V}'(y_n), \tilde{V}''(y_n)\right),$$

$$\rho \tilde{V}(z_n) = (rz_n - q)\,\tilde{V}'(z_n) + \tilde{U}(z_n) + \tilde{\mathcal{H}}^0\left(z_n, \tilde{V}'(z_n), \tilde{V}''(z_n)\right),$$

so passing to the limit we get for $i = 1, 2$

$$\rho \tilde{V}(x_0) = (rx_0 - q)\,\tilde{V}'(x_0) + \tilde{U}(x_0) + \tilde{\mathcal{H}}^0\left(x_0, \tilde{V}'(x_0), Q_i\right)$$

with the formal agreement that $\tilde{\mathcal{H}}^0\left(x_0, \tilde{V}'(x_0), -\infty\right) = 0$. Since $\tilde{V}'(x_0) > 0$, in this way $\tilde{\mathcal{H}}^0\left(x_0, \tilde{V}'(x_0), Q\right)$ is injective as function of $Q \in [-\infty, 0]$, so that we get what claimed.

Now we prove that such limit Q can never be $-\infty$. Assume by contradiction that for a given $x_0 \in (l, +\infty)$ we have $\tilde{V}''(x_n) \to -\infty$, for every sequence $\{x_n\}_{n\in\mathbb{N}} \subset \mathcal{A}$ such that $x_n \to x_0$. Take the function defined in $(l, +\infty)$ by

$$g(x) := \frac{1}{\rho}\left[\tilde{U}(x) + \tilde{V}'(x)(rx - q)\right].$$

For every $x \in (l, +\infty)$ we have $g(x) \leq \tilde{V}(x)$. Indeed, arguing as above and calling Q the limit of $\tilde{V}''(x_n)$ for every $x_n \to x$, $\{x_n\} \subset \mathcal{A}$, we get

$$\rho\tilde{V}(x) = (rx - q)\tilde{V}'(x) + \tilde{U}(x) + \tilde{\mathcal{H}}^0\left(x, \tilde{V}'(x), Q\right) \geq (rx - q)\tilde{V}'(x) + \tilde{U}(x),$$

where the inequality is strict on all the points of $\mathcal{A}$ and for points x such that $Q > -\infty$. Moreover $g(x_0) = \tilde{V}(x_0)$ because in the case of x_0 we are supposing $Q = -\infty$. Since $\tilde{V}$ is differentiable at x_0 we have

$$g(x) \leq \tilde{V}(x), \quad g(x_0) = \tilde{V}(x_0) \qquad \Longrightarrow \qquad \tilde{V}'(x_0) \in D^+g(x_0).$$

In particular this means that $D^+g(x_0) \neq \emptyset$. However, for every sequence $\{x_n\}_{n\in\mathbb{N}} \subset \mathcal{A}$ such that $x_n \to x_0^-$, we have that $\exists g'(x_n)$ and

$$g'(x_n) = \frac{1}{\rho}\left[\tilde{U}'(x_n) + \tilde{V}''(x_n)(rx_n - q) + r\tilde{V}'(x_n)\right],$$

so that

$$\lim_{n\to+\infty} g'(x_n) = -\infty.$$

This is a contradiction thanks to Lemma 1.3.24.

With this argument we have proved that $\tilde{V}''$, which exists almost everywhere on $(l, +\infty)$, can be extended to a continuous function h on the whole interval $(l, +\infty)$. Note that, differently from [Choulli, Taksar, Zhou; 2003], we cannot conclude that $\tilde{V}''$ exists on $(l, +\infty)$ and $\tilde{V}'' = h$. Indeed if $\tilde{V}'$ was the Cantor function, we would get a contradiction of such a conclusion at this stage. However we can say that, for any compact set $[a, b] \subset (l, +\infty)$, there exists $\delta_{a,b} > 0$ such that

$$\tilde{\mathcal{H}}^0(x, V'(x), h(x)) \geq \delta_{a,b}, \quad x \in [a, b].$$

Let us define the function

$$k(x) := \rho\tilde{V}(x) - (rx - q)\tilde{V}'(x) - \tilde{U}(x), \quad x \in (l, +\infty);$$

this function is equal to $\tilde{\mathcal{H}}^0(x, \tilde{V}'(x), \tilde{V}''(x))$ on $\mathcal{A}$ and moreover it is continuous on $[a, b]$, so that we have also

$$k(x) \geq \delta_{a,b}, \quad x \in [a, b]. \tag{1.90}$$

Note that

$$\left\{x \in [a,b] \,\Big|\, \frac{2k(x)}{\sigma\lambda x \tilde V'(x)} = 1\right\} = \left\{x \in [a,b] \,\Big|\, -\frac{\sigma\lambda x \tilde V'(x)}{2[k(x)-\sigma\lambda x \tilde V'(x)]} = 1\right\},$$

$$\left\{x \in [a,b] \,\Big|\, \frac{2k(x)}{\sigma\lambda x \tilde V'(x)} < 1\right\} \bigcap \left\{x \in [a,b] \,\Big|\, -\frac{\sigma\lambda x \tilde V'(x)}{2[k(x)-\sigma\lambda x \tilde V'(x)]} > 1\right\} = \emptyset,$$

$$\left\{x \in [a,b] \,\Big|\, \frac{2k(x)}{\sigma\lambda x \tilde V'(x)} \leq 1\right\} \bigcup \left\{x \in [a,b] \,\Big|\, -\frac{\sigma\lambda x \tilde V'(x)}{2[k(x)-\sigma\lambda x \tilde V'(x)]} \geq 1\right\} = [a,b].$$

Moreover

$$-\frac{\lambda^2 \tilde V'(x)^2}{2\,k(x)} = \frac{2[k(x)-\sigma\lambda x \tilde V'(x)]}{\sigma^2 x^2} \tag{1.91}$$

on the set

$$\left\{x \in [a,b] \,\Big|\, \frac{2k(x)}{\sigma\lambda x \tilde V'(x)} = 1\right\} = \left\{x \in [a,b] \,\Big|\, -\frac{\sigma\lambda x \tilde V'(x)}{2[k(x)-\sigma\lambda x \tilde V'(x)]} = 1\right\}. \tag{1.92}$$

Therefore the function $f : [a,b] \to \mathbb{R}$

$$f(x) := \begin{cases} -\dfrac{\lambda^2 \tilde V'(x)^2}{2\,k(x)}, & \text{if } \quad \dfrac{2k(x)}{\sigma\lambda x \tilde V'(x)} \leq 1, \\[2ex] \dfrac{2[k(x)-\sigma\lambda x \tilde V'(x)]}{\sigma^2 x^2}, & \text{if } \quad -\dfrac{\sigma\lambda x \tilde V'(x)}{2[k(x)-\sigma\lambda x \tilde V'(x)]} \geq 1, \end{cases}$$

is well-defined on $[a,b]$; moreover it is continuous (thanks to (1.91) and (1.92)) and negative (thanks to (1.90)). Let us consider the equation

$$v''(x) = f(x), \;\; x \in (a,b),$$

with boundary conditions

$$v(a) = \tilde V(a), \;\; v(b) = \tilde V(b). \tag{1.93}$$

This equation admits a unique $C^2([a,b];\mathbb{R})$ solution W satisfying the boundary conditions (1.93). One could check that W is a viscosity solution to (1.77) satisfying the boundary conditions (1.93). Actually, by the standard theory of viscosity solutions for elliptic equations, W is the unique viscosity solution to (1.77) satisfying the boundary conditions (1.93) (see, e.g., Theorem 3.3 in [Crandall, Ishii, Lions; 1992]; conditions (3.13), (3.14) therein can be easily proved for our equation: see also the proof of Lemma 7.1, Chapter 4, of [Fleming, Soner; 1993]). Since also $\tilde V$ is a viscosity solution to (1.77) with boundary conditions (1.93), we have $\tilde V = W$, so that $\tilde V \in C^2([a,b];\mathbb{R})$. By the arbitrariness of the compact set $[a,b]$ the proof is complete. $\square$

Remark 1.3.26. The proof above uses some arguments taken from the proof of Theorem 2, pp. 1958–1960, in [Choulli, Taksar, Zhou; 2003] even if it needs new ideas since *here we do not have uniform ellipticity* of the second order term. In Theorem 5.1 of [Zariphopoulou; 1994], pp. 78–82, a similar regularity result is proved for a similar HJB equation with a different technique (restriction to bounded intervals where the equation is proved to be uniformly elliptic). In any case the result of [Zariphopoulou; 1994] cannot be applied as it is to this case. We also note that the above proof indeed states the C^2 interior regularity for every concave viscosity solution to HJB equation (1.77) in $(l, +\infty)$. □

Remark 1.3.27. We can get more regularity of the value function. In particular observe that from the HJB equation, for any $x_0 \in (l, +\infty)$, if

$$-\frac{\lambda \tilde{V}'(x_0)}{\sigma x_0 \tilde{V}''(x_0)} < 1,$$

then in a suitable neighborhood $I(x_0)$ of x_0

$$\tilde{V}''(x) = -\frac{\frac{\lambda^2}{2}\left[\tilde{V}'(x)\right]^2}{\rho\tilde{V}(x) - (rx - q)\,\tilde{V}'(x) - \tilde{U}(x)}, \tag{1.94}$$

so $\tilde{V}''$ is continuously differentiable in $I(x_0)$. Similarly if

$$-\frac{\lambda \tilde{V}'(x_0)}{\sigma x_0 \tilde{V}''(x_0)} > 1$$

(or even when $\tilde{V}''(x_0) = 0$), then in a suitable neighborhood $I(x_0)$

$$\tilde{V}''(x) = \frac{2}{\sigma^2 x^2}\left[\rho\tilde{V}(x) - (rx + \sigma\lambda x - q)\,\tilde{V}'(x) - \tilde{U}(x)\right],$$

so $\tilde{V}''$ is differentiable in $I(x_0)$. □

Now we give a Corollary which will be very useful in proving the verification theorem for the case $rl > q$ stated in next subsection.

Corollary 1.3.28. *The value function is strictly concave and satisfies*

$$\tilde{V}'(x) > 0, \qquad \tilde{V}''(x) < 0,$$

for $x \in (l, +\infty)$. Moreover, if $rl > q$ and $\tilde{U}(l) > -\infty$, we have

$$\tilde{V}''(x) \longrightarrow -\infty, \qquad \textit{when} \quad x \longrightarrow l^+, \tag{1.95}$$

and, if $U'(l^+)$ is finite,

$$(x - l)\left[\tilde{V}''(x)\right]^2 \longrightarrow \frac{\lambda^2\left[\tilde{V}'(l^+)\right]^2}{4(rl - q)}, \qquad \textit{when} \quad x \longrightarrow l^+. \tag{1.96}$$

Proof. The fact that $\tilde{V}'(x) > 0$ for $x > l$ comes from concavity and strict monotonicity. Now let us suppose by contradiction that there exists $x_0 > l$ such that $\tilde{V}''(x_0) = 0$. In this case the maximum of $\tilde{\mathcal{H}}^0_{cv}\left(x_0, \tilde{V}'(x_0), \tilde{V}''(x_0); \theta\right)$ is reached for $\theta = 1$, and so we have taking the HJB equation (1.77) for x in a sufficiently small neighborhood $I(x_0)$ of x_0

$$\rho\tilde{V}(x) - (rx - q)\tilde{V}'(x) - \tilde{U}(x) = \sigma\lambda x\tilde{V}'(x) + \frac{1}{2}\sigma^2 x^2 \tilde{V}''(x).$$

Call now, for $x \in I(x_0)$,

$$h(x) = \frac{1}{2}\sigma^2 x^2 \tilde{V}''(x) = \rho\tilde{V}(x) - (rx + \sigma\lambda x - q)\tilde{V}'(x) - \tilde{U}(x).$$

Clearly h has a local maximum at x_0 and is twice differentiable at x_0 thanks to Remark 1.3.27. So it must be $h'(x_0) = 0$ and $h''(x_0) \le 0$. Now

$$\begin{aligned} h'(x) &= (\rho - r - \sigma\lambda)\tilde{V}'(x) - \tilde{U}'(x) - \tilde{V}''(x)(rx + \sigma\lambda x - q) \\ h''(x) &= (\rho - 2r - 2\sigma\lambda)\tilde{V}''(x) - \tilde{U}''(x) - \tilde{V}'''(x)(rx + \sigma\lambda x - q) \end{aligned}$$

and therefore, using that $\tilde{V}''(x_0) = 0$, we obtain

$$\begin{aligned} h'(x_0) &= (\rho - r - \sigma\lambda)\tilde{V}'(x_0) - U'(x_0) \\ h''(x_0) &= -U''(x_0) - \tilde{V}'''(x_0)(rx_0 + \sigma\lambda x_0 - q). \end{aligned}$$

Since x_0 is also a maximum for $\tilde{V}''$, it is clearly $\tilde{V}'''(x_0) = 0$ and consequently, by Hypothesis 1.3.2-(i), $h''(x_0) = -U''(x_0) > 0$, a contradiction.

Suppose now that $rl > q$. We are going to prove (1.95) under the assumption $\tilde{U}(l) > -\infty$. Observe that, for $x > l$,

$$\rho\tilde{V}(x) - (rx - q)\tilde{V}'(x) - \tilde{U}(x) = \tilde{\mathcal{H}}^0\left(x, \tilde{V}'(x), \tilde{V}''(x)\right).$$

Recall that by Remark 1.3.15 we have that $\tilde{V}'(l^+)$ is finite. Take any sequence $x_n \to l^+$ such that $\tilde{V}''(x_n) \to Q \in [-\infty, 0]$. Then, passing to the limit for $n \to +\infty$ in the HJB equation above, we get

$$\rho\tilde{V}(l) - (rl - q)\tilde{V}'\left(l^+\right) - \tilde{U}(l) = \tilde{\mathcal{H}}^0\left(l, \tilde{V}'\left(l^+\right), Q\right). \tag{1.97}$$

On the other hand by concavity we know that, for $x \ge l$,

$$\tilde{V}(x) \le \tilde{V}(l) + \tilde{V}'\left(l^+\right)(x - l).$$

Let $\delta > 0$; applying Lemma 1.3.23 with the function

$$g(x) = \begin{cases} \tilde{V}(l) + \tilde{V}'(l^+)(x - l) & \text{when} \quad x \ge l, \\ \tilde{V}(l) + (\tilde{V}'(l^+) + \delta)(x - l) & \text{when} \quad x < l, \end{cases}$$

we find f_ε defined on $\mathbb{R}$ such that

- $f_\varepsilon(x) \geq \tilde{V}(l) + \tilde{V}'(l^+)(x-l)$ for $x \geq l$;
- $f_\varepsilon(l) = \tilde{V}(l)$;
- $f'_\varepsilon(l) = \tilde{V}'(l^+)$;
- $f''_\varepsilon(l) \leq -\varepsilon^{-1}$.

Then we have, for $x \geq l$,

$$0 = \tilde{V}(l) - f_\varepsilon(l) \geq \tilde{V}(x) - f_\varepsilon(x),$$

so that, being $\tilde{V}$ a subsolution to the HJB equation (1.77) at $x = l$,

$$\rho\tilde{V}(l) - (rl-q) f'_\varepsilon(l) - \tilde{U}(l) \leq \tilde{\mathcal{H}}^0\left(l, f'_\varepsilon(l), f''_\varepsilon(l)\right).$$

By the arbitrariness of ε this gives

$$\rho\tilde{V}(l) - (rl-q)\tilde{V}'\left(l^+\right) - \tilde{U}(l) = 0. \tag{1.98}$$

Therefore, comparing (1.97) and (1.98), we get

$$\tilde{\mathcal{H}}^0\left(l, \tilde{V}'\left(l^+\right), Q\right) = 0 \quad \Longrightarrow \quad Q = -\infty.$$

The claim follows by a standard argument on subsequences.

Finally we prove (1.96) under the assumption $\tilde{U}'(l^+) < +\infty$. First observe that, for x in a suitable right neighborhood of l, we must have as a consequence of (1.95)

$$\frac{\lambda\tilde{V}'(x)}{\sigma x\tilde{V}''(x)} < 1,$$

so that by (1.94) we obtain

$$(x-l)\left[\tilde{V}''(x)\right]^2 = \lambda^4\left[\tilde{V}'(x)\right]^4 \cdot \frac{(x-l)}{\left[\rho\tilde{V}(x) - (rx-q)\tilde{V}'(x) - \tilde{U}(x)\right]^2}$$

To calculate the limit of the second factor we use de l'Hôpital rule. The ratio of the derivatives is (using (1.94) to rewrite it)

$$\frac{1}{2\left[\rho\tilde{V}(x) - (rx-A)\tilde{V}'(x) - U(x)\right]\left[(\rho-r)\tilde{V}'(x) - (rx-A)\tilde{V}''(x) - U'(x)\right]}$$
$$= \frac{1}{4\lambda^2\left[\tilde{V}'(x)\right]^2} \cdot \frac{-\tilde{V}''(x)}{(\rho-r)\tilde{V}'(x) - (rx-A)\tilde{V}''(x) - U'(x)}.$$

Since $U'(l^+)$ is finite the limit of the second factor is clearly $\frac{1}{rl-q}$, so the claim is proved. $\square$

Remark 1.3.29. Both the convergences (1.95) and (1.96) come from the boundary condition for the HJB equation, i.e. the subsolution inequality at the boundary (which depends on the structure of the second order superdifferential at the boundary). In particular we can say that if $rl > q$ and $\tilde{U}(l) > -\infty$ the value function $\tilde{V}$ solves the HJB equation (1.77) on $[l, +\infty)$ with the usual agreement that $\tilde{\mathcal{H}}^0\left(l, \tilde{V}'(l^+), -\infty\right) = 0$.

Similar results can be proved in the case when $rl = q$, but we will not need them since in that case we will only study a special case where explicit solutions are available (Subsection 1.3.7).

Finally we note that Corollary 1.3.28 holds for every concave constrained viscosity solution to the equation (1.77) on $[l, +\infty)$. □

1.3.6 The verification theorem and the optimal policy when $rl > q$

In this subsection we prove a verification theorem and the existence of optimal feedbacks when $rl > q$ and $\tilde{U}(l)$, $\tilde{U}'(l^+)$ are finite. We start by a lemma.

Lemma 1.3.30. *Let $rl > q$ and let $\tilde{U}(l), \tilde{U}'(l)$ be finite. Set*

$$\tilde{G}(x) = \begin{cases} \left(-\dfrac{\lambda \tilde{V}'(x)}{\sigma x \tilde{V}''(x)}\right) \wedge 1, & \text{when} \quad x > l, \\ 0, & \text{when} \quad x \le l. \end{cases} \tag{1.99}$$

1. *For every $x \ge l$ the closed loop equation*

$$\begin{cases} dX(t) = \left[\left(r + \sigma\lambda \tilde{G}(X(t))\right) X(t) - q\right] dt + \sigma \tilde{G}(X(t)) X(t) dB(t), \quad t \ge T, \\ X(T) = x, \end{cases} \tag{1.100}$$

   has a unique strong solution $X_{\tilde{G}}(\cdot\,; T, x)$.

2. *For every $t \ge T$ we have $X_{\tilde{G}}(t; T, x) \ge l$ almost surely.*

3. *$X_{\tilde{G}}(\cdot; x, T) \in C_{\mathcal{P}}([T, +\infty); L^p(\Omega; \mathbb{P}))$ for every $p \ge 1$.*

Proof.

1. For every $t \ge T$, consider the probability $\tilde{\mathbb{P}}_t = \exp\left(-\lambda B(t) - \frac{1}{2}\lambda^2 (t - T)\right) \cdot \mathbb{P}$ and let $\tilde{B}(s) = B(s) + \lambda(s - T)$, $s \in [T, t]$, the Brownian motion with respect to $\tilde{\mathbb{P}}_t$ over $[T, t]$ given by Girsanov's Theorem. We will show that the following equation

$$\begin{cases} dX(\tau) = (rX(\tau) - q)\, dt + \sigma \tilde{G}(X(\tau)) X(\tau)\, d\tilde{B}(\tau), \\ X(T) = x. \end{cases} \tag{1.101}$$

admits a unique strong solution X on the probabiity space $(\Omega, \mathcal{F}_t, \tilde{\mathbb{P}}_t)$ over $[T, t]$. If this is true, then, by definition of $\tilde{B}$, we see that X is the unique strong solution to (1.100) on the probability space $(\Omega, \mathcal{F}, \mathbb{P})$ over $[T, t]$. Since this can be done for arbitrary $t \geq T$, we would get the claim.

So, let us show that, for every $t \geq T$, (1.101) has a unique strong solution on the space $(\Omega, \mathcal{F}_t, \tilde{\mathbb{P}}_t)$.

- *Existence of a weak solution to* (1.101). Note that $\tilde{G}$ is continuous since $\tilde{V}', \tilde{V}''$ are continuous and since $\tilde{V}''(l^+) = -\infty$ while $\tilde{V}'(l^+) < +\infty$. Moreover $\tilde{G}$ is clearly bounded. Thus applying Theorem 2.4, p. 163, of [Ikeda, Watanabe; 1981] we get the existence of a weak solution.

- *Pathwise uniqueness for* (1.101). We want to apply Yamada-Watanabe Theorem (see

  - [Yamada, Watanabe; 1971] or

  - Theorem 3.5-(ii), p. 390, of [Revuz, Yor; 1999] or

  - Proposition 2.13, p. 291, of [Karatzas, Shreve; 1991])

  with the weakened assumptions of [Revuz, Yor; 1999], Exercise 3.14, p. 397.

  – *Estimate in a right-neighborhood of l.* Note that by Remark 1.3.15, we know that $\tilde{V}'(l^+)$ is finite and that $rl > q$ implies $l > 0$. Therefore we have

  $$\tilde{G}(l) = 0 = \lim_{x \to l^+} \left(-\frac{\lambda \tilde{V}'(x)}{\sigma x \tilde{V}''(x)} \right)$$
  $$= \frac{2}{\sigma \lambda l \tilde{V}'(l^+)} [\rho \tilde{V}(l) - (rl - A)\tilde{V}'(l^+) - \tilde{U}(l)].$$

  So, by Remark 1.3.27, there exists $\varepsilon \in (0,1)$ such that, for every $x \in [l, l+\varepsilon)$,

  $$\tilde{G}(x) = -\frac{\lambda \tilde{V}'(x)}{\sigma x \tilde{V}''(x)} = \frac{2}{\sigma \lambda x \tilde{V}'(x)} \left[\rho \tilde{V}(x) - (rx - q)\tilde{V}'(x) - \tilde{U}(x) \right].$$

  Thus, taking into account that $\tilde{V}'(\cdot) \in [\tilde{V}'(l+\varepsilon), \tilde{V}'(l^+)]$ in the intercal $[l, l+\varepsilon)$ with $\tilde{V}(l^+), \tilde{V}'(l+\varepsilon) \in (0, +\infty)$, we can find

$K_1, K_2 > 0$ such that, for any $x, y \in [l, l+\varepsilon)$, $x \geq y$,

$$\begin{aligned}
\left|\tilde{G}(x)\,x - \tilde{G}(y)\,y\right| &= 2\left|\frac{\rho\tilde{V}(x) - (rx-q)\tilde{V}'(x) - \tilde{U}(x)}{\sigma\lambda\tilde{V}'(x)}\right. \\
&\qquad \left. - \frac{\rho\tilde{V}(y) - (ry-q)\tilde{V}'(y) - \tilde{U}(y)}{\sigma\lambda\tilde{V}'(y)}\right| \\
&\leq \left|\left[\frac{2}{\sigma\lambda\tilde{V}'(x)} - \frac{2}{\sigma\lambda\tilde{V}'(y)}\right]\left[\rho\tilde{V}(x) - (rx-q)\tilde{V}'(x) - \tilde{U}(x)\right]\right| \\
&\quad + \left|\frac{2}{\sigma\lambda\tilde{V}'(y)}\right|\left|\left[\rho\tilde{V}(x) - (rx-q)\tilde{V}'(x) - \tilde{U}(x)\right]\right. \\
&\qquad \left. - \left[\rho\tilde{V}(y) - (ry-q)\tilde{V}'(y) - \tilde{U}(y)\right]\right| \\
&\leq K_1\,|x-y| + K_2\left|\tilde{V}'(x) - \tilde{V}'(y)\right| \\
&\leq K_1|x-y| + K_2\left|\int_y^x \tilde{V}''(\xi)\,d\xi\right|,
\end{aligned}$$

so by (1.96) we get, for some $K_H > 0$, for any $x, y \in [l, l+\varepsilon)$,

$$\left|\tilde{G}(x)\,x - \tilde{G}(y)\,y\right|^2 \leq K_H\,|x-y|\,.$$

– *Estimate out of a right-neighborhood of l.* Arguing as before, for some $K_L > 0$, we get the following estimate on $[l+\varepsilon, +\infty)$, for every $x \in [l+\varepsilon, +\infty)$ and for every $y \in [x-1, x+1] \cap [l+\varepsilon, +\infty)$,

$$\begin{aligned}
&\left|\tilde{G}(x)\,x - \tilde{G}(y)\,y\right|^2 \\
&\quad = 4\left|\left(\frac{\rho\tilde{V}(x) - (rx-q)\tilde{V}'(x) - \tilde{U}(x)}{\sigma\lambda\tilde{V}'(x)} \wedge x\right)\right. \\
&\qquad \left. - \left(\frac{\rho\tilde{V}(y) - (ry-q)\tilde{V}'(y) - \tilde{U}(y)}{\sigma\lambda\tilde{V}'(y)} \wedge y\right)\right|^2 \\
&\leq K_L\left[1 + \left|\frac{\rho\tilde{V}(x) - (rx-q)\tilde{V}'(x) - \tilde{U}(x)}{\tilde{V}'(x)\tilde{V}'(x+1)}\right|^2 + \left|\frac{1}{\tilde{V}'(x+1)}\right|^2\right]|x-y|^2\,.
\end{aligned}$$

Recall that the drift coefficient of the equation is Lipschitz continuous and that G was set identically zero on the set $(-\infty, l]$. Therefore we can apply the cited result of [Revuz, Yor; 1999], using for the diffusion coefficient the estimate, holding for every $x \in [l+\varepsilon, +\infty)$ and for every $y \in [x-1, x+1] \cap [l, +\infty)$,

$$\left|\tilde{G}(x)x - \tilde{G}(y)y\right|^2 \leq \left[1 + g(x)[\tilde{G}(x)x]^2\right]\rho(|x-y|),$$

where

$$g(x) = \begin{cases} 0, & \text{on } [l, l+\varepsilon), \\ \frac{1}{[G(x)x]^2}\left[\left|\frac{\rho\tilde{V}(x)-(rx-q)\tilde{V}'(x)-\tilde{U}(x)}{\tilde{V}'(x)\tilde{V}'(x+1)}\right| + \left|\frac{1}{\tilde{V}'(x+1)}\right|\right]^2, & \text{on } [l+\varepsilon, +\infty), \end{cases}$$

$$\rho(r) = K_H|r| + K_L|r|^2.$$

So we get pathwise uniqueness for the equation.

2. The fact that $X_{\tilde{G}}(t) \geq l$ almost surely follows from the fact that $\tilde{G}(x) = 0$ when $x \leq l$; this can be proved for example arguing by contradiction.

3. Take the control
$$\theta_{\tilde{G}}(t) := \tilde{G}(X_{\tilde{G}}(t; T, x);$$
then we clearly have $X(t; T, x, \theta_{\tilde{G}}(\cdot)) = X_{\tilde{G}}(t; T, x)$, for every $t \geq T$. The claim follows then by Theorem 1.3.1. □

Remark 1.3.31. In this remark we want to analyze the behaviour of the "optimal" diffusion $X_{\tilde{G}}$ (actually we will see that it is optimal in the next result) at the boundary l. $X_{\tilde{G}}$ is the unique solution to the stochastic differential equation

$$dX(t) = [rX(t) - q + \sigma\tilde{G}(X(t))X(t)]dt + \sigma\tilde{G}\left(X\left(t\right)\right)X(t)dB(t).$$

On the other hand we can rewrite the diffusion setting $Y(t) := X(t) - l$, so that Y is the solution of the stochastic differential equation

$$dY(t) = [rY(t) + b + \Sigma(Y(t))]dt + \Sigma\left(Y\left(t\right)\right)dB(t),$$

where
$$b := rl - q, \quad \Sigma(y) := \sigma\tilde{G}(y + l) \cdot (y + l).$$

First of all notice that, since $b > 0$ and $\Sigma(0) = 0$, the boundary 0 is not absorbing for Y. Therefore, since Y is continuous, Lemma 19.8 of [Kallenberg; 1997] shows that the boundary 0 is instantaneous (or reflecting, see [Kallenberg; 1997], p. 380) for the diffusion Y and so the boundary l is reflecting for the diffusion $X_{\tilde{G}}$.
Moreover thanks to (1.96) we have

$$\lim_{y\to 0^+}\left|\frac{\Sigma(y)}{\sqrt{y}}\right| = \lim_{x\to l^+}\left|\frac{\sigma\tilde{G}(x)x}{\sqrt{x-l}}\right| = \lim_{x\to l^+}\left|\frac{\lambda\tilde{V}'(x)}{\tilde{V}''(x)\sqrt{x-l}}\right| = 2\sqrt{b}.$$

Following [Feller; 1952], straightforward computations show that Feller's classification of the boundary 0 for Y (and so of the boundary l for $X_{\tilde{G}}$) is determined by the behaviour at 0, for generic $\eta > 0$ fixed, of the integrals

$$u(y) = \int_y^\eta dz\left[\int_z^\eta \frac{2}{\Sigma(s)^2}e^{-B(s)}\right]e^{B(z)}, \; v(y) = \int_y^\eta dz\left[\int_z^\eta e^{B(s)}ds\right]\frac{2}{\Sigma(z)^2}e^{-B(z)},$$

where
$$B(z) = \int_z^\eta \frac{2(rv + b + \Sigma(v))}{\Sigma(v)^2}dv.$$

Taking into account that for any $\varepsilon > 0$ we can find a sufficiently small $\eta > 0$ such that

$$\sqrt{(4b-\varepsilon)y} \leq \Sigma(y) \leq \sqrt{(4b+\varepsilon)y}, \quad y \in [0,\eta],$$

straightforward computations show that $u(0) < +\infty$, $v(0) < +\infty$. This means that the boundary l is regular. The fact that $u(0) < +\infty$ means that the boundary l is accessible for $X_{\tilde{G}}$, i.e.

$$\mathbb{P}\{\exists\, t \geq T \mid X_{\tilde{G}}(t;T,x) = l\} > 0, \qquad \forall x > l,$$

see [Feller; 1954]. $\square$

Theorem 1.3.32. *Let $rl > q$ and $\tilde{U}(l)$, $\tilde{U}'(l)$ be finite. Then, for every $x \geq l$, the control strategy $\theta^*(\cdot) \in \tilde{\Theta}_{ad}(x)$ such that*

$$\theta^*(t) = \tilde{G}\left(X_{\tilde{G}}(t;T,x)\right),$$

where $\tilde{G}$ is given by (1.99) *and $X_{\tilde{G}}(\cdot;T,x)$ is the unique strong solution to* (1.100), *is the unique optimal strategy at x.*

Proof. We cannot proceed with the standard proof of the verification theorem in the regular case (see on this, e.g., [Yong, Zhou; 1999], p. 268) since the function $\tilde{V}$ is not C^2 up to the boundary. Thus we use an approximation procedure. Given any $\varepsilon > 0$ we define a function $\tilde{V}_\varepsilon \in C^2(\mathbb{R})$ such that

- $\tilde{V}_\varepsilon(x) = \tilde{V}(x)$ in $[l+\varepsilon, +\infty)$;
- $\tilde{V}_\varepsilon(x) = a_1 + b_1 x + c_1 x^2$ in $\left(\frac{l+\frac{q}{r}}{2}, l+\varepsilon\right)$, where

$$\begin{aligned}
c_1 &= \frac{1}{2}\tilde{V}''(l+\varepsilon),\\
b_1 &= \tilde{V}'(l+\varepsilon) - \tilde{V}''(l+\varepsilon)(l+\varepsilon),\\
a_1 &= \tilde{V}(l+\varepsilon) - \tilde{V}'(l+\varepsilon)(l+\varepsilon) + \frac{1}{2}\tilde{V}''(l+\varepsilon)(l+\varepsilon)^2;
\end{aligned}$$

- $\tilde{V}_\varepsilon'(x) \geq 0$ in $\mathbb{R}$ and $\tilde{V}_\varepsilon'(x) = 0$ for $x \leq \frac{q}{r}$.

To define $\tilde{V}_\varepsilon$ on $\left[\frac{q}{r}, \frac{l+\frac{q}{r}}{2}\right]$ it is enough to take a suitable third degree polynomial. Since for $x \in [l, l+\varepsilon]$

$$\begin{aligned}
\tilde{V}_\varepsilon(x) - \tilde{V}(x) &= \tilde{V}(l+\varepsilon) - \tilde{V}(x) - \tilde{V}'(l+\varepsilon)(l+\varepsilon-x) + \frac{1}{2}\tilde{V}''(l+\varepsilon)(l+\varepsilon-x)^2\\
\tilde{V}_\varepsilon'(x) - \tilde{V}'(x) &= \tilde{V}'(l+\varepsilon) - \tilde{V}'(x) - \tilde{V}''(l+\varepsilon)(l+\varepsilon-x)\\
\tilde{V}_\varepsilon''(x) - \tilde{V}''(x) &= \tilde{V}''(l+\varepsilon) - \tilde{V}''(x),
\end{aligned}$$

using that $\varepsilon\tilde{V}''(l+\varepsilon) \to 0$ as $\varepsilon \to 0$ (see (1.96)), we have

$$\tilde{V}_\varepsilon \longrightarrow \tilde{V}, \qquad \tilde{V}_\varepsilon' \longrightarrow \tilde{V}', \qquad \text{uniformly in } [l,+\infty), \tag{1.102}$$

and

$$\tilde{\mathcal{H}}^0\left(x, \tilde{V}_\varepsilon'(x), \tilde{V}_\varepsilon''(x)\right) \longrightarrow \tilde{\mathcal{H}}^0\left(x, \tilde{V}'(x), \tilde{V}''(x)\right), \text{ uniformly in } [l,+\infty). \tag{1.103}$$

We claim that $\tilde{V}_\varepsilon$ solves in $\mathbb{R}$ the HJB equation

$$\rho\tilde{V}_\varepsilon(x) - (rx-q)\tilde{V}_\varepsilon'(x) - \tilde{\mathcal{H}}^0\left(x, \tilde{V}_\varepsilon'(x), \tilde{V}_\varepsilon''(x)\right) = g_\varepsilon(x) \tag{1.104}$$

where $g_\varepsilon \to \tilde{U}$ uniformly in $[l,+\infty)$ as $\varepsilon \to 0$ while $g_\varepsilon(x) \to -\infty$ for every $x < l$.

Indeed, (1.102), (1.103) and Remark 1.3.29 imply immediately that $g_\varepsilon \to \tilde{U}$ uniformly in $[l,+\infty)$. Moreover it is clear by its definition that $\tilde{V}_\varepsilon(x) \to -\infty$ for every $x < l$ and that $\tilde{V}_\varepsilon'(x) = 0$ for every $x < \frac{q}{r}$. Since $\tilde{V}_\varepsilon'(x) \geq 0$ and $\tilde{\mathcal{H}}^0\left(x, \tilde{V}_\varepsilon'(x), \tilde{V}_\varepsilon''(x)\right) \geq 0$, then we have

$$g_\varepsilon(x) \leq \rho\tilde{V}_\varepsilon(x), \qquad \forall x < l,$$

and so the claim.

Take $x \geq l$, $\theta(\cdot) \in \tilde{\Theta}_{ad}(x)$ and set $X(t) = X(t;T,x,\theta(\cdot))$. Consider the function

$$(t,x) \longmapsto e^{-\rho(t-T)}\tilde{V}_\varepsilon(x). \tag{1.105}$$

Since $\tilde{V}$ is concave, by construction $\tilde{V}_\varepsilon'$ is bounded. Then, taking into account that also $\theta(\cdot)$ is bounded and that $X \in C_{\mathcal{P}}([T,+\infty); L^2(\Omega,\mathbb{P}))$ (Theorem 1.3.1), we get that the process

$$t \mapsto \int_T^t \sigma\theta(s)X(s)\tilde{V}_\varepsilon'(X(s))dB(s)$$

is a martingale. Therefore we can apply Dynkin's formula to $X(\cdot)$ with the function (1.105), getting, for $t_1 \geq T$,

$$\begin{aligned}\mathbb{E}&\left[e^{-\rho(t_1-T)}\tilde{V}_\varepsilon(X(t_1)) - \tilde{V}_\varepsilon(x)\right] \\ &= \mathbb{E}\left[\int_T^{t_1} e^{-\rho(t-T)}\left[-\rho\tilde{V}_\varepsilon(X(t)) + [\mathcal{L}^{\theta(t)}\tilde{V}_\varepsilon](X(t))\right]dt\right],\end{aligned}$$

so by (1.104)

$$\begin{aligned}\mathbb{E}&\left[e^{-\rho(t_1-T)}\tilde{V}_\varepsilon(X(t_1)) - \tilde{V}_\varepsilon(x)\right] \\ &= \mathbb{E}\Big[\int_T^{t_1} e^{-\rho(t-T)}\Big[-g_\varepsilon(X(t)) - \tilde{\mathcal{H}}^0\left(X(t), \tilde{V}_\varepsilon'(X(t)), \tilde{V}_\varepsilon''(X(t))\right) \\ &\qquad - (rX(t)-q)\tilde{V}_\varepsilon'(X(t)) + [\mathcal{L}^{\theta(t)}\tilde{V}_\varepsilon](X(t))\Big]dt\Big],\end{aligned}$$

which implies

$$\begin{aligned}
\tilde{V}_\varepsilon(x) &= \mathbb{E}\left[\int_T^{t_1} e^{-\rho(t-T)} g_\varepsilon(X(t))\,dt + e^{-\rho(t_1-T)}\tilde{V}_\varepsilon(X(t_1))\right] \\
&\quad + \mathbb{E}\Big[\int_T^{t_1} e^{-\rho(t-T)}\Big[\tilde{\mathcal{H}}^0\left(X(t), \tilde{V}'_\varepsilon(X(t)), \tilde{V}''_\varepsilon(X(t))\right) \\
&\qquad\qquad - \tilde{\mathcal{H}}^0_{cv}\left(X(t), \tilde{V}'_\varepsilon(X(t)), \tilde{V}''_\varepsilon(X(t)); \theta(t)\right)\Big]dt\Big].
\end{aligned}$$

Sending $t_1 \to +\infty$ we get $\mathbb{E}\left[e^{-\rho(t_1-T)}\tilde{V}_\varepsilon(X(t_1))\right] \to 0$ by using (1.62), the last statement of Proposition 1.3.11 and estimating $\mathbb{E}\left[X(t)^\beta\right]$ as in the proof of Proposition 1.3.7. Therefore, by dominated convergence for the term with g_ε (recall that $g_\varepsilon \to \tilde{U}$ uniformly on $[l, +\infty)$ and that $\tilde{U}$ is finite at l and satisfies (1.62)) and by monotone convergence for the term with $\tilde{\mathcal{H}}^0 - \tilde{\mathcal{H}}^0_{cv}$ (note that $\tilde{\mathcal{H}}^0 \geq \tilde{\mathcal{H}}^0_{cv}$),

$$\begin{aligned}
\tilde{V}_\varepsilon(x) &= \mathbb{E}\left[\int_T^{+\infty} e^{-\rho(t-T)} g_\varepsilon(X(t))\,dt\right] \\
&\quad + \mathbb{E}\Big[\int_T^{+\infty} e^{-\rho(t-T)}\Big[\tilde{\mathcal{H}}^0\left(X(t), \tilde{V}'_\varepsilon(X(t)), \tilde{V}''_\varepsilon(X(t))\right) \\
&\qquad\qquad - \tilde{\mathcal{H}}^0_{cv}\left(X(t), \tilde{V}'_\varepsilon(X(t)), \tilde{V}''_\varepsilon(X(t)); \theta(t)\right)\Big]dt\Big]. \qquad (1.106)
\end{aligned}$$

Now we take $\theta(\cdot) \in \tilde{\Theta}_{ad}(x)$ and send $\varepsilon \to 0^+$ in the above formula. We have by the proof above

$$\tilde{V}_\varepsilon(x) \longrightarrow \tilde{V}(x), \qquad \mathbb{E}\left[\int_T^{+\infty} e^{-\rho(t-T)} g_\varepsilon(X(t))\,dt\right] \longrightarrow \tilde{J}(x;\theta(\cdot)),$$

and

$$\begin{aligned}
&\mathbb{E}\left[\int_T^{+\infty} e^{-\rho(t-T)}\tilde{\mathcal{H}}^0\left(X(t), \tilde{V}'_\varepsilon(X(t)), \tilde{V}''_\varepsilon(X(t))\right)dt\right] \longrightarrow \\
&\qquad\qquad \mathbb{E}\left[\int_T^{+\infty} e^{-\rho(t-T)}\tilde{\mathcal{H}}^0\left(X(t), \tilde{V}'(X(t)), \tilde{V}''(X(t))\right)dt.\right]
\end{aligned}$$

This means that also the limit

$$\lim_{\varepsilon\to 0^+}\mathbb{E}\left[\int_T^{+\infty} e^{-\rho(t-T)}\tilde{\mathcal{H}}^0_{cv}\left(X(t), \tilde{V}'_\varepsilon(X(t)), \tilde{V}''_\varepsilon(X(t)); \theta(t)\right)dt\right]$$

exists. Take now the closed loop strategy

$$\theta^x_{\tilde{G}}(t) = \tilde{G}\left(X_{\tilde{G}}(t;T,x)\right),$$

where $\tilde{G}$ is given by (1.99) and $X_{\tilde{G}}(\cdot;T,x)$ is the unique strong solution to (1.100). If we prove that, setting

$$X_{\tilde{G}}(t) := X_{\tilde{G}}(t;T,x), \quad \tilde{\mathcal{H}}^0\left(l, \tilde{V}'(l), V''(l)\right) := 0,$$

we have

$$\lim_{\varepsilon\to 0^+}\mathbb{E}\left[\int_T^{+\infty} e^{-\rho(t-T)}\tilde{\mathcal{H}}^0_{cv}\left(X_{\tilde{G}}(t),\tilde{V}'_\varepsilon(X_{\tilde{G}}(t)),\tilde{V}''_\varepsilon(X_{\tilde{G}}(t));\theta^*(t)\right)dt\right]$$
$$=\mathbb{E}\left[\int_T^{+\infty} e^{-\rho(t-T)}\tilde{\mathcal{H}}^0\left(X_{\tilde{G}}(t),\tilde{V}'(X_{\tilde{G}}(t)),\tilde{V}''(X_{\tilde{G}}(t))\right)dt\right], \quad (1.107)$$

then, passing to the limit in (1.106), we obtain

$$\tilde{V}(x)=\tilde{J}\left(x;\theta^x_{\tilde{G}}(\cdot)\right),$$

and so the optimality of $\theta^x_{\tilde{G}}(\cdot)$. To prove (1.107) it is enough to observe that

$$\tilde{\mathcal{H}}^0_{cv}\left(x,\tilde{V}'_\varepsilon(x),\tilde{V}''_\varepsilon(x);\tilde{G}(x)\right)\longrightarrow$$
$$\longrightarrow\ \tilde{\mathcal{H}}^0\left(x,\tilde{V}'(x),\tilde{V}''(x)\right)=\begin{cases}\tilde{\mathcal{H}}^0_{cv}\left(x,\tilde{V}'(x),\tilde{V}''(x);\tilde{G}(x)\right), & \text{if } x>l,\\ 0, & \text{if } x=l,\end{cases} \quad (1.108)$$

uniformly as $\varepsilon\to 0^+$. Indeed $\tilde{G}(l)=0$, so that the left-handside in (1.108) is equal to 0 for $x=l$. Moreover, for $x\in(l,+\infty)$,

$$\tilde{\mathcal{H}}^0_{cv}\left(x,\tilde{V}'_\varepsilon(x),\tilde{V}''_\varepsilon(x);\tilde{G}(x)\right)-\tilde{\mathcal{H}}^0_{cv}\left(x,\tilde{V}'(x),\tilde{V}''(x);\tilde{G}(x)\right)$$
$$=\sigma\lambda\tilde{G}(x)\,x\left[\tilde{V}'_\varepsilon(x)-\tilde{V}'(x)\right]+\frac{1}{2}\sigma^2\tilde{G}(x)^2x^2\left[\tilde{V}''_\varepsilon(x)-\tilde{V}''(x)\right].$$

The first term goes to 0 uniformly as $\varepsilon\to 0^+$ thanks to (1.102) while the second is, for ε sufficiently small and $x\in(l,l+\varepsilon)$ (for $x\geq l+\varepsilon$ it is zero),

$$\frac{1}{2}\lambda^2\left(\tilde{V}'(x)\right)^2\frac{\tilde{V}''_\varepsilon(x)-\tilde{V}''(x)}{\left[\tilde{V}''(x)\right]^2}.$$

Since

$$\frac{\tilde{V}''_\varepsilon(x)-\tilde{V}''(x)}{\left[\tilde{V}''(x)\right]^2}$$

is negative and greater than $\left[\tilde{V}''(l+\varepsilon)\right]^{-1}$, the convergence (1.108) follows; so we get (1.107) and the optimality of $\theta^*(\cdot)$.

The uniqueness follows from the strict concavity of $\tilde{U}$ arguing as in the proof of Proposition 1.3.12: one takes two different optimal strategies $\theta_1(\cdot)$ and $\theta_2(\cdot)$ with corresponding trajectories $X_1(\cdot)$ and $X_2(\cdot)$ and one proves that for any $\gamma\in[0,1]$ there exists an admissible strategy θ_γ whose associated trajectory is $\gamma X_1(\cdot)+(1-\gamma)X_2(\cdot)$. Then the strict concavity of $\tilde{U}$ implies that $\tilde{J}(x,\theta_\gamma(\cdot))>\gamma\tilde{J}(x,\theta_1(\cdot))+(1-\gamma)\tilde{J}(x,\theta_2(\cdot))=\tilde{V}(x)$, a contradiction. □

Remark 1.3.33. If $rl > q$ and $\tilde{U}(x) = \gamma^{-1}\left(x - \frac{q}{r}\right)^{\gamma}$ then, arguing as in the proof of Proposition 1.3.34, one can see that the function

$$v(x) = \gamma^{-1}\left(\rho - \gamma r - \frac{\lambda^2\gamma}{2(1-\gamma)}\right)^{-1}\left(x - \frac{q}{r}\right)^{\gamma}$$

is a regular solution to the HJB equation (1.77) in $(l, +\infty)$ when $\lambda \leq \sigma(1-\gamma)$. However this function v is not a constrained viscosity solution since it does not satisfy (1.95) that comes from the boundary condition (see on this Remark 1.3.29), and so v is not the value function.

In next section we will see that we have $v = \tilde{V}$ when $rl = q$. Therefore, from the arguments of next section, it follows that when $rl > q$ the function v is the value function if the state constraint $x \geq l$ is replaced by $x \geq \frac{q}{r}$, so we clearly have $v \geq \tilde{V}$.

Finally we observe that the proof of the above Theorem 1.3.32 works if $\tilde{V}$ is replaced by any concave constrained viscosity solution to equation (1.77) in $[l, +\infty)$. So, as a byproduct of this theorem, we get that the value function is the unique concave constrained viscosity solution to equation (1.77) in $[l, +\infty)$. □

1.3.7 An example when $rl = q$ with explicit solution

In the case of $rl = q$ and $\tilde{U}'(l^+) = +\infty$ it is possible to prove a general verification theorem on the line of Theorem 1.3.32. We do not do it here for brevity but we study a special case where, differently from the case $rl > q$, the explicit form of the value function and of the optimal couples is available. The utility function is given by

$$\tilde{U}(x) = \frac{(x-l)^{\gamma}}{\gamma}, \qquad \gamma \in (-\infty, 0) \cup (0, 1). \tag{1.109}$$

This utility function is defined for any $x \geq l$, if $\gamma \in (0,1)$, and for any $x > l$ if $\gamma \in (-\infty, 0)$; therefore the set of admissible strategies is never empty thanks to Lemma 1.3.4. Moreover it always satisfies Hypothesis 1.3.2. Notice that, considering the utility as a function of $x - l$, the above specification represents constant relative risk aversion preferences. The case of logarithmic utility may be treated in the same way but we do not do it for brevity.

As in Subsection 1.2.8 we look for a solution to the HJB equation (1.77) of the form

$$v(x) = C\frac{(x-l)^{\gamma}}{\gamma}, \qquad \gamma \in (-\infty, 0) \cup (0, 1), \tag{1.110}$$

for a suitable constant C. Substituting into the HJB equation (1.77) we see that it must be

$$C=\left(\rho-\gamma r-\frac{\lambda^2\gamma}{2\left(1-\gamma\right)}\right)^{-1}, \tag{1.111}$$

under the conditions

$$\rho>\gamma r+\frac{\lambda^2\gamma}{2(1-\gamma)}, \qquad \lambda\leq\sigma\left(1-\gamma\right). \tag{1.112}$$

The first condition is necessary in order to grant the finiteness of the value function. Indeed for $\gamma\in(-\infty,0)$ this finiteness is obvious, while for $\gamma\in(0,1)$ this condition guarantees (1.62) and so the finiteness of the value function. The second condition, as in Section 1.2.8, guarantees that the maximum point in the Hamiltonian is smaller than 1, so the no borrowing constraint is never active: this allows to keep $\tilde{\mathcal{H}}^0$ in the form which is suitable to find the explicit solution. When $\lambda>\sigma(1-\gamma)$ it is not difficult to see that $\tilde{V}(x)<v(x)$ for any $x>l$ using the fact that v is the value function of a problem with larger control set whose optimal trajectory is not admissible for our problem.

Define the feedback map

$$G(x)=\frac{\lambda}{\sigma\left(1-\gamma\right)}\frac{x-l}{x}; \tag{1.113}$$

the associated closed loop equation is

$$\begin{cases} dX\left(t\right)=\left(\dfrac{\lambda^2}{1-\gamma}+r\right)\left(X\left(t\right)-l\right)dt+\frac{\lambda}{1-\gamma}\left(X\left(t\right)-l\right)dB\left(t\right), \\ X(T)=x. \end{cases} \tag{1.114}$$

Such equation is linear and has a unique strong solution $X_G\left(\cdot;T,x\right)$. Moreover $X_G\left(\cdot;T,x\right)>l$ almost surely.

The main result of this subsection is the following.

Theorem 1.3.34. *Let conditions* (1.112) *be verified and the utility function $\tilde{U}$ be given by* (1.109) *with $\gamma\in(-\infty,0)\cup(0,1)$. Then*

(i) v given in (1.110)*, with C given by* (1.111)*, is the value function, i.e.*

$$\tilde{V}\left(x\right)=\gamma^{-1}\left[\rho-\gamma r-\frac{\lambda^2\gamma}{2\left(1-\gamma\right)}\right]^{-1}\left(x-l\right)^{\gamma}, \quad x\geq l \quad (x>l \text{ when } \gamma<0);$$

(ii) the feedback control

$$\theta^x_{\tilde{G}}(t):=\frac{\lambda}{\sigma\left(1-\gamma\right)}\cdot\frac{X_G(t;T,x)-l}{X_G(t;T,x)} \tag{1.115}$$

is admissible and optimal for the problem.

Proof. Note that the function v is not smooth up to the boundary also in this case. Nevertheless, since the boundary is abosorbing in this case, we can give a quite standard proof of the verification.

If $\gamma \in (0,1)$ and $x = l$, then the only admissible strategy is $\theta(\cdot) \equiv 0$, so that

$$v(l) = \frac{\tilde{U}(l)}{\rho} = \tilde{J}(l,0) = \tilde{V}(l).$$

So let us suppose $x > l$. We know that the function v given in (1.110) satisfies the following HJB equation

$$\rho v(y) - \frac{(y-l)^{\gamma}}{\gamma} - v'(y)\,(ry - A) - \tilde{\mathcal{H}}^0\left(y, v'(y), v''(y)\right) = 0, \quad y > l. \quad (1.116)$$

Take $\theta(\cdot) \in \tilde{\Theta}_{ad}(x)$ with the associated state trajectory $X(\cdot) := X(\cdot; T, x, \theta(\cdot))$. Define, with the convention $\inf \emptyset = +\infty$,

$$\tau_l := \inf\{t \geq T \mid X(t) = l\}, \quad \tau_{l+\varepsilon} := \inf\{t \geq T \mid X(t) = l + \varepsilon\},\ \varepsilon \in (0, x-l).$$

It is easy to see that $\tau_{l+\varepsilon} \uparrow \tau_l$ almost surely, when $\varepsilon \downarrow 0$. Note that v' is bounded on $[l+\varepsilon, +\infty)$ for every $\varepsilon \in (0,x)$, so that the process

$$t \mapsto \int_T^{t \wedge \tau_{l+\varepsilon}} \sigma\theta(s) X(s) v'(X(s)) dB(s)$$

is a martingale. Applying Dynkin's formula to the process $X(\cdot)$ with the function $(t,x) \mapsto e^{-\rho(t-T)} v(x)$, as in the proof of Theorem 1.3.32 we get, for any $t_1 > T$,

$$\begin{aligned} v(x) = \mathbb{E}\left[\int_T^{\tau_{l+\varepsilon}\wedge t_1} e^{-\rho(t-T)} \tilde{U}\left(X\left(t\right)\right) dt + e^{-\rho((\tau_{l+\varepsilon}\wedge t_1)-T)} v\left(X\left(\tau_{l+\varepsilon}\wedge t_1\right)\right)\right] \\ + \mathbb{E}\int_T^{\tau_{l+\varepsilon}\wedge t_1} e^{-\rho(t-T)} \big[\tilde{\mathcal{H}}^0\left(X(t), v'\left(X(t)\right), v''\left(X(t)\right)\right) \\ - \tilde{\mathcal{H}}^0_{cv}\left(X(t), v'\left(X(t)\right), v''\left(X(t)\right); \theta(t)\right)\big] dt. \quad (1.117) \end{aligned}$$

Since $\tilde{\mathcal{H}}^0 \geq \tilde{\mathcal{H}}^0_{cv}$, we can write, for any $t_1 > T$,

$$v(x) \geq \mathbb{E}\left[\int_T^{\tau_{l+\varepsilon}\wedge t_1} e^{-\rho(t-T)} U\left(X\left(t\right)\right) dt + e^{-\rho((\tau_{l+\varepsilon}\wedge t_1)-T)} v\left(X\left(\tau_{l+\varepsilon}\wedge t_1\right)\right)\right]. \quad (1.118)$$

Now we distinguish two cases.

- $\gamma < 0$. In this case, if $\theta(\cdot)$ is such that $\mathbb{P}\{\tau_l < +\infty\} > 0$, we clearly have

$$v(x) \geq -\infty = \tilde{J}(x; \theta(\cdot)).$$

Thus we can suppose that $\theta(\cdot)$ is such that $\tau_l = +\infty$ almost surely. Suppose now that $\theta(\cdot)$ is such that

$$\limsup_{t\to+\infty} \mathbb{E}\left[e^{-\rho(t-T)} v(X(t))\right] < 0.$$

This means that there exist $\delta > 0$, $\bar{t} > T$ such that, for every $t \geq \bar{t}$

$$\mathbb{E}\left[v(X(t))\right] \leq -\delta e^{\rho(t-T)}.$$

Since $\rho v(y) \geq \tilde{U}(y)$ for every $y > l$, we get, for every $t \geq \bar{t}$

$$\mathbb{E}\left[\tilde{U}(X(t))\right] \leq -\frac{\delta}{\rho} e^{\rho(t-T)}.$$

Since $\tilde{U} \leq 0$, this means that

$$\tilde{J}(x,\theta(\cdot)) \leq \int_{\bar{t}}^{+\infty} e^{-\rho(t-T)} \mathbb{E}\left[\tilde{U}(X(t))\right] dt \leq \int_{\bar{t}}^{+\infty} -\frac{\delta}{\rho} dt = -\infty \leq v(x).$$

Therefore we are reduced to suppose without loss of generality that $\theta(\cdot)$ is such that

$$\limsup_{t\to+\infty} \mathbb{E}\left[e^{-\rho(t-T)} v(X(t))\right] = 0.$$

Since $\tilde{U}$, v are negative, we can pass (1.118) to the limsup for $\varepsilon \downarrow 0$ getting by Fatou's Lemma

$$v(x) \geq \mathbb{E}\left[\int_T^{t_1} e^{-\rho(t-T)} \tilde{U}\left(X\left(t\right)\right) dt + e^{-\rho(t_1-T)} v\left(X\left(t_1\right)\right)\right]$$

Taking now the limsup for $t_1 \to +\infty$ we get

$$v(x) \geq \tilde{J}(x,\theta(\cdot)) + \limsup_{t_1\to+\infty} \mathbb{E}\left[e^{-\rho(t_1-T)} v(X(t_1))\right] = \tilde{J}(x,\theta(\cdot)).$$

By the arbitrariness of $\theta(\cdot)$ we get in this case $v(x) \geq \tilde{V}(x)$.

- $\gamma \in (0,1)$. Starting from (1.118), the passage to the limit for $\varepsilon \downarrow 0$ produces in this case by dominated convergence (due to the growth properties of $\tilde{U}, v$ and to the integrability properties of X)

$$v(x) \geq \mathbb{E}\left[I_{\{\tau_l=+\infty\}}\left[\int_T^{t_1} e^{-\rho(t-T)} \tilde{U}\left(X\left(t\right)\right) dt + e^{-\rho(t_1-T)} v\left(X\left(t_1\right)\right)\right]\right]$$
$$\mathbb{E}\left[I_{\{\tau_l<+\infty\}}\left[\int_T^{t_1\wedge\tau_l} e^{-\rho(t-T)} \tilde{U}\left(X\left(t\right)\right) dt + e^{-\rho((t_1\wedge\tau_l)-T)} v\left(X\left(t_1\wedge\tau_l\right)\right)\right]\right] \tag{1.119}$$

Note that on the set $\{\tau_l < +\infty\}$ we have

$$
\begin{aligned}
e^{-\rho(\tau_l - T)} v\left(X\left(\tau_l\right)\right) &= e^{-\rho(\tau_l - T)} v\left(l\right) = e^{-\rho(\tau_l - T)} \frac{\tilde{U}\left(l\right)}{\rho} \\
&= \int_{\tau_l}^{+\infty} e^{-\rho(t-T)} \tilde{U}\left(l\right) dt = \int_{\tau_l}^{+\infty} e^{-\rho(t-T)} \tilde{U}(X(t)) dt. \quad (1.120)
\end{aligned}
$$

So, letting $t_1 \to +\infty$ in (1.119) and taking into account (1.120), again we get by dominated convergence

$$
v(x) \geq \tilde{J}(x, \theta(\cdot)).
$$

Since $\theta(\cdot)$ is arbitrary, we get $v(x) \geq \tilde{V}(x)$.

On the other hand we see that, taking $\theta^x_{\tilde{G}}(\cdot)$ in (1.117), the second term in the right handside is zero, so that we get (1.118) with the equality. In this case the solution

$$
X_{\tilde{G}}(\cdot) := X(\cdot; T, x, \theta^x_{\tilde{G}}(\cdot)) = X_{\tilde{G}}(\cdot; T, x)
$$

is known (it is a geometric Brownian motion). We can check that, also in the case $\gamma < 0$, by the integrability properties of $X_{\tilde{G}}$ and the structure of U, v, every convergence works good with the limit in place of the limsup and with the equality in place of the inequality. So passing to the limit (1.118) for $\varepsilon \downarrow 0$, $t_1 \to +\infty$ we get

$$
v(x) = J_T(x, \theta^x_{\tilde{G}}(\cdot)),
$$

so that we have

$$
V(x) \leq v\left(x\right) = J_T(x, \theta^x_{\tilde{G}}(\cdot)) \leq V\left(x\right).
$$

Therefore we see that $v(x) = V(x)$ and that $\theta^*(\cdot)$ is optimal. □

Remark 1.3.35. In the case $rl = q$, $U'(l) = +\infty$, the problem could be faced by duality/martingale method, e.g. as in the papers [Bouchard, Pham; 2004] and [Blanchett-Scaillet, El Karoui, Jeanblanc, Martellini; 2008]. □

1.3.8 Analysis of the optimal policies

In this subsection we discuss the properties of the optimal policies described in Subsections 1.3.6 and 1.3.7 in the cases

- $rl > q$, $U(l)$ and $U'(l)$ finite;
- $rl = q$, $U'(l) = +\infty$.

First of all observe that in both cases the optimal feedback map is given by the function G of (1.99). When $G < 1$, this function is (similarly to the Merton model) the product of the payoff for every unit of risk $\frac{\lambda}{\sigma}$ and of the quantity $-\frac{\tilde{V}'}{x\tilde{V}''}$, i.e. the Arrow-Pratt measure of risk tolerance of the indirect utility function $\tilde{V}$ (the value function).

This implies that the optimal feedback map is increasing with the payoff per unit of risk and with the relative risk aversion of $\tilde{V}$, while the relation between the optimal policy and the level of wealth is known only implicitly, unless we know the explicit expression of $\tilde{V}$.

In the case of $rl > q$, $U(l)$ and $U'(l)$ finite, even taking a CRRA utility function the explicit form of the value function is not available. As seen in Remark 1.3.33 the natural candidate solution to the HJB equation does not satisfy the required boundary condition. This comes from the presence of the state constraints $x \geq l$ and from the fact that in this case the control $\theta(\cdot) \equiv 0$ bring the state from the boundary $x = l$ in the interior of the state region.

So, even starting from initial wealth equal to the solvency level l, the set of admissible strategies does not reduce to the trivial one (investment in the riskless asset forever) but allows to the fund manager to reinvest in the risky asset.

The possibility to exit from the boundary l (if the wealth process starts from or reaches it) is given by the fact that the capital amount l invested in the riskless asset will generate a return per unit of time rl. Hence the accrued return will produce disposable wealth to be invested in the risky asset and the wealth process can exit from the trivial state $X(\cdot) \equiv l$.

When $rl = q$, $U'(l) = +\infty$, and the utility function is in CRRA form (under constraints on the parameters), an explicit form of $\tilde{V}$ is available and it is exactly the natural candidate solution to the HJB equation. Indeed, here the situation at the boundary is different. The control $\theta(\cdot) \equiv 0$ leaves forever the state in the boundary $x = l$, so when the initial wealth x equals l the unique admissible allocation strategy is given by investing all the wealth in the riskless asset forever, and no risky investment is allowed. On the contrary, when initial wealth x is strictly greater than l the fund wealth will never reach the solvency level.

Concerning the case $rl = q$ treated in Subsection 1.3.7 the explicit form of the value function allows us to make a further consideration. According to the common sense, the portfolio selection rule (1.113) suggests to increase the fraction invested in the risky asset if the wealth level grows, and diminish the share invested in it if the fund level decreases. Indeed, this kind of policy seems to be reasonable with the social target of a pension fund, whose manager must be interested in protecting the wealth level and in caring about the risk the portfolio strategies involve.

Finally, we observe that within our model (whether the case of $rl > q$, $U(l)$ and $U'(l)$ finite, or the case of $rl = q$ and $U'(l) = +\infty$) we have similar results if we assume that the portfolio strategy $\theta(\cdot)$ belongs to $[0, \theta_0]$ with $\theta_0 < 1$, i.e. if the pension fund is forced not to invest the total amount of its wealth in the risky asset. Sometimes this constraint is imposed by the supervisory authority.

Chapter 2

Adding the surplus: an infinite-dimensional approach

In this chapter we extend the model described and studied in Chapter 1 adding a surplus term in the expression for the benefits. The main references for this chapter are the papers [Federico; 2008] and [Federico; WP]. The introduction of such a term is relevant from a financial point of view. Indeed usually the pension funds provide for their members also a surplus term as benefit besides a minimum guarantee term. Sometimes such a term depends on a contract subscribed in advance between the fund and the members; this contract can be viewed as a function depending on the past wealth of the fund in the "last" period: the higher is the performance of the fund in this period, the higher is the surplus paid to the fund members in retirement. From a mathematical point of view this unavoidably leads to a delay term in the state equation, making the problem considerably more difficult to treat. Delay problems have basically an infinite-dimensional nature. Sometimes the structure of the problem is such that it can be reduced to a finite-dimensional problem (see, e.g., for the stochastic case

- [Elsanousi, Øksendal, Sulem; 2000],
- [Larrsen; 2002],
- [Larrsen, Risebro; 2003],
- [Øksendal, Sulem; 2001],
- [Øksendal, Zhang; 2008]).

However this is not our case, so we choose to treat the problem by representation in infinite dimension (for this kind of approach we refer also to [Vinter, Kwong; 1981] in the deterministic case and to

- [Chojnowska-Michalik, 1978],
- [Da Prato, Zabczyk; 1996],
- [Federico, Øksendal; 2009],

- [Gozzi, Marinelli; 2004],

- [Gozzi, Marinelli, Savin],

in the stochastic case). This approach transforms the non-markovian problem with delay in an infinite-dimensional markovian problem, allowing to apply the dynamic programming techniques; the price to pay is that, as we said, the problem becomes infinite-dimensional.

The main results of this chapter are the rewriting of the problem in infinite dimension, the proof of the continuity of the value function in the infitine-dimensional setting and the proof that the value function is a constrained viscosity solution of the infinite-dimensional HJB equation in the special case when the boundary is absorbing. Due to the intrinsic difficulties of the problem, we have to stop our analysis just at a this viscosity stage.

2.1 The model with surplus

Within the model described in Section 1.1 and Section 1.3, we add a surplus term in the expression $b(\cdot)$ of the benefits flow. We use here the same notations as the ones used in Section 1.1 and Section 1.3, but we get rid of the subscript ˜ over the symbols appearing in Section 1.3. Moreover we will use the symbol $x(\cdot)$ in place of $X(\cdot)$ for the state variable: we will use the symbol $X(\cdot)$ to represent the state in the infinite-dimensional setting.

As we said, many pension funds provide for their members a surplus premium over the minimum guarantee. Very often the surplus contract is related to a performance index of the fund growth in the last period $[t-T,t]$: the idea behind is that the fund pays something more than the minimum guarantee to its member in retirement, if the fund growth was good in the period during which they were adhering to the fund. Therefore in general it is natural to set a contract which is mathematically represented by a path dependent function $S\left(t, x(\cdot)|_{[t-T,T]}\right)$. We choose as expression for the surplus term the function

$$S\left(t, x(\cdot)|_{[t-T,t]}\right) = f_0\big(x(t) - x(t-T)\big), \tag{2.1}$$

where $f_0 : \mathbb{R} \to [0,+\infty)$ is increasing, convex and Lipschitz continuous with Lipschitz constant K_0. With the form (2.1) for the surplus, the equation (1.2) for the wealth process $x(\cdot)$ becomes a stochastic delay differential equation:

$$dx(t) = \left[(r + \sigma\lambda\theta(t))\, x(t) - q\right] dt - f_0\left(x(t) - x(t-T)\right) dt + \sigma\theta(t)x(t)dB(t).$$

Remark 2.1.1. Of course other expressions for the surplus contract are possible. First of all we notice that without difficulties everything can be extended to the case of surplus function of this kind

$$S\left(t, x(\cdot)|_{[t-T,t]}\right) = f_0\big(x(t) - \kappa x(t-T)\big),$$

where $\kappa \geq 0$.

However, we stress that it is meaningful to consider a surplus term which is a function of the ratio $\frac{x(t)}{x(t-T)}$ rather than of the difference $x(t) - x(t-T)$, i.e.

$$S(t, x(\cdot)|_{[t-T,t]}) = f_0\left(\frac{x(t)}{x(t-T)}\right).$$

For example, referring to (1.4) for the model of the contributions flow, the cumulative benefit expression could be

$$b(t) = g + S\left(t, x(\cdot)|_{[t-T,t]}\right) = \bar{c}\cdot\alpha w \cdot \frac{e^{\left[\delta+\xi\left(\frac{1}{T}\log\frac{x(t)}{x(t-T)}-\delta\right)^+\right]T} - 1}{\delta + \xi\left(\frac{1}{T}\log\frac{x(t)}{x(t-T)} - \delta\right)^+},$$

where $\xi \in (0,1)$ is the retrocession rate, i.e. the share of the fund return exceeding δ awarded to the fund members. In this case the fund corresponds as return rate to its members in retirement a minimum guarantee rate δ plus a fraction $\xi \in (0,1)$ of the fund return rate (in the period during which these members were adhering to the fund) exceeding δ. This kind of contract leads to technical complications, due to existence problems for the state equation and, mainly, to the unavoidable loss of concavity of the value function, but it seems more meaningful from a financial point of view; we hope to investigate this contract in future works. □

Again the objective functional (expressing the total expected discounted utility coming from the wealth) which we want to maximize is given by

$$\mathbb{E}\left[\int_T^{+\infty} e^{-\rho(t-T)}U\left(x\left(t\right)\right)dt\right], \tag{2.2}$$

where $\rho > 0$ is the discount individual rate and U is a utility function satisfying Hypothesis 1.3.2.

Remark 2.1.2. In this case the manager's point of view in the optimization has a direct consequence on the benefits of the workers (and not only an indirect one as discussed in Subsection 1.1.5): since the surplus term depends on the performance of the fund, a good management of the fund with regard to the functional above will have a good correspective in the surplus payments. □

2.2 The stochastic control problem with delay

Now we come to a precise formulation of the problem. First of all notice that the initial time $t = T$ has been chosen as the first time of operations of the fund. However it also makes sense, in order to apply a dynamic programming approach, to look to a pension fund that is already running after a given amount of time $s \geq T$ so to establish an optimal decision policy from s on.

Consider the convex sets

$$\mathcal{C} := \left\{ \eta = (\eta_0, \eta_1(\cdot)) \in [l, +\infty) \times C([-T, 0); \mathbb{R}) \;\Big|\; \lim_{\zeta \to 0} \eta_1(\zeta) = \eta_0 \right\} \tag{2.3}$$

and

$$\mathcal{D} := \left\{ \eta \in \mathcal{C} \mid \eta_1(\cdot) \geq l(T + \cdot) \right\}, \tag{2.4}$$

where $l(\cdot)$ (the solvency level) is given according to (1.7). We have $\mathcal{D} \subset \mathcal{C} \subset E$, where

$$E := \left\{ (x_0, x_1(\cdot)) \in \mathbb{R} \times C([-T, 0); \mathbb{R}) \;\Big|\; \lim_{\zeta \to 0} x_1(\zeta) = x_0 \right\}. \tag{2.5}$$

The space E is a Banach space when endowed with the norm

$$\|(x_0, x_1(\cdot))\|_E = |x_0| + \sup_{\zeta \in [-T, 0)} |x_1(\zeta)|.$$

Dealing with a delay equation we have to specify not only the present, but also the past for the initial datum. We allow the initial past-present η belonging to the space $\mathcal{C}$ (see Remark 2.2.2 below). Now set $\eta \in \mathcal{C}$ and consider the following stochastic delay differential equation for the dynamics of the wealth

$$\begin{cases} dx(t) = \left[(r + \sigma\lambda\theta(t))\, x(t) - q \right] dt - f_0 \left(x(t) - x(t - T) \right) dt + \sigma\theta(t) x(t) dB^s(t), \\ x(s) = \eta_0, \; x(s + \zeta) = \eta_1(\zeta), \qquad \zeta \in [-T, 0). \end{cases} \tag{2.6}$$

Theorem 2.2.1. *For any $(\mathcal{F}_t^s)_{t \geq s}$-progressively measurable $[0, 1]$-valued process $\theta(\cdot)$*

- *equation* (2.6) *admits on the filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t^s)_{t \in [s, +\infty)}, \mathbb{P})$, a unique strong solution;*
- *this solution belongs to the space $C_{\mathcal{P}}([s, +\infty); L^p(\Omega, \mathbb{P}))$ of the p-mean continuous progressively measurable processes for any $p \in [1, +\infty)$.*

Proof. See Theorem 6.16, Chapter 1, in [Yong, Zhou; 1999]. □

We denote the unique strong solution to (2.6) by $x\,(t; s, x, \theta(\cdot))$.

Remark 2.2.2. Of course, from a financial point of view, the natural space of initial data for the initial time $s \geq T$ would be

$$\mathcal{D}_s := \{(\eta_0, \eta_1(\cdot)) \in [l, +\infty) \times C([-T, 0]; \mathbb{R}) \mid \lim_{\zeta \to 0} \eta_1(\zeta) = \eta_0,\ \eta_1(\cdot) \geq l(s + \cdot)\} \subset \mathcal{C},$$

which is, in particular, time-dependent. Nevertheless it makes sense (and it is convenient from a mathemathical point of view) to enlarge the set of initial data for initial time $s \geq T$ to the wider class $\mathcal{C}$ in order to consider a set of initial data not time-dependent. □

2.2.1 The set of admissible strategies

In the framework above we define the set of the admissible control strategies for initial time $s \geq T$ and initial past-present $\eta \in \mathcal{C}$ by

$$\Theta_{ad}(s, \eta) := \{\theta(\cdot) \text{ prog. meas. w.r.t. } (\mathcal{F}_t^s)_{t \geq s} \mid x(t; s, \eta, \theta(\cdot)) \geq l,\ \ t \geq s\}. \quad (2.7)$$

Lemma 2.2.3. *Let $\eta \in \mathcal{C}$; then $\Theta_{ad}(T, \eta) \neq \emptyset$ if and only if the null strategy $\theta(\cdot) \equiv 0$ (corresponding to the riskless investment of the whole wealth at every time) is admissible.*

Proof. Of course if $\theta(\cdot) \equiv 0$ belongs to $\Theta_{ad}(s, x)$, then $\Theta_{ad}(s, x) \neq \emptyset$. Again the proof of the converse implication is an application of the Girsanov Theorem A.1.1. Let $\theta(\cdot)$ be an admissible control for initial time T and initial past-present $\eta \in \mathcal{C}$; by Girsanov's Theorem, for $n \in \mathbb{N}$, $n \geq 2$, we can write the dynamics of $x(\cdot) := x(\cdot; T, \eta, \theta(\cdot))$ in the interval $[T, nT]$, under the probability $\tilde{\mathbb{P}}_{nT}$ given by Girsanov's Theorem, as

$$dx(t) = [rx(t) - q]\, dt - f_0(x(t) - x(t-T))dt + \sigma\theta(t)x(t)d\tilde{B}(t),$$

where $\tilde{B}(t) := B(t) + \lambda t$ is a Brownian motion under $\tilde{\mathbb{P}}_{nT}$ in the interval $[T, nT]$. Since $x(\cdot) \in C_{\mathcal{P}}([T, +\infty); L^2(\Omega))$, we can pass to the expectations getting (taking also into account the convexity of f_0 and Jensen's inequality)

$$\begin{aligned} d\tilde{\mathbb{E}}_{nT}[x(t)] &= \Big[r\tilde{\mathbb{E}}_{nT}[x(t)] - q\Big]\, dt - \tilde{\mathbb{E}}_{nT}\left[f_0(x(t) - x(t-T))\right] dt, \\ &\leq \Big[r\tilde{\mathbb{E}}_{nT}[x(t)] - q\Big]\, dt - f_0\left(\tilde{\mathbb{E}}_{nT}[x(t)] - \tilde{\mathbb{E}}_{nT}[x(t-T)]\right) dt, \end{aligned} \quad (2.8)$$

By assumption $x(t; T, \eta, \theta(\cdot)) \geq l$ for $t \geq T$, so that also $\tilde{\mathbb{E}}_{nT}[x(t; T, \eta, \theta(\cdot))] \geq l$ for $t \in [T, nT]$. But the dynamics of $y(\cdot) := x(\cdot; T, \eta, 0)$ is given by

$$dy(t) = [ry(t) - q]\, dt - f_0\left(y(t) - y(t-T)\right) dt, \quad (2.9)$$

Working in the interval $[T, 2T]$, we can see the delay term as a datum, so that (2.8) and (2.9) can be seen as in the theory of ordinary differential equations. Moreover

$$y(T) = \tilde{\mathbb{E}}_{nT}[x(T)] = \eta_0, \quad y(T+\zeta) = \tilde{\mathbb{E}}_{nT}[x(T+\zeta)] = \eta_1(\zeta), \;\; \zeta \in [-T, 0).$$

Thus we can apply the standard comparison criterion for ordinary differential equations to get

$$y(t) \geq \tilde{\mathbb{E}}_{nT}[x(t)] \geq l, \;\; \text{for } t \in [T, 2T]. \tag{2.10}$$

Then, in the interval $[2T, 3T]$, since f_0 is increasing, thanks to (2.10) we have

$$f_0\left(\tilde{\mathbb{E}}_{nT}[x(t)] - \tilde{\mathbb{E}}_{nT}[x(t-T)]\right) \geq f_0\left(\tilde{\mathbb{E}}_{nT}[x(t)] - y(t-T)\right),$$

so that

$$d\tilde{\mathbb{E}}_{nT}[x(t)] \leq \left[r\tilde{\mathbb{E}}_{nT}[x(t)] - q\right] dt - f_0\left(\tilde{\mathbb{E}}_{nT}[x(t)] - y(t-T)\right) dt. \tag{2.11}$$

Again we can see $y(t-T)$, $t \in [2T, 3T]$, as datum in (2.9) and (2.11). Taking also into account that, by (2.10), $y(2T) \geq \tilde{\mathbb{E}}_{nT}[x(2T)]$, we can again apply the standard comparison criterion for ordinary differential equation getting

$$y(t) \geq \tilde{\mathbb{E}}_{nT}[x(t)] \geq l, \;\; \text{for } t \in [2T, 3T].$$

Iterating the argument we get, for any $i = 1, ...n-1$,

$$y(t) \geq \tilde{\mathbb{E}}_{nT}[x(t; T, \eta, \theta(\cdot))] \geq l, \;\; \text{for } t \in [iT, (i+1)T].$$

The claim follows by the arbitrariness of n. □

Due to Lemma 2.2.5, we are led to introduce the following assumption.

Hypothesis 2.2.4.

$$f_0(l - l_0) \leq rl - q.$$

Lemma 2.2.5. *If Hypothesis 2.2.4 holds true, then the null strategy $\theta(\cdot) \equiv 0$ belongs to $\Theta_{ad}(T, \eta)$ for all $\eta \in \mathcal{D}$. In particular $\Theta_{ad}(T, \eta)$ is not empty for each $\eta \in \mathcal{D}$.*

Proof. Suppose that Hypotesis 2.2.4 holds true. Let us consider the state trajectory $x(\cdot)$ corresponding to the null strategy; at time t, supposing the constraint satified in the past and so in particular $x(t-T) \geq l_0$, we have, taking into account that f_0 is increasing,

$$\begin{aligned} dx(t) &= (rx(t) - q)dt - f_0(x(t) - x(t-T))dt \\ &\geq (rx(t) - q)dt - f_0(x(t) - l_0)dt. \end{aligned}$$

Thus, whenever $x(t) \le l$ (if there is the case), we have

$$dx(t) \ge (rl-q)dt - f_0(l - l_0)dt = (rl-q)dt - f_0(l-l_0)dt \ge 0.$$

Then a straight contradiction argument shows the claim. □

By Lemma 2.2.5 we will assume from now on that Hypothesis (2.2.4) holds true. Let us see the behaviour of the solution in some special cases: we give a lemma which will be useful afterwards.

Lemma 2.2.6. *Let $\eta \in \mathcal{D}$. Then:*
(i) *If $rl = q$ and $\eta_0 = l$, then $\Theta_{ad}(T,\eta) = \{0\}$ and $x(\cdot;T,\eta,0) \equiv l$.*
(ii) *If $\eta_0 > l$, then $x(\cdot;T,\eta,0) \ge l + \alpha$ for some $\alpha > 0$.*

Proof. (i) Let $\eta_0 = l = q/r$. By Lemma 2.2.5 we know that $0 \in \Theta_{ad}(T,\eta)$. With the same argument of Lemma 2.2.3 (taking into account that in this case $rl - q = f_0(l-l_0) = 0$), one can show that, if $\theta(\cdot)$ is an admissible strategy, then it has to be $x(t;T,\eta,\theta(\cdot)) \ge l$ and $\tilde{\mathbb{E}}_u[x(t;T,\eta,\theta(\cdot))] \equiv l$ for $t \in [T,u]$ (for arbitrary $u \ge T$), so $x(\cdot;T,\eta,\theta(\cdot)) \equiv l$ and we can say that $\theta(\cdot) \equiv 0$ is the unique admissible strategy.

(ii) Let us consider the state trajectory $x(\cdot)$ corresponding to the null strategy $\theta(\cdot) \equiv 0$; let us define $t_0 := \inf \left\{ t \ge T \mid x(t) = l + \frac{\eta_0 - l}{2} \right\} > T$. If $t_0 = +\infty$, we have concluded. Otherwise, set $\varepsilon := l(t_0 - T) - l_0$; the solvency level is strictly increasing in the interval $[0,T]$ by the assumptions done in Subsection 1.1.4, so that we have $\varepsilon > 0$. Let us suppose that $\frac{\eta_0 - l}{2} \le \varepsilon$. For every $t \ge t_0$ we have $x(t-T) \ge l(t_0 - T)$; therefore

$$\begin{aligned} f_0(x(t) - x(t-T)) &\le f_0(x(t) - l(t_0 - T)) \\ &= f_0(x(t) - (l_0 + \varepsilon)). \end{aligned}$$

Thus, whenever $x(t) \le l + \frac{\eta_0 - l}{2}$ (e.g. in t_0),

$$\begin{aligned} dx(t) &= [rx(t) - q]dt - f_0(x(t) - x(t-T))dt \\ &\ge [rl - q]dt - f_0\left(l + \frac{\eta_0 - l}{2} - (l_0 + \varepsilon)\right) dt \\ &\ge [rl - q]dt - f_0(l - l_0)\, dt \ge 0. \end{aligned}$$

Then $x(t) \ge l + \frac{\eta_0 - l}{2}$ for every $t \ge t_0$, and the claim follows.
In the case that $\frac{\eta_0 - l}{2} > \varepsilon$, let $t_1 := \inf \{ t \ge t_0 \, | x(t) = l + \varepsilon \}$; arguing as above we can show that $x(t) \ge l + \varepsilon$ for every $t \ge t_1$, and the claim follows also in this case. □

Remark 2.2.7. Thanks to Remark 1.3.3 the results of Lemma 2.2.3, Lemma 2.2.5 and Lemma 2.2.6 hold true also for initial time $s \ge T$. □

2.2.2 The value function and its properties

As we said in Section 2.1, the objective functional which we want to maximize over the set of admissible strategies $\theta(\cdot) \in \Theta_{ad}(s,\eta)$ is

$$J(s,\eta;\theta(\cdot)) := \mathbb{E}\left[\int_s^{+\infty} e^{-\rho t} U(x(t;s,\eta,\theta(\cdot)))\,dt\right], \quad s \geq T,$$

where $\rho > 0$ and U satisfies Hypothesis 1.3.2. First of all we show that the functional above is well-defined.

Proposition 2.2.8. *Let us suppose that Hypothesis 1.3.2 holds true for U and let $s \geq T$, $\eta \in \mathcal{C}$, $\theta(\cdot) \in \Theta_{ad}(s,\eta)$; then*

$$\mathbb{E}\left[\int_s^{+\infty} e^{-\rho(t-T)}[U^+(x(t))]\,dt\right] < +\infty.$$

Proof. Let us consider the process

$$\begin{cases} dx(t) = [(r+\sigma\lambda\theta(t))x(t) - q]\,dt - f_0(x(t) - x(t-T))dt + \sigma\theta(t)x(t)dB(t), \\ x(s) = \eta_0,\ x(s+\zeta) = \eta_1(\zeta), \qquad \zeta \in [-T,0), \end{cases}$$

and compare it with the one without surplus of Section 1.3:

$$\begin{cases} dy(t) = [(r+\sigma\lambda\theta(t))y(t) - q]dt + \sigma\theta(t)y(t)dB(t), \\ y(s) = \eta_0. \end{cases}$$

Therefore, by comparison criterion (see, e.g., [Karatzas, Shreve; 1991], Proposition 2.18), if $\theta(\cdot) \in \Theta_{ad}(s,\eta)$, then $l \leq x(\cdot) \leq y(\cdot)$. Thus the claim follows by Proposition 1.3.7. □

We define the value function

$$V(s,\eta) := \sup_{\theta(\cdot)\in\Theta_{ad}(s,\eta)} \mathbb{E}\left[\int_s^{+\infty} e^{-\rho(t-T)} U(x(t;s,\eta,\theta(\cdot)))\,dt\right], \quad s \geq T,\ \eta \in \mathcal{C}, \tag{2.12}$$

with the convention $\sup\emptyset = -\infty$.

Definition 2.2.9. (i) Let $s \geq T$, $\eta \in \mathcal{C}$. An optimal strategy for initial data (s,η) is a strategy $\theta^*(\cdot) \in \Theta_{ad}(s,\eta)$ such that for the corresponding trajectory $x^*(t) := x(t;s,\eta,\theta^*(\cdot))$ we have

$$V(s,\eta) = J(s,\eta;\theta^*(\cdot)) = \mathbb{E}\left[\int_s^{+\infty} e^{-\rho(t-T)} U(x^*(t))dt\right].$$

The couple $(\theta^*(\cdot), x^*(\cdot))$ is called an optimal pair.

(ii) Let $s \geq T$, $\eta \in \mathcal{C}$, $\varepsilon > 0$. An ε-optimal strategy for initial data (s, η) is a strategy $\theta^\varepsilon(\cdot) \in \Theta_{ad}(s, \eta)$ such that for the corresponding state trajectory $x^\varepsilon(t) := x(t; s, \eta, \theta^\varepsilon(\cdot))$ we have

$$V(s, \eta) - \varepsilon \leq J(s, \eta; \theta^\varepsilon(\cdot)) = \mathbb{E}\left[\int_s^{+\infty} e^{-\rho(t-T)} U(x^\varepsilon(t)) dt\right].$$

The couple $(\theta^\varepsilon(\cdot), x^\varepsilon(\cdot))$ is called an ε-optimal pair. □

We have the following result, giving the dependence of the value function with respect to the time variable.

Proposition 2.2.10. *Let $s \geq T$ and $\eta \in \mathcal{C}$; then*

$$V(s, \eta) = e^{-\rho(s-T)} V(T, \eta).$$

Proof. As in Proposition 1.3.10 □

By the previous result we are reduced to study the function $V(T, \eta)$; again with a slight abuse of notation we set

$$J(\eta; \theta(\cdot)) := J(T, \eta; \theta(\cdot)), \quad V(\eta) := V(T, \eta), \quad \Theta_{ad}(\eta) := \Theta_{ad}(T, \eta),$$

and we will write $B(\cdot)$ for $B^T(\cdot)$ and $\mathcal{F}_t$ for $\mathcal{F}_t^T$, $t \geq T$.

Proposition 2.2.11. *Let $\eta, \eta' \in \mathcal{C}$ such that $\eta_0 \leq \eta'_0$, $\eta_1(\cdot) \leq \eta'_1(\cdot)$. Then we have $V(\eta) \leq V(\eta')$.*

Proof. If $V(\eta) = -\infty$ we have to prove nothing. Otherwise let us consider a control $\theta(\cdot) \in \Theta_{ad}(\eta)$ and set $x(t) := x(t; T, \eta, \theta(\cdot)), x'(t) := x(t; T, \eta', \theta(\cdot))$. We have for the dynamics of $x'(\cdot)$ in $[T, 2T]$,

$$\begin{aligned}
dx'(t) &= [(r + \sigma\lambda\theta(t))x'(t) - q]\, dt - f_0(x'(t) - x'(t-T))dt + \sigma\theta(t)x'(t)dB(t) \\
&= [(r + \sigma\lambda\theta(t))x'(t) - q]\, dt - f_0(x'(t) - x(t-T))dt + \sigma\theta(t)x'(t)dB(t) \\
&\quad + \left[f_0(x'(t) - x(t-T)) - f_0(x'(t) - x'(t-T))\right] dt,
\end{aligned}$$

where, by monotonicity of f_0 and since $x(t-T) \leq x'(t-T)$ for $t \in [T, 2T]$,

$$f_0(x'(t) - x(t-T))dt - f_0(x'(t) - x'(t-T)) \geq 0. \tag{2.13}$$

Instead for the dynamics of $x(\cdot)$ in $[T, 2T]$,

$$dx(t) = [r + \sigma\lambda\theta(t)]x(t)dt - q\, dt - f_0(x(t) - x(t-T))dt + \sigma\theta(t)x(t)dB(t).$$

By comparison criterion (see [Karatzas, Shreve; 1991], Chapter 5, Proposition 2.18, and take into account (2.13)) we have $x'(t) \geq x(t)$ on $[T, 2T]$. We can iterate the argument (indeed the proof of the cited result of [Karatzas, Shreve; 1991] holds true also in the case of random coefficients) and conclude that $\theta(\cdot) \in \Theta_{ad}(\eta')$, i.e $\Theta_{ad}(\eta) \subset \Theta_{ad}(\eta')$, and $x'(t) \geq x(t)$ for every $t \geq T$, so we can conclude by monotonicity of U. □

Proposition 2.2.12. *1. Let $\eta \in \mathcal{D}$. We have the following statements regarding the lower finiteness of the value function:*

- *(i) If $U(l) > -\infty$, then $V(\eta) > -\infty$.*
- *(ii) If $U(l) = -\infty$ and $\eta_0 > l$, then $V(\eta) > -\infty$.*
- *(iii) If $U(l) = -\infty$, $rl - q = 0$ and $\eta_0 = l$, then $V(\eta) = -\infty$.*
- *(iv) If $U(l) = -\infty$, $rl - q > f_0(k)$ and $\eta_0 = l$, we have to distinguish two cases:*
  - *(iv.a) if U is integrable in l^+, then $V(\eta) > -\infty$;*
  - *(iv.b) if U is not integrable in l^+, then $V(\eta) = -\infty$.*

2. Let $\eta \in \mathcal{C}$. We have the following estimate regarding the upper finiteness of the value function: there exists $K > 0$ such that

$$V(\eta) \leq K(1 + \eta_0^\beta),$$

where β is given by Hypothesis 1.3.2-(ii).

Proof.

1-(i) Of course

$$V(\eta) \geq J(\eta; 0) \geq \int_T^{+\infty} e^{-\rho(t-T)} U(l) dt \geq \frac{U(l)}{\rho}$$

and this statement is proved.

1-(ii) By Lemma 2.2.6-(ii) we know that $x(\cdot; T, \eta, 0) \geq l + \alpha$, for some $\alpha > 0$. Therefore

$$V(\eta) \geq J(\eta, 0) \geq \int_T^{+\infty} e^{-\rho(t-T)} U(l + \alpha) dt = \frac{U(l + \alpha)}{\rho} > -\infty,$$

so also this statement is proved.

1-(iii) By Lemma 2.2.6-(i) we have $\Theta_{ad}(\eta) = \{0\}$ and $x(\cdot; T, \eta, 0) \equiv l$, so that $J(\eta; 0) = -\infty$ and therefore also $V(\eta) = -\infty$.

1-(iv.a) Let us suppose U integrable at l^+ and let $\eta_0 = l$; consider again the null strategy $\theta(\cdot) \equiv 0$, the corresponding state trajectory $x(\cdot) := x(\cdot\,; T, \eta, 0)$ and set

$$\varepsilon := \frac{rl - q - f_0(l - l_0)}{2} > 0.$$

Moreover let $\delta > 0$ such that $f_0(l - l_0 + \delta) < f_0(l - l_0) + \varepsilon$. Then, until

$$l \leq x(t) \leq l + \delta,$$

the dynamics of $x(t)$ is given by

$$\begin{aligned} dx(t) &= [rx(t) - q]\, dt - f_0(x(t) - x(t-T))dt \\ &\geq [rl - q]\, dt - f_0(l + \delta - l_0)dt \geq \varepsilon\, dt. \end{aligned}$$

So, until $l \leq x(t) \leq l + \delta$, we have $x(t) \geq l + \varepsilon(t - T)$ and then $x(t)$ remains up to the level $l + \delta$. Let $t_0 = \inf\{t \geq T \mid x(t) \geq l + \delta\}$; of course $t_0 < +\infty$. We can write

$$\begin{aligned} J(\eta; 0) &= \int_T^{+\infty} e^{-\rho(t-T)} U(x(t))dt \\ &\geq \int_T^{t_0} e^{-\rho(t-T)} U(l + \varepsilon(t-T))dt + \int_{t_0}^{+\infty} e^{-\rho(t-T)} U(x(t))dt. \end{aligned}$$

The finiteness of the second part of the objective functional (the part $\int_{t_0}^{+\infty}$) is obvious, since there we have $x(t) \geq l + \delta$, so that

$$\int_{t_0}^{+\infty} e^{-\rho(t-T)} U(x(t))dt \geq \frac{e^{-(t_0 - T)}}{\rho} U(l + \delta);$$

for the first one (the part $\int_T^{t_0}$) we have

$$\int_T^{t_0} e^{-\rho(t-T)} U(x(t))dt \geq \int_T^{t_0} e^{-\rho(t-T)} U(l + \varepsilon(t-T))dt;$$

by integrability of U and by the change of variable $\xi = l + \varepsilon(t - T)$, we get the finiteness also for this term.

1-(iv.b) Let us suppose U not integrable at l^+ and let $\eta_0 = l$; let $\theta(\cdot) \in \Theta_{ad}(\eta)$ and set $x(t) := x(t; T, \eta, \theta(\cdot))$. The dynamics of $x(\cdot)$ is given by

$$\begin{cases} dx(t) = [(r + \sigma\lambda\theta(t))x(t) - q]\, dt - f_0(x(t) - x(t-T))dt + \sigma\theta(t)x(t)dB(t), \\ x(T) = \eta_0 = l, \quad x(T + \zeta) = \eta_1(\zeta), \ \ \zeta \in [-T, 0). \end{cases}$$

Comparing it with the dinamycs of the problem without surplus

$$\begin{cases} dy(t) = [(r + \sigma\lambda\theta(t))y(t) - q]dt + \sigma\theta(t)y(t)dB(t), \\ y(T) = l, \end{cases}$$

we see, again by comparison criterion, that $l \leq x(\cdot) \leq y(\cdot)$. Therefore the claim follows by Proposition 1.3.11-(ii).

2. Let $\eta \in \mathcal{C}$ and $\theta(\cdot) \in \Theta_{ad}(\eta)$. Again, setting $x(t) := x(t; T, \eta, \theta(\cdot))$, we have $l \leq x(\cdot) \leq y(\cdot)$, where $y(\cdot)$ is given by

$$\begin{cases} dy(t) = [(r + \sigma\lambda\theta(t))y(t) - q]dt + \sigma\theta(t)y(t)dB(t), \\ y(T) = \eta_0, \end{cases}$$

Therefore the claim follows by Proposition 1.3.11. $\square$

Remark 2.2.13. By Proposition 2.2.12, we can see that, when $U(l) = -\infty$ and $\eta_0 = l$, the best cases to treat are the cases either when $rl - q = f_0(k) = 0$ or $rl - q > f_0(k) \geq 0$. The case $rl = f_0(k) > 0$ can be treated falling into the previous ones, but we should distinguish with regard to the structure of f_0 and $\eta_1(\cdot)$. We will do not treat this case for simplicity. $\square$

Let us define

$$D(V) := \{\eta \in \mathcal{C} \mid V(\eta) > -\infty\}$$

and

$$\mathcal{D}_0 := \{\eta \in \mathcal{D} \mid \eta_0 > l\};$$

by Proposition 2.2.12 we get the inclusion

$$\mathcal{D}_0 \subset D(V). \tag{2.14}$$

Moreover, again by Proposition 2.2.12,

$$U(l) > -\infty \implies \mathcal{D} \subset D(V).$$

Proposition 2.2.14. *The set $D(V)$ is convex and the function $\eta \mapsto V(\eta)$ is concave on $D(V)$.*

Proof. Fix $\eta, \eta' \in D(V)$; set also $\eta_\gamma := \gamma\eta + (1-\gamma)\eta'$, $\gamma \in [0,1]$; of course $\eta_\gamma \in \mathcal{C}$. We have to prove that

$$V(\eta_\gamma) \geq \gamma V(\eta) + (1-\gamma)V(\eta'). \tag{2.15}$$

Let $\varepsilon > 0$ and take $\theta(\cdot) \in \Theta_{ad}(\eta)$ and $\theta'(\cdot) \in \Theta_{ad}(\eta')$ ε-optimal for η, η' respectively and $x(\cdot), x'(\cdot)$ the corresponding state trajectories. Then

$$\begin{aligned}
\gamma V(\eta) + (1-\gamma)V(\eta') &\leq \gamma\big[J(\eta;\theta(\cdot)) + \varepsilon\big] + (1-\gamma)\big[J(\eta';\theta'(\cdot)) + \varepsilon\big] \\
&= \varepsilon + \gamma J(\eta;\theta(\cdot)) + (1-\gamma)J(\eta';\theta'(\cdot)) \\
&= \varepsilon + \gamma\mathbb{E}\left[\int_T^{+\infty} e^{-\rho(t-T)} U\big(x(t)\big)\,dt\right] \\
&+ (1-\gamma)\mathbb{E}\left[\int_T^{+\infty} e^{-\rho(t-T)} U\big(x'(t)\big)\,dt\right] \\
&= \varepsilon + \mathbb{E}\left[\int_T^{+\infty} e^{-\rho(t-T)}\left[\gamma U\big(x(t)\big) + (1-\gamma)U\big(x'(t)\big)\right]dt\right].
\end{aligned}$$

The concavity of U implies that

$$\gamma U\big(x(t)\big) + (1-\gamma)U\big(x'(t)\big) \leq U\big(\gamma x(t) + (1-\gamma)x'(t)\big), \qquad \forall t \geq T.$$

Consequently, if we set $x_\gamma(\cdot) := \gamma x(\cdot) + (1-\gamma)x'(\cdot)$, we get

$$\gamma V(\eta) + (1-\gamma)\, V(\eta') \le \varepsilon + \mathbb{E}\left[\int_T^{+\infty} e^{-\rho(t-T)} U\big(x_\gamma(t)\big)\, dt\right].$$

If there exists $\theta_\gamma(\cdot) \in \Theta(\eta_\gamma)$ such that $x_\gamma(\cdot) \le x(\cdot; T, \eta_\gamma, \theta_\gamma(\cdot))$, then we would have

$$\varepsilon + \mathbb{E}\left[\int_T^{+\infty} e^{-\rho(t-T)} U\big(x_\gamma(t)\big)\, dt\right] \le \varepsilon + J(\eta_\gamma; \theta_\gamma(\cdot)) \le \varepsilon + V(\eta_\gamma),$$

i.e.

$$\gamma V(\eta) + (1-\gamma) V(\eta') \le \varepsilon + V(\eta_\gamma)$$

and therefore, by the arbitrariness of ε, the claim (2.15) would be proved. We will show that

$$\theta_\gamma(t) := a(t)\theta(t) + d(t)\theta'(t),$$

where

$$a(\cdot) = \gamma \frac{x(\cdot)}{x_\gamma(\cdot)}, \quad d(\cdot) = (1-\gamma)\frac{x'(\cdot)}{x_\gamma(\cdot)},$$

is good. The admissibility of $\theta_\gamma(\cdot)$ is clear since:

(i) for every $t \ge T$ we have $\theta(t), \theta'(t) \in [0,1]$, and $a(t) + d(t) = 1$ so by convexity of $[0,1]$ we get $\theta_\gamma(t) \in [0,1]$;

(ii) by construction $x_\gamma(t) \ge l$ for any $t \ge s$, so that also $x_\gamma(t; T, \eta_\gamma, \theta_\gamma(\cdot)) \ge l$.

We can write for the dynamics of $x_\gamma(\cdot)$

$$\begin{aligned}
dx_\gamma(t) &= \gamma\, dx(t) + (1-\gamma)\, dx'(t) \\
&= \gamma\big[\, [(r+\sigma\lambda\theta(t))x(t) - q]\, dt - f_0(x(t) - x(t-T))dt + \theta(t)\sigma x(t) dB(t)\big] \\
&\quad + (1-\gamma)\big[\, [(r+\sigma\lambda\theta'(t))x'(t) - q]\, dt - f_0(x'(t) - x'(t-T))dt + \theta'(t)\sigma x'(t) dB(t)\big] \\
&= \big[r x_\gamma(t) - q + [\gamma\theta(t)x(t) + (1-\gamma)\theta'(t)x'(t)]\,\big]\, dt \\
&\quad + \sigma\, [\gamma\theta(t)x(t) + (1-\gamma)\theta'(t)x'(t)]\, dB(t) \\
&\quad - [\gamma f_0(x(t) - x(t-T)) + (1-\gamma) f_0(x'(t) - x'(t-T))]\, dt \\
&= \left[r x_\gamma(t) - q + \left[\gamma\theta(t)\frac{x(t)}{x_\gamma(t)} + (1-\gamma)\theta'(t)\frac{x'(t)}{x_\gamma(t)}\right] x_\gamma(t)\right] dt \\
&\quad + \sigma\left[\gamma\theta(t)\frac{x(t)}{x_\gamma(t)} + (1-\gamma)\theta'(t)\frac{x'(t)}{x_\gamma(t)}\right] x_\gamma(t) dB(t) \\
&\quad - [\gamma f_0(x(t) - x(t-T)) + (1-\gamma) f_0(x'(t) - x'(t-T))]\, dt \\
&\le [(r+\sigma\lambda\theta_\gamma(t))x_\gamma(t) - q]\, dt - f_0(x_\gamma(t) - x_\gamma(t-T))dt + \sigma\theta_\gamma(t)x_\gamma(t) dB(t)
\end{aligned}$$

(where the inequality follows by the convexity of the function f_0), with initial condition

$$x_\gamma(T) = \eta_{\gamma 0};\ x_\gamma(T+\zeta) = \eta_{\gamma 1}(\zeta), \qquad \zeta \in [-T, 0).$$

Instead by definition $x_\gamma(t; T, \eta_\gamma, \theta_\gamma(\cdot))$ satisfies the previous one with equality, so that, by comparison criterion (see [Protter; 2003], Chapter V, Theorem 54), we get $x(t; s, \eta_\gamma, \theta_\gamma(\cdot)) \ge x_\gamma(t)$. The claim follows. □

2.3 The delay problem rephrased in infinite dimension

In this section we will formulate an infinite-dimensional stochastic control problem equivalent to the one of the previous section. As we said at the beginning of the chapter we refer to [Vinter, Kwong; 1981] for this kind of approach in the deterministic case and to

- [Chojnowska-Michalik, 1978],
- [Da Prato, Zabczyk; 1996],
- [Gozzi, Marinelli; 2004],
- [Gozzi, Marinelli, Savin]

in the stochastic case.

Before to proceed we want to expose some considerations about what we are going to do. As well-known every delay problem can be reformulated as infinite-dimensional problem. Nevertheless we should take into account two problems in doing that.

(1) The reformulation of the problem in the infinite-dimensional setting is only formal: once it has been set, we must show that it works good as reformulation of the originary delay problem, i.e. we must prove two results.

   (i) We must study the infinite-dimensional problem in terms of existence and uniqueness of (some kind of) solutions to the state equation (Theorem 2.3.10 in the paper).

   (ii) We must prove that there is actually equivalence (in some "good" sense) between the originary delay problem and the infinite dimensional one (Theorem 3.2.3 in the paper).

(2) In order to get these results and to proceed beyond with the analysis of the problem (in particular, in the case of control problems, with the study of the infinite-dimensional HJB equation), we have to be careful in the choice of the spaces where to embed the problem. In particular it turns out to be very important (for the study of the evolution equation and of the HJB equation) to work, as much as possible, with an Hilbert setting, because this gives a good representation of the duality relationships in terms of inner product.

Taking into account the considerations above, we notice that our problem presents two problem with regard to the embedding in an Hilbert setting.

(i) The delay in the state equation appears concentrated at a point of the past.

(ii) The delay in the state equation appears in a nonlinear way.

These two features together make problematic the reformulation in an Hilbert space. Indeed the natural Hilbert space for the infinite-dimensional reformulation would be $L^2([-T,0];\mathbb{R})$, but unfortunately in this space the term $x_1(-T)$, $x_1 \in L^2([-T,0];\mathbb{R})$, which will turn out to be in our interest, does not make sense. If the delay appeared in a linear way, we could overcome this problem inserting this term in the linear unbounded operator of the evolution equation (as done, e.g., in [Gozzi, Marinelli; 2004, Gozzi, Marinelli, Savin]; indeed this term is well defined on the domain of that operator). Unfortunately this is not our case, so that we are forced to use a more refined framework, dealing also with a suitable subspace of the Hilbert space. We stress that this methodology should be suitable to approach *every delay problem presenting the features described above (even more)*, i.e. a delay having the following form

$$g\left(x(t), \int_{-T}^{0} x(t+\xi)d\mu(\xi)\right),$$

where g is a Lipschitz function and μ a generic measure on $[-T,0]$.

Now we come to a precise study of the problem along the lines of the considerations expressed above. Let us set

$$L^2_{-T} := L^2([-T,0];\mathbb{R}), \quad W^{1,2}_{-T} := W^{1,2}([-T,0];\mathbb{R})$$

and consider the Hilbert space

$$H = \mathbb{R} \times L^2_{-T},$$

with inner product

$$\langle x, y\rangle = x_0 y_0 + \int_{-T}^{0} x_1(\xi) y_1(\xi) d\xi$$

and norm

$$\|x\| = \left(|x_0|^2 + \int_{-T}^{0} |x_1(\xi)|^2 d\xi\right)^{1/2},$$

where $x_0, x_1(\cdot)$ denote respectively the $\mathbb{R}$-valued and the L^2_{-T}-valued components of the generic element $x \in H$; let also $(B(t))_{t\geq T}$ be the same Brownian motion on the probability space $(\Omega, \mathcal{F}, \mathbb{P})$ of the previous section. Let us consider the space E defined in (2.5); we have the estimate

$$\|x\|_H \leq (1+T)^{1/2} \|x\|_E, \quad x \in E,$$

so that we have the continuous and dense embedding

$$\begin{array}{rccc} \iota: & (E, \|\cdot\|_E) & \longrightarrow & (H, \|\cdot\|_H), \\ & x & \longmapsto & x. \end{array}$$

Within this section we intend that the spaces E and H are respectively endowed with the norm $\|\cdot\|_E$ and $\|\cdot\|_H$, which render them Banach spaces. Given an $(\mathcal{F}_t)_{t\geq T}$-progressively measurable and $[0,1]$-valued process $\theta(\cdot)$ and $x \in E$, we can consider the infinite dimensional H-valued stochastic evolution equation starting at time T

$$\begin{cases} dX(t) = AX(t)dt + \sigma\lambda\theta(t)\Phi X(t)dt - F(X(t))dt + \sigma\theta(t)\Phi X(t)dB(t), \\ X(T) = x \in E, \end{cases} \tag{2.16}$$

where

- $A : D(A) \subset H \to H$ is the unbounded linear operator defined by

$$(x_0, x_1(\cdot)) \mapsto (rx_0, x_1'(\cdot)),$$

  with

$$D(A) = \{(x_0, x_1(\cdot)) \in H \mid x_1(\cdot) \in W^{1,2}_{-T},\ x_0 = x_1(0)\};$$

  above by $x_1(0)$ we mean the evaluation at $\zeta = 0$ of the unique (absolutely) continuous representative of $x_1 \in W^{1,2}{-T}$.

- $F : E \to H$ is the nonlinear map

$$\begin{pmatrix} x_0 \\ x_1(\cdot) \end{pmatrix} \mapsto \begin{pmatrix} f(x_0, x_1(\cdot)) \\ 0 \end{pmatrix},$$

  where $f : E \to \mathbb{R}$, $(x_0, x_1(\cdot)) \mapsto f_0(x_0 - x_1(-T)) + q$.

- $\Phi : H \to H$ is the linear operator defined by $\Phi x := (x_0, 0)$.

It is well known that A is a closed densely defined operator and that it is the infinitesimal generator of a C_0-semigroup $(S(t))_{t\geq 0}$ on H; more precisely it is defined by

$$S(t)(x_0, x_1(\cdot)) = \left(x_0 e^{rt}, I_{[-T,0]}(t+\cdot)\, x_1(t+\cdot) + I_{[0,+\infty)}(t+\cdot)\, x_0 e^{r(t+\cdot)}\right);$$

in particular $S(t)$ maps E in itself and

$$S(t)(x_0, 0) = \left(x_0 e^{rt}, I_{[0,+\infty)}(t+\cdot)\, x_0 e^{r(t+\cdot)}\right).$$

Concerning the norm of the semigroup, we have the estimate

$$\|S(t)x\|_H^2 \le \left|x_0 e^{rt}\right|^2 + 2\int_{-T}^{0} \left|I_{[-T,0]}(t+\zeta)\, x_1(t+\zeta)\right|^2 d\zeta$$
$$+ 2\int_{-T}^{0} \left|I_{[0,+\infty)}(t+\zeta)\, x_0 e^{r(t+\zeta)}\right|^2 d\zeta \le (3+2T)e^{2rt}\|x\|_H^2;$$

so

$$\|S(t)\|_{\mathcal{L}(H)} \le M e^{\omega t}, \tag{2.17}$$

with $M = (3+2T)^{1/2}$ and $\omega = r$.

2.3.1 The state equation: existence and uniqueness of mild solutions

In this section we give the definition of *mild solution* of (2.16) and investigate the existence and uniqueness of such a solution. The main reference for the concept of mild solution to an evolution equation is [Da Prato, Zabczyk; 1992]. Basically, since the operator A is defined only on a subspace of H, we cannot expect to get a strong solution (see [Da Prato, Zabczyk; 1992], Chapter 6) to (2.16) in general. Indeed, this concept would require that the process X takes values in $D(A)$. In particular in our case, we cannot require a similar condition, even starting from the initial datum having good regularity. Indeed, consider for simplicity the equation with $\theta(\cdot) \equiv 1$. Then, also starting from an initial point $x \in D(A)$, the solution X (that formally we want representing the present and the past of the solution of the state equation (2.16), see Proposition 3.2.3) should get on its component $X_1(t)$ immediately the same regularity as the Brownian trajectories (i.e. $X_1(t)(\cdot)$ should be not differentiable almost everywhere in the interval $[(-t)\vee(-T), 0]$). So $X(t) \notin D(A)$ for every $t > 0$. Instead the concept of mild solution does not require this kind of regularity. In finite dimension it would represent the variation of constants formula for the solution of an ordinary differential equation and the two concepts (strong and mild solutions) would be equivalent. In infinite dimension it is a weaker concept of solution usually suitable to get existence and uniqueness of solutions.

Definition 2.3.1. A mild solution to (2.16) is an E-valued process $(X(t))_{t\ge 0}$ which satisfies, for $t \ge 0$, the integral equation

$$X(t) = S(t)x + \int_0^t \sigma\lambda\theta(\tau)S(t-\tau)\left[\Phi X(\tau)\right] d\tau$$
$$- \int_0^t S(t-\tau)F(X(\tau))d\tau + \int_0^t \sigma\theta(\tau)S(t-\tau)\left[\Phi X(\tau)\right] dB(\tau). \tag{2.18}$$

Remark 2.3.2. In this remark we want to explain the reasons for the choice of the space E to study (2.16). This space has three important properties:

- the expression $x_1(-T)$ is well-defined on the points $x \in E$ (while in general it is not well defined if $x \in H$);
- it is invariant for the semigroup $S(\cdot)$ and moreover $S(\cdot)$ is a C_0-semigroup also on the space $(E, \|\cdot\|_E)$;
- the value function (1.15) is defined on the points of this space.

The first two properties are essential for working with the state equation (2.16). The third one allows us to link the one-dimensional optimal control problem with delay defined in the previous section with the infinite-dimensional one we will define in this section. $\square$

Note that in the equation above the stochastic integration is one dimensional. In order to be able to manipulate this equation and link the L^2_{-T}-valued integration (stochastic or deterministic) with the canonical $\mathbb{R}$-valued integration, we need a technical digression. So, let $a, b \in \mathbb{R}$, $a < b$ and let G be a Banach space. Moreover let $\mathcal{S}_{\mathcal{P}}([a,b]; L^2(\Omega; G))$ be the space of the simple predictable processes, i.e. $X \in \mathcal{S}_{\mathcal{P}}([a,b]; L^2(\Omega; G))$ if

$$X(t) = X_0 I_{\{t_0\}}(t) + \sum_{k=0}^{n} X_k I_{(t_k, t_{k+1}]}(t),$$

where $a = t_0 < t_1 < \dots < t_n = b$ and the X_k's are G-valued random variables measurable with respect to $\mathcal{F}_{t_k}$; this space is dense in $L^2_{\mathcal{P}}([a,b]; L^2(\Omega; G))$ of the square-integrable progressively measurable processes, endowed with the norm

$$\|X\|_{L^2_{\mathcal{P}}([a,b];L^2(\Omega;G))} = \left(\int_a^b E\left[\|X(t)\|_G^2\right] \right)^{1/2}.$$

Now take $G = L^2_{-T}$; given $U \in L^2(\Omega; L^2[-T, 0])$, we can define, for $\zeta \in [-T, 0]$, the $\mathbb{R}$-valued random variable

$$Z(\zeta)(\omega) := \overline{\widetilde{U}(\omega)}(\zeta),$$

where we denote by the symbols $\bar{f}$, $\tilde{U}$ the pointwise well-defined representatives of generic $f \in L^2_{-T}$ and $U \in L^2(\Omega; L^2_{-T})$. With a slight abuse of notation, we will write $U(\zeta)$ for denoting the equivalence class in $L^2(\Omega; \mathbb{R})$ of the $\mathbb{R}$-valued random variable $Z(\zeta)$ defined above. Note that, choosing other representatives, i.e. defining

$$Z'(\zeta)(\omega) := \overline{\overline{\tilde{\tilde{U}}(\omega)}}(\zeta)$$

and denoting by $U'(\zeta)$ the equivalence class in $L^2(\Omega; \mathbb{R})$ of $Z'(\zeta)$, we would have $U(\zeta) = U'(\zeta)$ (in $L^2(\Omega; \mathbb{R})$), for a.e. $\zeta \in [-T, 0]$.

Lemma 2.3.3. *Let $U \in L^2(\Omega; L^2_{-T})$; then $U(\zeta) \in L^2(\Omega;\mathbb{R})$ for a.e. $\zeta \in [-T, 0]$.*

Proof. For $U \in L^2(\Omega; L^2_{-T})$, let $Z(\zeta)$ be the random variable defined above; we have

$$E\left[\|U\|^2_{L^2_{-T}}\right] < +\infty;$$

this implies

$$\int_\Omega \left[\int_{-T}^0 Z(\zeta)(\omega)^2 d\zeta\right] \mathbb{P}(d\omega) < +\infty,$$

namely

$$\int_{-T}^0 \left[\int_\Omega Z(\zeta)(\omega)^2 \mathbb{P}(d\omega)\right] d\zeta < +\infty;$$

thus

$$\int_{-T}^0 E[U(\zeta)^2] d\zeta < +\infty$$

and therefore, in particular,

$$E[U(\zeta)^2] < +\infty, \quad \zeta - a.e.$$

□

Lemma 2.3.4. *Let $X, Y \in L^2(\Omega; L^2_{-T}))$; suppose that, for a.e. $\zeta \in [-T, 0]$, we have $X(\zeta) = Y(\zeta)$ in $L^2(\Omega;\mathbb{R})$. Then $X = Y$ in $L^2(\Omega; L^2_{-T})$.*

Proof. Of course we have, for a.e. $\zeta \in [-T, 0]$,

$$(X - Y)(\zeta) = X(\zeta) - Y(\zeta), \quad \text{in } L^2(\Omega;\mathbb{R}).$$

Therefore, arguing with representatives as in the proof of the lemma above for the next first equality,

$$E\left[\|X - Y\|^2_{L^2_{-T}}\right] = \int_{-T}^0 E[(X - Y)(\zeta)^2] d\zeta = \int_{-T}^0 E[(X(\zeta) - Y(\zeta))^2] d\zeta = 0.$$

□

The desired link between the L^2_{-T}-valued integration and the $\mathbb{R}$-valued integration is given by the following result.

Lemma 2.3.5. *Let $X \in L^2_{\mathcal{P}}([a, b]; L^2(\Omega; L^2_{-T}))$. Then we have, for a.e. $\zeta \in [-T, 0]$, the following equalities in $L^2(\Omega;\mathbb{R})$:*

- $\left(\int_a^b X(t)dt\right)(\zeta) = \int_a^b X(t)(\zeta)dt;$
- $\left(\int_a^b X(t)dB(t)\right)(\zeta) = \int_a^b X(t)(\zeta)dB(t).$

Proof. For $X \in \mathcal{S}([a,b];L^2(\Omega;L^2_{-T}))$ the claim is obvious. Therefore, let $X \in L^2_{\mathcal{P}}([a,b];L^2(\Omega;L^2_{-T}))$, consider a sequence $(X_n) \subset \mathcal{S}([a,b];L^2(\Omega;L^2_{-T}))$ such that $X_n \to X$ in $L^2_{\mathcal{P}}([a,b];L^2(\Omega;L^2_{-T}))$. We have the equality in $L^2(\Omega;\mathbb{R})$, for $n \in \mathbb{N}$ and for a.e. $\zeta \in [-T,0]$,

$$\left(\int_a^b X_n(t)dt\right)(\zeta) = \int_a^b X_n(t)(\zeta)dt. \tag{2.19}$$

We want to pass to the limit this equality to get the first claim. Since $X_n \to X$ in $L^2_{\mathcal{P}}([a,b];L^2(\Omega;L^2_{-T}))$, we have

$$\int_a^b dt \int_{-T}^0 E\left[|X_n(t)(\zeta) - X(t)(\zeta)|^2\right] d\zeta = \int_a^b E\left[\|X_n(t) - X(t)\|^2_{L^2_{-T}}\right] dt \to 0,$$

so that

$$\int_{-T}^0 d\zeta \int_a^b E\left[|X_n(t)(\zeta) - X(t)(\zeta)|^2\right] dt \to 0.$$

Hence, without loss of generality, taking a subsequence if necessary, we can suppose that, for a.e. $\zeta \in [-T,0]$, we have the convergence $X_n(\cdot)(\zeta) \to X(\cdot)(\zeta)$ in the space $L^2([a,b];L^2(\Omega;\mathbb{R}))$, i.e., for a.e. $\zeta \in [-T,0]$, the convergence in $L^2(\Omega;\mathbb{R})$,

$$\int_a^b X_n(t)(\zeta)dt \to \int_a^b X(t)(\zeta)dt. \tag{2.20}$$

Moreover, since $X_n \to X$ in $L^2_{\mathcal{P}}([a,b];L^2(\Omega;L^2_{-T}))$, we have the convergence in $L^2(\Omega;L^2_{-T})$

$$\int_a^b X_n(t)dt \to \int_a^b X(t)dt.$$

Thus, by the latter convergence,

$$\int_{-T}^0 E\left[\left(\left(\int_a^b X(t)dt\right)(\zeta) - \left(\int_a^b X_n(t)dt\right)(\zeta)\right)^2\right] d\zeta \longrightarrow 0;$$

so, without loss of generality, taking if necessary a subsequence, we can conclude that, for a.e. $\zeta \in [-T,0]$,

$$\left(\int_a^b X_n(t)dt\right)(\zeta) \longrightarrow \left(\int_a^b X(t)dt\right)(\zeta), \quad \text{in } L^2(\Omega;\mathbb{R}). \tag{2.21}$$

Therefore the first claim follows combining (2.19), (2.20) and (2.21). The proof of the second claim proceeds in the same way. $\square$

We denote by $C_{\mathcal{P}}([a,b];L^2(\Omega;G))$ the space of progressively measurable G-valued processes and mean-square continuous. The space $C_{\mathcal{P}}([a,b];L^2(\Omega;G))$ endowed with the norm

$$\|X\|_{C_{\mathcal{P}}} = \left(\sup_{t\in[a,b]} E\left[\|X(t)\|^2_G\right]\right)^{1/2}$$

is a Banach space. For $n \in \mathbb{N}$ and $k = 1, ..., n$, set $t_k^n := a + k\frac{b-a}{n}$. It is easy to prove that, for $X \in C_{\mathcal{P}}([a,b]; L^2(\Omega; G))$, we have

- $\displaystyle\int_a^b X(t)dt = \lim_{n\to\infty} \frac{1}{n}\sum_{k=1}^n X(t_{k-1}^n)$,
- $\displaystyle\int_a^b X(t)dB(t) = \lim_{n\to\infty} \sum_{k=1}^n X(t_{k-1}^n)(B(t_k^n) - B(t_{k-1}^n))$,

where the equalities and the limits have to be intended in the space $L^2(\Omega; G)$. Also it is easy to prove that the (stochastic or deterministic) integral of a generic $X \in L^2_{\mathcal{P}}([a,b]; L^2(\Omega; G))$ (in particular the solution to a G-valued stochastic differential equation) belongs to the space $C_{\mathcal{P}}([a,b]; L^2(\Omega; G))$.

As we said, if $x \in E$, we have $\|x\|_H \leq (1+T)^{1/2}\|x\|_E$; therefore we have the continuous and dense embedding

$$\begin{array}{cccc} \iota: & L^2_{\mathcal{P}}\big([a,b]; L^2(\Omega; (E, \|\cdot\|_E))\big) & \longrightarrow & L^2_{\mathcal{P}}\big([a,b]; L^2(\Omega; (H, \|\cdot\|_H))\big), \\ & X & \longmapsto & X, \end{array}$$

and the continuous and dense embedding

$$\begin{array}{cccc} \tilde{\iota}: & C_{\mathcal{P}}\big([a,b]; L^2(\Omega; (E, \|\cdot\|_E))\big) & \longrightarrow & C_{\mathcal{P}}\big([a,b]; L^2(\Omega; (H, \|\cdot\|_H))\big), \\ & X & \longmapsto & X. \end{array}$$

We are going to prove a slight extension of the existence and uniqueness result we need. Let $g : (E, \|\cdot\|_E) \to \mathbb{R}$ be a Lipschitz continuous map, with Lipschitz constant C_g, and consider the following map associated with it

$$\begin{array}{cccc} G: & E & \longrightarrow & H, \\ & x & \longmapsto & \begin{pmatrix} g(x) \\ 0 \end{pmatrix}. \end{array} \tag{2.22}$$

Let $\theta(\cdot)$ be a fixed $(\mathcal{F}_t)_{t\geq 0}$-progressively measurable and $[0,1]$-valued process $\theta(\cdot)$, let $x \in E$, and consider the stochastic evolution equation

$$\begin{cases} dX(t) = AX(t)dt + \sigma\lambda\theta(t)\Phi X(t)dt - G(X(t))dt + \sigma\theta(t)\Phi X(t)dB(t), \\ X(0) = x \in E. \end{cases} \tag{2.23}$$

We notice that we cannot use the theory of [Da Prato, Zabczyk; 1992] to treat (2.23): the nearest result in that book is Theorem 7.19, but, differently from there, *here we do not require any dissipativity of G with respect to* $\langle\cdot,\cdot\rangle_H$, i.e. we do not require that, for some $\omega \geq 0$,

$$\langle\, G(x) - G(y),\, x - y\,\rangle_H \leq \omega\|x-y\|_H^2, \quad \forall x, y \in E.$$

We cannot require this condition because in particular the map F, which is in our interest, does not satifsy such a condition. ILet us show this fact. Given $\omega \geq 0$ we have

$$\begin{aligned}\langle (F-\omega I)(x) - (F-\omega I)(y), x-y\rangle_H \\ &= (x_0-y_0)\big[f_0(x_0-x_1(-T)) - f_0(y_0-y_1(-T))\big] \\ &\qquad - \omega\left[|x_0-y_0|^2 + \int_{-T}^0 (x_1(\zeta)-y_1(\zeta))^2 d\zeta\right].\end{aligned}$$

We claim that for every $\omega \geq 0$ there exist $x, y \in E$ such that the right hand-side in the previous expression is strictly positive. Fix $\omega \geq 0$ and let $y = (0,0)$, $x_0 > 0$; by convexity of f_0 we can make the expression $f_0(x_0 - x_1(-T)) - f_0(0)$ large as we want moving $x_1(-T)$ and, at the same time, to take fix x_0 and $|x_0|^2 + \int_{-T}^0 x_1(\zeta)^2 d\zeta$. This shows what claimed.

Lemma 2.3.6. *Consider, for $a, b \in \mathbb{R}$, $T \leq a < b$, the linear map*

$$\begin{array}{rccc}\chi: & L^2\big((\Omega, \mathcal{F}_a^T); E\big) & \longrightarrow & L^2_{\mathcal{P}}\big([a,b]; L^2(\Omega; H)\big), \\ & \psi & \longmapsto & S(\cdot - a)\psi.\end{array}$$

Then, for any $\psi \in L^2((\Omega, \mathcal{F}_a^T); E)$, we have $\chi(\psi) \in C_{\mathcal{P}}\big([a,b]; L^2(\Omega; E)\big)$. Moreover there exists a constant C_{b-a} such that

$$\|\chi(\psi)\|_{C_{\mathcal{P}}([a,b];L^2(\Omega;E))} \leq C_{b-a}\|\psi\|_{L^2(\Omega;E)}.$$

Proof. Let $\psi \in L^2((\Omega, \mathcal{F}_a^T); E)$; for $a \leq t \leq b$, we can choose $S(t-a)\psi$ taking values in E, because $S(t-a)$ maps E in itself. Moreover $S(\cdot)$ is a strongly continuous semigroup also on the space E, so that (see e.g. [Engel, Nagel; 2000]) there exists a constant C_{b-a} such that $\|S(t-a)\|_{\mathcal{L}(E)} \leq C_{b-a}$, $t \in [a,b]$. Therefore, for $t \in [a,b]$, $\|S(t-a)\psi\|_E \leq C_{b-a}\|\psi\|_E$, so that $S(t-a)\psi \in L^2(\Omega; E)$.

The fact that $\chi(\psi) \in C_{\mathcal{P}}\big([a,b]; L^2(\Omega; E)\big)$ follows by dominated convergence, since, by the property of strong continuity of the semigroup, if $t_0 \in [a,b]$ and $[a,b] \ni t \to t_0$, then $S(t-a)\psi \longrightarrow S(t_0-a)\psi$ in E (pointwise on Ω after having chosen a representative for the random variable ψ) and moreover $\|S(t-a)\psi\|_E \leq C_{b-a}\|\psi\|_E$.

The last statement follows again because $\|S(t-a)\psi\|_E \leq C_{b-a}\|\psi\|_E$. $\square$

Lemma 2.3.7. *For $a, b \in \mathbb{R}$, $T \leq a < b$ and for a given $(\mathcal{F}_t)_{t \geq T}$-progressively measurable and $[0,1]$-valued process $\theta(\cdot)$, consider the linear map*

$$\begin{array}{rccc}\gamma_\theta: & C_{\mathcal{P}}\big([a,b]; L^2(\Omega; E)\big) & \longrightarrow & C_{\mathcal{P}}\big([a,b]; L^2(\Omega; H)\big) \\ & X(\cdot) & \longmapsto & \int_a^{\cdot} \theta(\tau) S(\cdot - \tau)\left[\Phi X(\tau)\right] d\tau.\end{array}$$

Then, for any $X \in C_{\mathcal{P}}\big([a,b];L^2(\Omega;E)\big)$, we have $\gamma_\theta(X) \in C_{\mathcal{P}}\big([a,b];L^2(\Omega;E)\big)$. Moreover we have the estimate

$$\|\gamma_\theta(X)\|^2_{C_{\mathcal{P}}([a,b];L^2(\Omega;E))} \leq 4e^{2r(b-a)}(b-a)^2\|X\|^2_{C_{\mathcal{P}}([a,b];L^2(\Omega;E))}.$$

Proof. Let $X \in C_{\mathcal{P}}\big([a,b];L^2(\Omega;E)\big)$; for $a \leq t \leq b$, taking into account Lemma 2.3.5 and the equality $I_{[0,+\infty)}(t+\zeta-\tau) = I_{[\tau,+\infty)}(t+\zeta)$, we have, for a.e. $\zeta \in [-T,0]$, the following equalities:

$$\begin{aligned}
\begin{pmatrix} \gamma_\theta(X)_0(t) \\ \gamma_\theta(X)_1(t)(\zeta) \end{pmatrix} &= \int_a^t \theta(\tau)S(t-\tau)\left[\Phi X(\tau)\right]d\tau \\
&= \int_a^t \theta(\tau)S(t-\tau)\left(X_0(\tau),0\right)d\tau \\
&= \int_a^t \begin{pmatrix} \theta(\tau)X_0(\tau)e^{r(t-\tau)} \\ I_{[\tau,+\infty)}(t+\zeta)\theta(\tau)X_0(\tau)e^{r(t+\zeta-\tau)} \end{pmatrix} d\tau \\
&= \begin{pmatrix} \int_a^t \theta(\tau)X_0(\tau)e^{r(t-\tau)}d\tau \\ \\ \begin{cases} 0, & \text{if } t+\zeta \leq a \\ e^{r\zeta}\int_a^{t+\zeta}\theta(\tau)X_0(\tau)e^{r(t-\tau)}d\tau, & \text{if } t+\zeta \geq a \end{cases} \end{pmatrix};
\end{aligned}$$

this shows that, for any $t \in [a,b]$ and $X \in C_{\mathcal{P}}\big([a,b];L^2(\Omega;E)\big)$, we can take the random variable $\gamma_\theta(X)(t)$ taking values in E. Now we have to show that $\gamma_\theta(X)(t) \in L^2(\Omega;E)$. Taking into account Hölder's inequality, we have

$$\begin{aligned}
\mathbb{E}\left[\|\gamma_\theta(X)(t)\|_E^2\right] &\leq 2\mathbb{E}\Bigg[\left|\int_a^t \theta(\tau)X_0(\tau)e^{r(t-\tau)}d\tau\right|^2 \\
&\quad + \sup_{\zeta\in[-T,0)}\left|\int_a^{t+\zeta}\theta(\tau)X_0(\tau)e^{r(t-\tau)}d\tau\right|^2\Bigg] \\
&\leq 4(b-a)e^{2r(b-a)}\int_a^t \mathbb{E}[|X_0(\tau)|^2]d\tau < +\infty. \qquad (2.24)
\end{aligned}$$

By an estimate like the previous one we can get, for $t_0, t \in [a,b]$,

$$\mathbb{E}\left[\|\gamma_\theta(X)(t)-\gamma_\theta(X)(t_0)\|_E^2\right] \leq 4(b-a)e^{2r(b-a)}\int_{t\wedge t_0}^{t\vee t_0}\mathbb{E}[|X_0(\tau)|^2]d\tau.$$

getting $\gamma_\theta(X) \in C_{\mathcal{P}}\big([a,b];L^2(\Omega;E)\big)$, by mean-square continuity of $X_0(\cdot)$. The last statement follows arguing as in the estimate (2.24), but taking the supremum on $t \in [a,b]$ before to pass to the expectations. $\square$

Lemma 2.3.8. *Let $g : (E, \|\cdot\|) \to \mathbb{R}$ be a Lipschitz continuous map, with Lipschitz constant C_g, and consider the map associated with it*

$$\begin{array}{rccc} G: & E & \longrightarrow & H, \\ & x & \longmapsto & \begin{pmatrix} g(x) \\ 0 \end{pmatrix}; \end{array}$$

(of course, if $X(\cdot) \in C_{\mathcal{P}}\big([a,b]; L^2(\Omega;E)\big)$, then $g(X(\cdot)) \in C_{\mathcal{P}}([a,b]; L^2(\Omega;\mathbb{R}))$). For $T \le a < b$, consider the nonlinear map

$$\begin{array}{rccc} \gamma: & C_{\mathcal{P}}\big([a,b]; L^2(\Omega;E)\big) & \longrightarrow & C_{\mathcal{P}}\big([a,b]; L^2(\Omega;H)\big) \\ & X(\cdot) & \longmapsto & \int_a^{\cdot} S(\cdot-\tau)G(X(\tau))d\tau. \end{array}$$

Then, for any $X \in C_{\mathcal{P}}\big([a,b]; L^2(\Omega;E)\big)$, we have $\gamma(X) \in C_{\mathcal{P}}\big([a,b]; L^2(\Omega;E)\big)$. Moreover, for $X, Y \in C_{\mathcal{P}}\big([a,b]; L^2(\Omega;E)\big)$, we have the estimate

$$\|\gamma(X)-\gamma(Y)\|^2_{C_{\mathcal{P}}([a,b];L^2(\Omega;E))} \le 4C_g^2 e^{2r(b-a)}(b-a)^2\|X-Y\|^2_{C_{\mathcal{P}}([a,b];L^2(\Omega;E))}.$$

Proof. Let $X \in C_{\mathcal{P}}\big([a,b]; L^2(\Omega;E)\big)$; for $a \le t \le b$, taking into account Lemma 2.3.5 and the equality $I_{[0,+\infty)}(t+\zeta-\tau) = I_{[\tau,+\infty)}(t+\zeta)$, we have, for a.e. $\zeta \in [-T,0]$, the following equalities:

$$\begin{aligned} \begin{pmatrix} \gamma(X)_0(t) \\ \gamma(X)_1(t)(\zeta) \end{pmatrix} &= \int_a^t S(t-\tau)G(X(\tau))d\tau = \int_a^t S(t-\tau)\left(g(X(\tau)),0\right)d\tau \\ &= \int_a^t \begin{pmatrix} g(X(\tau))e^{r(t-\tau)} \\ I_{[\tau,+\infty)}(t+\zeta)g(X(\tau)))e^{r(t+\zeta-\tau)} \end{pmatrix} d\tau \\ &= \begin{pmatrix} \int_a^t g(X(\tau))e^{r(t-\tau)}d\tau \\ \\ \begin{cases} 0, & \text{if } t+\zeta \le a \\ e^{r\zeta}\int_a^{t+\zeta} g(X(\tau))e^{r(t-\tau)}d\tau, & \text{if } t+\zeta \ge a \end{cases} \end{pmatrix}; \end{aligned}$$

this shows that, for any $t \in [a,b]$ and $X \in C_{\mathcal{P}}\big([a,b]; L^2(\Omega;E)\big)$, we can take the random variable $\gamma(X)(t)$ taking values in E. Now we have to show that, for any $t \in [a,b]$, $\gamma(X)(t) \in L^2(\Omega;E)$; indeed we have

$$\begin{aligned} &\mathbb{E}\left[\|\gamma(X)(t)\|_E^2\right] \\ &\quad \le 2\mathbb{E}\left[\left|\int_a^t g(X(\tau))e^{r(t-\tau)}d\tau\right|^2 + \sup_{\zeta\in[-T,0)}\left|\int_a^{t+\zeta} g(X(\tau))e^{r(t-\tau)}d\tau\right|^2\right] \\ &\qquad \le 8e^{2r(b-a)}(b-a)\mathbb{E}\left[\int_a^t \left(C_g^2\|X(\tau)\|_E^2 + |g(0)|^2\right)d\tau\right] < +\infty. \end{aligned}$$

We can prove that $\gamma(X) \in C_{\mathcal{P}}\big([a,b]; L^2(\Omega;E)\big)$ arguing as in the proof of the previous Lemma and the last statement arguing as in the previous estimate, but taking the supremum on $t \in [a,b]$ before to pass to the expectations. $\square$

Lemma 2.3.9. *For $a, b \in \mathbb{R}$, $T \le a < b$ and for a fixed $(\mathcal{F}_t)_{t \ge T}$-progressively measurable and $[0,1]$-valued process $\theta(\cdot)$, consider the linear map*

$$\begin{array}{rccc} \tilde{\gamma}_\theta: & C_{\mathcal{P}}\big([a,b]; L^2(\Omega; E)\big) & \longrightarrow & C_{\mathcal{P}}\big([a,b]; L^2(\Omega; H)\big) \\ & X(\cdot) & \longmapsto & \int_a^{\cdot} \theta(\tau) S(\cdot - \tau)\, [\Phi X(\tau)]\, dB(\tau). \end{array}$$

Then, for any $X \in C_{\mathcal{P}}\big([a,b]; L^2(\Omega; E)\big)$, we have $\tilde{\gamma}_\theta(X) \in C_{\mathcal{P}}\big([a,b]; L^2(\Omega; E)\big)$. Moreover we have the estimate

$$\|\tilde{\gamma}_\theta(X)\|^2_{C_{\mathcal{P}}([a,b]; L^2(\Omega; E))} \le 10 e^{2r(b-a)} (b-a) \|X\|^2_{C_{\mathcal{P}}([a,b]; L^2(\Omega; E))}.$$

Proof. Let $X \in C_{\mathcal{P}}\big([a,b]; L^2(\Omega; E)\big)$; for $a \le t \le b$, taking into account Lemma 2.3.5 and the equality $I_{[0,+\infty)}(t + \zeta - \tau) = I_{[\tau,+\infty)}(t + \zeta)$, we have, for a.e. $\zeta \in [-T, 0]$, the following equalities:

$$\begin{aligned} \begin{pmatrix} \tilde{\gamma}_\theta(X)_0(t) \\ \tilde{\gamma}_\theta(X)_1(t)(\zeta) \end{pmatrix} &= \int_a^t \theta(\tau) S(t-\tau)\, [\Phi X(\tau)]\, dB(\tau) \\ &= \int_a^t S(t-\tau)\, (\theta(\tau) X_0(\tau), 0)\, dB(\tau) \\ &= \int_a^t \begin{pmatrix} \theta(\tau) X_0(\tau) e^{r(t-\tau)} \\ I_{[\tau,+\infty)}(t+\zeta) \theta(\tau) X_0(\tau) e^{r(t+\zeta-\tau)} \end{pmatrix} dB(\tau) \\ &= \begin{pmatrix} \int_a^t \theta(\tau) X_0(\tau) e^{r(t-\tau)} dB(\tau) \\ \\ \begin{cases} 0, & \text{if } t + \zeta \le a \\ e^{r\zeta} \int_a^{t+\zeta} \theta(\tau) X_0(\tau) e^{r(t-\tau)} dB(\tau), & \text{if } t + \zeta \ge a \end{cases} \end{pmatrix}; \end{aligned}$$

this shows that, for any $t \in [a,b]$ and $X \in C_{\mathcal{P}}\big([a,b]; L^2(\Omega; E)\big)$, we can take the random variable $\tilde{\gamma}_\theta(X)(t)$ taking values in E. Now we have to show that $\tilde{\gamma}_\theta(X)(t) \in L^2(\Omega; E)$. Taking into account the Doob's inequality for continuous square-integrable martingales and the Itô's isometry, we have

$$\begin{aligned} &\mathbb{E}\left[\|\tilde{\gamma}_\theta(X)(t)\|_E^2\right] \\ &\le 2\mathbb{E}\left[\left|\int_a^t \theta(\tau) X_0(\tau) e^{r(t-\tau)} dB(\tau)\right|^2 + \sup_{\zeta \in [-T,0)} \left|\int_a^{t+\zeta} \theta(\tau) X_0(\tau) e^{r(t-\tau)} dB(\tau)\right|^2\right] \\ &\le 10\, \mathbb{E}\left[\left|\int_a^t \theta(\tau) X_0(\tau) e^{r(t-\tau)} dB(\tau)\right|^2\right] \le 10\, e^{2r(b-a)} \int_a^t \mathbb{E}[|X_0(\tau)|^2] d\tau < +\infty. \end{aligned}$$

Then we can prove that $\tilde{\gamma}_\theta(X) \in C_{\mathcal{P}}\big([a,b]; L^2(\Omega; E)\big)$ and the last statement arguing as in Lemma 2.3.7. $\square$

Now we are ready to prove the desired existence and uniqueness result.

Theorem 2.3.10. *For every $x \in E$ and for every given $(\mathcal{F}_t)_{t\geq 0}$-progressively measurable and $[0,1]$-valued process $\theta(\cdot)$, equation* (2.23) *admits a unique mild solution in the space $C_{\mathcal{P}}\big([0,+\infty); L^2(\Omega; E)\big)$, i.e. there exists a unique $X \in C_{\mathcal{P}}\big([0,+\infty); L^2(\Omega; E)\big)$ such that*

$$\begin{aligned} X(t) &= S(t)x + \int_0^t \sigma\lambda\theta(\tau)S(t-\tau)\left[\Phi X(\tau)\right]d\tau - \int_0^t S(t-\tau)G(X(\tau))d\tau \\ &\quad + \int_0^t \sigma\theta(\tau)S(t-\tau)\left[\Phi X(\tau)\right]dB(\tau). \end{aligned}$$

Proof. Let $a \geq T$ and $\psi \in L^2((\Omega, \mathcal{F}_a^T); E)$; for $b > a$ such that

$$e^{r(b-a)}\left[(2C_g + \sigma\lambda)(b-a) + 10^{1/2}\sigma(b-a)^{1/2}\right] < 1, \tag{2.25}$$

consider the map

$$\begin{array}{llll} \Gamma_{\psi,\theta}: & C_{\mathcal{P}}\big([a,b]; L^2(\Omega; E)\big) & \longrightarrow & C_{\mathcal{P}}\big([a,b]; L^2(\Omega; E)\big), \\ & X & \longmapsto & \chi(\psi) + \sigma\lambda\gamma_\theta(X) - \gamma(X) + \sigma\tilde{\gamma}_\theta(X); \end{array}$$

by the previous Lemmata and by (2.25) this map is a contraction on the space $C_{\mathcal{P}}\big([a,b]; L^2(\Omega; E)\big)$ and so it admits a unique fixed point in this space. This fixed point is the solution to equation (2.23) in the interval $[a,b]$ for initial time a and initial condition ψ. So starting from T we can find a unique $X \in C_{\mathcal{P}}\big([T, T+(b-a)]; L^2(\Omega; E)\big)$ such that, for $t \in [T, T+(b-a)]$,

$$\begin{aligned} X(t) &= S(t-T)x + \sigma\lambda\int_T^t \theta(\tau)S(t-\tau)\left[\Phi X(\tau)\right]d\tau - \int_T^t S(t-\tau)G(X(\tau))d\tau \\ &\quad + \sigma\int_T^t \theta(\tau)S(t-\tau)\left[\Phi X(\tau)\right]dB(\tau). \end{aligned}$$

Then we can iterate the argument by $(b-a)$-steps (notice that the achievement of (2.25) depends only on the difference $b-a$) and get the solution in the whole interval $[T,+\infty)$ by the semigroup property of $S(\cdot)$. □

Remark 2.3.11. In the proof of Theorem 2.3.10 we could obtain the solution directly on the whole interval $[T,+\infty)$ using the *exponential* norm on the space $C_{\mathcal{P}}([T,+\infty); L^2(\Omega; E))$; indeed, if we endow this space with the norm

$$\|X\|_{C_{\mathcal{P}}\left([T,+\infty); L^2(\Omega;E)\right)} = \sup_{t\in[a,b]} e^{-\lambda t}\left(\mathbb{E}\left[\|X(t)\|_E^2\right]\right)^{1/2},$$

then we could find $\lambda > 0$ such that $\Gamma_{\psi,\theta}$ becomes a contraction on the space $C_{\mathcal{P}}\big([T,+\infty); L^2(\Omega; E)\big)$ and get the fixed point directly in this space. □

2.3.2 Equivalence between the stochastic delay problem and the infinite dimensional problem

At the beginning of this section we have defined the infinite dimensional equation (2.16) and we have proved an existence and uniqueness result for such equation. Now, to give sense to our approach, we want to link (2.16) with (2.6). The link is given by the following result.

Proposition 2.3.12. *Let $\theta(\cdot)$ be an $(\mathcal{F}_t)_{t\geq T}$-progressively measurable and $[0,1]$-valued process and $x \in E$. Let $x(\cdot)$ be the unique solution to (2.6) with data $x, \theta(\cdot)$ in the space $C_{\mathcal{P}}([0,+\infty); L^2(\Omega))$ and let $X(\cdot)$ be the unique mild solution to (2.16) with data $x, \theta(\cdot)$ in the space $C_{\mathcal{P}}([0,+\infty); L^2(\Omega; E))$. Then*

$$X(\cdot) = (x(\cdot), x(\cdot + \zeta)|_{\zeta\in[-T,0]}).$$

Proof. Let $x(\cdot)$ be the solution to (2.6) and let $X(\cdot) := (x(\cdot), x(\cdot+\zeta)|_{\zeta\in[-T,0]})$. Then X belongs to the space $C_{\mathcal{P}}([T,+\infty); L^2(\Omega; E))$ because the function

$$[T,+\infty) \to L^2(\Omega;\mathbb{R}),\ t \mapsto x(t)$$

is continuous and therefore uniformly continuous on the compact subsets of $[T,+\infty)$. So we have to prove that $X(t) = (X_0(t), X_1(t)) = (x(t), x(t+\zeta)|_{\zeta\in[-T,0]})$ satisfies (2.16) on both the components. For the first one we have to verify that, for any $t \geq T$,

$$\begin{aligned} X_0(t) &= e^{r(t-T)}x_0 + \sigma\lambda \int_T^t e^{r(t-\tau)}\theta(\tau)X_0(\tau)d\tau \\ &- \int_T^t e^{r(t-\tau)} \left[f_0\left(X_0(\tau) - X_1(\tau)(-T)\right) + q\right] d\tau + \sigma \int_T^t e^{r(t-\tau)}\theta(\tau)X_0(\tau)dB(\tau), \end{aligned}$$

i.e. that

$$\begin{aligned} x(t) &= e^{r(t-T)}x_0 + \sigma\lambda \int_T^t e^{r(t-\tau)}\theta(\tau)x(\tau)d\tau \\ &\quad - \int_T^t e^{r(t-\tau)} \left[f_0\left(x(\tau) - x(\tau - T)\right) + q\right] d\tau + \sigma \int_T^t e^{r(t-\tau)}\theta(\tau)x(\tau)dB(\tau); \end{aligned}$$

but this comes from the assumption that $x(\cdot)$ is a solution to (2.6).

For the second component, taking into account that

$$I_{[0,+\infty)}(t + \cdot - \tau) = I_{[\tau,+\infty)}(t + \ \cdot)$$

and Lemma 2.3.4, Lemma 2.3.5, we have to verify, for any $t \geq T$, the equalities in $L^2(\Omega;\mathbb{R})$ for a.e. $\zeta \in [-T,0]$

$$\begin{aligned} X_1(t)(\zeta) &= I_{[0,T]}(t+\zeta)x_1(t+\zeta-T) + I_{[T,+\infty)}(t+\zeta)x_0 e^{r(t+\zeta-T)} \\ &\quad +\sigma\lambda \int_T^t I_{[\tau,+\infty)}(t+\zeta)\theta(\tau)X_0(\tau)e^{r(t+\zeta-\tau)}d\tau \\ &\quad -\int_T^t I_{[\tau,+\infty)}(t+\zeta)\left[f_0\left(X_0(\tau)-X_1(\tau)(-T)\right)+q\right]e^{r(t+\zeta-\tau)}d\tau \\ &\quad +\sigma \int_T^t I_{[\tau,+\infty)}(t+\zeta)\theta(\tau)X_0(\tau)e^{r(t+\zeta-\tau)}dB(\tau); \end{aligned}$$

i.e., for any $t \geq T$, the equalities in $L^2(\Omega,\mathbb{R})$ for a.e. $\zeta \in [-T,0]$

$$\begin{aligned} x(t+\zeta) &= I_{[0,T]}(t+\zeta)x_1(t+\zeta-T) + I_{[T,+\infty)}(t+\zeta)x_0 e^{r(t+\zeta-T)} \\ &\quad +\sigma\lambda \int_T^t I_{[\tau,+\infty)}(t+\zeta)\theta(\tau)x(\tau)e^{r(t+\zeta-\tau)}d\tau \\ &\quad -\int_T^t I_{[\tau,+\infty)}(t+\zeta)\left[f_0\left(x(\tau)-x(\tau-T)\right)+q\right]e^{r(t+\zeta-\tau)}d\tau \\ &\quad +\sigma \int_T^t I_{[\tau,+\infty)}(t+\zeta)\theta(\tau)x(\tau)e^{r(t+\zeta-\tau)}dB(\tau). \end{aligned} \tag{2.26}$$

For $\zeta \in [-T,0]$ such that $t+\zeta \in [0,T]$, (2.26) reduces to

$$x(t+\zeta) = x_1(t+\zeta-T)$$

and this is true by the initial condition of (2.6); instead for $\zeta \in [-T,0]$ such that $t+\zeta \geq T$, (2.26) reduces to

$$\begin{aligned} x(t+\zeta) &= x_0 e^{r(t+\zeta-T)} + \sigma\lambda \int_T^{t+\zeta} \theta(\tau)x(\tau)e^{r(t+\zeta-\tau)}d\tau \\ &\quad -\int_T^{t+\zeta} \left[f_0\left(x(\tau)-x(\tau-T)\right)+q\right]e^{r(t+\zeta-\tau)}d\tau \\ &\quad +\sigma\int_T^{t+\zeta} \theta(\tau)x(\tau)e^{r(t+\zeta-\tau)}dB(\tau); \end{aligned}$$

setting $u := t+\zeta$ this equality becomes, for $u \geq T$,

$$\begin{aligned} &x(u) = x_0 e^{r(u-T)} + \sigma\lambda \int_T^u \theta(\tau)x(\tau)e^{r(u-\tau)}d\tau \\ &\quad -\int_T^u \left[f_0\left(x(\tau)-x(\tau-T)\right)+q\right]e^{r(u-\tau)}d\tau + \sigma\int_T^u \theta(\tau)x(\tau)e^{r(u-\tau)}dB(\tau); \end{aligned}$$

again this is true because $x(\cdot)$ solves (2.6). □

Thanks to the previous equivalence result, we can rewrite our optimization problem in the infinite-dimensional setting: consider the class of equations, for

$\theta(\cdot) \in \Theta_{ad}(x)$, $x \in \mathcal{C}$,

$$\begin{cases} dX(t) = AX(t)dt + \sigma\lambda\theta(t)\Phi X(t)dt - F(X(t))dt + \sigma\theta(t)\Phi X(t)dB(t), \\ X(T) = x. \end{cases}$$

and denote the mild solution to the previous equation by $X(t;T,x,\theta(\cdot))$. Thanks to Proposition 3.2.3 the objective functional defined in Section 2.2.2 can be rewritten as

$$J\left(x;\theta\left(\cdot\right)\right) = \mathbb{E}\left[\int_T^{+\infty} e^{-\rho(t-T)} U\left(X_0\left(t;T,x,\theta(\cdot)\right)\right) dt\right].$$

2.3.3 Continuous dependence on initial data

In this subsection we investigate the continuous dependence on the initial data for the mild solution to equation (2.16) with respect to the $\|\cdot\|_E$-norm and with respect to the $\|\cdot\|_H$-norm. Moreover we will give also a result of continuity with respect to the control strategy $\theta(\cdot)$.

Proposition 2.3.13. *In the hypotheses of Theorem 2.3.10, let $x, y \in E$ be two initial data for the equation and denote by $X(x), X(y)$ the solution associated respectively to x, y. Then, for $u \geq T$, there exists a constant $K_u > 0$ such that*

$$\|X(x) - X(y)\|_{C_{\mathcal{P}}([T,u];L^2(\Omega;E))} \leq K_u \|x - y\|_E.$$

Proof. Let us consider the map

$$\begin{array}{rccc} \Gamma_\theta : & L^2((\Omega, \mathcal{F}_a^T); E) \times C_{\mathcal{P}}\big([a,b]; L^2(\Omega;E)\big) & \longrightarrow & C_{\mathcal{P}}\big([a,b]; L^2(\Omega;E)\big), \\ & (\psi, X) & \longmapsto & \chi(\psi) + \sigma\lambda\gamma_\theta(X) - \gamma(X) + \sigma\tilde{\gamma}_\theta(X); \end{array}$$

we have already proved in Theorem 2.3.10 that, for $(b-a)$ small enough, there exists $C < 1$ such that

$$\|\Gamma_\theta(\psi, X) - \Gamma_\theta(\psi, Y)\|_{C_{\mathcal{P}}([a,b];L^2(\Omega;E))} \leq C\, \|X - Y\|_{C_{\mathcal{P}}([a,b];L^2(\Omega;E))}.$$

Moreover, by Lemma 2.3.6, we know that

$$\begin{aligned} \|\Gamma_\theta(\psi, X) - \Gamma_\theta(\psi', X)\|_{C_{\mathcal{P}}([a,b];L^2(\Omega;E))} &= \|\chi(\psi - \psi')\|_{C_{\mathcal{P}}([a,b];L^2(\Omega;E))} \\ &\leq C_{b-a}\|\psi - \psi'\|_{L^2(\Omega;E)}. \end{aligned}$$

Thus, denoting by $X(\psi), X(\psi')$ the solution in $[a,b]$ to our equation, starting from ψ, ψ' respectively, we have

$$\begin{aligned} \|X(\psi) - X(\psi')\|_{C_{\mathcal{P}}([a,b];L^2(\Omega;E))} &= \|\Gamma_\theta(\psi, X(\psi)) - \Gamma_\theta(\psi', X(\psi'))\|_{C_{\mathcal{P}}([a,b];L^2(\Omega;E))} \\ &\leq C_{b-a}\|\psi - \psi'\|_{L^2(\Omega;E)} \\ &\quad + C\|X(\psi) - X(\psi')\|_{C_{\mathcal{P}}([a,b];L^2(\Omega;E))}, \end{aligned}$$

therefore

$$\|X(\psi) - X(\psi')\|_{C_{\mathcal{P}}([a,b];L^2(\Omega;E))} \leq (1-C)^{-1} C_{b-a} \|\psi - \psi'\|_{L^2(\Omega;E)}.$$

Thus, starting from T with initial conditions x, y and iterating by $(b-a)$-steps until reaching u, we get the claim. $\square$

Thanks to Proposition 2.3.12, we can prove a continuous dependence on initial data result with respect to the H-norm under the null control:

Proposition 2.3.14. *Let $x, y \in E$ be two initial data for the equation (2.16) and denote by $X(x), X(y)$ the mild solutions associated respectively to x, y, both under the null control $\theta(\cdot) \equiv 0$. Then, for each $u \geq T$, there exists a constant $K_u > 0$ such that*

$$\|X(x) - X(y)\|_{C([T,u];(E,\|\cdot\|_H))} \leq K_u \|x - y\|_H.$$

Proof. Notice that we cannot proceed as in Proposition 2.3.13 because the function F is not Lipschitz continuous with respect to the H-norm. Take $T \leq a < b$ such that $b - a \leq T$ and let γ be the map defined by

$$\begin{array}{rccc} \gamma: & C([a,b];(E,\|\cdot\|_H)) & \longrightarrow & C([a,b];(E,\|\cdot\|_H) \\ & X(\cdot) & \longmapsto & \int_a^{\cdot} S(\cdot - \tau) F(X(\tau)) d\tau. \end{array}$$

As in the proof of Lemma 2.3.8, for $t \in [a,b]$,

$$\begin{pmatrix} \gamma(X)_0(t) \\ \gamma(X)_1(t)(\zeta) \end{pmatrix} = \begin{pmatrix} \int_a^t f(X(\tau)) e^{r(t-\tau)} d\tau \\ \\ \begin{cases} 0, & \text{if } t + \zeta \leq a, \\ e^{r\zeta} \int_a^{t+\zeta} f(X(\tau)) e^{r(t-\tau)} d\tau, & \text{if } t + \zeta \geq a; \end{cases} \end{pmatrix};$$

let K_0 be the Lipschitz constant of f_0. For every $X, Y \in C([a,b];(E,\|\cdot\|_H))$ we have the following estimate with respect to the H-norm:

$$\begin{aligned}
&\sup_{t\in[a,b]} \|\gamma(X)(t)-\gamma(Y)(t)\|_H^2 = \sup_{t\in[a,b]} \left[\left| \int_a^t e^{r(t-\tau)} \left(f(X(\tau)) - f(Y(\tau)) \right) d\tau \right|^2 \right. \\
&\left. + \int_{(a-t)}^0 \left| e^{r\zeta} \int_a^{t+\zeta} e^{r(t-\tau)} \left(f(X(\tau))) - f(Y(\tau)) \right) d\tau \right|^2 d\zeta \right] \\
&\le \sup_{t\in[a,b]} \left[(t-a) \int_a^t e^{2r(t-\tau)} \left| f(X(\tau)) - f(Y(\tau)) \right|^2 d\tau \right. \\
&\left. + \int_{(a-t)}^0 \left[e^{2r\zeta}(t+\zeta-a) \int_a^{t+\zeta} e^{2r(t-\tau)} \left| f(X(\tau)) - f(Y(\tau)) \right|^2 d\tau \right] d\zeta \right] \\
&\le 2(b-a)e^{2r(b-a)} K_0^2 \int_a^b \left[|X_0(\tau)-Y_0(\tau)|^2 + |X_1(\tau)(-T) - Y_1(\tau)(-T)|^2 \right] d\tau \\
&+2T(b-a)e^{2r(b-a)} K_0^2 \int_a^b \left[|X_0(\tau)-Y_0(\tau)|^2 + |X_1(\tau)(-T) - Y_1(\tau)(-T)|^2 \right] d\tau \\
&= 2(b-a)e^{2r(b-a)}(1+T) K_0^2 \int_a^b \left[|X_0(\tau)-Y_0(\tau)|^2 + |X_1(\tau)(-T) - Y_1(\tau)(-T)|^2 \right] d\tau.
\end{aligned}$$

Now, if we set $X := X(x)$, $Y := X(y)$, taking into account that

$$-T \le \tau - a \le b - a - T \le 0,$$

by Proposition 3.2.3 we have

$$X_1(\tau)(-T) = X_1(a)(\tau - a - T), \quad Y_1(\tau)(-T) = Y_1(a)(\tau - a - T).$$

So we get

$$\begin{aligned}
\sup_{t\in[a,b]} \|\gamma(X)(t)-\gamma(Y)(t)\|_H^2 &\le 2(b-a)^2 e^{2r(b-a)}(1+T)K_0^2 \|X-Y\|^2_{C([a,b];H)} \\
&\quad +2(b-a)e^{2r(b-a)}(1+T)\, K_0^2\, \|X(a)-Y(a)\|_H^2 \\
&\le 2(b-a)^2 e^{2r(b-a)}(1+T)\, K_0^2\, \|X-Y\|^2_{C([a,b];H)} \\
&\quad +2(b-a)e^{2r(b-a)}(1+T)\, K_0^2\, \|X-Y\|^2_{C([a,b];H)};
\end{aligned}$$

thus, for small enough $b-a$,

$$\|\gamma(X)-\gamma(Y)\|_{C([a,b];(E,\|\cdot\|_H))} \le C\|X-Y\|_{C([a,b];(E,\|\cdot\|_H))}$$

for some $0 < C < 1$. Now the claim follows arguing as in the proof of Proposition 2.3.13, taking into account that, if χ is the map defined in Lemma 2.3.6 and $\psi, \psi' \in E$, then by (2.17)

$$\|\chi(\psi-\psi')\|_{C([a,b];(E,\|\cdot\|_H))} \le (3+2T)^{1/2} e^{r(b-a)} \|\psi-\psi'\|_H.$$

□

By the previous we get the following very useful result.

Corollary 2.3.15. *Let $x, y \in E$ two initial data for the equation (2.16) with null control $\theta(\cdot) \equiv 0$ and denote by $X(x), X(y)$ the mild solution associated respectively to x, y. Then, for $u \geq T$, there exists a constant $K_u > 0$ such that*

$$\sup_{t\in[T,u]} |X_0(x)(t) - X_0(y)(t)| \leq K_u \|x - y\|_H.$$

□

Now we give a result of continuous dependence of solutions with respect to the strategies.

Proposition 2.3.16. *In the hypotheses of Theorem 2.3.10, let $(\theta_n)_{n\geq 1}, \theta$ be $(\mathcal{F}_t)_{t\geq T}$-progressively measurable and $[0,1]$-valued processes such that*

$$(\theta_n - \theta)^2 \overset{*}{\rightharpoonup} 0 \quad \text{in } L^\infty(\Omega \times [T, u])$$

for some $u \geq T$, and denote by $X(\theta_n), X(\theta)$ respectively the solutions to the equation associated with θ_n, θ. Then

$$X(\theta_n) \to X(\theta), \quad \text{in } C_{\mathcal{P}}([T, u]; L^2(\Omega; E)).$$

Proof. Let us denote by $X(\psi, \theta), X(\psi_n, \theta_n)$ the solutions to the equation in $[a, b]$ starting from ψ, ψ_n and with controls θ, θ_n respectively; then we have, by Proposition 2.3.14, for suitable $K > 0$ and for small enough $(b - a)$,

$$\begin{aligned} &\|X(\psi, \theta) - X(\psi_n, \theta_n)\|_{C_{\mathcal{P}}([a,b];L^2(\Omega;E))} \\ &\qquad \leq K\|\psi - \psi_n\|_{L^2(\Omega;E)} + \|X(\psi, \theta) - X(\psi, \theta_n)\|_{C_{\mathcal{P}}([a,b];L^2(\Omega;E))}. \end{aligned} \tag{2.27}$$

Let us consider the map

$$\begin{array}{llcc} \Gamma_\psi: & L^2([a, b]; L^2(\Omega; E)) \times C_{\mathcal{P}}([a, b]; L^2(\Omega; E)) & \longrightarrow & C_{\mathcal{P}}([a, b]; L^2(\Omega; E)), \\ & (\theta, X) & \longmapsto & \chi(\psi) + \sigma\lambda\gamma_\theta(X) - \gamma(X) + \sigma\tilde{\gamma}_\theta(X); \end{array}$$

we know that, for small enough $(b - a)$, there exists $C < 1$ such that, for any $n \in \mathbb{N}$,

$$\|\Gamma_\psi(\theta_n, X) - \Gamma_\psi(\theta_n, Y)\|_{C_{\mathcal{P}}([a,b];L^2(\Omega;E))} \leq C \|X - Y\|_{C_{\mathcal{P}}([a,b];L^2(\Omega;E))}.$$

So,

$$\begin{aligned} &\|X(\psi, \theta) - X(\psi, \theta_n)\|_{C_{\mathcal{P}}([a,b];L^2(\Omega;E))} \\ &\qquad \leq \|\Gamma_\psi(\theta, X(\psi, \theta)) - \Gamma_\psi(\theta_n, X(\psi, \theta))\|_{C_{\mathcal{P}}([a,b];L^2(\Omega;E))} \\ &\qquad + \|\Gamma_\psi(\theta_n, X(\psi, \theta)) - \Gamma_\psi(\theta_n, X(\psi, \theta_n))\|_{C_{\mathcal{P}}([a,b];L^2(\Omega;E))} \\ &\qquad \leq \|\Gamma_\psi(\theta, X(\psi, \theta)) - \Gamma_\psi(\theta_n, X(\psi, \theta))\|_{C_{\mathcal{P}}([a,b];L^2(\Omega;E))} \\ &\qquad\qquad + C\|X(\psi, \theta) - X(\psi, \theta_n)\|_{C_{\mathcal{P}}([a,b];L^2(\Omega;E))}, \end{aligned}$$

i.e.,

$$\begin{aligned}&\|X(\psi,\theta)-X(\psi,\theta_n)\|_{C_{\mathcal{P}}([a,b];L^2(\Omega;E))}\\&\qquad\leq(1-C)^{-1}\|\Gamma_\psi(\theta,X(\psi,\theta))-\Gamma_\psi(\theta_n,X(\psi,\theta))\|_{C_{\mathcal{P}}([a,b];L^2(\Omega;E))}.\end{aligned}$$

On the other hand

$$\begin{aligned}&\|\Gamma_\psi(\theta,X(\psi,\theta))-\Gamma_\psi(\theta_n,X(\psi,\theta))\|^2_{C_{\mathcal{P}}([a,b];L^2(\Omega;E))}\\&\leq 2\left[\lambda^2\sigma^2\sup_{t\in[a,b]}E\left[\|\gamma_{\theta-\theta_n}(X(\psi,\theta))(t)\|_E^2\right]+\sigma^2\sup_{t\in[a,b]}E\left[\|\tilde\gamma_{\theta-\theta_n}(X(\psi,\theta))(t)\|_E^2\right]\right]\end{aligned}$$

and, as in the proof of Lemma 2.3.7 and Lemma 2.3.9,

$$\begin{aligned}&\sup_{t\in[a,b]}E\left[\|\gamma_{\theta-\theta_n}(X(\psi,\theta))(t)\|_E^2\right]\\&\qquad\leq 4\sup_{t\in[a,b]}E\left[\left|\int_a^t(\theta(\tau)-\theta_n(\tau))X_0(\psi,\theta)(\tau)e^{r(t-\tau)}d\tau\right|^2\right]\\&\qquad\qquad\leq 4e^{2r(b-a)}(b-a)\int_a^b E\left[|\theta(\tau)-\theta_n(\tau)|^2|X_0(\psi,\theta)(\tau)|^2\right]d\tau,\end{aligned}$$

$$\begin{aligned}&\sup_{t\in[a,b]}E\left[\|\tilde\gamma_{\theta-\theta_n}(X(\psi,\theta))(t)\|_E^2\right]\\&\qquad\leq 10\sup_{t\in[a,b]}E\left[\left|\int_a^t(\theta(\tau)-\theta_n(\tau))X_0(\psi,\theta)(\tau)e^{r(t-\tau)}dB(\tau)\right|^2\right]\\&\qquad\qquad\leq 10e^{2r(b-a)}\int_a^b E\left[|\theta(\tau)-\theta_n(\tau)|^2|X_0(\psi,\theta)(\tau)|^2\right]d\tau.\end{aligned}$$

So, starting from T, taking $n\to\infty$ and iterating by $(b-a)$-steps (and taking in account the inequality (2.27) in the iterations, because the starting point changes), we get the claim. □

Remark 2.3.17. We notice that the condition $\theta_n\to\theta$ in measure on $\Omega\times[T,u]$, for example, is enough to have $(\theta_n-\theta)^2\overset{*}{\rightharpoonup}0$ in $L^\infty(\Omega\times[T,u])$.
On the other hand we notice also that the convergence $\theta_n\overset{*}{\rightharpoonup}\theta$ in $L^\infty(\Omega\times[T,u])$ is not enough to have the desired convergence $(\theta_n-\theta)^2\overset{*}{\rightharpoonup}0$. Let us consider, for instance, this framework: the space $L^\infty([0,1],\lambda)\cong(L^1([0,1],\lambda))^*$, where λ denotes the Lebesgue measure on $[0,1]$; then, by the Riemann-Lebesgue Lemma, we have the convergence $\sin(nx)\overset{*}{\rightharpoonup}0$, but of course the convergence $\sin^2(nx)\overset{*}{\rightharpoonup}0$ is not true. □

2.4 Continuity of the value function

In this section we prove the continuity of the value function with respect to the $\|\cdot\|_H$-norm in the interior part of its domain and then we extend it to a $\|\cdot\|_H$-continuous function defined on an open set of $(H, \|\cdot\|_H)$. Moreover, in the special case of absorbing boundary, i.e. when $rl = q$ and $U(l) > -\infty$ (see Proposition 2.4.6-(6)), we prove that the value function is $\|\cdot\|_H$-continuous up to the boundary. Actually, for brevity, we will prove the lower semicontinuity of the value function at the boundary and only the idea of the proof of the upper semicontinuity. We will indend the topological notions in the $\|\cdot\|_H$ norm. Nevertheless we will use subscripsts to stress in which topological space we are working.

2.4.1 Continuity in the interior of the domain

In this subsection we study the continuity of the value function in the interior part of its domain. Let us recall that the domain of V is the set

$$D(V) := \{x \in \mathcal{C} \mid V(x) > -\infty\}$$

and that, by Proposition 2.2.12, we have $\mathcal{D}_0 \subset D(V)$, where $\mathcal{D}_0$ was defined in (2.14).

Lemma 2.4.1. *Let $x \in \mathcal{D}_0$; then there exists $\varepsilon > 0$ such that V is bounded from below on $B_{(E,\|\cdot\|_H)}(x, \varepsilon)$.*

Proof. Consider the null strategy $\theta(\cdot) \equiv 0$ and set $X(t) := X(t; T, x, 0)$; by Lemma 2.2.6-(2) there exists $\beta > 0$ such that $X_0(t) \geq l + \beta$. Set, for $y \in B_{(E,\|\cdot\|_H)}(x, \varepsilon)$, $Y(t) := Y(t; T, y, 0)$. Take $\varepsilon < \beta/2K_{2T}$, where $K_{2T} > 0$ is the constant given in the estimate of Corollary 2.3.15; by the same Corollary, if $\|x - y\|_H \leq \varepsilon$, then

$$\sup_{t \in [T, 2T]} |X_0(t) - Y_0(t)| \leq \beta/2,$$

so that we have $Y_0(t) \geq l + \beta/2$ on $[T, 2T]$. Thus, for $y \in B_{(E,\|\cdot\|_H)}(x, \varepsilon)$,

$$\mathbb{E}\left[\int_T^{2T} e^{-\rho(t-T)} U(Y_0(t))dt\right] \geq \frac{1 - e^{-\rho T}}{\rho} U(l + \beta/2).$$

Let $z = (z_0, z_1(\cdot))$, where $z_0 = l + \beta/2$ and $z_1 : [-T, 0) \to \mathbb{R}$ is the constant function $\zeta \mapsto l + \beta/2$. By Proposition 3.2.3 we have $Y(2T) \geq z$, in the sense of Proposition 2.2.11. By the semigroup property of the mild solution $Y(\cdot)$, we have, for $t \geq 2T$, $Y(t) = X(t; 2T, Y(2T), 0)$, so that $Y_0(t) \geq X_0(t; 2T, z, 0)$, for $t \geq 2T$. Since $z \in \mathcal{D}_0$, by Lemma 2.2.6-(2) and Proposition 3.2.3 we have

$X_0(t; 2T; z, 0) \geq l + \beta_z$, where $\beta_z > 0$ does not depend on $y \in B_{(E,\|\cdot\|_H)}(x, \varepsilon)$. Thus we can write

$$\begin{aligned} V(y) &\geq \mathbb{E}\left[\int_T^{2T} e^{-\rho(t-T)} U(Y_0(t))dt\right] + \mathbb{E}\left[\int_{2T}^{+\infty} e^{-\rho(t-T)} U(Y_0(t))dt\right] \\ &= \frac{1 - e^{-\rho T}}{\rho} U(l + \beta/2) + \frac{1 - e^{-\rho T}}{\rho} U(l + \beta_z/2). \end{aligned}$$

The estimate holds uniformly without regard to the point $y \in B_{(E,\|\cdot\|_H)}(x, \varepsilon)$ chosen, so that the claim is proved. $\square$

By the previous result we have proved in particular that $\text{Int}_{(E,\|\cdot\|_H)}(D(V))$ is not empty and that $\mathcal{D}_0 \subset \text{Int}_{(E,\|\cdot\|_H)}(D(V))$. The proof of the following Lemma can be found e.g. in [Ekeland, Temam; 1976], Chapter 1, Corollary 2.4.

Lemma 2.4.2. *Let Z be a topological vector space and $f : Z \to \bar{\mathbb{R}}$ concave and proper. Let us define*

$$D(f) := \{z \in Z \mid f(z) > -\infty\};$$

if f is bounded from below on a neighborhood of some $z_0 \in D(f)$, then $Int(D(f))$ is not empty and f is continuous on $Int(D(f))$.

Thus from Lemma 2.4.1 and Lemma 2.4.2 above we can get the following result.

Corollary 2.4.3. *The value function V is continuous on $Int_{(E,\|\cdot\|_H)}(D(V))$.* $\square$

For semplicity of notation we set

$$\mathcal{V} := \text{Int}_{(E,\|\cdot\|_H)}(D(V)). \tag{2.28}$$

We want to extend V to a continuous function defined on an open set of $(H, \|\cdot\|_H)$ containing $\mathcal{V}$. Notice that, if $\mathcal{A}$ is an open set of $(H, \|\cdot\|_H)$, then $\mathcal{A} \cap E$ is $\|\cdot\|_H$-dense in $\mathcal{A}$.

Proposition 2.4.4. *There exist an open set $\mathcal{O}$ of $(H, \|\cdot\|_H)$ and a continuous function $\bar{V} : \mathcal{O} \to \mathbb{R}$ such that:*

1. *$\mathcal{O} \supset \mathcal{V}$ and $\bar{V}|_{\mathcal{V}} = V$.*

2. *$\mathcal{V} = \mathcal{O} \cap E$ and $\mathcal{O} = Int_{(H,\|\cdot\|_H)}\left(Clos_{(H,\|\cdot\|_H)}(\mathcal{V})\right)$.*

3. *$\mathcal{O}$ is convex and $\bar{V}$ is concave on $\mathcal{O}$.*

Proof. 1. Let $x \in \mathcal{V}$; we know that V is continuous at x in E, so that, by concavity, it is Lipschitz continuous on $B_{(E,\|\cdot\|_H)}(x,\rho_x) = \{y \in E \mid \|y - x\|_H < \rho_x\} \subset \mathcal{V}$ for suitable $\rho_x > 0$ (see again Corollary 2.4, Chapter 1, of [Ekeland, Temam; 1976]). $B_{(E,\|\cdot\|_H)}(x,\rho_x)$ is dense in $B_{(H,\|\cdot\|_H)}(x,\rho_x)$, since $B_{(E,\|\cdot\|_H)}(x,\rho_x) = B_{(H,\|\cdot\|_H)}(x,\rho_x) \cap E$. Thus we can extend by continuity V to a continuous function V_x on $B_{(H,\|\cdot\|_H)}(x,\rho_x) = \{y \in H \mid \|y - x\|_H < \rho_x\}$. Of course we can repeat this construction for all the points of $\mathcal{V}$. These extensions are compatible each other, in the sense that, if $x, x' \in \mathcal{V}$ and $y \in B_{(H,\|\cdot\|_H)}(x,\rho_x) \cap B_{(H,\|\cdot\|_H)}(x',\rho_{x'})$, then $V_x(y) = V_{x'}(y)$. Indeed, by density, we can take a sequence $(x_n) \subset B_{(E,\|\cdot\|_H)}(x,\rho_x) \cap B_{(E,\|\cdot\|_H)}(x',\rho_{x'})$ such that $x_n \to y$; on this sequence we have $V_x(x_n) = V(x_n) = V_{x'}(x_n)$, therefore, taking the limit for $n \to \infty$, we get $V_x(y) = V_{x'}(y)$ by continuity of $V_x, V_{x'}$.

Let us define the open set $\mathcal{O}$ of $(H, \|\cdot\|_H)$ by

$$\mathcal{O} := \bigcup_{x \in \mathcal{V}} B_{(H,\|\cdot\|_H)}(x,\rho_x); \tag{2.29}$$

thanks to the compatibility argument it remains defined on $\mathcal{O}$ a continuous function $\bar{V}$. Of course

$$\mathcal{V} = \bigcup_{x \in \mathcal{V}} B_{(E,\|\cdot\|_H)}(x,\rho_x) \subset \mathcal{O} \tag{2.30}$$

and, by construction, $\bar{V}|_{\mathcal{V}} = V$.

2. Let $x \in \mathcal{V}$; then, of course $x \in \mathcal{O}$ and $x \in E$, so that $\mathcal{V} \subset \mathcal{O} \cap E$. Conversely let $x \in \mathcal{O} \cap E$; then, since $x \in \mathcal{O}$, from (2.29) we have $x \in B_{(H,\|\cdot\|_H)}(z,\rho_z)$ for some $z \in \mathcal{V}$ and, on the other hand, since $x \in E$, we have $x \in B_{(H,\|\cdot\|_H)}(z,\rho_z) \cap E = B_{(E,\|\cdot\|_H)}(z,\rho_z)$, so that, by (2.30), we have $x \in \mathcal{V}$ and therefore we can conclude that also $\mathcal{O} \cap E \subset \mathcal{V}$.
About the second statement, thanks to the fact that E is dense in H and that $\mathcal{O}$ is open in $(H, \|\cdot\|_H)$, we can write

$$\begin{aligned} \mathrm{Int}_{(H,\|\cdot\|_H)}\left(\mathrm{Clos}_{(H,\|\cdot\|_H)}(\mathcal{V})\right) &= \mathrm{Int}_{(H,\|\cdot\|_H)}\left(\mathrm{Clos}_{(H,\|\cdot\|_H)}(\mathcal{O} \cap E)\right) \\ &= \mathrm{Int}_{(H,\|\cdot\|_H)}\left(\mathrm{Clos}_{(H,\|\cdot\|_H)}(\mathcal{O})\right) = \mathcal{O}. \end{aligned} \tag{2.31}$$

3. The convexity of $\mathcal{O}$ follows by (2.31) and by the fact that $\mathcal{V}$ is convex. The concavity of $\bar{V}$ follows by its continuity and by concavity of V on $\mathcal{V}$. □

Hereafter, with a slight abuse of notation, we still indicate the extended value function $\bar{V}$ on $\mathcal{O}$ by V.

2.4.2 Continuity at the boundary

In this subsection we study the continuity properties of value function at the boundary. We start with a topological lemma which makes clear the link between the boundary $\mathrm{Fr}_{(H,\|\cdot\|_H)}(\mathcal{O})$ of $\mathcal{O}$ in $(H, \|\cdot\|_H)$ and the boundary $\mathrm{Fr}_{(E,\|\cdot\|_H)}(\mathcal{V})$ of $\mathcal{V}$ in $(E, \|\cdot\|_H)$. Notice that $\mathrm{Fr}_{(E,\|\cdot\|_H)}(\mathcal{V})$ is not empty: indeed, by Proposition 2.2.12, we have $\{x \in \mathcal{D} \mid x_0 = l\} \subset \mathrm{Fr}_{(E,\|\cdot\|_H)}(\mathcal{V})$.

Lemma 2.4.5. *We have the following statements:*

1. $E \setminus \mathcal{V} = (H \setminus \mathcal{O}) \cap E$.

2. $\mathit{Clos}_{(E,\|\cdot\|_H)}(\mathcal{V}) = \mathit{Clos}_{(H,\|\cdot\|_H)}(\mathcal{O}) \cap E$.

3. $\mathit{Fr}_{(E,\|\cdot\|_H)}(\mathcal{V}) = \mathit{Fr}_{(H,\|\cdot\|_H)}(\mathcal{O}) \cap E$.

Proof. 1. By Proposition 2.4.4-(2) we know that $\mathcal{V} = \mathcal{O} \cap E$. Thus we can write

$$(H \setminus \mathcal{O}) \cap E = (H \cap E) \setminus (\mathcal{O} \cap E) = E \setminus \mathcal{V}.$$

2. Let $x \in \mathrm{Clos}_{(E,\|\cdot\|_H)}(\mathcal{V})$; then we can find a sequence $(x_n) \subset \mathcal{V}$ such that $x_n \xrightarrow{\|\cdot\|_H} x$; of course $(x_n) \subset \mathcal{O}$, so that $x \in \mathrm{Clos}_{(H,\|\cdot\|_H)}(\mathcal{O})$; on the other hand $x \in E$, since $x \in \mathrm{Clos}_{(E,\|\cdot\|_H)}(\mathcal{V})$, so that $x \in \mathrm{Clos}_{(H,\|\cdot\|_H)}(\mathcal{O}) \cap E$.

Conversely, let $x \in \mathrm{Clos}_{(H,\|\cdot\|_H)}(\mathcal{O}) \cap E$; we know (Proposition 2.4.4) that

$$\mathrm{Clos}_{(H,\|\cdot\|_H)}(\mathcal{O}) = \mathrm{Clos}_{(H,\|\cdot\|_H)}(\mathcal{V}).$$

Thus there exists a sequence $(x_n) \subset \mathcal{V}$ such that $x_n \xrightarrow{\|\cdot\|_H} x$; together with the assumption $x \in E$, this shows that $x \in \mathrm{Clos}_{(E,\|\cdot\|_H)}(\mathcal{V})$.

3. Notice that $E \setminus \mathcal{V}$ and $H \setminus \mathcal{O}$ are closed respectively in $(E, \|\cdot\|_H)$ and $(H, \|\cdot\|_H)$. Let $x \in \mathrm{Fr}_{(E,\|\cdot\|_H)}(\mathcal{V})$; this means that $x \in \mathrm{Clos}_{(E,\|\cdot\|_H)}(\mathcal{V}) \cap (E \setminus \mathcal{V})$. Thanks to the point (2) of this proposition we get $x \in \mathrm{Clos}_{(H,\|\cdot\|_H)}(\mathcal{O}) \cap (H \setminus \mathcal{O}) \cap E$, i.e. $x \in \mathrm{Fr}_{(H,\|\cdot\|_H)}(\mathcal{O}) \cap E$.
Conversely, let $x \in \mathrm{Fr}_{(H,\|\cdot\|_H)}(\mathcal{O}) \cap E$; this means that $x \in \big(\mathrm{Clos}_{(H,\|\cdot\|_H)}(\mathcal{O}) \cap E\big) \cap \big((H \setminus \mathcal{O}) \cap E\big)$, so that, by the point (2), $x \in \mathrm{Clos}_{(E,\|\cdot\|_H)}(\mathcal{V}) \cap (E \setminus \mathcal{V})$, i.e. $x \in \mathrm{Fr}_{(E,\|\cdot\|_H)}(\mathcal{V})$. $\square$

Refining the assumptions on the parameters of the model the boundary becomes absorbing.

Proposition 2.4.6. *Let $rl = q$, $U(l) > -\infty$ and let $\mathcal{C} \subset E$ be the convex set defined in (2.3). Then the following statements hold:*

1. *We have $D(V) = \{x \in \mathcal{C} \mid 0 \in \Theta_{ad}(x)\} = \mathit{Clos}_{(E,\|\cdot\|_H)}(\mathcal{V})$.*

2. *Let $x \in D(V)$; then either there exists $\beta > 0$ such that $X_0(t; T, x, 0) \geq l + \beta$ for all $t \geq T$ (first case) or there exists $s \in [T, 2T)$ such that $X_0(s; T, x, 0) = l$ (second case). In the second case $X_0(t; T, x, 0) = l$, for every $t \geq s$.*

3. *We have $x \in \mathcal{V}$ if and only if $x \in D(V)$ and there exists $\beta > 0$ such that $X(t;T,x,0) \geq l + \beta$, for all $t \geq T$.*

4. *We have $x \in Fr_{(E,\|\cdot\|_H)}(\mathcal{V})$ if and only if $x \in D(V)$ and there exists $s \in [T, 2T)$ such that $X(s;T,x,0) = l$. In this case $X_0(t;T,x,0) = l$, for $t \geq s$.*

5. *We have $Fr_{(E,\|\cdot\|_H)}(\mathcal{V}) = \{x \in \mathcal{C} \mid \Theta_{ad}(x) = \{0\}\}$.*

6. *The boundary $Fr_{(E,\|\cdot\|_H)}(\mathcal{V})$ is absorbing for the problem, in the sense that, if $x \in Fr_{(E,\|\cdot\|_H)}(\mathcal{V})$, then we have $X(t;T,x,0) \in Fr_{(E,\|\cdot\|_H)}(\mathcal{V})$ for all $t \geq T$.*

Proof. 1. Recall that, by definition, $D(V) = \{x \in \mathcal{C} \mid V(x) > -\infty\}$. By the assumption $U(l) > -\infty$, we get $x \in D(V)$ if and only if $x \in \mathcal{C}$ and $\Theta_{ad}(x) \neq \emptyset$ and, thanks to Lemma 2.2.3, this occurs if and only if $x \in \mathcal{C}$ and $0 \in \Theta_{ad}(x)$, so that $D(V) = \{x \in \mathcal{C} \mid 0 \in \Theta_{ad}(x)\}$.

Since by definition $\mathcal{V} = \mathrm{Int}_{(E,\|\cdot\|_H)}(D(V))$, we can say that

$$\{x \in \mathcal{C} \mid 0 \in \Theta_{ad}(x)\} = D(V) \subset \mathrm{Clos}_{(E,\|\cdot\|_H)}(\mathcal{V}),$$

so that it remains to prove the inclusion

$$\mathrm{Clos}_{(E,\|\cdot\|_H)}(\mathcal{V}) \subset D(V) = \{x \in \mathcal{C} \mid 0 \in \Theta_{ad}(x)\}.$$

So let us take $x \in \mathrm{Clos}_{(E,\|\cdot\|_H)}(\mathcal{V}) = \mathcal{V} \cup \mathrm{Fr}_{(E,\|\cdot\|_H)}(\mathcal{V})$. If $x \in \mathcal{V}$, then, by definition of $\mathcal{V}$, $x \in D(V)$. So let us suppose $x \in \mathrm{Fr}_{(E,\|\cdot\|_H)}(\mathcal{V})$; we want to prove that $x \in \{x \in \mathcal{C} \mid 0 \in \Theta_{ad}(x)\}$, i.e. $X_0(\cdot;T,x,0) \geq l$. Let us suppose, by contradiction, that, for some $t \geq T$, $\varepsilon > 0$, we have $X_0(t;T,x,0) \leq l - \varepsilon$. By definition of $\mathrm{Fr}_{(E,\|\cdot\|_H)}(\mathcal{V})$, there exists $y \in \mathcal{V} \subset D(V)$ such that $\|x - y\|_H < \varepsilon/2K_t$, where K_t is the constant in the estimate of Corollary 2.3.15; so, by the same result, we would have $X_0(t;T,y,0) \leq l - \varepsilon/2$, i.e., by Lemma 2.2.3, $\Theta_{ad}(y) = \emptyset$ and thus $y \notin D(V)$. Therefore the contradiction arises and the claim is proved.

2. Let $x \in D(V)$. If $x_0 = l$, then we have $X(T;T,x,0) = l$ and we are in the second case. So let $x \in D(V)$ be such that $x_0 > l$. Of course, since $x_1(\zeta) \to x_0$ when $\zeta \to 0^-$, we can find $\varepsilon > 0$ such that $x_1(\zeta) \geq l$ for $\zeta \in [-\varepsilon, 0)$. Let us suppose that we are not in the second case, i.e. $X_0(t;T,x,0) > 0$ for every $t \in [T, 2T)$; then there exists some $\alpha > 0$ such that $X_0(t;T,x,0) \geq l + \alpha$, for $t \in [T, 2T - \varepsilon]$. We want to show that then we are in the first case. By the semigroup property of the mild solution $X(\cdot;T,x,0)$ we have, for $t \geq 2T - \varepsilon$, $X_0(t;T,x,0) = X_0(t;2T-\varepsilon, X(2T-\varepsilon;T,x,0),0)$; since $X_1(2T-\varepsilon;T,x,0)(\cdot) \geq l$ and $X_0(2T - \varepsilon;T,x,0) \geq l + \alpha$, we have $X(2T - \varepsilon;T,x,0) \in \mathcal{D}_0$. Therefore, by Lemma 2.2.6-(2), there exists $\alpha' > 0$ such that $X_0(t;2T - \varepsilon, X(2T -$

$\varepsilon; T, x, 0), 0) \geq l + \alpha'$, for all $t \geq 2T - \varepsilon$, so that we get what desired taking $\beta = \alpha \wedge \alpha'$.

About the second part of the statement, notice that, when $X_0(t; T, x, 0) = l$, thanks to Proposition 3.2.3, we have from the state equation with delay $dX_0(t; T, x, 0) \leq rl - q = 0$. On the other hand $x \in D(V)$, so that we know, by the point (1), that $0 \in \Theta_{ad}(x)$; thus also $dX_0(t; T, x, 0) \geq 0$. Therefore we get $dX_0(t; T, x, 0) = 0$ and the proof is complete.

3. If $x \in D(V)$ is such that $X_0(\cdot; T, x, 0) \geq l + \beta$ for some $\beta > 0$, then, arguing as in Lemma 2.4.1, we get $x \in \mathcal{V}$. Conversely let $x \in \mathcal{V}$; of course $x \in D(V)$; thus we have to prove that $X_0(t; T, x, 0) > 0$ for $t \in [T, 2T)$ and then, by the point (2), we get the claim. So let us suppose by contradiction that, for some $s \in [T, 2T)$, we have $X_0(s; T, x, 0) = l$; since $x \in \mathcal{V}$, there exists $\varepsilon > 0$ such that $B_{(E, \|\cdot\|_H)}(x, \varepsilon) \subset \mathcal{V} \subset D(V) = \{x \in \mathcal{C} \mid 0 \in \Theta_{ad}(x)\}$ and in particular it has to be $x_0 > l$. On the other hand we can choose $y \in B_{(E, \|\cdot\|_H)}(x, \varepsilon)$ such that $y_1(\zeta) = x_1(\zeta)$, for $\zeta \in [-T, s + \frac{2T-s}{2} - 2T]$, and $y_0 < x_0$. Working in the interval $[T, s + \frac{2T-s}{2}]$ we can forget to be concerned with a delay equation and consider the term $x(t-T)$ in the equation as a datum. Thus, thanks to Proposition 3.2.3, we have for the dynamics of $X_0(\cdot; T, x, 0)$ and $X_0(\cdot; T, y, 0)$ in the interval $[T, s + \frac{2T-s}{2}]$

$$\begin{cases} dX_0(t; T, x, 0)) = [rX_0(t; T, x, 0) - q]\, dt - f_0\left(X_0(t; T, x, 0) - x_1(t - 2T)\right) dt, \\ X_0(T; T, x, 0) = x_0; \end{cases}$$

$$\begin{cases} dX_0(t; T, y, 0)) = [rX_0(t; T, y, 0) - q]\, dt - f_0\left(X_0(t; T, y, 0) - y_1(t - 2T)\right) dt, \\ X_0(T; T, y, 0) = y_0. \end{cases}$$

These two dynamics refer to the same ordinary differential equation on the interval $[T, s + \frac{2T-s}{2}]$, since $y_1(\zeta) = x_1(\zeta)$, for $\zeta \in [-T, s + \frac{2T-s}{2} - 2T]$. Such differential equation satisfies the classic hypothesis of the Cauchy's Theorem for ordinary differential equations. Therefore, by uniqueness, since $y_0 < x_0$, the solution starting at y_0 has to stay strictly below the solution starting at x_0; in particular, since $X_0(s; T, x, 0) = l$, we get $X_0(s; T, y, 0) < l$, i.e. $0 \notin \Theta_{ad}(y)$ and the contradiction arises.

4. If we denote by the symbol $\mathring{\cup}$ the disjoint union, we have

$$\mathrm{Clos}_{(E, \|\cdot\|_H)}(\mathcal{V}) = \mathcal{V} \mathring{\cup} \mathrm{Fr}_{(E, \|\cdot\|_H)}(\mathcal{V}),$$

since $\mathcal{V}$ is open. By the point (2) we can split $D(V)$ as

$$D(V) = \big\{x \in D(V) \mid \exists\, \beta > 0 \text{ s.t. } X_0(\cdot; T, x, 0) \geq l + \beta\big\} \\ \mathring{\cup} \big\{x \in D(V) \mid \exists\, s \in [T, 2T) \text{ s.t. } X_0(s; T, x, 0) = l\big\}.$$

On the other hand, by the point (3), we can say that

$$\mathcal{V} = \{x \in D(V) \mid \exists\, \beta > 0 \text{ s.t. } X_0(\cdot; T, x, 0) \geq l + \beta\},$$

so that, by the point (1) it has to be

$$\mathrm{Fr}_{(E,\|\cdot\|_H)}(\mathcal{V}) = \{x \in D(V) \mid \exists\, s \in [T, 2T) \text{ s.t. } X_0(s; T, x, 0) = l\},$$

i.e. the claim.

5. Let us suppose that $x \in \mathrm{Fr}_{(E,\|\cdot\|_H)}(\mathcal{V})$. We know, by the point (1), that $0 \in \Theta_{ad}(x)$; moreover, by the point (4), there exists $s \in [T, 2T)$ such that $X_0(s; T, x, 0) = l$. Let us suppose that $\theta(\cdot) \in \Theta_{ad}(x)$. Then, arguing as in the proof of Lemma 2.2.3, we get, using the same notation therein, that $\tilde{\mathbb{E}}_{2T}\left[X_0(s; T, x, \theta(\cdot))\right] \leq X_0(s; T, x, 0) = l$. On the other hand, since $\theta(\cdot) \in \Theta_{ad}(x)$, it has to be $X_0(s; T, x, \theta(\cdot)) \geq l$ almost surely; therefore we can say that $X_0(s; T, x, \theta(\cdot)) = l$ almost surely, so that $\mathrm{Var}\left[X_0(s, T, x, \theta(\cdot))\right] = 0$ and this fact can occur only if $\theta(t) \equiv 0$ for $t \in [T, s]$. Afterwards, for $t \geq s$, of course it must be $\theta(t) \equiv 0$. Therefore we have proved that

$$\mathrm{Fr}_{(E,\|\cdot\|_H)}(\mathcal{V}) \subset \{x \in D(V) \mid \Theta_{ad}(x) = \{0\}\}.$$

Conversely let us suppose that $x \in D(V)$ is such that $\Theta_{ad}(x) = \{0\}$. If, by contradiction, $x \in \mathcal{V}$, then we can find $\varepsilon > 0$ such that $B_{(E,\|\cdot\|_H)}(x, \varepsilon) \subset \mathcal{V}$. Let us consider the constant strategy $\theta \equiv 1$ and let us define the stopping time

$$\tau := \inf\{t \geq T \mid X(t; T, x, 1) \notin B_{(E,\|\cdot\|_H)}(x, \varepsilon)\}.$$

The trajectories of $X(\cdot; T, x, 1)$ are $\|\cdot\|_E$-continuous and therefore they are also $\|\cdot\|_H$-continuous, so that $\tau > T$ almost surely. Now define the strategy

$$\theta(t) = \begin{cases} 1, & \text{if } t \leq \tau, \\ 0, & \text{if } t > \tau. \end{cases}$$

By definition $X(\tau; T, x, 1) \in \mathcal{V}$, so that in particular $0 \in \Theta_{ad}(X(\tau; T, x, 1))$; therefore we have $\theta(\cdot) \neq 0$ and $\theta(\cdot) \in \Theta_{ad}(x)$, so that a contradiction arises.

6. Let $t \geq T$; by the point (5), if $x \in \mathrm{Fr}_{(E,\|\cdot\|_H)}(\mathcal{V})$, then the only admissible strategy is the null one; therefore it has to be also $\Theta_{ad}(X(t; T, x, 0)) = \{0\}$. Applying again the point (5) we get $X(t; T, x, 0) \in \mathrm{Fr}_{(E,\|\cdot\|_H)}(\mathcal{V})$, i.e. the claim. $\square$

Proposition 2.4.7. *Let $U(l) > -\infty$, $rl = q$. Then the (extended) value function*

$$V : \mathcal{O} \cup Fr_{(E,\|\cdot\|_H)}(\mathcal{V}) \to \mathbb{R}$$

is continuous at the boundary $Fr_{(E,\|\cdot\|_H)}(\mathcal{V})$.

Proof. (i) Here we prove the lower semicontinuity at the boundary. So let $x \in \mathrm{Fr}_{(E,\|\cdot\|_H)}(\mathcal{V})$ and notice that, since V is continuous on $\mathcal{O}$ and $\mathcal{V}$ is dense in $\mathcal{O}$, without loss of generality we can prove that

$$V(x) \leq \liminf_{\substack{y \to x \\ y \in \mathrm{Clos}_{(E,\|\cdot\|_H)}(\mathcal{V})}} V(y).$$

Set $X(t) := X(t, T, x, 0)$; by Proposition 2.4.6-(5) we know that the only admissible strategy is the null one and, by Proposition 2.4.6-(4), there exists $s \in [T, 2T)$ such that $X_0(t) = l$ for every $t \geq s$; therefore

$$V(x) = J(x, 0) = \int_T^s e^{-\rho(t-T)} U(X_0(t))dt + \int_s^{+\infty} e^{-\rho(t-T)} U(l)dt.$$

Take $\varepsilon > 0$, $y \in B_{(E,\|\cdot\|_H)}(x, \varepsilon) \cap \mathrm{Clos}_{(E,\|\cdot\|_H)}(\mathcal{V})$ and set $Y(t) := X(t; T, y, 0)$. Of course $0 \in \Theta_{ad}(y)$ and

$$\begin{aligned} V(y) \geq J(y, 0) &= \int_T^s e^{-\rho(t-T)} U(Y_0(t))dt + \int_s^{+\infty} e^{-\rho(t-T)} U(Y_0(t))dt \\ &\geq \int_T^s e^{-\rho(t-T)} U(Y_0(t))dt + \int_s^{+\infty} e^{-\rho(t-T)} U(l)dt. \end{aligned}$$

By Corollary 2.3.15 we get $|X_0(t) - Y_0(t)| \leq K_s \, \varepsilon$, for $t \in [T, s]$, where K_s is the constant given in the same Corollary. Therefore we get the claim by uniform continuity of U.

(ii) As we said, we give only a sketch of the proof of the upper semicontinuity at the boundary. Again, without loss of generality, we can reduce the problem to prove

$$V(x) \geq \limsup_{\substack{y \to x \\ y \in \mathrm{Clos}_{(E,\|\cdot\|_H)}(\mathcal{V})}} V(y).$$

The heart of the idea is that our value function is obviously smaller of the value function of the corresponding problem without surplus, i.e. the value function which is the object of Section 1.3. Here wel call this value function without surplus V^0. We know, by Proposition 1.3.14, that V^0 is continuous up to l.

Let $x \in \mathrm{Fr}_{(E,\|\cdot\|_H)}(\mathcal{V})$ be such that $x_0 = l$; then, by Proposition 2.4.6-(2), we get $V(x) = U(l)/\rho = V^0(x_0)$. On the other hand, since for each $x \in \mathcal{C}$ we have

$V(x) \leq V^0(x_0)$, we must have

$$\limsup_{\substack{y \to x \\ y \in \mathrm{Clos}_{(E,\|\cdot\|_H)}(\mathcal{V})}} V(y) \leq \limsup_{\substack{y \to x \\ y \in \mathrm{Clos}_{(E,\|\cdot\|_H)}(\mathcal{V})}} V^0(y_0) = V^0(x_0) = V(x),$$

so that the claim is proved in this case.

Now let $x \in \mathrm{Fr}_{(E,\|\cdot\|_H)}(\mathcal{V})$ be such that $x_0 > l$; then, by Proposition 2.4.6, we can write, for some $s \in [T, 2T)$,

$$V(x) = \int_T^s e^{-\rho(t-T)} U(X_0(t; T, x, 0)dt + \int_s^{+\infty} e^{-\rho(t-T)} U(l)dt.$$

Let us take a sequence $(y_n) \subset \mathrm{Clos}_{(E,\|\cdot\|_H)}(\mathcal{V})$ such that $y_n \stackrel{\|\cdot\|_H}{\longrightarrow} x$; we can write, by Dynamic Programming Principle (see Proposition 2.5.7 and Remark 2.5.8),

$$\begin{aligned} & V(y_n) \\ \leq & \sup_{\theta(\cdot) \in \Theta_{ad}(y_n)} \mathbb{E}\left[\int_T^s e^{-\rho(t-T)} U(X_0(t; T, x, 0)dt + e^{-\rho(s-T)} V(X(s; T, y_n, \theta(\cdot)))\right] \\ \leq & \sup_{\theta(\cdot) \in \Theta_{ad}(y_n)} \mathbb{E}\left[\int_T^s e^{-\rho(t-T)} U(X_0(t; T, x, 0)dt + e^{-\rho(s-T)} V^0(X_0(s; T, y_n, \theta(\cdot)))\right]. \end{aligned}$$

Then, using Girsanov's Theorem A.1.1, the convexity of the map f_0, Corollary 2.3.15 and the continuity of V^0, one could prove that the limsup of the last term in the previous inequality is less than $V(x)$, concluding the proof. □

Remark 2.4.8. As we have seen also in Subsection 1.2.4, the proof of the continuity of the value function is not an easy topic in the context of control problems with state constraints. The fact that the value function is concave was very useful to prove the continuity result in the interior of the domain in this section. Unfortunately the concavity property does not say anything about the continuity of the value function at the boundary and we have to work directly with estimates on the state equation to prove continuity properties. As we have seen, when the boundary is absorbing we can estimate the value function at the boundary due to the fact that the only admissible strategy in this case is the null one. This fact together with the estimate $V \leq V^0$ allowed us to prove the continuity at the boundary when it is absorbing. When the boundary is not absorbing is absolutely not clear how we could proceed to prove continuity properties. □

2.5 The Hamilton-Jacobi-Bellman equation

In this section we write and study the infinite-dimensional Hamilton-Jacobi-Bellman equation associated with the value function. Unless differently speci-

fied we intend the topological notions referred to the $\|\cdot\|_H$-norm.
For $x \in H$, let us define

$$\begin{array}{rccc} \Sigma(x): & \mathbb{R} & \longrightarrow & H. \\ & a & \longmapsto & a\,(\Phi x) \end{array}$$

For an operator $Q \in \mathcal{L}(H)$, we can use the representation on the components $\mathbb{R}$ and L^2_{-T},

$$Q = \begin{pmatrix} Q_{11} & Q_{12} \\ Q_{21} & Q_{22} \end{pmatrix},$$

where $Q_{11} \in \mathcal{L}(\mathbb{R}) \cong \mathbb{R}$, $Q_{12} \in \mathcal{L}(L^2_{-T};\mathbb{R})$, $Q_{21} \in \mathcal{L}(\mathbb{R}, L^2_{-T})$ and $Q_{22} \in \mathcal{L}(L^2_{-T})$.

Let us define, for $x = (x_0, x(\cdot)) \in E, p \in D(A^*)$, $Q \in \mathcal{L}(H)$ and $\theta \in [0,1]$, the function

$$\begin{aligned} \mathcal{H}_{cv}(x,p,Q;\theta) \;&:=\; U(x_0) + \frac{1}{2}\theta^2\sigma^2 \mathrm{Tr}[Q\Sigma(x)\Sigma(x)^*] + \langle \sigma\lambda\theta\Phi x - F(x), p\rangle \\ &=\; U(x_0) - f(x)p_0 + \frac{1}{2}\theta^2\sigma^2 x_0^2 Q_{11} + \sigma\lambda\theta x_0 p_0. \end{aligned}$$

Formally the HJB equation associated with V in the space H is

$$\rho v(x) = \langle x, A^* v_x(x)\rangle + \mathcal{H}(x, v_x(x), v_{xx}(x)), \tag{2.32}$$

where

$$\mathcal{H}(x,p,Q) := \sup_{\theta\in[0,1]} \mathcal{H}_{cv}(x,p,Q;\theta).$$

Notice that we can write

$$\mathcal{H}(x,p,Q) = U(x_0) - f(x)p_0 + \sup_{\theta\in[0,1]} \left\{ \frac{1}{2}\theta^2\sigma^2 x_0^2 Q_{11} + \sigma\lambda\theta x_0 p_0 \right\}$$

As in Section 1.2 and Section 1.3, we define

$$\mathcal{H}^0_{cv}(x_0,p_0,Q_{11};\theta) := \frac{1}{2}\theta^2\sigma^2 x_0^2 Q_{11} + \sigma\lambda\theta x_0 p_0. \tag{2.33}$$

When $p_0 \geq 0$, $Q_{11} \leq 0$, $p_0^2 + Q_{11}^2 > 0$, has a unique maximum point over $\theta \in [0,1]$ given by

$$\theta^* = -\frac{\lambda p_0}{\sigma x Q_{11}} \wedge 1$$

(where we mean that, for $Q_{11} = 0$, $\theta^* = 1$) and

$$\begin{aligned} \mathcal{H}^0(x_0,p_0,Q_{11}) := \sup_{\theta\in[0,1]} \mathcal{H}^0_{cv}\,(x_0,p_0,Q_{11};\theta) \\ = \begin{cases} -\frac{\lambda^2 p_0^2}{2Q_{11}}, & \text{if} \quad \theta^* < 1, \\ \sigma\lambda x_0 p_0 + \frac{1}{2}\sigma^2 x_0^2 Q_{11}, & \text{if} \quad \theta^* = 1. \end{cases} \end{aligned}$$

When $p_0 = Q_{11} = 0$, each $\theta \in [0, 1]$ is a maximum point and we have

$$\mathcal{H}^0 (x_0, p_0, Q_{11}) = 0.$$

Thus (2.32) can be rewritten as

$$\rho v(x) = \langle x, A^* v_x(x) \rangle_H + U(x_0) - f(x) v_{x_0}(x) + \mathcal{H}^0 (x_0, v_{x_0}(x), v_{x_0 x_0}(x)). \quad (2.34)$$

Remark 2.5.1. We want to underline some specific features of the above HJB equation.

- It is defined on the points of E, due the presence of f which is defined on this space.
- It is fullynonlinear, as the nonlinearity involves also the second derivative.
- The linear term is unbounded.
- $f(\cdot)$ is not continuous with respect to $\| \cdot \|_H$.
- The terms associated with the control θ involve only the derivatives with respect to the real component of v: therefore one may hope to prove a verification theorem giving optimal feedback strategies even only having regularity properties of the value function with respect to the real component (see the argument used in Chapter 3). This makes clear the importance of Proposition 2.4.4 which splits the real and the infinite dimensional component in the argument of the value function, leaving open the possibility of studying the regularity of this function only with respect to the real component. Of course at this stage this possibility is only theoretical, since we have given only an abstract extension of the value function on the set $\mathcal{O}$; we have no constructive information about this extended function at the points of $\mathcal{O} \setminus \mathcal{V}$, because it is not viewed at these points as optimum of an infinite-dimensional control problem. $\square$

Hamilton-Jacobi equations in infinite dimension are treated in literature basically by means of three different approaches.

- *Regular solutions.* This approach has been started by Barbu and Da Prato (see

  - [Barbu, Da Prato; 1981],

  - [Barbu, Da Prato; 1983a],

  -[Barbu, Da Prato; 1983b])

and then developed by many other authors. This theory is based on the deep study of the linear part of the Hamilton-Jacobi equation when the nonlinear part does not appear (i.e. the equation reduces to the so called Kolmogorov equations studied in Chapter 9 of [Da Prato, Zabczyk; 1992]) and then on the use of perturbation methods for the treatment of the nonlinear part.

- *Mild solutions and backward SDEs.* For this approach we mainly refer to the paper [Fuhrman, Tessitore; 2004]. It provides existence and uniqueness of mild solutions of the HJB equation by means of forward and backward infinite-dimensional stochastic evolution equations.

- *Viscosity Solutions.* The theory of viscosity solutions to Hamilton-Jacobi equations in infinite dimension has been started by Crandall and Lions for the first-order case in a series of papers (we refer in particular to [Crandall, Lions; 1990] and [Crandall, Lions; 1991] for the case of unbounded linear term). The viscosity approach to infinite-dimensional HJB equations coming from control problems with state constraint is studied in [Cannarsa, Gozzi, Soner; 1991] and [Kocan, Soravia; 1998] in the deterministic case. There are not many papers treating the second-order case: we mainly refer to
  - [Gozzi, Rouy, Swiech; 2000],
  - [Gozzi, Swiech; 2000],
  - [Gozzi, Sritharan, Swiech; 2005],
  - [Ishii; 1993],
  - [Kelome, Swiech; 2003],
  - [Lions; 1988], [Lions; 1989a], [Lions; 1989b],
  - [Swiech; 1994].

The first two methods work in the case of semilinear equations, i.e. when the equation is linear on the second derivative. In particular the second methods makes addressable also cases of nonlinearities in the first derivatives leaving out by the fisrt method as well as remove trace conditions on the second derivatives required by the first method.

The viscosity approach works also when the equation is fully nonlinear, i.e. when the nonlinearity involves also the second derivative. So, since as observed our equation is fully nonlinear, we choose to treat the HJB by a viscosity approach. The nearest paper seems to be [Kelome, Swiech; 2003], where a second-order fully nonlinear equation with a similar unbounded operator is studied and an existence-uniqueness result is proved. Nevertheless our problem is more difficult with respect to the one studied in [Kelome, Swiech; 2003], due to its particular features: the equation is defined in a subspace of the Hilbert space, it is a boundary problem due to the state constraint and the Hamiltonian is not continuous on the state variable with respect to the norm of the Hilbert space.

2.5.1 Rewriting the problem with a maximal monotone operator

Dealing with viscosity solution, in order to be able to get a uniqueness result for the equation, it is very important to have a maximal dissipative operator as linear operator in the equation. Indeed this fact allows to take as test functions also the radial functions (see Definition 2.5.4-(ii)). This class of functions is needed to obtain good test functions on which the viscosity property can be used to prove a comparison result. Therefore, in order to work with a maximal dissipative operator, we rewrite the state equation as

$$\begin{cases} dX(t) = \tilde{A}X(t)dt + \left[\left(r+\frac{1}{2}\right) + \sigma\lambda\theta(t)\right]\Phi X(t)dt - F(X(t))dt + \sigma\theta(t)\Phi X(t)dB(t), \\ X(T) = x, \end{cases} \tag{2.35}$$

where $\tilde{A} = A - \left(r+\frac{1}{2}\right)\Phi$; of course X is a mild solution to (2.35) if and only if it is a mild solution to (2.16). We also rewrite the HJB equation (2.32) as

$$\rho v(x) = \langle x, \tilde{A}^* v_x(x)\rangle + \tilde{\mathcal{H}}(x, v_x(x), v_{xx}(x)) \tag{2.36}$$

where, for $x \in E$, $p \in D(\tilde{A}^*)$, $Q \in \mathcal{L}(H)$,

$$\tilde{\mathcal{H}}(x,p,Q) = \sup_{\theta\in[0,1]} \tilde{\mathcal{H}}_{cv}(x,p,Q;\theta),$$

and

$$\tilde{\mathcal{H}}_{cv}(x,p,Q;\theta) = U(x_0) + \left(r+\frac{1}{2}\right)x_0p_0 - f(x)p_0 + \mathcal{H}^0_{cv}(x_0,p_0,Q_{11};\theta),$$

where $\mathcal{H}^0_{cv}$ was defined in (2.33). Notice that $D(\tilde{A}^*) = D(A^*)$ and

$$\tilde{A}^* = A^* - \left(r+\frac{1}{2}\right)\Phi^* = A^* - \left(r+\frac{1}{2}\right)\Phi.$$

The following proposition gives the desired properties of the operator $\tilde{A}$; we will use the dissipativity of $\tilde{A}$ to obtain a Dynkin type formula with inequality for radial functions.

Proposition 2.5.2. *The operator $\tilde{A}$ is maximal dissipative.*

Proof. (i) We have, for $x \in D(\tilde{A}) = D(A)$, taking into account that $x_1(0) = x_0$,

$$\langle \tilde{A}x, x\rangle = -\frac{1}{2}x_0^2 + \int_{-T}^0 x_1'(\xi)x_1(\xi)d\xi = -\frac{1}{2}x_0^2 + \left[\frac{x_1(\cdot)^2}{2}\right]_{-T}^0 = -x_1(-T)^2 \le 0,$$

so that $\tilde{A}$ is dissipative.

(ii) In order to prove that $\tilde{A}$ is maximal we have to prove that the image of $(\tilde{A} - I)$ is the whole space H; this means that, for each $y = (y_0, y_1(\cdot)) \in H$, we must be able to find $x = (x_0, x_1(\cdot)) \in D(\tilde{A})$ such that

$$\begin{cases} -\frac{3}{2}x_0 = y_0, \\ x_1'(\cdot) - x_1(\cdot) = y_1(\cdot) \quad \text{a.e.}; \end{cases}$$

this means that we must be able to solve the first order ordinary problem of finding $f \in W^{1,2}_{-T}$ such that, for given $g \in L^2_{-T}$,

$$\begin{cases} f' - f = g, \\ f(0) = \alpha \in \mathbb{R}; \end{cases}$$

As in the classical case the solution is given by the variation of constants formula

$$f(t) = \left(\alpha - \int_{-T}^{0} g(\xi) e^{-\xi} d\xi \right) e^t + \int_{-T}^{t} g(\xi) e^{t-\xi} d\xi.$$

□

2.5.2 Test functions and Dynkin type formulae

In this section we define two sets of functions which will play an important role in the definition of viscosity solution to the HJB equation and we prove Dynkin type formulae for these functions applied to the process X mild solution to

$$\begin{cases} dX(t) = \tilde{A}X(t)dt + \left[\left(r + \frac{1}{2}\right) + \sigma\lambda\theta(t)\right] \Phi X(t)dt - G(X(t))dt + \sigma\theta(t)\Phi X(t)dB(t), \\ X(T) = x \in E, \end{cases} \tag{2.37}$$

where $G : E \to H$, $x \mapsto (g(x), 0)$, and $g : (E, \|\cdot\|_E) \to \mathbb{R}$ is Lipschitz continuous. We have proved in Theorem 2.3.10, for any given $[0,1]$-valued and $(\mathcal{F}_t)_{t \geq T}$-progressively measurable process $\theta(\cdot)$, existence and uniqueness of a mild solution to equation (2.37) in the class $C_{\mathcal{P}}\big([T, +\infty); L^2(\Omega, (E, \|\cdot\|_E))\big)$.

Let us start with an approximation result which will be useful for our goal.

Lemma 2.5.3. *Let $X(\cdot)$ be the mild solution to equation (2.37) and let $\tilde{A}_n$ be the Yosida approximations for the operator $\tilde{A}$. Then the stochastic differential equation*

$$\begin{cases} dX_n(t) = \tilde{A}_n X_n(t)dt + \left[\left(r + \frac{1}{2}\right) + \sigma\lambda\theta(t)\right] \Phi X(t)dt - G(X(t))dt + \sigma\theta(t)\Phi X(t)dB(t), \\ X_n(T) = x, \end{cases} \tag{2.38}$$

admits a unique strong solution in $C_{\mathcal{P}}([T,+\infty);L^2(\Omega,E))$*. Moreover, for any* $t\in[T,+\infty)$,

$$X_n(t)\to X(t),\ \ \textit{in } L^2(\Omega;H), \tag{2.39}$$

and, for any $a\in[T,+\infty)$,

$$X_n\to X,\ \ \textit{in } \ L^2(\Omega\times[T,a];H). \tag{2.40}$$

Proof. First of all notice that, in the equation (2.38), $X(\cdot)$ is a given process. Let $\tilde{S}_n$ be the uniformly continuous semigroups on H genenerated by the bounded operators $\tilde{A}_n$ and $\tilde{S}$ that generated by the unbounded operator $\tilde{A}$. $\tilde{A}$ is maximal dissipative, so that it generates a contractive semigroup (see Corollary II.3.20 in [Engel, Nagel; 2000]); therefore, for any $t\geq 0$,

$$\|\tilde{S}(t)\|_{\mathcal{L}(H)}\leq 1. \tag{2.41}$$

Moreover also the semigroups $\tilde{S}_n$ are contractive (see Theorem II.3.5 again in [Engel, Nagel; 2000]), so that, for any $t\geq 0$,

$$\|\tilde{S}_n(t)\|_{\mathcal{L}(H)}\leq 1. \tag{2.42}$$

Moreover, for any $t\geq 0$ and $x\in H$, we have the convergence, for $n\to\infty$,

$$\|\tilde{S}(t)x-\tilde{S}_n(t)x\|_H\to 0. \tag{2.43}$$

The proof of the existence and uniqueness for the strong solution to (2.38) follows e.g. by Proposition 6.4 of [Da Prato, Zabczyk; 1992], because of the boundedness of $\tilde{A}_n$. Of course a strong solution is also a mild solution, therefore we can write

$$\begin{aligned} X_n(t) &= \tilde{S}_n(t-T)x+\int_T^t\left[\left(r+\frac{1}{2}\right)+\sigma\lambda\theta(\tau)\right]\tilde{S}_n(t-\tau)\Phi X(\tau)d\tau \\ &\quad -\int_T^t\tilde{S}_n(t-\tau)G(X(\tau))d\tau+\sigma\int_T^t\theta(\tau)\tilde{S}_n(t-\tau)\Phi X(\tau)dB(\tau). \end{aligned}$$

Therefore, setting $K:=r+\frac{1}{2}+\sigma(\lambda+1)$,

$$\begin{aligned} \|X(t)-X_n(t)\|_H &\leq \|\tilde{S}(t-T)x-\tilde{S}_n(t-T)x\|_H \\ &\quad +K\int_T^t\|(\tilde{S}(t-\tau)-\tilde{S}_n(t-\tau))\Phi X(\tau)\|_H d\tau \\ &\quad +\int_T^t\|(\tilde{S}(t-\tau)-\tilde{S}_n(t-\tau))\,G(X(\tau))\|_H d\tau. \end{aligned} \tag{2.44}$$

Let C_g be the Lipschitz constant of the map $g:(E,\|\cdot\|_E)\to\mathbb{R}$; then

$$\begin{aligned} \mathbb{E}\left[\int_T^t\|G(X(\tau))\|_H^2 d\tau\right] &\leq \mathbb{E}\left[\int_T^t\|G(X(\tau))\|_E^2 d\tau\right]\leq\mathbb{E}\left[\int_T^t|g(X(\tau))|^2 d\tau\right] \\ &\leq 2\mathbb{E}\left[\int_T^t\left(C_g^2\|X(\tau)\|_E^2+|g(0)|^2\right)d\tau\right]<+\infty. \end{aligned} \tag{2.45}$$

The first and the second term of the right-handside of (2.44) can be dominated in $L^2(\Omega;\mathbb{R})$ thanks to (2.41), (2.42) and by Hölder's inequality; the third one can be dominated thanks to (2.41), (2.42), (2.45) and by Hölder's inequality. Moreover the right-handside of (2.44) converges pointwise to 0, when $n \to \infty$, thanks to (2.43). Therefore (2.39) follows by dominated convergence from (2.44) taking the expectations and letting $n \to \infty$.

Integrating (2.44) on $[T,a]$, taking the expectations and then letting $n \to \infty$ we get in the same way (2.40) by dominated convergence. □

Definition 2.5.4. (i) We call $\mathcal{T}_1$ the set of functions $\psi \in C^2(H)$ such that $\psi_x(x) \in D(A^*)$ for any $x \in H$ and $\psi, \psi_x, \tilde{A}^*\psi_x, \psi_{xx}$ are uniformly continuous.

(ii) We call $\mathcal{T}_2$ the set of functions $g \in C^2(H)$ which are radial and nondecreasing, i.e.

$$g(x) = g_0(\|x\|), \quad g_0 \in C^2([0,+\infty);\mathbb{R}), \quad g_0' \geq 0,$$

and g, g_x, g_{xx} are uniformly continuous. □

Let us define, for $\theta \in [0,1]$, the operator $\mathcal{L}^\theta$ on $\mathcal{T}_1$ by

$$\begin{aligned}[\mathcal{L}^\theta\psi](x) \quad := \quad & -\rho\psi(x) + \langle x, \tilde{A}^*\psi_x(x)\rangle + \left[r + \frac{1}{2} + \sigma\lambda\theta\right]\langle \Phi x, \psi_x(x)\rangle \\ & + \langle G(x), \psi_x(x)\rangle + \frac{1}{2}\sigma^2\theta^2 \operatorname{Tr}[\Sigma(x)\Sigma(x)^*\psi_{xx}(x)]. \end{aligned} \tag{2.46}$$

Lemma 2.5.5 (Dynkin's formula (i))**.** *Let $\psi \in \mathcal{T}_1$, let $X(\cdot)$ be the solution to (2.37) and let τ be a bounded stopping time; then we have*

$$\mathbb{E}\left[e^{-\rho(\tau-T)}\psi(X(\tau)) - \psi(x)\right] = \mathbb{E}\left[\int_T^\tau e^{-\rho(t-T)}[\mathcal{L}^{\theta(t)}\psi](X(t))dt\right].$$

Moreover the same formula holds true also for stopping time almost surely finite such that, for some $r > 0$, $X(t) \in B(x,r)$ for $t \leq \tau$.

Proof. *First step.* Let us suppose that τ takes a finite number of finite values; we can apply the Dynkin's formula to the approximating processes X_n of Lemma 2.5.3 (see Theorem 4.7 of [Da Prato, Zabczyk; 1992]) with the function $e^{-\rho(t-T)}\psi(x)$ to get

$$\begin{aligned}\mathbb{E}\left[e^{-\rho(\tau-T)}\psi(X_n(\tau)) - \psi(x)\right] = \mathbb{E}\Bigg[\int_T^\tau e^{-\rho(t-T)}\Big(& -\rho\psi(X_n(t)) \\ + \langle \tilde{A}_n X_n(t), \psi_x(X_n(t))\rangle + \left[r + \frac{1}{2} + \sigma\lambda\theta(t)\right] & \langle \Phi X(t), \psi_x(X_n(t))\rangle \\ - \langle G(X(t)), \psi_x(X_n(t))\rangle + \frac{1}{2}\sigma^2\theta(t)^2 \operatorname{Tr}[\Sigma(X(t))\Sigma(X(t))^*\psi_{xx}(X_n(t))] & \Big)dt\Bigg];\end{aligned}$$

we want to get the first claim letting $n \to \infty$ and taking into account Lemma 2.5.3 and the continuity properties of ψ and its derivatives. By (2.39) we have $X_n(\tau) \to X(\tau)$ in $L^2(\Omega; H)$, so that we have the desired convergence in the left handside thanks to Lemma A.1.2. For the right handside we have the desired convergences of the bounded terms thanks to (2.40) and Lemma A.1.2, taking also into account the estimate (2.45) for the term containing G.

The only non trivial convergence in the right hand-side is that one concerned with the unbounded linear term, i.e.

$$\mathbb{E}\left[\int_T^\tau e^{-\rho(t-T)}\langle \tilde{A}_n X_n(t), \psi_x(X_n(t))\rangle dt\right] \longrightarrow \\ \longrightarrow \mathbb{E}\left[\int_T^\tau e^{-\rho(t-T)}\langle X(t), \tilde{A}^*\psi_x(X(t))\rangle dt\right].$$

Without loss of generality we can suppress in the following argument the term $e^{-\rho(t-T)}$ which is bounded. Let $x \in H$ and $(y_n) \subset D(\tilde{A}^*)$, $y \in D(\tilde{A}^*)$; we have

$$\begin{aligned} |\langle \tilde{A}_n x_n, y_n\rangle - \langle x, \tilde{A}^* y\rangle| &\leq |\langle \tilde{A}_n x_n, y_n\rangle - \langle \tilde{A}_n x, y_n\rangle| + |\langle \tilde{A}_n x, y_n\rangle - \langle \tilde{A}_n x, y\rangle| \\ &\quad + |\langle \tilde{A}_n x, y\rangle - \langle x, \tilde{A}^* y\rangle| \\ &= |\langle (x_n - x), \tilde{A}_n^* y_n\rangle| \\ &\quad + |\langle x, \tilde{A}_n^*(y_n - y)\rangle| + |\langle x, (\tilde{A}_n^* - \tilde{A}^*)y\rangle|. \end{aligned}$$

let us indicate by $\|\cdot\|_{\mathcal{L}}$ the operator norm for the linear bounded operators

$$\left(D(\tilde{A}^*), \|\cdot\|_{D(\tilde{A}^*)}\right) \to (H, \|\cdot\|_H).$$

We have the convergence $\tilde{A}_n^* v \to \tilde{A}^* v$ for any $v \in D(\tilde{A}^*)$ (see Proposition 4.13 in [Li, Yong; 1995], Lemma 3.4-(ii) and Corollary B.12 in [Engel, Nagel; 2000]). So, by the Banach-Steinhaus Theorem A.2.3 we have $\|\tilde{A}_n^*\|_{\mathcal{L}} \leq C$ for some $C > 0$. Thus

$$\begin{aligned} &|\langle \tilde{A}_n x_n, y_n\rangle - \langle x, \tilde{A}^* y\rangle| \\ &\qquad \leq C\Big[\|x_n - x\|_H \, \|y_n\|_{D(\tilde{A}^*)} + \|x\|_H \, \|y_n - y\|_{D(\tilde{A}^*|}\Big] + |\langle x, (\tilde{A}_n^* - \tilde{A}^*)y\rangle|. \end{aligned}$$

Therefore, by Hölder's inequality,

$$\begin{aligned} &\mathbb{E}\left[\int_T^\tau \Big|\langle \tilde{A}_n X_n(t), \psi_x(X_n(t))\rangle - \langle X(t), \tilde{A}^*\psi_x(X(t))\rangle\Big| dt\right] \\ &\qquad \leq C\,\mathbb{E}\left[\int_T^\tau \|X_n(t) - X(t)\|_H^2 dt\right] \mathbb{E}\left[\int_T^\tau \|\psi_x(X_n(t))\|_{D(\tilde{A}^*)}^2 dt\right] \\ &\qquad + C\,\mathbb{E}\left[\int_T^\tau \|X(t)\|_H^2 dt\right] \mathbb{E}\left[\int_T^\tau \|\psi_x(X_n(t)) - \psi_x(X(t))\|_{D(\tilde{A}^*)}^2 dt\right] \\ &\qquad\qquad + \mathbb{E}\left[\int_T^\tau \|X(t)\|_H \|(\tilde{A}_n^* - \tilde{A}^*)\psi_x(X(t))\|_{D(\tilde{A}^*)} dt\right]. \end{aligned}$$

The first and the second term of the right handside go to 0 thanks to (2.40) and Lemma A.1.2 (recall that we are assuming $\psi_x, \tilde{A}^*\psi_x$ uniformly continuous). The third one converges pointwise to 0 and the integrand is dominated by

$$(C + \|\tilde{A}^*\|_{\mathcal{L}}) \, \|X(\cdot)\|_H \, \|\psi_x(X(\cdot))\|_{D(\tilde{A}^*)},$$

which is integrable thanks to Hölder's inequality, so that we can conclude by dominated convergence.

Second step. Now let τ be a bounded stopping time; of course we may find a suquence (τ_n) of stopping time taking a finite number of finite values such that $\tau_n \uparrow \tau$ almost surely; by the first step for these stopping time the claim holds true, therefore we can pass to the limit and again conclude the proof in this case by dominated convergence.

Third step. Let τ be a stopping time almost surely finite such that, for some $r > 0$, $X(t) \in B(x, r)$ for $t \le \tau$; we can find a sequence of bounded stopping times (τ_n) such that $\tau_n \uparrow \tau$; for these stopping times the claim holds true by the second step; moreover notice that, by the equivalence result of Proposition 3.2.3, if $X(t) \in B(x, r)$ for $t \le \tau$, then $X(t) \in B_{(E,\|\cdot\|_E)}(x, r')$ for some r'; therefore, taking also into account the estimate (2.45), we can get the claim for τ passing to the limit by dominated convergence thanks to Lemma A.2.4. □

Now, for $\theta \in [0, 1]$, let us define the operator $\mathcal{G}^\theta$ on $\mathcal{T}_2$ by

$$\begin{aligned}[][\mathcal{G}^\theta g](x) \;:=\; & -\rho g(x) + \left[r + \frac{1}{2} + \sigma\lambda\theta\right] \langle \Phi x, g_x(x)\rangle \\ & + \langle G(x), g_x(x)\rangle + \frac{1}{2}\sigma^2\theta^2 \operatorname{Tr}[g_{xx}\Sigma(x)\Sigma(x)^*]. \end{aligned} \tag{2.47}$$

Lemma 2.5.6 (Dynkin's formula (ii))**.** *Let $g \in \mathcal{T}_2$, let $X(\cdot)$ be the solution to (2.37) and τ a stopping time almost surely finite such that, for some $r > 0$, $X(t) \in B(x, r)$ for $t \le \tau$; then we have*

$$\mathbb{E}\left[e^{-\rho(\tau - T)} g(X(\tau)) - g(x)\right] \le \mathbb{E}\left[\int_T^\tau e^{-\rho(t-T)} [\mathcal{G}^{\theta(t)} g](X(t)) dt\right].$$

Proof. The proof follows the same line of the proof of Lemma 2.5.5; but in this case we cannot have the convergence for the term $\langle \tilde{A}_n X_n(u), g_x(X_n(u))\rangle$; nevertheless we have

$$\langle \tilde{A}_n x_n, g_x(x_n)\rangle = \frac{g_0'(\|x_n\|)}{\|x_n\|} \langle \tilde{A}_n x, x_n\rangle;$$

since $\tilde{A}$ is dissipative we have $\langle \tilde{A}_n x_n, x_n\rangle \le 0$ for any $n \in \mathbb{N}$ and so the claim follows taking the limsup. □

2.5.3 The Dynamic Programming Principle and the value function as viscosity solution to the HJB equation

In this section we will work to investigate the properties of the value function as viscosity solution to the HJB equation (2.32) associated with it. The link between the HJB equation and the value function is given by the Dynamic Programming Principle.

Proposition 2.5.7 (Dynamic Programming Principle)**.** *Let $x \in D(V)$ and let $\left(\tau^{\theta(\cdot)}\right)_{\theta(\cdot)\in\Theta_{ad}(x)}$ be a family of stopping times with respect to $\mathcal{F}^T$ such that $\tau^{\theta(\cdot)} \in [T, +\infty)$ almost surely. Then, setting $X^{\theta(\cdot)}(t) := X(t; T, x, \theta(\cdot))$, we have*

$$V(x) = \sup_{\theta(\cdot)\in\Theta_{ad}(x)} \mathbb{E}\left[\int_T^{\tau^{\theta(\cdot)}} e^{-\rho(t-T)} U\left(X_0^{\theta(\cdot)}(t)\right) dt + e^{-\rho(\tau^{\theta(\cdot)}-T)} V\left(X^{\theta(\cdot)}(\tau^{\theta(\cdot)})\right)\right].$$

□

Remark 2.5.8. We do not give the proof of the statement of the previous Proposition. For a proof of this kind of statement in the infinite dimensional case we refer to [Gozzi, Sritharan, Swiech; 2005], when the value function is continuous and the state unconstrained. Similar arguments can be used to prove the result in our case, taking into account the separability of our space H. We want to stress that in Proposition 2.4.7 we proved the continuity of the value function at the boundary only using the "easy" inequality of the Dynamic Programming Principle, which can be proved without measurable selection arguments, so that we could use the continuity of the value function to prove, without loss of generality, the Dynamic Programming Principle. □

Now we give a definition of viscosity solution for the equation (2.32); we will prove that the value function solves (2.32) in this sense. Recall that the set $\mathcal{V}$ was defined in (2.28) and the set $\mathcal{O}$ was defined in Proposition 2.4.4. Recall also that $\mathcal{V} = \mathcal{O} \cap E$.

Definition 2.5.9. (i) A continuous function $v : \mathcal{O} \longrightarrow \mathbb{R}$ is called a viscosity subsolution to equation (2.32) on $\mathcal{V}$ if, for any triple $(x_M, \psi, g) \in \mathcal{V} \times \mathcal{T}_1 \times \mathcal{T}_2$ such that x_M is a local maximum point of $v - \psi - g$, we have

$$\rho v(x_M) \leq \langle x_M, \tilde{A}^*\psi_x(x_M)\rangle + \tilde{\mathcal{H}}(x_M, \psi_x(x_M) + g_x(x_M), \psi_{xx}(x_M) + g_{xx}(x_M)).$$

(ii) A continuous function $v : \mathcal{O} \longrightarrow \mathbb{R}$ is called a viscosity supersolution to equation (2.32) on $\mathcal{V}$ if, for any triple $(x_m, \psi, g) \in \mathcal{V} \times \mathcal{T}_1 \times \mathcal{T}_2$ such that x_m is a local minimum point of $v - \psi + g$, we have

$$\rho v(x_m) \geq \langle x_m, \tilde{A}^*\psi_x(x_m)\rangle + \tilde{\mathcal{H}}(x_m, \psi_x(x_m) - g_x(x_m), \psi_{xx}(x_m) - g_{xx}(x_m)).$$

(iii) A continuous function $v : \mathcal{O} \longrightarrow \mathbb{R}$ is called a viscosity solution to equation (2.32) on $\mathcal{V}$ if it is both a viscosity subsolution and a viscosity supersolution. □

Proposition 2.5.10. *The value function V is a viscosity solution to (2.32) on $\mathcal{V}$.*

Proof. (i) Let x_m, ψ, g be as in Definition 2.5.9-(ii); without loss of generality we assume

$$V(x_m) = \psi(x_m) - g(x_m). \tag{2.48}$$

Let $B_{(H,\|\cdot\|_H)}(x_m, \varepsilon) \subset \mathcal{O}$ be such that

$$V(x) \geq \psi(x) - g(x), \quad \forall x \in B_{(H,\|\cdot\|_H)}(x_m, \varepsilon). \tag{2.49}$$

Fix a constant control $\theta \in [0,1]$ and let $X(t) := X(t; T, x_m, \theta)$ be the state solution for our problem associated with the control θ and starting from x_m at time T; set

$$\tau^\theta := \inf\{t \geq T \mid X(t) \notin B_{(H,\|\cdot\|_H)}(x_m, \varepsilon)\};$$

this is of course a stopping time. Moreover the trajectories of $X(\cdot)$ are continuous in $(E, \|\cdot\|_E)$, therefore in $(H, \|\cdot\|_H)$, so that $\tau^\theta > T$ almost surely. By (2.48) and (2.49) we get, for $T \leq t \leq \tau^\theta$,

$$e^{-\rho(t-T)} V(X(t)) - V(x_m) \geq e^{-\rho(t-T)} \left(\psi(X(t)) - g(X(t))\right) - \left(\psi(x_m) - g(x_m)\right). \tag{2.50}$$

Let $h > T$ and set $\tau_h^\theta := \tau^\theta \wedge h$; by the dynamic programming principle we get, for all $\theta \in [0,1]$,

$$V(x_m) \geq \mathbb{E}\left[\int_T^{\tau_h^\theta} e^{-\rho(t-T)} U(X_0(t)) dt + e^{-\rho(\tau_h^\theta - T)} V(X(\tau_h^\theta))\right]. \tag{2.51}$$

So, by (2.50) and (2.51) we get, for all $\theta \in [0,1]$,

$$\begin{aligned} 0 \geq \mathbb{E}&\left[\int_T^{\tau_h^\theta} e^{-\rho(t-T)} U(X_0(t)) dt + e^{-\rho(\tau_h^\theta - T)} V(X(\tau_h^\theta)) - V(x_m)\right] \\ &\geq \mathbb{E}\left[\int_T^{\tau_h^\theta} e^{-\rho(t-T)} U(X_0(t)) dt \right. \\ &\left. + e^{-\rho(\tau_h^\theta - T)} \Big(\psi(X(\tau_h^\theta)) - g(X(\tau_h^\theta))\Big) - \left(\psi(x_m) - g(x_m)\right)\right]. \end{aligned} \tag{2.52}$$

Now we can apply the Dynkin formulae to the function

$$\varphi(t, x) = e^{-\rho(t-T)}(\psi(x) - g(x))$$

and put the result in (2.52) getting, for all $\theta \in [0,1]$,

$$0 \geq \mathbb{E}\left[\int_T^{\tau_h^\theta} e^{-\rho(t-T)} \Big(U(X_0(t)) + [\mathcal{L}^\theta \psi](X(t)) - [\mathcal{G}^\theta g](X(t))\Big) dt\right],$$

i.e., for all $\theta \in [0,1]$,

$$0 \geq \mathbb{E}\Bigg[\int_T^{\tau_h^\theta} e^{-\rho(t-T)} \Big(-\rho(\psi(X(t)) - g(X(t))) + \langle x, \tilde{A}^* \psi_x(X(t)) \rangle \\ + \tilde{\mathcal{H}}_{cv}(X(t), \psi_x(X(t)) - g_x(X(t)), \psi_{xx}(X(t)) - g_{xx}(X(t))); \theta \Big) dt \Bigg].$$

Therefore, for all $\theta \in [0,1]$, we can write

$$0 \geq \mathbb{E}\Big[\frac{1}{h-T} \int_T^h I_{[T,\tau^\theta]}(t) e^{-\rho(t-T)} \Big(-\rho(\psi(X(t)) - g(X(t))) + \langle x, \tilde{A}^* \psi_x(X(t)) \rangle \\ + \tilde{\mathcal{H}}_{cv}(X(t), \psi_x(X(t)) - g_x(X(t)), \psi_{xx}(X(t)) - g_{xx}(X(t))); \theta \Big) dt \Big].$$

Now, using the continuity properties of ψ, g and their derivatives and of $\tilde{\mathcal{H}}_{cv}$, taking into account that $\tau^\theta > T$ almost surely and passing to the limit for $h \to T$, we get by dominated convergence, for all $\theta \in [0,1]$,

$$0 \geq -\rho(\psi(x_m) - g(x_m)) + \langle x_m, \tilde{A}^* \psi_x(x_m) \rangle \\ + \tilde{\mathcal{H}}_{cv}\left(x_m, \psi_x(x_m) - g_x(x_m), \psi_{xx}(x_m) - g_{xx}(x_m); \theta \right),$$

i.e., taking into account (2.48) and passing to the supremum on $\theta \in [0,1]$,

$$\rho V(x_m) \geq \langle x_m, \tilde{A}^* \psi_x(x_m) \rangle + \tilde{\mathcal{H}}\left(x_m, \psi_x(x_m) - g_x(x_m), \psi_{xx}(x_m) - g_{xx}(x_m) \right).$$

Notice that, passing to the limit in $\tilde{\mathcal{H}}_{cv}$, we have to use that $X(t) \to x_m$ in $(E, \|\cdot\|_E)$ almost surely, as $t \downarrow T$, since $\tilde{\mathcal{H}}_{cv}$ is not continuous with respect to $\|\cdot\|_H$ on the variable x, due to the presence in $\tilde{\mathcal{H}}_{cv}$ of the term f, which is not continuous with respect to $\|\cdot\|_H$.
Therefore we have proved that V is a supersolution on $\mathcal{V}$.

(ii) Let x_M, ψ, g be as in Definition 2.5.9-(i); without loss of generality we assume

$$V(x_M) = \psi(x_M) + g(x_M). \tag{2.53}$$

Let $B_{(H,\|\cdot\|_H)}(x_M, \varepsilon') \subset \mathcal{O}$ be such that

$$V(x) \leq \psi(x) + g(x), \quad \forall x \in B_{(H,\|\cdot\|_H)}(x_M, \varepsilon'). \tag{2.54}$$

We have to prove that

$$\rho V(x_M) \leq \langle x, \tilde{A}^* \psi_x(x_M) \rangle + \tilde{\mathcal{H}}\left(x_M, \psi_x(x_M) + g_x(x_M), \psi_{xx}(x_M) + g_{xx}(x_M) \right).$$

Let us suppose by contradiction that there exists ν such that

$$0 < \nu \leq \rho V(x_M) - \langle x_M, \tilde{A}^* \psi_x(x_M) \rangle \\ - \tilde{\mathcal{H}}\left(x_M, \psi_x(x_M) + g_x(x_M), \psi_{xx}(x_M) + g_{xx}(x_M) \right).$$

By the continuity propreties of ψ, g and their derivatives and of $\tilde{\mathcal{H}}$, we can find $\varepsilon > 0$ such that, for any $x \in B_{(E,\|\cdot\|_E)}(x_M, \varepsilon)$,

$$0 < \nu/2 \leq \rho V(x) - \langle x, \tilde{A}^* \psi_x(x) \rangle - \tilde{\mathcal{H}} \left(x, \psi_x(x) + g_x(x), \psi_{xx}(x) + g_{xx}(x) \right). \tag{2.55}$$

Notice that to state (2.55) we have to take the ball in the space $(E, \|\cdot\|_E)$, since $\tilde{\mathcal{H}}$ is not continuous with respect to $\|\cdot\|_H$ on the variable x, due to the presence of the term f, which is not continuous with respect to $\|\cdot\|_H$.
Without loss of generality, since $\|\cdot\|_H \leq (1+T)^{1/2} \|\cdot\|_E$, taking a smaller ε if necessary, we can suppose that $B_{(E,\|\cdot\|_E)}(x_M, \varepsilon) \subset B_{(H,\|\cdot\|_H)}(x_M, \varepsilon')$ and therefore, taking into account (2.54), that

$$V(x) \leq \psi(x) + g(x), \quad \text{for any } x \in B_{(E,\|\cdot\|_E)}(x_M, \varepsilon). \tag{2.56}$$

Consider a generic control $\theta(\cdot) \in \Theta_{ad}(x_M)$ and set $X(t) := X(t; T, x_M, \theta(\cdot))$; let us define the stopping time

$$\tau^\theta := \inf \left\{ t \geq T \mid X(t) \notin B_{(E,\|\cdot\|_E)}(x_M, \varepsilon) \right\} \wedge (2T).$$

The trajectories of $X(\cdot)$ are continuous in $(E, \|\cdot\|_E)$, so that we have

$$T < \tau^\theta \leq 2T$$

almost surely. Now we can apply (2.55) to $X(t)$, for $t \in [T, \tau^\theta]$, and get

$$\begin{aligned} 0 < \nu/2 \quad &\leq \quad \rho V(X(t)) - \langle X(t), \tilde{A}^* \psi_x(X(t)) \rangle \\ &\quad - \tilde{\mathcal{H}} \left(X(t), \psi_x(X(t)) + g_x(X(t)), \psi_{xx}(X(t)) + g_{xx}(X(t)) \right); \end{aligned}$$

we multiply by $e^{-\rho(t-T)}$, integrate on $[T, \tau^\theta]$ and take the expectations getting, also taking into account (2.56),

$$\begin{aligned} 0 \quad &< \quad \frac{\nu}{2} \mathbb{E} \left[\int_T^{\tau^\theta} e^{-\rho(t-T)} dt \right] \\ &\leq \quad \mathbb{E} \Bigg[\int_T^{\tau^\theta} e^{-\rho(t-T)} \Big(\rho(\psi(X(t)) + g(X(t))) - \langle X(t), \tilde{A}^* \psi_x(X(t)) \rangle \\ &\qquad - \tilde{\mathcal{H}} \left(X(t), \psi_x(X(t)) + g_x(X(t)), \psi_{xx}(X(t)) + g_{xx}(X(t)) \right) \Big) dt \Bigg]. \end{aligned}$$

We claim that there exists a constant $\delta > 0$, independent on the control $\theta(\cdot)$ chosen, such that

$$\frac{\nu}{2} \mathbb{E} \left[\int_T^{\tau^\theta} e^{-\rho(t-T)} dt \right] \geq \delta;$$

we will prove this fact in Lemma 2.5.11.
So, assuming what claimed above, we can write, taking into account (2.56),

$$\delta + \mathbb{E}\left[\int_T^{\tau^\theta} e^{-\rho(t-T)} U(X_0(t))\, dt\right] \\ \leq \mathbb{E}\left[\int_T^{\tau^\theta} e^{-\rho(t-T)} \left[-\mathcal{L}^{\theta(t)}\psi(X(t)) - \mathcal{G}^{\theta(t)} g(X(t))\right] dt\right].$$

Now we apply the Dynkin formulae to X on $[T, \tau^\theta]$ with the function

$$\varphi(t,x) = e^{-\rho(t-T)}(\psi(x) + g(x))$$

and, comparing with the previous inequality, we get

$$\psi(x_M) + g(x_M) - \mathbb{E}\left[e^{-\rho(\tau^\theta - T)}\left(\psi(X(\tau^\theta)) + g(X(\tau^\theta))\right)\right] \\ \geq \delta + \mathbb{E}\left[\int_T^{\tau^\theta} e^{-\rho(t-T)} U(X(t))\, dt\right];$$

from the previous inequality, taking into account (2.53) and (2.56), we get

$$V(x_M) - \mathbb{E}[e^{-\rho(\tau^\theta - T)} V(X(\tau^\theta))] \quad \geq \quad \delta + \mathbb{E}\left[\int_T^{\tau^\theta} e^{-\rho(t-T)} U(X_0(t))\, dt\right];$$

on the other hand, if we choose a $\delta/2$-optimal control $\theta(\cdot) \in \Theta_{ad}(x_M)$, we get

$$\begin{aligned} V(x_M) - \delta/2 \quad &\leq \quad \mathbb{E}\left[\int_T^{\tau^\theta} e^{-\rho(t-T)} U(X_0(t))\, dt + \int_{\tau^\theta}^{+\infty} e^{-\rho(t-T)} U(X_0(t)) dt\right] \\ &\leq \quad \mathbb{E}\left[\int_T^{\tau^\theta} e^{-\rho(t-T)} U(X_0(t))\, dt + e^{-\rho(\tau^\theta - T)} V(X(\tau^\theta))\right]. \end{aligned}$$

So we have proved by contradiction that V is even a viscosity subsolution. $\square$

Lemma 2.5.11. *Let τ^θ defined as in the part (ii) of the proof of Proposition 2.5.10.Then there exists $\alpha > 0$ such that, for each $\theta(\cdot) \in \Theta_{ad}(\bar{x})$,*

$$\mathbb{E}\left[\int_T^{\tau^\theta} e^{-\rho(t-T)} dt\right] \geq \alpha.$$

Proof. Here, in order to semplify the notation, we write $\bar{x}$ for x_M.
First step. Let $\theta(\cdot) \in \Theta_{ad}(\bar{x})$ and define the stopping time

$$\tau_\theta := \inf\{t \geq T \mid |X_0(t) - \bar{x}_0| \geq \varepsilon/4\} \wedge (2T);$$

in this step we show that there exists $\beta > 0$ such that, for every $\theta(\cdot) \in \Theta_{ad}(\bar{x})$,

$$\mathbb{E}\left[\int_T^{\tau_\theta} e^{-\rho(t-T)}dt\right] \geq \beta.$$

For the controls such that $\mathbb{P}\{\tau_\theta < 2T\} < 1/2$, we have the estimate

$$\mathbb{E}\left[\int_T^{\tau_\theta} e^{-\rho(t-T)}dt\right] \geq \frac{1}{2}\left[\frac{1-e^{-\rho T}}{\rho}\right].$$

Therefore we can suppose without loss of generality that $\mathbb{P}\{\tau_\theta < 2T\} \geq 1/2$. Recalling Proposition (3.2.3) and setting $x(t) := x(t; T, \bar{x}, \theta(\cdot))$ for the solution to (2.6), we have

$$\tau_\theta = \inf\{t \geq T \mid |x(t) - \bar{x}_0| \geq \varepsilon/4\} \wedge (2T).$$

Now we can apply the classical Dynkin's formula to the one-dimensional process $y(\cdot) := x(\cdot) - \bar{x}_0$ with the function $\psi(t,y) = e^{-\rho(t-T)}y^2$ on $[T, \tau_\theta]$ and get

$$\begin{aligned}
\mathbb{E}\left[e^{-\rho(\tau_\theta - T)}\left(x(\tau_\theta) - \bar{x}_0\right)^2\right] &= \mathbb{E}\left[\int_T^{\tau_\theta}\left(-\rho e^{-\rho(t-T)}(x(t)-\bar{x}_0)^2\right)dt\right] \\
&+ \mathbb{E}\left[\int_T^{\tau_\theta} e^{-\rho(t-T)}(x(t)-\bar{x}_0)(r+\sigma\lambda\theta(t))x(t)dt\right] \\
&- \mathbb{E}\left[\int_T^{\tau_\theta} e^{-\rho(t-T)}(x(t)-\bar{x}_0)[f_0\left(x(t)-x(t-T)\right)+q]dt\right] \\
&+ \mathbb{E}\left[\int_T^{\tau_\theta} e^{-\rho(t-T)}\sigma^2\theta(t)^2x(t)^2dt\right];
\end{aligned}$$

now, taking into account that before τ_θ we have

$$|x(t) - \bar{x}_0| < \varepsilon/4, \quad |x(t-T)| \leq \|\bar{x}\|_E + \varepsilon/4$$

and that $|\theta(t)| \leq 1$, we can write, passing to the modulus on the right hand-side and taking into account that f_0 is Lipschitz continuous with Lipschitz constant K_0,

$$\begin{aligned}
\mathbb{E}\left[e^{-\rho(\tau_\theta - T)}\left(x(\tau_\theta) - \bar{x}_0\right)^2\right] &\leq \mathbb{E}\left[\int_T^{\tau_\theta}\rho e^{-\rho(t-T)}\frac{\varepsilon^2}{16}\,dt\right] \\
&+ \mathbb{E}\left[\int_T^{\tau_\theta} e^{-\rho(t-T)}\frac{\varepsilon}{4}(r+\sigma\lambda)\left(|\bar{x}_0|+\frac{\varepsilon}{4}\right)\,dt\right] \\
&+ \mathbb{E}\left[\int_T^{\tau_\theta} e^{-\rho(t-T)}\frac{\varepsilon}{4}\,\left[K_0\left(|\bar{x}_0|+\frac{\varepsilon}{4}+\left(\frac{\varepsilon}{4}+\|\bar{x}\|_E\right)\right)+q\right]\,dt\right] \\
&+ \mathbb{E}\left[\int_T^{\tau_\theta} e^{-\rho(t-T)}\sigma^2\left(|\bar{x}_0|+\frac{\varepsilon}{4}\right)^2dt\right],
\end{aligned}$$

so that, for some $K > 0$,

$$\mathbb{E}\left[e^{-\rho(\tau_\theta - T)}\left(x(\tau_\theta) - \bar{x}_0\right)^2\right] \leq K\,\mathbb{E}\left[\int_T^{\tau_\theta} e^{-\rho(t-T)}dt\right]. \tag{2.57}$$

Recalling that $(x(\tau_\theta)-\bar{x}_0)^2 \le \varepsilon^2/16$ on $[T,\tau_\theta]$ and $(x(\tau_\theta)-\bar{x}_0)^2 = \varepsilon^2/16$ on $\{\tau_\theta < 2T\}$, and considering that

$$e^{-\rho(\tau_\theta - T)} = 1 - \rho \int_T^{\tau_\theta} e^{-\rho(t-T)} dt,$$

we can write by (2.57)

$$\frac{\varepsilon^2}{16} P\{\tau_\theta < 2T\} - \frac{\rho\varepsilon^2}{16} \mathbb{E}\left[\int_T^{\tau_\theta} e^{-\rho(t-T)}dt\right] \le K\mathbb{E}\left[\int_T^{\tau_\theta} e^{-\rho(t-T)}dt\right],$$

so that, by the assumption $\mathbb{P}\{\tau_\theta < +\infty\} \ge 1/2$,

$$\frac{1}{32}\varepsilon^2 \le \left(K + \frac{\rho\varepsilon^2}{16}\right) \mathbb{E}\left[\int_T^{\tau_\theta} e^{-\rho(t-T)}dt\right]$$

and we get what claimed with $\beta = \frac{\varepsilon^2}{32}\left(K + \frac{\rho\varepsilon^2}{16}\right)^{-1}$.

Second step. The function $(\bar{x})_1(\cdot)$ is uniformly continuous and we have $\bar{x}_0 = \lim_{\zeta\to 0^-} \bar{x}_1(\zeta)$; let $\omega(\cdot)$ be its modulus of uniform continuity. Let $x(t) := x(t;T,\bar{x},\theta(\cdot))$ be the solution to (2.6). If we denote by $\omega'(\cdot)$ the modulus of uniform continuity of the trajectory $s \mapsto x(s)$ on $[0,\tau_\theta]$ (thus depending on the trajectory, i.e. on the point of the probability space), we have

$$\omega'(\eta) = \sup_{\substack{|s-s'|<\eta \\ s,s'\in[0,\tau_\theta]}} |x(s)-x(s')|;$$

but, by definition of τ_θ, if $|s-s'|<\eta$,

$$|x(s)-x(s')| \le \begin{cases} \omega(\eta), & \text{if } s,s' \in [0,T], \\ \varepsilon/2, & \text{if } s,s' \in [T,\tau_\theta], \\ \omega(\eta)+\varepsilon/2, & \text{if } 0 \le s < T < s' \le \tau_\theta; \end{cases}$$

therefore $\omega'(\eta) \le \omega(\eta) + \varepsilon/2$ without regard to the trajectory. Thus take $c > 0$ such that $\omega(c) < \varepsilon/4$; we get, for $T \le t \le \tau_\theta \wedge (T+c)$,

$$\begin{aligned} \|X(t)-\bar{x}\|_E &= \left[\sup_{\zeta\in[-T,0)} \left| X_1(t)(\zeta) - \bar{x}_1(\zeta)\right|\right] + \left|X_0(t)-\bar{x}_0\right| \\ &\le \left[\sup_{\zeta\in[-T,0)} \left| x(t+\zeta) - x(T+\zeta)\right|\right] + \varepsilon/4 \le \omega(c) + \frac{\varepsilon}{2} + \frac{\varepsilon}{4} < \varepsilon. \end{aligned}$$

Therefore we have $\tau^\theta \geq \tau_\theta \wedge (T+c)$. Thus we can write

$$\begin{aligned}
\mathbb{E}\left[\int_T^{\tau^\theta} e^{-\rho(t-T)}dt\right] &\geq \mathbb{E}\left[\int_T^{T+c} I_{\{\tau_\theta \geq T+c\}} e^{-\rho(t-T)}dt\right] \\
&\quad +\mathbb{E}\left[\int_T^{\tau_\theta} I_{\{\tau_\theta < T+c\}} e^{-\rho(t-T)}dt\right] \\
&= \frac{1-e^{-\rho c}}{\rho}\,\mathbb{P}\left\{\tau_\theta \geq T+c\right\} \\
&\quad +\mathbb{E}\left[\int_T^{\tau_\theta} I_{\{\tau_\theta < T+c\}} e^{-\rho(t-T)}dt\right].
\end{aligned}$$

We claim that the last term is greater than a strictly positive number independent on $\theta(\cdot)$ (that is enough to make our proof complete); indeed let us suppose by contradiction that there exists a sequence $(\theta_n(\cdot))$ such that

$$\frac{1-e^{-\rho c}}{\rho}\mathbb{P}\left\{\tau_{\theta_n} \geq T+c\right\} + \mathbb{E}\left[\int_T^{\tau_{\theta_n}} I_{\{\tau_{\theta_n} < T+c\}} e^{-\rho(t-T)}dt\right] \longrightarrow 0;$$

then, of course, we would have

$$\mathbb{P}\left\{\tau_{\theta_n} \geq T+c\right\} \longrightarrow 0 \tag{2.58}$$

and also

$$\mathbb{E}\left[\int_T^{\tau_{\theta_n}} I_{\{\tau_{\theta_n} < T+c\}} e^{-\rho(t-T)}dt\right] \longrightarrow 0. \tag{2.59}$$

We consider

$$\mathbb{E}\left[\int_T^{\tau_{\theta_n}} e^{-\rho(t-T)}dt\right] - \mathbb{E}\left[\int_T^{\tau_{\theta_n}} I_{\{\tau_{\theta_n} < T+c\}} e^{-\rho(t-T)}dt\right] \tag{2.60}$$

and rewrite it as

$$\mathbb{E}\left[\int_T^{2T} I_{[T,\tau_{\theta_n}]}(t)\left(1 - I_{\{\tau_{\theta_n} < T+c\}}\right) e^{-\rho(t-T)}dt\right];$$

by (2.58) the integrand converges to 0 in measure $\mathbb{P} \times dt$ on $\Omega \times [T, 2T]$; so, by dominated convergence,

$$\mathbb{E}\left[\int_T^{2T} I_{[T,\tau_{\theta_n}]}(t)\left(1 - I_{\{\tau_{\theta_n} < T+c\}}\right) e^{-\rho(t-T)}dt\right] \longrightarrow 0,$$

i.e. also the expression in (2.60) goes to 0. Taking into account (2.59), we should conclude that also

$$\mathbb{E}\left[\int_T^{\tau_{\theta_n}} e^{-\rho(t-T)}dt\right] \longrightarrow 0,$$

but this convergence contradicts the first step. □

Now we give a definition of constrained viscosity solution on

$$\mathrm{Clos}_{(H,\|\cdot\|_H)}(\mathcal{O}) \cap E = \mathrm{Clos}_{(E,\|\cdot\|_H)}(\mathcal{V}).$$

Recall that we have proved in Lemma 2.4.5-(3) that

$$\mathrm{Fr}_{(E,\|\cdot\|_H)}(\mathcal{V}) = \mathrm{Fr}_{(H,\|\cdot\|_H)}(\mathcal{O}) \cap E.$$

Definition 2.5.12. A continuous function $v : \mathcal{O} \cup \mathrm{Fr}_{(E,\|\cdot\|_H)}(\mathcal{V}) \to \mathbb{R}$, is said a constrained viscosity solution to equation (2.32) on $\mathrm{Clos}_{(E,\|\cdot\|_H)}(\mathcal{V})$ if v is a viscosity solution to (2.32) on $\mathcal{V}$ and if

(i) (supersolution property at the boundary) for any triple

$$(x_m, \psi, g) \in \mathrm{Fr}_{(E,\|\cdot\|_H)}(\mathcal{V}) \times \mathcal{T}_1 \times \mathcal{T}_2$$

such that x_m is a local minimum point for $v - \psi + g$ on $\mathcal{O} \cup \mathrm{Fr}_{(E,\|\cdot\|_H)}(\mathcal{V})$, we have

$$\rho v(x_m) \geq \langle x_m, \tilde{A}^* \psi_x(x_m)\rangle + \tilde{\mathcal{H}}_{cv}(x_m, \psi(x_m) - g(x_m), \psi_{xx}(x_m) - g_{xx}(x_m); 0);$$

(ii) (subsolution property at the boundary) for any triple

$$(x_M, \psi, g) \in \mathrm{Fr}_{(E,\|\cdot\|_H)}(\mathcal{V}) \times \mathcal{T}_1 \times \mathcal{T}_2$$

such that x_M is a local maximum point for $v - \psi - g$ on $\mathcal{O} \cup \mathrm{Fr}_{(E,\|\cdot\|_H)}(\mathcal{V})$, we have

$$\rho v(x_M) \leq \langle x_M, \tilde{A}^* \psi_x(x_M)\rangle + \tilde{\mathcal{H}}(x_M, \psi(x_M) - g(x_M), \psi_{xx}(x_M) - g_{xx}(x_M)).$$

□

We can give the main result.

Theorem 2.5.13. *Let $U(l) > -\infty$ and $rl = q$. Then the value function V is a constrained viscosity solution to (2.32) on Clos$_{(E,\|\cdot\|_H)}(\mathcal{V})$.*

Proof. We have proved in Proposition 2.5.10 that V is a viscosity solution on $\mathcal{V}$. The proof of the viscosity properties at the points of the boundary $\mathrm{Fr}_{(E,\|\cdot\|_H)}(\mathcal{V})$ follows the same line of the proof of Proposition 2.5.10. Notice that in this case the proof is even easier: in this case the stopping times are constant, since we are constrained to choose $\theta = 0$ in the case of the supersolution and $\theta(\cdot) \equiv 0$ (see Proposition 2.4.6-(5)) in the case of the subsolution. □

Remark 2.5.14. Usually in stochastic control problems with state contraints the definition of solution for the HJB equation which works good to get the uniqueness is the definition of constrained viscosity solution. It was introduced by Soner, see [Soner; 1986], in the deterministic case and successfully developed and applied in the stochastic case in other papers, for instance

[Katsoulakis; 1994] and [Zariphopoulou; 1994]. It consists in requiring the subsolution property at the boundary.

In [Ishii, Loreti; 2002] it is also required the supersolution property at the boundary replacing $\mathcal{H}$ with

$$\mathcal{H}_{in}(\cdot) = \sup_{\theta \in \mathcal{A}} \mathcal{H}_{cv}(\cdot; \theta),$$

where $\mathcal{A}$ is the subset of the control space for which the diffusion term of the state equation vanishes and the correspondent drift term directs inside to the state space (under the assumption that $\mathcal{A}$ is not empty). This last condition is similar to our supersolution condition at the boundary, but in our case the drift term is such that the state remains on the boundary of the state space.

Finally we notice that in this case of absorbing boundary we could use at the boundary also a Dirichlet type condition. Indeed, since the boundary is absorbing and the only admissible strategy is the null one, the value function is theoretically computable at the boundary points. In particular the value function is really explicitely computable at the boundary points

$$x = (x_0, x_1(\cdot)) \in \mathrm{Fr}_{(E,\|\cdot\|_H)}(\mathcal{V})$$

such that $x_0 = l$: indeed for these points x we clearly have the expression $V(x) = U(l)/\rho$. $\square$

Remark 2.5.15. We want to point out that, if we tried to prove a subsolution viscosity property at the boundary in the general case (when the boundary is not necessarily absorbing) for the upper semicontinuous envelope of the value function, we would be in trouble. Indeed, denoting by V^* this upper semicontinuous envelope, we cannot apply the Dynamic Programming Principle to V^* starting from $x_M \in \mathrm{Fr}_{(E,\|\cdot\|_H)}(\mathcal{V})$. Then a possible technique to proceed consists in working on a sequence $(x_n) \subset \mathcal{V}$ approximating x_M, applying the Dynamic Programming Principle starting from these points x_n and passing to the limit. The trouble for this approach consists in the fact that the term f in the equation is not continuous with respect to $\|\cdot\|_H$, whereas the upper semicontinuous envelope is defined with respect to this norm. The same problem would come up if we tried to prove a supersolution viscosity property for the lower semicontinuous envelope. Therefore we need to have semicontinuity properties for the value function at the boundary, in order to be able to apply the Dynamic Programming Principle directly on these points to get viscosity properties there. This consideration motivates the choice to work with an absorbing boundary, in order to get the continuity for the value function at the boundary (Proposition 2.4.7). Finally our opinion is that the lower semicontinuity of the value function at the boundary holds also in the case when the boundary is not absorbing, but of course the proof of this fact would require a more subtle argument. □

2.6 Comments, future developements and an example

In the context of Chapter 1 the surplus term did not appear, so that the problem was one-dimensional without delay. In the present context the delay term makes the problem considerably more difficult and an infinite-dimensional approach seems to be necessary in order to make the problem markovian and apply the dynamic programming techniques. As we have seen, while the topics treated in this chapter are quite standard in a context without delay, adding the delay make them immediately nontrivial. On the other hand we need them in order to proceed with the study of the problem.

The main results of this chapter are

- the rewriting of the problem in an infinite dimensional setting (Section 2.3);
- the study of the infinite dimensional state equation (Theorem 2.3.10);
- the proof of the equivalence between the delay one dimensional problem and the abstract infinite dimensional one (Proposition 2.3.12);
- the proof of the continuity of the value function in the infinite dimensional setting (Corollary 2.4.3 and Proposition 2.4.7).
- the proof that the value function is a constrained viscosity solution (in the sense given in Definition 2.5.12) of the associated infinite-dimensional HJB equation (Theorem 2.5.13).

The investigation leaves open several topics, such as whether the given definition of constrained viscosity solution is strong enough to guarantee the uniqueness or not, the regularity properties of the value function and a verification theorem giving optimal feedback strategies for the problem (see also the next chapter on this topic). The nearest result seems to be the one concerned uniqueness. The other topics seem to be not addressable with the standard techniques until now available, so that new techniques seem to be necessary.

We want to comment also on the possible numerical approach to the problem. The numerical approach to stochastic control problems can be probabilistic or analytic. The probabilistic approach consists in the study of the convergence of a suitable sequence of discretized control problems. This is the so called method of Markov's chains, described in [Kushner; 1977], which basically consists in discretizing the time set, the state space and the control space of the original control problems. In [Kushner; 1977] also the case of delay systems is treated. More recent references on this approach for stochastic delay control problems are the papers [Fischer, Nappo; 2008], [Fischer, Nappo; 2007], [Kushner; 2006]. However another possible approach is the analytic one, which considers directly the numerical study of the differential problem, i.e. of the HJB equation. This is done, e.g., in [Barles, Souganidis; 1991] in the context of viscosity solutions. Therein it is proved that a suitable approximation scheme for the equation leads to a good approximation for the value function, providing that a comparison result holds true for the viscosity solutions of the HJB equation. For approximation schemes for HJB equations arising in stochastic control problems with delay see also [Chang, Peng, Pemy; 2008].

We want to spend some words more about the applications of delay equations in Economics and Finance. Problems with delay arise naturally in many situations, although sometimes, also due to their intrinsic difficulty, they are not investigated with fair depth. We refer to [Kolmanovskii, Shaikhet; 1996] for an overview on the works on this mathematical subject and their applications. We mention the following papers.

- Economics: the so called time-to-build problems; see, e.g., the papers
  - [Asea, Zak; 1999],
  - [Bambi; 2008],
  - [Federico, Goldys, Gozzi; 2009a],
  - [Kydland, Prescott; 1982].
- Advertising models: see [Gozzi, Marinelli; 2004], [Gozzi, Marinelli, Savin].
- Finance: see, e.g., [Øksendal, Sulem; 2001].
- Pension funds: see [Gabay, Grasselli; 2009]. Therein the delay appears in an exogenous process, so that the problem can be treated by a duality approach. Here it is worth to stress that *the delay arises quite naturally in the context of pension funds*, due to the lag between the time of entrance of the worker into the fund and the time of exit from it.

Moreover, we should say that the infinite dimensional framework has been succesfully used in many applied contexts. In particular it has been applied in Finance to model and study problems arising in the context of bond markets. Indeed, starting from the Heath-Jarrow-Morton model and the Musiela parametrization (see [Heath, Jarrow, Morton; 1992], [Musiela; 1993]), many authors have set problems related to forward rates in infinite dimensional spaces (see, e.g.,

- [Ekeland, Taflin; 2005],
- [Goldys, Musiela, Sondermman; 2000],
- [Filipovic; 2001],
- [Kelome, Swiech; 2003],
- [Vargiolu; 1999]).

Before to conclude the chapter we want to give an example in order to show how the infinite dimensional approach to stochastic optimal control problems with delay can work in order to solve explicitely the problem in some "good" cases.

2.6.1 An example of a solvable stochastic optimal control problem with delay by means of the infinite-dimensional approach

Here we give an example (out of the assumptions of the chapter). We think that, on the line of this example, it would be worth to try to estabilish conditions ensuring the solvability of stochastic optimal control problems with delay by means of their infinite-dimensional representation, in the spirit of what is done in [Larrsen, Risebro; 2003] by means of a finite-dimensional approach.

In the same setting of the chapter, suppose to be involved with the management of an investment fund having the following rules.

- At time $t = 0$ the community of members endows the fund with an initial amount $\eta_0 > 0$.
- At each time $t \geq 0$ the fund must satisfy the capital requirement

$$x(t) + \int_{t-1}^{t} x(t+\xi)d\xi \geq l, \quad \mathbb{P} - \text{a.s.}, \tag{2.61}$$

  where $0 \leq l \leq \eta_0$ and where it is set $x(\xi) = \eta_1(\xi) \geq l$ for $\xi \in [-1, 0)$ (it might be, e.g., $\eta_1(\cdot) \equiv l$ or $\eta_1(\cdot) \equiv \eta_0$).
- At each time $t \geq 0$ the fund pays to the members the instantaneous return of its wealth with respect to the market spot rate r, i.e. the quantity $rx(t)$.
- At each time $t \geq 0$ there is a flow of money between the members and the fund that is related to the trend of the fund in the last year. Precisely the fund pays the (positive or negative) quantity $x(t) - x(t-1)$ to its members.

- The manager can invest in the risky asset and in the riskless one, but short selling of the risky asset are not allowed.
- The manager fee at time $t \geq 0$ is by contract proportional to

$$x(t) + \int_{t-1}^{t} x(t+\xi)d\xi - l.$$

Since in this example we do not require a state constraint on the current value of the fund, we prefer to define the investment strategy in terms of the amount $\pi(\cdot)$ invested in the risky asset, because this gives a better representation of the investment constrain. Indeed it can be simply expressed as $\pi(\cdot) \geq 0$, in place of $\theta(\cdot)x(\cdot) \geq 0$ that would involve also the state. We assume that the strategies $\pi(\cdot)$ must be progressively mesaurable and moreover that must satisfy the integrability condition $\pi(\cdot) \in L^2_{loc}([0,+\infty); L^2(\Omega))$. Using this representation for the investment and the assumptions above, the dynamics of the fund starting at time 0 from the generic initial "present-past" state $x = (x_0, x_1(\cdot))$ is

$$\begin{cases} dx(t) = \sigma\lambda\pi(t)dt - (x(t) - x(t-1))dt + \sigma\pi(t)dB(t), \\ x(0) = x_0,\ x(\xi) = x_1(\xi),\ \xi \in [-1,0), \end{cases} \tag{2.62}$$

which is again a stochastic controlled dynamics with delay in the state. We denote the solution to (2.62) by $x(\cdot;(x_0,x_1(\cdot)),\pi(\cdot))$.

Let $H = \mathbb{R} \times L^2([-1,0];\mathbb{R})$, $x := (x_0, x_1(\cdot)) \in H$ and denote by $\Pi_{ad}(x)$ the set of admissible strategies starting form x, i.e.

$$\Pi_{ad}(x) := \{\pi(\cdot) \in L^2_{loc}([0,+\infty); L^2(\Omega;[0,+\infty))) \mid$$
$$\pi(\cdot) \text{ is progressively mesaurable and } x(t;(x_0,x_1(\cdot)),\pi(\cdot)) \text{ satisfy (2.61) } \forall t \geq 0\}.$$

One could prove that $\Pi_{ad}(x) \supset \{0\}$ for every $x \in \mathcal{D}$ where

$$\mathcal{D} = \{x \in H \mid x_0 \geq l, x_1(\cdot) \geq l\},$$

so in particular for the initial datum $\eta = (\eta_0, \eta_1(\cdot))$ which indeed belongs to $\mathcal{D}$.

We can rewrite the problem in the space H as in Section 2.3. In this case many topics are much easier thanks to the specific structure of the problem. In particular here the delay is linear, so that we can insert it in the linear operator, and work on the whole space H. The (unbounded) linear operator, denoted by $\tilde{A}$, in this case is defined on

$$D(\tilde{A}) = \{(x_0, x_1(\cdot)) \in H \mid x_1(\cdot) \in W^{1,2}([-1,0];\mathbb{R}),\ x_0 = x_1(0)\}$$

by

$$\tilde{A}(x_0, x_1(\cdot)) = (-x_0 + x_1(-T), x_1'(\cdot)).$$

The operator $\tilde{A}$ is the generator of a strongly continuous semigroup $(\tilde{S}(t))_{t\geq 0}$ on H (see, e.g., [Hale, Verduyn; 1993]). It turns out that there exists a unique mild solution in the sense of Definition 2.3.1 to the H-valued SDE

$$\begin{cases} dX(t) = \tilde{A}X(t)dt + \sigma\lambda\pi(t)\cdot\hat{n}\ dt + \sigma\pi(t)\cdot\hat{n}\ dB(t), \\ X(0) = x \in H, \end{cases} \tag{2.63}$$

where $\hat{n} = (1,0) \in H$. Notice that in this case, since we do not need to work with a subspace of H as state space, the initial datum is allowed to be in H. We denote the mild solution to (2.63) by $X(\cdot; x, \pi(\cdot))$. Moreover, arguing as in the paper, we could see that such mild solution $X(\cdot) := X(\cdot; x, \pi(\cdot))$ has the property that $(X_0(t), X_1(t)) = \big(x(t), x(t+\xi)|_{\xi\in[-1,0]}\big)$ for every $t \geq 0$, where $x(\cdot)$ is the solution to (2.62), i.e. the equivalence result holds.

Consider $h = (1,1) \in H$. Then the capital requirement can be expressed as $\langle X(t), h\rangle_H \geq l$ for every $t \geq 0$. Then, setting

$$\mathcal{G} = \{x \in H \mid \langle x, h\rangle_H \geq l\},$$

we can see that $\mathcal{D} \subset \mathcal{G}$ and that $\Pi_{ad}(x) \supset \{0\}$. Moreover the boundary

$$\partial\mathcal{G} = \{x \in H \mid \langle x, h\rangle_H = l\}$$

is absorbing in the sense that

$$x \in \partial\mathcal{G} \implies \Pi_{ad}(x) = \{0\};\ \ \langle X(t; x, 0), h\rangle_H = l,\ \ \forall t \geq 0. \tag{2.64}$$

To show this fact we introduce the concept of *weak solution* to (2.63). A process $X(\cdot)$ is called a weak solution of (2.63) if, for every $t \geq 0$ and $a \in D(A^*)$,

$$\begin{aligned} \langle X(t), a\rangle_H = \langle x, a\rangle_H + \int_0^t \langle X(s), \tilde{A}^* a\rangle_H ds \\ + \int_0^t \sigma\lambda\pi(s)\langle \hat{n}, a\rangle_H ds + \int_0^t \sigma\pi(s)\langle \hat{n}, a\rangle_H dB(s), \end{aligned}$$

where $\tilde{A}^*$ is the adjoint of $\tilde{A}$ defined on

$$D(\tilde{A}^*) = \{(y_0, y_1(\cdot)) \in H \mid y_1(\cdot) \in W^{1,2}([-1,0];\mathbb{R}),\ y_1(-1) = y_0\}$$

by

$$\tilde{A}^*(y_0, y_1(\cdot)) = (-y_0 + y_1(0), -y_1'(\cdot)).$$

Theorem 6.5, Chapter 6, in [Da Prato, Zabczyk; 1992] states that X is a mild solution to (2.63) if and only if it is a weak solution. Therefore we have that

$X(\cdot) := X(\cdot; x, \pi(\cdot))$ is a weak solution to (2.63), so that in particular

$$\langle X(t), h\rangle_H = \langle x, h\rangle_H + \int_0^t \langle X(s), \tilde{A}^* h\rangle_H ds + \int_0^t \sigma\lambda\pi(s)\langle \hat{n}, h\rangle_H ds + \int_0^t \sigma\pi(s)\langle \hat{n}, h\rangle_H dB(s),$$

Notice that $h \in D(\tilde{A}^*)$ and $\tilde{A}^* h = 0$. Therefore, if $x \in \partial\mathcal{G}$, we have

$$\langle X(t; x, 0), h\rangle_H = \langle x, h\rangle_H = l, \ \ \forall t \geq 0.$$

On the other hand, starting from $x \in \partial\mathcal{G}$, due to the stochastic term in the equation, any other strategy $\pi(\cdot)$ different from the null one would bring the state to violate the requirement $\langle X(t), h\rangle_H \geq l$. Therefore (2.64) is proved.

Now suppose that the utility function of the manager is $U(z) = \frac{z^\gamma}{\gamma}$, $z \geq 0$, $\gamma \in (0,1)$, and that his individual discount factor is $\rho > 0$. Then the natural aim of the manager would be to maximize the expected intertemporal discounted utility coming from his fee, i.e. the functional

$$\mathbb{E}\left[\int_0^{+\infty} e^{-\rho t} \ \frac{(\langle X(t), h\rangle_H - l)^\gamma}{\gamma} dt\right].$$

We can write formally the infinite dimensional HJB equation for this problem, that is

$$\rho v(x) = \langle x, A^* v_x(x)\rangle_H + \frac{(\langle x, h\rangle_H - l)^\gamma}{\gamma} - \frac{\lambda^2}{2}\frac{v_{x_0}(x)^2}{v_{x_0 x_0}(x)}, \tag{2.65}$$

Due to the specific structure of the equation, we may guess that a solution to the HJB equation has the form

$$v(x) = C\frac{(\langle x, h\rangle_H - l)^\gamma}{\gamma}. \tag{2.66}$$

Indeed we can check that, if

$$C = \left(\rho - \frac{\lambda^2}{2(1-\gamma)}\right)^{-1}, \tag{2.67}$$

the function v given by (2.66) is a solution to (2.65).

If we add the assumption

$$\rho > \frac{\lambda^2}{2(1-\gamma)}, \tag{2.68}$$

we see that v is positive and concave, therefore it is a very good candidate to be the value function V of the problem. To prove that it is really the value function and to give an optimal feedback strategy for the problem, we need to argue with verification techniques in infinite dimension. In order to do that we need a Dynkin type formula for v applied to the mild solution $X(\cdot) := X(\cdot; x, \pi(\cdot))$, $x \in \text{Int}(\mathcal{G})$, of (2.63). Indeed, it is possible to prove that, for every stopping time $\tau \geq 0$ almost surely finite, and such that $\langle X(t), h\rangle_H \geq l + \varepsilon$ for some $\varepsilon > 0$ for every $t \in [0, \tau]$, we have

$$\mathbb{E}\left[e^{-\rho\tau} v(X(\tau)) - v(x)\right] = \mathbb{E}\left[\int_0^\tau e^{-\rho t}[\mathcal{L}^{\pi(t)} v](X(t))dt\right], \tag{2.69}$$

where

$$[\mathcal{L}^\pi v](x) := -\rho v(x) + \sigma\lambda\pi v_{x_0}(x) + \frac{1}{2}\sigma^2\pi^2 v_{x_0x_0}(x).$$

Now take $x \in \text{Int}(\mathcal{G})$ (the claim is trivial for $x \in \partial\mathcal{G}$, due to (2.64)) and take any strategy $\pi(\cdot) \in \Pi_{ad}(x)$. Consider the mild solution $X(\cdot) := X(\cdot; x, \pi(\cdot))$ to (2.63) associated to this strategy. Let $t_0 > 0$, $\varepsilon > 0$ and let τ_{ε,t_0} be the stopping time defined by

$$\tau_{\varepsilon,t_0} = \inf\{t \geq 0 \mid \langle X(t), h\rangle_H \leq l + \varepsilon\} \wedge t_0.$$

Applying (2.69) we get

$$\mathbb{E}\left[e^{-\rho\tau_{\varepsilon,t_0}} v(X(\tau_{\varepsilon,t_0})) - v(x)\right] = \mathbb{E}\left[\int_0^{\tau_{\varepsilon,t_0}} e^{-\rho t}[\mathcal{L}^{\pi(t)} v](X(t))dt\right].$$

Since v is solution of the HJB equation we get, ordering the terms,

$$v(x) = \mathbb{E}\left[e^{-\rho\tau_{\varepsilon,t_0}} v(X(\tau_{\varepsilon,t_0}))\right] + \mathbb{E}\left[\int_0^{\tau_{\varepsilon,t_0}} e^{-\rho t}\ \frac{(\langle X(t), h\rangle_H - l)^\gamma}{\gamma} dt\right]$$
$$+\mathbb{E}\left[\int_0^{\tau_{\varepsilon,t_0}} e^{-\rho t}\left(-\frac{\lambda^2}{2}\frac{v_{x_0}(X(t))^2}{v_{x_0x_0}(X(t))} - \sigma\lambda\pi(t)v_{x_0}(X(t)) - \frac{1}{2}\sigma^2\pi(t)^2 v_{x_0x_0}(X(t))\right)dt\right]. \tag{2.70}$$

Since v is positive, we see that the first term in the right handside is positive; moreover, due to the concavity and to the monotonicity of v with respect to the variable x_0, we see that also the third term in the right handside of (2.70) is positive. Therefore

$$v(x) \geq \mathbb{E}\left[\int_0^{\tau_{\varepsilon,t_0}} e^{-\rho t}\ \frac{(\langle X(t), h\rangle_H - l)^\gamma}{\gamma} dt\right] \tag{2.71}$$

Taking the limsup for $t_0 \to +\infty$, $\varepsilon \downarrow 0$, we get by Fatou's Lemma

$$v(x) \geq \mathbb{E}\left[\int_0^{+\infty} e^{-\rho t}\ \frac{(\langle X(t), h\rangle_H - l)^\gamma}{\gamma} dt\right],$$

By the arbitrariness of $\pi(\cdot)$ we have shown that $v(x) \geq V(x)$.

Now take the feedback map

$$G(x) = \frac{\lambda}{\sigma(1-\gamma)}\left(\langle x, h\rangle_H - l\right).$$

It is straightforward to see that the closed loop equation associated to this feedback map, i.e.

$$\begin{cases} dX(t) = \tilde{A}X(t)dt + \frac{\lambda^2}{1-\gamma}\left(\langle X(t), h\rangle_H - l\right)\cdot \hat{n}\ dt + \frac{\lambda}{1-\gamma}\left(\langle X(t), h\rangle_H - l\right)\cdot \hat{n}\ dB(t), \\ X(0) = x \in H, \end{cases} \tag{2.72}$$

admits a unique mild solution X_G. Define the feedback strategy

$$\pi_G(t) := G(X_G(t)). \tag{2.73}$$

We can see that $\pi_G(\cdot) \in \Pi_{ad}(x)$. Moreover, taking (2.70) with $\pi_G(\cdot)$ we see that the third term in the right handside vanishes, so that (2.70) reduces to

$$v(x) = \mathbb{E}\left[e^{-\rho\tau_{\varepsilon,t_0}} v(X_G(\tau_{\varepsilon,t_0}))\right] + \mathbb{E}\left[\int_0^{\tau_{\varepsilon,t_0}} e^{-\rho t}\ \frac{(\langle X_G(t), h\rangle_H - l)^\gamma}{\gamma} dt\right].$$

By dominated convergence we get

$$v(x) = \mathbb{E}\left[\int_0^{+\infty} e^{-\rho t}\ \frac{(\langle X_G(t), h\rangle_H - l)^\gamma}{\gamma} dt\right].$$

With the argument above we have proved the following.

Theorem 2.6.1. *Let assumption* (2.68) *hold. The function v defined in* (2.66)*, with C given by* (2.67)*, is the value function, i.e. $v = V$ and the feedback strategy $\pi_G(\cdot)$ defined in* (2.73) *is optimal for the problem.* □

Finally we observe that (2.72) considered in weak form shows that the optimal trajectory X_G is such that the one dimensional process

$$Z(\cdot) := \langle X_G(\cdot), h\rangle_H - l$$

is a geometric Brownian motion. Therefore, if $x \in \mathrm{Int}(\mathcal{G})$ the optimal state never reaches the boundary $\partial\mathcal{G}$; if $x \in \partial\mathcal{G}$ the optimal state, as we have shown also before, remains on the boundary $\partial\mathcal{G}$. According to Feller's boundary classification, we can say that $\partial\mathcal{G}$ is a *natural boundary* for the infinite dimensional diffusion X_G.

Chapter 3

Optimal Control of DDEs with State Constraints

The object of this chapter is the study of a deterministic control problem with diffused delay in the state. The main references for this chapter are the papers [Federico, Goldys, Gozzi; 2009a] and [Federico, Goldys, Gozzi; 2009b]. The motivation for introducing such a problem within the present thesis is purely mathematical. Indeed this problem represents a first starting point to approach problems like the one described in Chapter 2. The problem which is the object of Chapter 2 is very specific and it was not possible to solve it completely. The problem stated and studied in the present chapter has some features very similar to the ones of the problem of Chapter 2 and allows a (theoretically) complete solution.

Here we study a class of optimal control problems with state constraints where the state equation is a differential equation with delays. This class includes some problems arising in economics, in particular the so-called models with time to build, see [Asea, Zak; 1999], [Bambi; 2008], [Kydland, Prescott; 1982]. As in Chapter 2 we embed the problem in a suitable Hilbert space H and consider the associated Hamilton-Jacobi-Bellman equation. This kind of infinite-dimensional HJB equation has not been previously studied and is difficult due to the presence of state constraints and the lack of smoothing properties of the state equation.

The main result is the proof of a C^1 regularity result for a class of first order infinite dimensional HJB equations associated to the optimal control of deterministic delay equations arising in economic models.

The C^1 regularity of solutions of the HJB equations arising in deterministic optimal control theory is a crucial issue to solve in a satisfactory way the control problems. Indeed, even in finite dimension, in order to obtain the optimal strategies in feedback form one usually needs the existence of an appropriately defined gradient of the solution. It is possible to prove verification theorems and representation of optimal feedbacks in the framework of viscosity solutions, even if the gradient is not defined in classical sense (see e.g. [Bardi, Capuzzo-Dolcetta; 1997, Yong, Zhou; 1999]), but this is usually not satisfactory in applied problems since the closed loop equation becomes very hard to treat in such cases.

The need of C^1 regularity results for HJB equations is particularly important in infinite dimension since in this case verification theorems in the framework of viscosity solutions are rather weak and in any case not applicable to problems with state constraints (see e.g [Fabbri, Gozzi, Swiech, Li, Yong; 1995]). To the best of our knowledge C^1 regularity results for first order HJB equation have been proved by method of convex regularization introduced by Barbu and Da Prato [Barbu, Da Prato; 1983b] and then developed by various authors (see e.g.

- [Barbu, Da Prato; 1985a],
- [Barbu, Da Prato; 1985b],
- [Barbu, Da Prato, Popa; 1983],
- [Barbu, Precupanu; 1986],
- [Di Blasio; 1985],
- [Di Blasio; 1991],
- [Faggian; 2005],
- [Gozzi; 1989],
- [Gozzi; 1991]).

All these results do not hold in the case of state constraints and, even without state constraints, do not cover problems where the state equation is a nonlinear differential equation with delays. In the papers [Cannarsa, Di Blasio; 1995, Cannarsa, Di Blasio; 1993, Faggian; 2008] a class of state constraints problems is treated using the method of convex regularization but the C^1 type regularity is not proved.

Using the approach of Chapter 2, we embed the problem in an infinite dimensional control problem in the Hilbert space $H = \mathbb{R} \times L^2([-T,0];\mathbb{R})$, where intuitively speaking $\mathbb{R}$ describes the "present" and $L^2([-T,0];\mathbb{R})$ describes the "past" of the system. In this paper we consider the associated Hamilton-Jacobi-Bellman (HJB) equation in H that has not been previously studied. Such a HJB is difficult due to the presence of state constraints and the lack of smoothing properties of the state equation.

We prove that the value function is continuous in a sufficiently big open set of H (Proposition 3.2.10), that it solves in the viscosity sense the associated HJB equation (Theorem 3.3.3) and it has continuous classical derivative in the direction of the "present" (Theorem 3.3.5). This regularity result is enough to define the formal optimal feedback strategy in classical sense, since the objective functional only depends on the "present". When such a strategy effectively exists and is admissible we prove (Theorem 3.3.11) that it must be optimal for the problem: this is not trivial since we do not have the full gradient of the value function and so we need to use a verification theorem for viscosity solution which is new in this context. We then give (Proposition 3.3.15) sufficient conditions under which the formal optimal feedback exists and is admissible. Finally we show some approximation results that allow to apply our main theorem to obtain ε-optimal strategies for a wider class of problems (Propositions 3.4.2, 3.4.5, 3.4.7).

The method we use to prove regularity is completely different from the one of convex regularization mentioned above. Indeed, it is based on a finite dimensional result of Cannarsa and Soner [Cannarsa, Soner; 1989] (see also [Bardi, Capuzzo-Dolcetta; 1997], pag. 80) that exploits the concavity of the data and the strict convexity of the Hamiltonian to prove the continuous differentiability of the viscosity solution of the HJB equation. Generalizing to the infinite dimensional case such result is not trivial as the definition of viscosity solution in this case strongly depends (via the unbounded differential operator A contained in the state equation) on the structure of the problem. In particular we need to establish specific properties of superdifferential that are given in Subsection 3.2.3.

We believe that such a method could be also used to analyze other problems featuring concavity of the data and strict convexity of the Hamiltonian.

3.1 Setup of the control problem and preliminary results

In this section we will formally define the control delay problem giving some possible motivations for it. As in Chapter 2 we will use the notations

$$L^2_{-T} := L^2([-T,0];\mathbb{R}), \quad W^{1,2}_{-T} := W^{1,2}([-T,0];\mathbb{R}).$$

We will denote by H the Hilbert space

$$H := \mathbb{R} \times L^2_{-T},$$

endowed with the inner product

$$\langle \cdot,\cdot \rangle_H = \langle \cdot,\cdot \rangle_{\mathbb{R}} + \langle \cdot,\cdot \rangle_{L^2_{-T}},$$

and the norm

$$\|\cdot\|_H^2 = |\cdot|_{\mathbb{R}}^2 + \|\cdot\|_{L^2_{-T}}^2.$$

We will denote by $\eta := (\eta_0, \eta_1(\cdot))$ the generic element of this space. For convenience we set also

$$H_+ := (0, +\infty) \times L^2_{-T}, \qquad H_{++} := (0, +\infty) \times \{\eta_1(\cdot) \in L^2_{-T} \mid \eta_1(\cdot) \geq 0 \; a.e.\}.$$

Remark 3.1.1. We notice that the economic motivations we are mainly interested in (see [Asea, Zak; 1999, Bambi; 2008], [Kydland, Prescott; 1982] and Remark 3.1.7 above) require to study the optimal control problem with the initial condition in H_{++}. However, the set H_{++} is not convenient to work with, since its interior with respect to the $\|\cdot\|_H$-norm is empty. That is why we enlarge the problem and allow the initial state belonging to the class H_+. □

For $\eta \in H_+$, we consider an optimal control of the following differential delay equation:

$$\begin{cases} x'(t) = rx(t) + f_0\left(x(t), \int_{-T}^0 a(\xi)x(t+\xi)d\xi\right) - c(t), \\ x(0) = \eta_0, \; x(s) = \eta_1(s), \; s \in [-T, 0), \end{cases} \tag{3.1}$$

with state constraint $x(\cdot) > 0$ and control constraint $c(\cdot) \geq 0$. We set up the following assumptions on the functions a, f_0.

Hypothesis 3.1.2.

- $a(\cdot) \in W^{1,2}_{-T}$ is such that $a(\cdot) \geq 0$ and $a(-T) = 0$;
- $f_0 : [0, \infty) \times \mathbb{R} \to \mathbb{R}$ is jointly concave, nondecreasing with respect to the second variable, Lipschitz continuous with Lipschitz constant C_{f_0}, and

$$f_0(0, y) > 0, \quad \forall y > 0. \tag{3.2}$$

□

Remark 3.1.3. In the papers

- [Asea, Zak; 1999],
- [Bambi; 2008],
- [Kydland, Prescott; 1982]

the pointwise delay is used. We cannot treat exactly this case for technical reason that are explained in Remark 3.4.6 below. However we have the freedom of choosing the function a in a wide class and this allows to take account of various economic phenomena. Moreover we can approximate the pointwise delay with suitable sequence of functions $\{a_n\}$ getting convergence of the value function and constructing ε-optimal strategies (see Subsections 3.4.2 and 3.4.3). □

From now on we will assume that f_0 is extended to a Lipschitz continuous map on $\mathbb{R}^2$ setting

$$f_0(x,y) := f_0(0,y), \quad \text{for } x < 0.$$

For technical reasons, which will be clear in Subsection 3.2.2, we work with the case $r > 0$, noting that nevertheless the case $r \leq 0$ can be treated by shifting the linear part of the state equation. Indeed in this case we can rewrite the state equation taking for example as new coefficient for the linear part $\tilde{r} = 1$ and shifting the nonlinear term defining $\tilde{f}_0(x,y) = f_0(x,y) - (1-r)x$.

We say that a function $x : [-T,\infty) \to \mathbb{R}^+$ is a solution to equation (3.1) if $x(t) = \eta_1(t)$ for $t \in [-T,0)$ and

$$x(t) = \eta_0 + \int_0^t rx(s)ds + \int_0^t f_0\left(x(s), \int_{-T}^0 a(\xi)x(s+\xi)d\xi\right)ds - \int_0^t c(s)ds, \quad t \geq 0.$$

Theorem 3.1.4. *For any given $\eta \in H_+$, $c(\cdot) \in L^1_{loc}([0,+\infty);\mathbb{R}^+)$, equation* (3.1) *admits a unique solution that is absolutely continuous on* $[0,+\infty)$.

Proof. Let $K = \sup_{\xi\in[-T,0]} a(\xi)$. For any $t \geq 0$, $z^1, z^2 \in C([-T,t];\mathbb{R})$, we have

$$\begin{aligned}
&\int_0^t \Bigg[r|z_1(s) - z_2(s)| \\
&\qquad + \left| f_0\left(z_1(s), \int_{-T}^0 a(\xi)z_1(s+\xi)\right) - f_0\left(z_2(s), \int_{-T}^0 a(\xi)z_2(s+\xi)\right)\right| \Bigg] ds \\
&\leq \int_0^t \left[r|z_1(s) - z_2(s)| + C_{f_0}\left[|z_1(s) - z_2(s)| + K\int_{-T}^0 |z_1(s+\xi) - z_2(s+\xi)|d\xi\right]\right] ds \\
&\qquad \leq \int_0^t \left[(r + C_{f_0})|z_1(s) - z_2(s)| + C_{f_0}K\int_{-T}^t |z_1(\xi) - z_2(\xi)|d\xi\right] ds \\
&\qquad \leq (r + C_{f_0})\int_0^t |z_1(s) - z_2(s)|ds + tC_{f_0}K\int_{-T}^t |z_1(\xi) - z_2(\xi)|d\xi \\
&\qquad\qquad \leq [(r + C_{f_0}) + tC_{f_0}K]\int_{-T}^t |z_1(\xi) - z_2(\xi)|d\xi \\
&\qquad\qquad \leq [(r + C_{f_0}) + tC_{f_0}K](t+T)^{1/2}\left(\int_{-T}^t |z_1(\xi) - z_2(\xi)|^2 d\xi\right)^{1/2}.
\end{aligned}$$

Due to the previous inequality, the claim is a straight consequence of Theorem 3.2, pag. 246, of [Bensoussan, Da Prato, Delfour, Mitter; 2007]. $\square$

We denote by $x(\cdot;\eta,c(\cdot))$ the unique solution of (3.1) with initial point $\eta \in H_+$ and under the control $c(\cdot)$. We emphasize that this solution actually satisfies pointwise only the integral equation associated with (3.1); it satisfies (3.1) in differential form only for almost every $t \in [0,+\infty)$.

For $\eta \in H_+$ we define the class of the admissible controls starting from η as

$$\mathcal{C}(\eta) := \{c(\cdot) \in L^1_{loc}([0,+\infty);\mathbb{R}^+) \mid x(\cdot;\eta,c(\cdot)) > 0\}.$$

Setting $x(\cdot) := x(\cdot, ;\eta, c(\cdot))$, the problem consists in maximizing the functional

$$J(\eta; c(\cdot)) := \int_0^{+\infty} e^{-\rho t} \left[U_1(c(t)) + U_2(x(t))\right] dt, \quad \rho > 0,$$

over the set of the admissible strategies.

The following will be standing assumptions on the utility functions U_1, U_2, holding throughout the whole paper.

Hypothesis 3.1.5.

(i) $U_1 \in C([0,+\infty);\mathbb{R}) \cap C^2((0,+\infty);\mathbb{R})$, $U_1' > 0$, $U_1'(0^+) = +\infty$, $U_1'' < 0$ and U_1 is bounded.

(ii) $U_2 \in C((0,+\infty);\mathbb{R})$ is increasing, concave, bounded from above. Moreover

$$\int_0^{+\infty} e^{-\rho t} U_2\left(e^{-C_{f_0} t}\right) dt > -\infty. \tag{3.3}$$

$\square$

Since U_1, U_2 are bounded from above, the previous functional is well-defined for any $\eta \in H_+$ and $c(\cdot) \in \mathcal{C}(\eta)$. We set

$$\bar{U}_1 := \lim_{s\to+\infty} U_1(s), \quad \bar{U}_2 := \lim_{s\to+\infty} U_2(s).$$

Remark 3.1.6. We give some comments on Hypothesis 3.1.5 and on the structure of the utility in the objective functional.

1. The assumption that U_1, U_2 are bounded from above is done for simplicity to avoid too many technicalities. It guarantees that the value function is bounded from above and this fact simplifies arguments in the following parts of this paper. Similarly the assumption that U_1 is bounded from below guarantees that the value function is bounded from below. We think that it is possible to replace such assumptions with more general conditions relating the growth of U_1, U_2, the value of ρ and the parameters of the state equation. Typically, such a condition requires that ρ is sufficiently large.

2. All utility functions bounded from below satisfy (3.3). Also

$$U_2(x) = \log(x), \quad U_2(x) = x^\gamma, \ \gamma > -\frac{\rho}{C_{f_0}},$$

satisfy (3.3). Note also that (3.3) is equivalent to

$$\int_0^{+\infty} e^{-\rho t} U_2\left(\xi e^{-C_{f_0} t}\right) dt > -\infty, \quad \forall \xi > 0.$$

3. When $r < 0$, then in (3.3) we have to replace C_{f_0} with $|r| + C_{f_0}$.

4. In Subsection 3.3.3 we will assume that U_2 is not integrable at 0^+. In the case of power utility, if we want both the no integrability condition and (3.3) holding, we have to require $\rho > C_{f_0}$ and $-\frac{\rho}{C_{f_0}} \leq \gamma < -1$.

5. If we assume that

$$\exists \delta > 0 \text{ such that } rx + f_0(x, 0) \geq 0, \forall x \in (0, \delta], \tag{3.4}$$

then the assumption (3.3) can be suppressed. Since (3.2) implies $f_0(0,0) \geq 0$, (3.4) occurs for example if $x \mapsto rx + f_0(x, 0)$ is nondecreasing, therefore in particular if $r \geq 0$ and f_0 depends only on the second variable. □

Remark 3.1.7. The control problem described above covers also the following optimal consumption problem. We may think of the dynamics defined in (3.1) as the dynamics of the bank account driven by a contract which takes into account the past history of the accumulation of wealth. Such a situation arises when the bank offers to the customer an interest rate r smaller than the market spot rate r^M and as a compensation, it provides a premium on the past of the wealth (this may happen e.g. for pension funds). Then the following equation is a possible simple model of the evolution of the bank account under such a contract:

$$\begin{cases} x'(t) = rx(t) + g_0\left(\int_{-T}^0 a(\xi) x(t+\xi) d\xi\right) - c(t), \\ x(0) = \eta_0, \ x(s) = \eta_1(s), \ s \in [-T, 0), \end{cases}$$

where $g_0 : \mathbb{R} \to \mathbb{R}$ is a concave, Lipschitz continuous and strictly increasing function such that $g_0(0) \geq 0$. Dependence on the past is an incentive for the customer to keep his investments with the bank for a longer period of time in order to receive gains produced by the term g_0. Here we assume the point of view of the customer and we think it is interesting to study the behaviour of the optimal consumption in this setting, comparing it with the one coming from the classical case, which corresponds to set $r = r^M$, $g \equiv 0$.

We think also that our technique could be adapted to cover optimal advertising model with nonlinear memory effects (see [Gozzi, Marinelli, Savin] on this subject in a stochastic framework). □

For $\eta \in H_+$ the value function of our problem is defined by

$$V(\eta) := \sup_{c(\cdot)\in\mathcal{C}(\eta)} J(\eta, c(\cdot)), \tag{3.5}$$

with the convention $\sup \emptyset = -\infty$. The domain of the value function is the set

$$\mathcal{D}(V) := \{\eta \in H_+ \mid V(\eta) > -\infty\}.$$

Due to the assumptions on U_1, U_2 we directly get that $V \leq \frac{1}{\rho}(\bar{U}_1 + \bar{U}_2)$.

3.1.1 Preliminary results

In this subsection we investigate some first qualitative properties of the delay state equation and of the value function.

Lemma 3.1.8 (Comparison)**.** *Let $\eta \in H_+$ and let $c(\cdot) \in L^1_{loc}([0,+\infty);\mathbb{R}^+)$. Let $x(t)$, $t \geq 0$, be an absolutely continuous function satisfying almost everywhere the differential inequality*

$$\begin{cases} x'(t) \leq rx(t) + f_0\left(x(t), \int_{-T}^0 a(\xi)x(t+\xi)d\xi\right) - c(t), \\ x(0) \leq \eta_0,\ x(s) \leq \eta_1(s), \text{ for a.e. } s \in [-T,0). \end{cases}$$

Then $x(\cdot) \leq x(\cdot;\eta,c(\cdot))$.

Proof. Set $\bar{a} := \sup_{\xi\in[-T,0]} |a(\xi)|$, $y(\cdot) := x(\cdot;\eta,c(\cdot))$ and $h(\cdot) := [x(\cdot)-y(\cdot)]^+$. We must show that $h(\cdot) = 0$. Let $\varepsilon > 0$ be fixed and such that

$$\varepsilon C_{f_0}\bar{a}Te^{\varepsilon(r+C_{f_0})} \leq 1/2 \tag{3.6}$$

and let $M := \max_{t\in[0,\varepsilon]} h(t)$. By monotonicity of f_0 we get

$$f_0\left(x(t), \int_{-T}^0 a(\xi)x(t+\xi)d\xi\right) \leq f_0\left(x(t), \int_{-T}^0 a(\xi)y(t+\xi)d\xi + \bar{a}TM\right),\ t \in [0,\varepsilon]. \tag{3.7}$$

Define, for $n \in \mathbb{N}$,

$$\varphi_n(x) := \begin{cases} 0, & \text{for } x \leq 0, \\ nx^2, & \text{for } x \in (0, 1/2n], \\ x - 1/4n, & \text{for } x > 1/2n. \end{cases}$$

The sequence $(\varphi_n)_{n\in\mathbb{N}} \subset C^1(\mathbb{R};\mathbb{R})$ is such that

$$\begin{cases} \varphi_n(x) = \varphi_n'(x) = 0, \ \text{ for every } x \in (-\infty,0], \ n \in \mathbb{N}, \\ 0 \le \varphi_n'(x) \le 1, \ \text{ for every } x \in \mathbb{R}, \ n \in \mathbb{N}, \\ \varphi_n(x) \to x^+, \quad \text{uniformly on } x \in \mathbb{R}, \\ \varphi_n'(x) \to 1, \ \text{for } x \in (0,+\infty). \end{cases}$$

We can write for $t \in [0,\varepsilon]$, taking into account (3.7),

$$\begin{aligned} \varphi_n(x(t)-y(t)) &= \varphi_n(x(0)-\eta_0) + \int_0^t \varphi_n'(x(s)-y(s))[x'(s)-y'(s)]ds \\ &\le \int_0^t \varphi_n'(x(s)-y(s)) \Bigg[r(x(s)-y(s)) \\ &+ f_0\left(x(s), \int_{-T}^0 a(\xi)x(s+\xi)d\xi\right) - f_0\left(y(s), \int_{-T}^0 a(\xi)y(s+\xi)d\xi\right)\Bigg]ds \\ &\le \int_0^t \varphi_n'(x(s)-y(s)) \Bigg[r(x(s)-y(s)) \\ &+ f_0\left(x(s), \int_{-T}^0 a(\xi)y(s+\xi)d\xi + \bar{a}TM\right) - f_0\left(y(s), \int_{-T}^0 a(\xi)y(s+\xi)d\xi\right)\Bigg]ds \\ &\le \int_0^t \varphi_n'(x(s)-y(s))\Big[(r+C_{f_0})|x(s)-y(s)| + C\bar{a}TM\Big]ds. \end{aligned}$$

Letting $n \to \infty$ we get

$$h(t) \le \int_0^t (r+C_{f_0})\, h(s)ds + C_{f_0}\bar{a}TMt \le \int_0^t (r+C_{f_0})\, h(s)ds + C_{f_0}\bar{a}TM\varepsilon.$$

Therefore by Gronwall's Lemma we get

$$h(t) \le \varepsilon C_{f_0}\bar{a}TMe^{\varepsilon(r+C_{f_0})}, \quad \text{for } t \in [0,\varepsilon],$$

and by (3.6)

$$h(t) \le \frac{M}{2}, \quad \text{for } t \in [0,\varepsilon].$$

This shows that $M = 0$, i.e. that $h = 0$ on $[0,\varepsilon]$. Iterating the argument, since ε is fixed, we get $h \equiv 0$ on $[0,+\infty)$, i.e. the claim. □

Proposition 3.1.9. *We have*

$$H_{++} \subset \mathcal{D}(V), \qquad \mathcal{D}(V) = \{\eta \in H_+ \mid 0 \in \mathcal{C}(\eta)\}.$$

Proof. Let $\eta \in H_{++}$ and set $x(\cdot) := x(\cdot;\eta,0)$. By assumption, $x(0) = \eta_0 > 0$ and until $x(t) > 0$ we have

$$x'(t) = rx(t) + f_0\left(x(t), \int_{-T}^0 a(\xi)x(t+\xi)d\xi\right) \ge rx(t) + f_0(x(t),0).$$

Since $f_0(0,0) \geq 0$ and $f_0(\cdot,0)$ is Lipschitz continuous (with Lipschitz constant C_{f_0}), we get

$$x'(t) \geq -C_{f_0}\, x(t), \quad \text{until } x(t) > 0.$$

This fact forces to be

$$\inf\{t \geq 0 \mid x(t) = 0\} = +\infty,$$

and $x(t) \geq \eta_0 e^{-C_{f_0} t}$ for any $t \geq 0$, which proves the inclusion $H_{++} \subset \mathcal{D}(V)$ thanks to (3.3).

Now let $\eta \in \mathcal{D}(V)$; then, by definition of $\mathcal{D}(V)$, there exists $c(\cdot) \in \mathcal{C}(\eta)$. By Lemma 3.1.8 $0 \in \mathcal{C}(\eta)$, so that we have the inclusion $\mathcal{D}(V) \subset \{\eta \in H_+ \mid 0 \in \mathcal{C}(\eta)\}$. Conversely let $\eta \in H_+$ be such that $0 \in \mathcal{C}(\eta)$. Then, by definition of $\mathcal{C}(\eta)$, we have

$$\inf_{t\in[0,T]} x(t;\eta,0) \geq \xi > 0.$$

Repeating the argument used above, we get $x(t;\eta,0) \geq \xi e^{-C_{f_0}(t-T)}$ for $t \geq T$, so that $\eta \in \mathcal{D}(V)$ and the proof is complete. $\square$

Remark 3.1.10. It is straightforward to see that the proof of Proposition 3.1.9 above works if we replace the assumption (3.3) with the assumption (3.4). $\square$

Definition 3.1.11. (i) Let $\eta \in \mathcal{D}(V)$. An admissible control $c^*(\cdot) \in \mathcal{C}(\eta)$ is said to be optimal for the initial state η if $J(\eta;c^*(\cdot)) = V(\eta)$. In this case the corresponding state trajectory $x^*(\cdot) := x(\cdot;\eta,c^*(\cdot))$ is said to be an optimal trajectory and the couple $(x^*(\cdot),c^*(\cdot))$ is said an optimal couple.

(ii) Let $\eta \in \mathcal{D}(V)$, $\varepsilon > 0$; an admissible control $c^\varepsilon(\cdot) \in \mathcal{C}(\eta)$ is said ε-optimal for the initial state η if $J(\eta;c^\varepsilon(\cdot)) > V(\eta) - \varepsilon$. In this case the corresponding state trajectory $x^\varepsilon(\cdot) := x(\cdot;\eta,c^\varepsilon(\cdot))$ is said an ε-optimal trajectory and the couple $(x^\varepsilon(\cdot),c^\varepsilon(\cdot))$ is said an ε-optimal couple. $\square$

Proposition 3.1.12. *The set $\mathcal{D}(V)$ is convex and the value function V is concave on $\mathcal{D}(V)$.*

Proof. Let $\eta,\bar{\eta} \in \mathcal{D}(V)$ and set, for $\lambda \in [0,1]$, $\eta_\lambda = \lambda\eta + (1-\lambda)\bar{\eta}$. For $\varepsilon > 0$, let $c^\varepsilon(\cdot) \in \mathcal{C}(\eta)$, $\bar{c}^\varepsilon(\cdot) \in \mathcal{C}(\bar{\eta})$ be two controls ε-optimal for the initial states $\eta,\bar{\eta}$ respectively. Set

$$x(\cdot) := x(\cdot,\eta,c^\varepsilon(\cdot)), \quad \bar{x}(\cdot) := x(\cdot;\eta,\bar{c}^\varepsilon(\cdot)), \quad c^\lambda(\cdot) := \lambda c^\varepsilon(\cdot) + (1-\lambda)\bar{c}^\varepsilon(\cdot).$$

Finally set $x_\lambda(\cdot) := \lambda x(\cdot) + (1-\lambda)\bar{x}(\cdot)$. Let us write the dynamics for $x_\lambda(\cdot)$:

$$\begin{aligned} x'_\lambda(t) &= \lambda x'(t) + (1-\lambda)\bar{x}'(t) \\ &= \lambda \left[rx(t) + f_0\left(x(t), \int_{-T}^0 a(\xi)x(t+\xi)d\xi \right) - c^\varepsilon(t) \right] \\ &\quad +(1-\lambda)\left[r\bar{x}(t) + f_0\left(\bar{x}(t), \int_{-T}^0 a(\xi)\bar{x}(t+\xi)d\xi \right) - \bar{c}^\varepsilon(t) \right] \\ &\le rx_\lambda(t) + f_0\left(x_\lambda(t), \int_{-T}^0 a(\xi)x_\lambda(t+\xi)d\xi \right) - c^\lambda(t), \end{aligned}$$

where the inequality follows from the concavity of f_0 with initial condition η_λ. Let $x(\cdot;\eta_\lambda, c^\lambda(\cdot))$ be a solution of the equation

$$x'(t) = rx(t) + f_0\left(x(t), \int_{-T}^0 a(\xi)x(t+\xi)d\xi \right) - c^\lambda(t).$$

Since by construction $x_\lambda(\cdot) > 0$, by Lemma 3.1.8 we have

$$x(\cdot;\eta_\lambda, c^\lambda(\cdot)) \ge x_\lambda(\cdot) > 0.$$

This shows that $c^\lambda(\cdot) \in \mathcal{C}(\eta_\lambda)$. By concavity of U_1, U_2 and by monotonicity of U_2 we get

$$V(\eta_\lambda) \ge J(\eta_\lambda; c^\lambda(\cdot)) \ge \lambda J(\eta; c^\varepsilon(\cdot)) + (1-\lambda)J(\eta; \bar{c}^\varepsilon(\cdot)) > \lambda V(\eta) + (1-\lambda)V(\bar{\eta}) - \varepsilon.$$

Since ε is arbitrary, we get the claim. □

From the assumptions of monotonicity of the utility functions and from Lemma 3.1.8 we obtain the following result.

Proposition 3.1.13. *The function $\eta \mapsto V(\eta)$ is nondecreasing in the sense that*

$$\eta_0 \ge \bar{\eta}_0,\ \eta_1(\cdot) \ge \bar{\eta}_1(\cdot) \Longrightarrow V(\eta_0, \eta_1(\cdot)) \ge V(\bar{\eta}_0, \bar{\eta}_1(\cdot)).$$

□

Indeed, the value function is strictly increasing in the first variable.

Proposition 3.1.14. *We have the following statements:*

1. $V(\eta) < \frac{1}{\rho}(\bar{U}_1 + \bar{U}_2)$ *for any* $\eta \in H_+$.
2. $\lim_{\eta_0 \to +\infty} V(\eta_0, \eta_1(\cdot)) = \frac{1}{\rho}(\bar{U}_1 + \bar{U}_2)$, *for all* $\eta_1(\cdot) \in L^2_{-T}$.
3. *V is strictly increasing with respect to the first variable.*

Proof. 1. Let $\eta \in \mathcal{D}(V)$ and set

$$\bar{a} := \sup_{\xi\in[-T,0]} a(\xi), \quad p := \sup_{\xi\in[0,T]} x(\xi;\eta,0), \quad q := \int_{-T}^{0} \eta_1^+(\xi)d\xi.$$

Let $c(\cdot) \in \mathcal{C}(\eta)$ and set $x(\cdot) := x(\cdot;\eta,c(\cdot))$; by comparison criterion we have $x(t) \le p$ in $[0,T]$.

Since f_0 is Lipschitz continuous, there exists $C > 0$ such that

$$f_0(x,y) \le C(1+|x|+|y|), \quad \forall x,y \in \mathbb{R}.$$

Therefore, for $t \in [0,T]$, we can write, considering the state equation in integral form,

$$x(t) \le \eta_0 + rTp + TC(1+|p|+|\bar{a}(Tp+q)|) - \int_0^T c(\tau)d\tau.$$

Define

$$K := \eta_0 + rTp + TC(1+|p|+|\bar{a}(Tp+q)|);$$

since $c(\cdot) \in \mathcal{C}(\eta)$, we have $x(t) > 0$ in $[0,T]$, so that

$$\int_0^T c(\tau)d\tau \le K.$$

Denoting by m the Lebesgue measure, this means that

$$m\{\tau \in [0,T] \mid c(\tau) \le 2K/T\} \ge T/2.$$

Therefore (in the next inequality, since $e^{-\rho t}$ is decreasing, we suppose without loss of generality that $c(\cdot) \le 2K/T$ on $\left[\frac{T}{2},T\right]$)

$$\begin{aligned}
&\int_0^{+\infty} e^{-\rho t}U_1(c(t))dt \\
&\quad \le \int_0^{T/2} e^{-\rho t}U_1(c(t))dt + \int_{T/2}^{T} e^{-\rho t}U_1(2K/T)dt + \int_T^{+\infty} e^{-\rho t}U_1(c(t))dt \\
&\qquad\qquad \le \frac{\bar{U}_1}{\rho} - \int_{T/2}^{T} e^{-\rho t}\Big(\bar{U}_1 - U_1(2K/T)\Big)dt.
\end{aligned}$$

Since the quantity $\bar{U}_1 - U_1(2K/T)$ is strictly positive and does not depend on $c(\cdot)$, the claim is proved.

2. For given $\eta_1(\cdot) \in L^2_{-T}$, let $K > 0$, $M > 0$ and let us define the control

$$c(t) := \begin{cases} M, & \text{if } t \in [0,K], \\ 0, & \text{if } t > K. \end{cases}$$

Take $\eta_0 > 0$. Since f_0 is Lipschitz continuous and nondecreasing with respect to the second variable, we can see that, until it is positive, $x(t;(\eta_0,\eta_1(\cdot)),c(\cdot))$ satisfies the differential inequality

$$\begin{cases} x'(t) \geq -C(1+x(t)+q) - M, \\ x(0) = \eta_0, \end{cases}$$

for some $C > 0$, where

$$q := \left(\sup_{\xi\in[-T,0]} a(\xi)\right)\left(\int_{-T}^{0} \eta_1^-(\xi)d\xi\right).$$

This actually shows that, for any $M > 0$, $K > 0$, $R > 0$, we can find η_0 such that $c(\cdot) \in \mathcal{C}(\eta_0,\eta_1(\cdot))$ and $x(\cdot;(\eta_0,\eta_1(\cdot)),c(\cdot)) \geq R$ on $[0,K]$. By the arbitrariness of M, K, R the claim is proved.

3. Fix $\eta_1(\cdot)$; we know that $\eta_0 \mapsto V(\eta_0,\eta_1(\cdot))$ is concave and increasing. If it is not strictly increasing, then it has to be constant on an half line $[k,+\infty)$, but this contradicts the first two claims. $\square$

3.2 The delay problem rephrased in infinite dimension

Our aim is to apply the dynamic programming technique in order to solve the control problem described in the previous section. However, this approach requires a markovian setting. That is why we will reformulate the problem as an infinite-dimensional control problem. Let $\hat{n} = (1,0) \in H_+$ and let us consider, for $\eta \in H$ and $c(\cdot) \in L^1([0,+\infty);\mathbb{R}^+)$, the following evolution equation in the space H:

$$\begin{cases} X'(t) = AX(t) + F(X(t)) - c(t)\hat{n}, \\ X(0) = \eta \in H_+. \end{cases} \tag{3.8}$$

In the equation above:

- $A : \mathcal{D}(A) \subset H \longrightarrow H$ is an unbounded operator defined by

  $$A(\eta_0,\eta_1(\cdot)) := (r\eta_0,\eta_1'(\cdot))$$

  on

  $$\mathcal{D}(A) := \{\eta \in H \mid \eta_1(\cdot) \in W^{1,2}_{-T},\ \eta_1(0) = \eta_0\};$$

- $F : H \longrightarrow H$ is a Lipschitz continuous map defined by

  $$F(\eta_0,\eta_1(\cdot)) := (f(\eta_0,\eta_1(\cdot)),0),$$

  where $f(\eta_0,\eta_1(\cdot)) := f_0\left(\eta_0, \int_{-T}^0 a(\xi)\eta_1(\xi)d\xi\right)$.

As in Chapter 2 A is the infinitesimal generator of a strongly continuous semigroup $(S(t))_{t\geq 0}$ on H with explicit expression given by

$$S(t)(\eta_0, \eta_1(\cdot)) = \left(\eta_0 e^{rt}, I_{[-T,0]}(t+\cdot)\, \eta_1(t+\cdot) + I_{[0,+\infty)}(t+\cdot)\, \eta_0 e^{r(t+\cdot)}\right).$$

3.2.1 The state equation: existence and uniqueness of mild solutions and equivalence with the delay problem

In this subsection we give a definition of the *mild solution* to (3.8), prove the existence and uniqueness of such a solution and the equivalence between the one dimensional delay problem and the infinite dimensional one.

Definition 3.2.1. A mild solution of (3.8) is a function $X \in C([0,+\infty); H)$ which satisfies the integral equation

$$X(t) = S(t)\eta + \int_0^t S(t-\tau)F(X(\tau))d\tau + \int_0^t c(\tau)S(t-\tau)\hat{n}\, d\tau. \tag{3.9}$$

Theorem 3.2.2. *For any $\eta \in H$, there exists a unique mild solution of (3.8).*

Proof. Due to the Lipschitz continuity of F and to (2.17), the proof follows from the Fixed Point Theorem. □

We denote by $X(\cdot;\eta,c(\cdot)) = (X_0(\cdot;\eta,c(\cdot)), X_1(\cdot;\eta,c(\cdot)))$ the unique solution to (3.8) for the initial state $\eta \in H$ and under the control $c(\cdot) \in L^1([0,+\infty);\mathbb{R}^+)$. The following equivalence result justifies our approach.

Proposition 3.2.3. *Let $\eta \in H_+$, $c(\cdot) \in \mathcal{C}(\eta)$ and let $x(\cdot)$, $X(\cdot)$ be respectively the unique solution to (3.1) and the unique mild solution to (3.8) starting from η and under the control $c(\cdot)$. Then, for any $t \geq 0$, we have the equality in H*

$$X(t) = \left(x(t), x(t+\xi)_{\xi\in[-T,0]}\right).$$

Proof. Let $x(\cdot)$ be a solution of (3.1) and let $Z(\cdot) := (x(\cdot), x(\cdot+\zeta)|_{\zeta\in[-T,0]})$. Then $Z(\cdot)$ belongs to the space $C([0,+\infty); H)$ because the function $[0,+\infty) \ni t \mapsto x(t) \in \mathbb{R}$ is (absolutely) continuous. Therefore, it remains to prove that $Z(t) = (Z_0(t), Z_1(t))$ satisfies (3.8) and then the claim will follow by uniqueness. For the first component we have to verify that, for any $t \geq 0$,

$$Z_0(t) = e^{rt}\eta_0 + \int_0^t e^{r(t-\tau)} f_0\left(Z_0(\tau), \int_{-T}^0 a(\xi)Z_1(\tau)(\xi)d\xi\right)d\tau - \int_0^t e^{r(t-\tau)}c(\tau)d\tau,$$

i.e. that

$$x(t) = e^{rt}\eta_0 + \int_0^t e^{r(t-\tau)} f_0\left(x(\tau), \int_{-T}^0 a(\xi)x(\tau+\xi)d\xi\right)d\tau - \int_0^t e^{r(t-\tau)}c(\tau)d\tau,$$

but this follow from the assumption that $x(\cdot)$ is a solution to (3.1).
For the second component, taking into account that

$$I_{[0,+\infty)}(t+\cdot-\tau) = I_{[\tau,+\infty)}(t+\ \cdot),$$

we have to verify, for any $t \geq 0$, for a.e. $\zeta \in [-T, 0]$,

$$\begin{aligned}
Z_1(t)(\zeta) \quad = \quad & I_{[-T,0]}(t+\zeta)\eta_1(t+\zeta) + I_{[0,+\infty)}(t+\zeta)\eta_0 e^{r(t+\zeta)} \\
& + \int_0^t I_{[\tau,+\infty)}(t+\zeta)\ e^{r(t+\zeta-\tau)} f_0\left(Z_0(\tau), \int_{-T}^0 a(\xi) Z_1(\tau)(\xi) d\xi\right) d\tau \\
& - \int_0^t I_{[\tau,+\infty)}(t+\zeta)\ e^{r(t+\zeta-\tau)} c(\tau) d\tau,
\end{aligned}$$

i.e., for any $t \geq 0$, for a.e. $\zeta \in [-T, 0]$,

$$\begin{aligned}
x(t+\zeta) \quad = \quad & I_{[-T,0]}(t+\zeta)\eta_1(t+\zeta) + I_{[0,+\infty)}(t+\zeta)\ \eta_0 e^{r(t+\zeta)} \\
& + \int_0^t I_{[\tau,+\infty)}(t+\zeta)\ e^{r(t+\zeta-\tau)} f_0\left(x(\tau), \int_{-T}^0 a(\xi) x(\tau+\xi) d\xi\right) d\tau \\
& - \int_0^t I_{[\tau,+\infty)}(t+\zeta)\ e^{r(t+\zeta-\tau)} c(\tau) d\tau. \qquad (3.10)
\end{aligned}$$

For $\zeta \in [-T, 0]$ such that $t+\zeta \in [-T, 0]$, (3.10) reduces to

$$x(t+\zeta) = \eta_1(t+\zeta)$$

and this is true since η_1 is the initial condition of (3.1). If $\zeta \in [-T, 0]$ is such that $t+\zeta \geq 0$, then (3.10) reduces to

$$\begin{aligned}
x(t+\zeta) = \eta_0 e^{r(t+\zeta)} + \int_0^{t+\zeta} e^{r(t+\zeta-\tau)} f_0\left(x(\tau), \int_{-T}^0 a(\xi) x(\tau+\xi) d\xi\right) d\tau \\
- \int_0^{t+\zeta} e^{r(t+\zeta-\tau)} c(\tau) d\tau.
\end{aligned}$$

Setting $u := t+\zeta$ this equality becomes, for $u \geq 0$,

$$x(u) = x_0 e^{ru} + \int_0^t e^{r(u-\tau)} f_0\left(x(\tau), \int_{-T}^0 a(\xi) x(\tau+\xi) d\xi\right) d\tau - \int_0^t e^{r(u-\tau)} c(\tau) d\tau.$$

Again this is true because $x(\cdot)$ solves (3.1). □

3.2.2 Continuity of the value function

In this subsection we prove a continuity property of the value function that will be useful to investigate the geometry of its superdifferential in the next subsection. To this end we recall that the generator A of the semigroup $(S(t))_{t\geq 0}$ has bounded inverse in H given by

$$\begin{array}{rccc} A^{-1}: & (H, \|\cdot\|_H) & \longrightarrow & (\mathcal{D}(A), \|\cdot\|_H), \\ & \eta & \longmapsto & \left(\frac{\eta_0}{r}, \frac{\eta_0}{r} - \int_{\cdot}^{0} \eta_1(\xi) d\xi\right). \end{array}$$

It is well known that A^{-1} is compact in H. It is also clear that A^{-1} is an isomorphism of H onto $\mathcal{D}(A)$ endowed with the graph norm.

We define the $\|\cdot\|_{-1}$-norm on H by

$$\|\eta\|_{-1} := \|A^{-1}\eta\|.$$

In the next proposition we characterize the adjoint operator A^* and its domain $\mathcal{D}(A^*)$.

Proposition 3.2.4. *Let $\eta = (\eta_0, \eta_1(\cdot)) \in H$. Then $\eta \in \mathcal{D}(A^*)$ if and only if*

$$(\eta_0, \eta_1(\cdot)) \in \mathcal{D} := \left\{\eta \in H \mid \eta_1 \in W^{1,2}_{-T},\ \eta_1(-T) = 0\right\}.$$

Moreover, if this is the case, then

$$A^*\eta = (r\eta_0 + \eta_1(0), -\eta_1'(\cdot)). \tag{3.11}$$

Proof. Let $(\eta_0, \eta_1(\cdot)) \in \mathcal{D}$. Then, for $\zeta \in \mathcal{D}(A)$,

$$\langle A\zeta, \eta\rangle = r\zeta_0\eta_0 + \int_{-T}^{0} \zeta_1'(s)\eta_1(s)\, ds = r\zeta_0\eta_0 + \zeta_0\eta_1(0) - \int_{-T}^{0} \zeta_1(s)\eta_1'(s)\, ds,$$

thus $\zeta \mapsto \langle A\zeta, \eta\rangle$ is continuous on $\mathcal{D}(A)$ with respect to the norm $\|\cdot\|$, i.e. $\eta \in D(A^*)$ and

$$A^*\eta = (r\eta_0 + \eta_1(0), -\eta_1'(\cdot)).$$

Therefore, $\eta \in \mathcal{D}(A^*)$ and (3.11) holds. To show that $\mathcal{D}(A^*) = \mathcal{D}$ note first that for $t \leq T$

$$S^*(t)(\eta_0, \eta_1(\cdot)) = \left(e^{rt}\left(\eta_0 + \int_{-t}^{0} \eta_1(\xi)e^{r\xi} d\xi\right), \eta_1(\cdot - t)I_{[-T,0]}(\cdot - t)\right). \tag{3.12}$$

Clearly, $\mathcal{D}$ is dense in H and it is easy to check that $S^*(t)\mathcal{D} \subset \mathcal{D}$ for any $t \geq 0$. Hence, by Theorem 1.9 on p. 8 of [Davies; 1980] $\mathcal{D}$ is dense in $\mathcal{D}(A^*)$ endowed with the graph norm. Finally, using (3.11) it is easy to show that $\mathcal{D}$ is closed in the graph norm of A^* and we find that $\mathcal{D}(A^*) = \mathcal{D}$. □

Lemma 3.2.5. *The map F is Lipschitz continuous with respect to $\|\cdot\|_{-1}$.*

Proof. Due to the Lipschitz continuity of f_0, it is sufficient to prove that

$$|\eta_0| + \left| \int_{-T}^{0} a(\xi)\eta_1(\xi)d\xi \right| \le C_{a(\cdot)} \|\eta\|_{-1}, \quad \forall \eta \in H. \tag{3.13}$$

Indeed, since $|\eta_0| \le r\|\eta\|_{-1}$ $(0, a(\cdot)) \in \mathcal{D}(A^*)$, we find that

$$\begin{aligned} \left| \int_{-T}^{0} a(\xi)\eta_1(\xi)d\xi \right| &= |\langle (0, a(\cdot)), \eta \rangle| = |\langle (0, a(\cdot)), AA^{-1}\eta \rangle| \\ &= |\langle A^*(0, a(\cdot)), A^{-1}\eta \rangle| \le \|A^*(0, a(\cdot))\| \cdot \|\eta\|_{-1}. \end{aligned}$$

So, since $|\eta_0| \le r\|\eta\|_{-1}$, we get (3.13) with $C_{a(\cdot)} = r + \|A^*(0, a(\cdot))\|$. □

Remark 3.2.6. The condition $a(-T) = 0$ is in general necessary for the previous result. Indeed, consider for example the case $a(\cdot) \equiv 1$. Then the sequence

$$\eta^n = (\eta_0^n, \eta_1^n(\cdot)), \quad \eta_0^n := 0, \ \eta_1^n(\cdot) := I_{[-T, -T+1/n]}(\cdot), \quad n \ge 1,$$

is such that

$$\left| \int_{-T}^{0} a(\xi)\eta_1^n(\xi)d\xi \right| = 1 \ \ \forall n \ge 1, \quad \|\eta^n\|_{-1} \to 0 \text{ when } n \to \infty,$$

so that (3.13) cannot be satisfied. If for example, $f_0(r, u) = u$, the previous result does not hold. □

Lemma 3.2.7. *Let $X(\cdot), \bar{X}(\cdot)$ be the mild solutions to (3.8) starting respectively from $\eta, \bar{\eta} \in H$ and both under the null control. Then there exists a constant $C > 0$ such that*

$$\|X(t) - \bar{X}(t)\|_{-1} \le C\|\eta - \bar{\eta}\|_{-1}, \quad \forall t \in [0, T].$$

In particular

$$|X_0(t) - \bar{X}_0(t)| \le rC\|\eta - \bar{\eta}\|_{-1}, \quad \forall t \in [0, T].$$

Proof. From (3.9) we can write, for all $t \in [0, T]$,

$$X(t) - \bar{X}(t) = S(t)(\eta - \bar{\eta}) + \int_0^t S(t - \tau) \left[F(X(\tau)) - F(\bar{X}(\tau)) \right] d\tau,$$

so that

$$A^{-1}(X(t) - \bar{X}(t)) = S(t)A^{-1}(\eta - \bar{\eta}) + \int_0^t S(t - \tau)\, A^{-1} \left[F(X(\tau)) - F(\bar{X}(\tau)) \right] d\tau,$$

i.e., taking into account Lemma 3.2.5, there exists some $K > 0$ such that

$$\|X(t) - \bar{X}(t)\|_{-1} \le K \left(\|\eta - \bar{\eta}\|_{-1} + \int_0^t \|X(\tau) - \bar{X}(\tau)\|_{-1} d\tau \right)$$

and the claim follows by Gronwall's Lemma. □

Proposition 3.2.8. *The set $\mathcal{D}(V)$ is open in the space $(H, \|\cdot\|_{-1})$.*

Proof. Let $\bar{\eta} \in \mathcal{D}(V)$, $\eta \in H_+$ and set $\bar{X}(\cdot) := X(\cdot; \bar{\eta}, 0)$, $X(\cdot) := X(\cdot; \eta, 0)$. By Proposition 3.1.9 we have $\bar{X}(t) \geq \xi > 0$ for $t \in [0, T]$. For any $\varepsilon \in \left(0, \frac{\xi}{2rC}\right)$ and any η such that $\|\eta - \bar{\eta}\|_{-1} < \varepsilon$, Lemma 3.2.7 yields $X_0(t) \geq \xi/2$ for $t \in [0, T]$. Arguing as in Proposition 3.1.9 we get $X_0(t) \geq \frac{\xi}{2} e^{-K(t-T)}$ for $t \geq T$. Thus we have the claim. □

Remark 3.2.9. Note that $\mathcal{D}(V)$ is open also with respect to $\|\cdot\|_H$. □

Proposition 3.2.10. *The value function is continuous with respect to $\|\cdot\|_{-1}$ on $\mathcal{D}(V)$. Moreover*

$$(\eta_n) \subset \mathcal{D}(V), \ \ \eta_n \rightharpoonup \eta \in \mathcal{D}(V) \implies V(\eta_n) \to V(\eta). \tag{3.14}$$

Proof. The function V is concave and, thanks to the proof of Lemma 3.2.8, it is $\|\cdot\|_{-1}$-locally bounded from below at the points of $\mathcal{D}(V)$. Therefore the first claim follows by a classic result of convex analysis (see [Ekeland, Temam; 1976], Chapter 1, Corollary 2.4).

The claim (3.14) follows by the first claim and since A^{-1} is compact. □

3.2.3 Properties of superdifferential

In this subsection we focus on the properties of the superdifferential of concave and $\|\cdot\|_{-1}$-continuous functions. This will be very useful in proving a regularity result for the value function. Recall that, if v is a function defined on some open set $\mathcal{O}$ of H, the subdifferential and the superdifferential of v at a point $\bar{\eta} \in \mathcal{O}$ are the convex and closed sets defined respectively by

$$D^- v(\bar{\eta}) := \left\{ \zeta \in H \;\middle|\; \liminf_{\eta \to \bar{\eta}} \frac{v(\eta) - v(\bar{\eta}) - \langle \eta - \bar{\eta}, \zeta \rangle_H}{\|\eta - \bar{\eta}\|} \geq 0 \right\},$$

$$D^+ v(\bar{\eta}) := \left\{ \zeta \in H \;\middle|\; \limsup_{\eta \to \bar{\eta}} \frac{v(\eta) - v(\bar{\eta}) - \langle \eta - \bar{\eta}, \zeta \rangle_H}{\|\eta - \bar{\eta}\|} \leq 0 \right\}.$$

It is well-known that, if $D^+ v(\eta) \cap D^- v(\eta) \neq \emptyset$, then $D^+ v(\eta) \cap D^- v(\eta) = \{\zeta\}$, v is differentiable at η and $\nabla v(\eta) = \zeta$. Moreover the set of the *reachable gradients* is defined as

$$D^* v(\bar{\eta}) := \left\{ \zeta \in H \;\middle|\; \exists \eta_n \to \bar{\eta} \text{ such that } \exists \nabla v(\eta_n), \ \nabla v(\eta_n) \to \zeta \right\}.$$

If $\mathcal{O}$ is convex and open and $v : \mathcal{O} \to \mathbb{R}$ is concave, then the set $D^+ v$ is not empty at any point of $\mathcal{O}$ and

$$D^+ v(\bar{\eta}) = \left\{ \zeta \in H \;\middle|\; v(\eta) - v(\bar{\eta}) \leq \langle \eta - \bar{\eta}, \zeta \rangle_H, \ \ \forall \eta \in \mathcal{O} \right\} = \overline{co}\left(D^* v(\bar{\eta})\right). \tag{3.15}$$

In this case, if $D^+ v(\bar{\eta}) = \{\zeta\}$, then v is differentiable at η and $\nabla v(\eta) = \zeta$.

Lemma 3.2.11. *The following statements hold:*

1. $A^{-1}(\mathcal{D}(V))$ *is a convex open set of* $(\mathcal{D}(A), \|\cdot\|_H)$.

2. $\mathcal{O} := Int_{(H,\|\cdot\|_H)}\left(Clos_{(H,\|\cdot\|_H)}\left(A^{-1}(\mathcal{D}(V))\right)\right)$ *is a convex open of* $(H, \|\cdot\|_H)$.

3. $\mathcal{O} \supset A^{-1}(\mathcal{D}(V))$ *and* $\mathcal{D}(V) = \mathcal{O} \cap \mathcal{D}(A)$.

Proof. The first and the second statement are obvious. We prove the third one. Of course, since $A^{-1}(\mathcal{D}(V))$ is open in $(\mathcal{D}(A), \|\cdot\|_H)$, we can find $(\varepsilon_x)_{x\in A^{-1}(\mathcal{D}(V))}$, $\varepsilon_x > 0$, such that

$$A^{-1}(\mathcal{D}(V)) = \bigcup_{x\in A^{-1}(\mathcal{D}(V))} B_{(\mathcal{D}(A),\|\cdot\|_H)}(x, \varepsilon_x).$$

By this representation of $A^{-1}(\mathcal{D}(V))$ we can see that

$$\mathcal{O} = \bigcup_{x\in A^{-1}(\mathcal{D}(V))} B_{(H,\|\cdot\|_H)}(x, \varepsilon_x).$$

Therefore we get both the claims of the third statement. □

Proposition 3.2.12. *Let* $v : \mathcal{D}(V) \to \mathbb{R}$ *be a concave function continuous with respect to* $\|\cdot\|_{-1}$. *Then*

1. $v = u \circ A^{-1}$, *where* $u : \mathcal{O} \subset H \to \mathbb{R}$ *is a concave* $\|\cdot\|_H$*-continuous function.*

2. $D^+v(\eta) \subset \mathcal{D}(A^*)$, *for any* $\eta \in \mathcal{D}(V)$.

3. $D^+u(A^{-1}\eta) = A^*D^+v(\eta)$, *for any* $\eta \in \mathcal{D}(V)$. *In particular, since* A^* *is injective,* v *is differentiable at* η *if and only if* u *is differentiable at* $A^{-1}\eta$.

4. *If* $\zeta \in D^*v(\eta)$, *then there exists a sequence* $\eta_n \to \eta$ *such that there exist* $\nabla v(\eta_n)$ *and the convergences* $\nabla v(\eta_n) \to \zeta$, $A^*\nabla v(\eta_n) \rightharpoonup A^*\zeta$ *hold true.*

Proof. Within this proof, for $\eta \in \mathcal{D}(V)$, we set $\eta' := A^{-1}\eta$. Since A^{-1} is one-to-one, there is a one-to-one correspondence between the elements $\eta \in \mathcal{D}(V)$ and $\eta' \in A^{-1}(\mathcal{D}(V))$.

1. Let us define the function $u_0 : A^{-1}(\mathcal{D}(V)) \to \mathbb{R}$ by

$$u_0(\eta') := v(\eta).$$

Thanks to the assumptions on v, we see that u_0 is a concave continuous function on $(A^{-1}(\mathcal{D}(V)), \|\cdot\|_H)$. By the third statement of Lemma 3.2.11 we see that $A^{-1}(\mathcal{D}(V))$ is $\|\cdot\|_H$-dense in $\mathcal{O}$. Since v is concave it is locally Lipschitz continuous, so that can be extended to a concave $\|\cdot\|_H$-continuous function u defined on $\mathcal{O}$. This function u satisfies the claim by construction.

2. Let $\bar\eta \in \mathcal{D}(V)$, $\zeta \in D^+v(\bar\eta)$. Then

$$v(\eta) - v(\bar\eta) \le \langle \eta - \bar\eta, \zeta\rangle_H, \quad \forall \eta \in \mathcal{D}(V),$$

i.e.

$$u(\eta') - u(\bar\eta') \le \langle A(\eta' - \bar\eta'), \zeta\rangle_H, \quad \forall \eta' \in A^{-1}(\mathcal{D}(V)).$$

Thus the function

$$\begin{array}{rccc} T_\zeta : & (\mathcal{D}(A), \|\cdot\|_H) & \longrightarrow & \mathbb{R}, \\ & \eta' & \longmapsto & \langle A\eta', \zeta\rangle_H, \end{array}$$

is lower semicontinuous at $\bar\eta'$. It is also linear and therefore it is continuous on $(\mathcal{D}(A), \|\cdot\|)$, so that we can conclude that $\zeta \in \mathcal{D}(A^*)$.

3. Let $\bar\eta \in \mathcal{D}(V)$, $\zeta \in D^+v(\bar\eta)$. Then

$$v(\eta) - v(\bar\eta) \le \langle \eta - \bar\eta, \zeta\rangle, \quad \forall \eta \in \mathcal{D}(V),$$

i.e.

$$u(\eta') - u(\bar\eta') \le \langle A(\eta' - \bar\eta'), \zeta\rangle_H = \langle (\eta' - \bar\eta'), A^*\zeta\rangle_H, \quad \forall \eta' \in A^{-1}(\mathcal{D}(V)),$$

so that $A^*\zeta \in D^+u(\bar\eta')$, that gives $D^+u(A^{-1}\eta) \supset A^*D^+v(\eta)$.

Conversely let $\bar\eta' \in \mathcal{A}^{-1}(\mathcal{D}(V))$ and $\zeta' \in D^+u(\bar\eta')$. Then

$$u(\eta') - u(\bar\eta') \le \langle A(\eta' - \bar\eta'), \zeta'\rangle_H, \quad \forall \eta' \in A^{-1}(\mathcal{D}(V)),$$

i.e.

$$v(\eta) - v(\bar\eta) \le \langle A^{-1}(\eta - \bar\eta), \zeta'\rangle_H = \langle (\eta - \bar\eta), (A^{-1})^*\zeta'\rangle_H, \quad \forall \eta \in \mathcal{D}(V).$$

Since $(A^{-1})^* = (A^*)^{-1}$, we get $(A^*)^{-1}\zeta' \in D^+v(\bar\eta)$, that gives also the opposite inclusion $D^+u(A^{-1}\eta) \subset A^*D^+v(\eta)$.

4. Let $\bar\eta \in \mathcal{D}(V)$ and $\zeta \in D^*v(\bar\eta)$. Due to (3.15), we can find a sequence $\eta_n \to \bar\eta$ such that $\nabla v(\eta_n)$ exists for any $n \in \mathbb{N}$ and $\nabla v(\eta_n) \to \zeta$. Thanks to the third claim we can say that also $\nabla u(\bar\eta'_n)$ exists and $\nabla u(\bar\eta'_n) = A^*\nabla v(\eta_n)$. The sequence $\nabla u(\bar\eta'_n)$ is bounded, due to the fact that the set-valued map $\eta' \mapsto D^+u(\bar\eta')$ is locally bounded. Therefore from any subsequence we can extract a subsubsequence weakly converging to some element $\zeta' \in H$. A^* is a closed operator, so that it is also a weakly closed operator. Therefore we can conclude that $\zeta \in \mathcal{D}(A^*)$ and $\zeta' = A^*\zeta$. Since this holds for any subsequence, we can conclude that $A^*\nabla v(\eta_n) \rightharpoonup A^*\zeta$. □

3.3 Dynamic Programming

The dynamic programming principle states that, for any $\eta \in \mathcal{D}(V)$ and for any $s \geq 0$,

$$V(\eta) = \sup_{c(\cdot)\in\mathcal{C}(\eta)} \left[\int_0^s e^{-\rho t} \left(U_1(c(t) + U_2(X_0(t))\right) dt + e^{-\rho s} V(X(s;\eta,c(\cdot))) \right].$$

The HJB equation associated is

$$\rho v(\eta) = \langle \eta, A^* \nabla v(\eta) \rangle_H + f(\eta) v_{\eta_0}(\eta) + U_2(\eta_0) + \mathcal{H}(v_{\eta_0}(\eta)), \tag{3.16}$$

where $\mathcal{H}$ is the Legendre transform of U_1, i.e.

$$\mathcal{H}(\zeta_0) := \sup_{c\geq 0} \left(U_1(c) - \zeta_0 c \right), \quad \zeta_0 > 0.$$

Due to Hyphothesis 3.1.5-(i) and to Corollary 26.4.1 of [Rockafellar; 1970], we have that $\mathcal{H}$ is strictly convex on $(0, +\infty)$. Notice that, thanks to Proposition 3.1.14-(3),

$$D^+_{\eta_0} V(\eta) := \{ \zeta_0 \in \mathbb{R} \mid (\zeta_0, \zeta_1(\cdot)) \in D^+ V(\eta) \} \subset (0, \infty)$$

for any $\eta \in \mathcal{D}(V)$, i.e. where $\mathcal{H}$ is defined.

3.3.1 Viscosity solutions

First we study the HJB equation using the viscosity solutions approach. In order to follow this approach, we have to define a suitable set of regular test functions. This is the set

$$\tau := \Big\{ \varphi \in C^1(H) \mid \nabla\varphi(\cdot) \in \mathcal{D}(A^*),\ \eta_n \to \eta \Rightarrow A^*\nabla\varphi(\eta_n) \rightharpoonup A^*\nabla\varphi(\eta) \Big\}. \tag{3.17}$$

Let us define, for $c \geq 0$, the operator $\mathcal{L}^c$ on τ by

$$[\mathcal{L}^c \varphi](\eta) := -\rho\varphi(\eta) + \langle \eta, A^*\nabla\varphi(\eta)\rangle_H + f(\eta)\varphi_{\eta_0}(\eta) - c\varphi_{\eta_0}(\eta).$$

Lemma 3.3.1. *Let $\varphi \in \tau$, $c(\cdot) \in L^1([0,+\infty);\mathbb{R}^+)$ and set $X(t) := X(t;\eta,c(\cdot))$. Then the following identity holds for any $t \geq 0$:*

$$e^{-\rho t}\varphi(X(t)) - \varphi(\eta) = \int_0^t e^{-\rho s} [\mathcal{L}^{c(s)}\varphi](X(s)) ds.$$

Proof. The statement holds if we replace A with the Yosida approximations. Then we can pass to the limit and get the claim thanks to the regularity properties of the functions belonging to τ. ☐

Definition 3.3.2. (i) A continuous function $v : \mathcal{D}(V) \to \mathbb{R}$ is called a viscosity subsolution of (3.16) on $\mathcal{D}(V)$ if for any $\varphi \in \tau$ and any $\eta_M \in \mathcal{D}(V)$ such that $v - \varphi$ has a $\|\cdot\|_H$-local maximum at η_M we have

$$\rho v(\eta_M) \leq \langle \eta_M, A^*\nabla\varphi(\eta_M)\rangle_H + f(\eta_M)\varphi_{\eta_0}(\eta_M) + U_2(\eta_0) + \mathcal{H}(\varphi_{\eta_0}(\eta_M)).$$

(ii) A continuous function $v : \mathcal{D}(V) \to \mathbb{R}$ is called a viscosity supersolution of (3.16) on $\mathcal{D}(V)$ if for any $\varphi \in \tau$ and any $\eta_m \in \mathcal{D}(V)$ such that $v - \varphi$ has a $\|\cdot\|_H$-local minimum at η_m we have

$$\rho v(\eta_m) \geq \langle \eta_m, A^*\nabla\varphi(\eta_m)\rangle_H + f(\eta_m)\varphi_{\eta_0}(\eta_m) + U_2(\eta_0) + \mathcal{H}(\varphi_{\eta_0}(\eta_m)).$$

(iii) A continuous function $v : \mathcal{D}(V) \to \mathbb{R}$ is called a viscosity supersolution of (3.16) on $\mathcal{D}(V)$ if it is both a viscosity sub and supersolution.

We can prove the following:

Theorem 3.3.3. *The value function V is a viscosity solution of (3.16) on $\mathcal{D}(V)$.*

Proof. (i) We prove that V is a viscosity subsolution. Let $(\eta_M, \varphi) \in \mathcal{D}(V) \times \tau$ be such that $V - \varphi$ has a local maximum at η_M. Without loss of generality we can suppose $V(\eta_M) = \varphi(\eta_M)$. Let us suppose, by contradiction that there exists $\nu > 0$ such that

$$2\nu \leq \rho V(\eta_M) - \big(\langle \eta_M, A^*\nabla\varphi(\eta_M)\rangle_H + f(\eta_M)\varphi_{\eta_0}(\eta_M) + U_2(\eta_0) + \mathcal{H}(\varphi_{\eta_0}(\eta_M))\big).$$

Let us define the function

$$\tilde{\varphi}(\eta) := V(\eta_M) + \langle \nabla\varphi(\eta_M), \eta - \eta_M\rangle_H + \|\eta - \eta_M\|_{-1}^2.$$

We have

$$\nabla\tilde{\varphi}(\eta) = \nabla\varphi(\eta_M) + (A^*)^{-1}A^{-1}(\eta - \eta_M),$$

Thus $\tilde{\varphi}$ is a test function and we must have also

$$2\nu \leq \rho V(\eta_M) - \big(\langle \eta_M, A^*\nabla\tilde{\varphi}(\eta_M)\rangle_H + f(\eta_M)\tilde{\varphi}_{\eta_0}(\eta_M) + U_2(\eta_0) + \mathcal{H}(\tilde{\varphi}_{\eta_0}(\eta_M))\big).$$

By concavity of V we have

$$V(\eta_M) = \tilde{\varphi}(\eta_M), \qquad \tilde{\varphi}(\eta) \geq V(\eta) + \|\eta - \eta_M\|_{-1}^2, \quad \forall \eta \in \mathcal{D}(V).$$

By the continuity property of $\tilde{\varphi}$ we can find $\varepsilon > 0$ such that

$$\nu \leq \rho V(\eta) - \big(\langle \eta, A^*\nabla\tilde{\varphi}(\eta)\rangle_H + f(\eta)\tilde{\varphi}_{\eta_0}(\eta) + U_2(\eta_0) + \mathcal{H}(\tilde{\varphi}_{\eta_0}(\eta))\big),$$
$$\forall \eta \in B_\varepsilon := B_{(H,\|\cdot\|_H)}(\eta_M, \varepsilon).$$

Take a sequence $\delta_n > 0$, $\delta_n \to 0$ and, for any n, take a δ_n-optimal control $c_n(\cdot) \in \mathcal{C}_{ad}(\eta_M)$. Set $X^n(\cdot) := X(\cdot; \eta_M, c_n(\cdot))$ and define

$$t_n := \inf\{t \geq 0 \mid \|X^n(t) - \eta_M\| = \varepsilon\} \wedge 1.$$

Of course t_n is well-defined and belongs to $(0, 1]$. Moreover, by continuity of trajectories, $X^n(t) \in B_\varepsilon$, for $t \in [0, t_n)$. We distinguish two cases:

$$\limsup_n t_n = 0 \qquad \text{or} \qquad \limsup_n t_n > 0.$$

In the first case we can write

$$\begin{aligned}
\delta_n \;\geq\; & -\int_0^{t_n} e^{-\rho t} \left[U_1(c_n(t)) + U_2(X_0^n(t))\right] dt - \left(e^{-\rho t_n} V(X(t_n)) - V(\eta_M)\right) \\
\geq\; & -\int_0^{t_n} e^{-\rho t} \left[U_1(c_n(t)) + U_2(X_0^n(t))\right] dt \\
& - \left(e^{-\rho t_n}(\tilde{\varphi}(X^n(t_n))) - \tilde{\varphi}(\eta_M)\right) + e^{-\rho t_n} \|X^n(t_n) - \eta_M\|_{-1}^2 \\
=\; & -\int_0^{t_n} e^{-\rho t} \Big[U_1(c_n(t)) + U_2(X_0^n(t)) + [\mathcal{L}^{c_n(t)} \tilde{\varphi}](X^n(t))\Big] dt \\
& + e^{-\rho t_n} \|X^n(t_n) - \eta_M\|_{-1}^2 \\
\geq\; & -\int_0^{t_n} e^{-\rho t} \Big[U_2(X_0^n(t)) - \rho\tilde{\varphi}(X^n(t)) + \langle X^n(t), A^* \nabla\tilde{\varphi}(X^n(t))\rangle_H \\
& \qquad + f(X^n(t))\tilde{\varphi}_{\eta_0}(X^n(t)) + \mathcal{H}(\tilde{\varphi}_{\eta_0}(X^n(t)))\Big] dt + e^{-\rho t_n} \|X^n(t_n) - \eta_M\|_{-1}^2 \\
\geq\; & t_n \nu + e^{-\rho t_n} \|X^n(t_n) - \eta_M\|_{-1}^2,
\end{aligned}$$

thus it has to be

$$\|X^n(t_n) - \eta_M\|_{-1}^2 \to 0.$$

Let us show that this is impossible. The above convergence implies in particular that

$$|X_0^n(t_n) - (\eta_M)_0| \to 0. \tag{3.18}$$

Moreover, by definition of t_n, it has to be

$$|X_0^n(t) - (\eta_M)_0| \leq \varepsilon, \quad t \in [0, t_n]. \tag{3.19}$$

Since $t_n \to 0$, taking into account (3.19), we have also

$$\|X_1^n(t_n) - (\eta_M)_1\|_{L^2_{-T}} \to 0. \tag{3.20}$$

The convergences (3.18) and (3.20) are not compatible with the definition of t_n and the contradiction arises. In the second case we can suppose, eventually passing to a subsequence, that $t_n \to \bar{t} \in (0, 1]$. So we get as before

$$\delta_n \geq t_n \nu + e^{-\rho t_n} \|X^n(t_n) - \eta_M\|_{-1}^2 \geq t_n \nu;$$

since $\delta_n \to 0$ and $t_n \nu \to \bar{t}\nu$, again a contradiction arises.

(ii) The proof that V is a viscosity supersolution is standard. We refer to [Li, Yong; 1995]. □

3.3.2 Smoothness of viscosity solutions

In this subsection we show that the concave $\|\cdot\|_{-1}$-continuous viscosity solutions of (3.16) (so that in particular the value function V) are differentiable along the direction $\hat{n} = (1,0)$. For this purpose we need the following lemma.

Lemma 3.3.4. *Let $v : \mathcal{D}(V) \to \mathbb{R}$ a concave $\|\cdot\|_{-1}$-continuous function and suppose that $\bar{\eta} \in \mathcal{D}(V)$ is a differentiability point for v and that $\nabla v(\bar{\eta}) = \zeta$. Then*

1. *There exists a test function φ such that $v - \varphi$ has a local maximum at $\bar{\eta}$ and $\nabla\varphi(\bar{\eta}) = \zeta$.*
2. *There exists a test function φ such that $v - \varphi$ has a local minimum at $\bar{\eta}$ and $\nabla\varphi(\bar{\eta}) = \zeta$.*

Proof. Thanks to Proposition 3.2.12 and due to the concavity of v, the first statement is clearly satisfied by the function $\langle\cdot,\zeta\rangle_H$. We prove now the second statement, which is more delicate. We use the notation of Proposition 3.2.12. Thanks to the third claim of Proposition 3.2.12, we have $A^*\zeta \in D^+u(\bar{\eta}')$. This means that

$$u(\eta') - u(\bar{\eta}') - \langle \eta' - \bar{\eta}', A^*\zeta\rangle_H \geq -\|\eta' - \bar{\eta}'\|_H \cdot \varepsilon(\|\eta' - \bar{\eta}'\|_H),$$

where $\varepsilon : [0,+\infty) \to [0,+\infty)$ is an increasing function such that

$$\|\eta' - \bar{\eta}'\|_H \to 0 \Longrightarrow \varepsilon(\|\eta' - \bar{\eta}'\|_H) \to 0.$$

The previous inequality can be rewritten also as

$$u(\eta') - u(\bar{\eta}') - \langle A(\eta' - \bar{\eta}'), \zeta\rangle_H \geq -\|\eta' - \bar{\eta}'\|_H \cdot \varepsilon(\|\eta' - \bar{\eta}'\|_H).$$

Passing to v this reads as

$$v(\eta) - v(\bar{\eta}) - \langle \eta - \bar{\eta}, \zeta\rangle_H \geq -\|\eta - \bar{\eta}\|_{-1} \cdot \varepsilon\left(\|\eta - \bar{\eta}\|_{-1}\right),$$

where $\varepsilon\left(\|\eta - \bar{\eta}\|_{-1}\right) \to 0$, when $\|\eta - \bar{\eta}\|_{-1} \to 0$. We look for a test function of this form:

$$\varphi(\eta) = v(\bar{\eta}) + \langle \eta - \bar{\eta}, \zeta\rangle_H + g\left(\|\eta - \bar{\eta}\|_{-1}\right),$$

where $g : [0,+\infty) \to [0,+\infty)$ is a suitable increasing C^1 function such that $g(0) = 0$. Notice that $\varphi(\bar{\eta}) = v(\bar{\eta})$, so that, in order to prove that $v - \varphi$ has a local minimum at $\bar{\eta}$, we have to prove that $\varphi \leq v$ in a neighborhood of $\bar{\eta}$.

Let us define the function

$$g(r) := \int_0^{2r} \varepsilon(s)ds \geq \int_r^{2r} \varepsilon(s)ds \geq r\varepsilon(r).$$

Then

$$\begin{aligned}\varphi(\eta) &= v(\bar{\eta}) + \langle \eta - \bar{\eta}, \zeta \rangle_H - g(\|\eta - \bar{\eta}\|_{-1}) \\ &\le v(\bar{\eta}) + \langle \eta - \bar{\eta}, \zeta \rangle_H - \|\eta - \bar{\eta}\|_{-1} \cdot \varepsilon(\|\eta - \bar{\eta}\|_{-1}) \\ &\le v(\eta).\end{aligned}$$

Moreover

$$\nabla \varphi(\eta) = \begin{cases} \zeta - (A^*)^{-1} \dfrac{\varepsilon(2\|\eta - \bar{\eta}\|_{-1})}{\|\eta - \bar{\eta}\|_{-1}} A^{-1}(\eta - \bar{\eta}), & \text{if } \eta \neq \bar{\eta}, \\ \zeta, & \text{if } \eta = \bar{\eta}, \end{cases}$$

i.e. φ is a test function and $\nabla \varphi(\bar{\eta}) = \zeta$. □

Now we can state and prove the main result.

Theorem 3.3.5. *Let v be a concave $\|\cdot\|_{-1}$-continuous viscosity solution of (3.16) on $\mathcal{D}(V)$. Then v is differentiable along the direction $\hat{n} = (1, 0)$ at any point $\eta \in \mathcal{D}(V)$ and the function $\eta \mapsto v_{\eta_0}(\eta)$ is continuous on $\mathcal{D}(V)$.*

Proof. Let $\eta \in \mathcal{D}(V)$ and $\zeta, \xi \in D^* v(\bar{\eta})$. Thanks to Proposition 3.2.12, there exist sequences (η_n), $(\tilde{\eta}_n)$ such that:

- $\eta_n \to \eta$, $\tilde{\eta}_n \to \eta$;
- $\nabla v(\eta_n)$ and $\nabla v(\tilde{\eta}_n)$ exist for all $n \in \mathbb{N}$;
- $A^* \nabla v(\eta_n) \rightharpoonup A^* \zeta$ and $A^* \nabla v(\tilde{\eta}_n) \rightharpoonup A^* \xi$.

Thanks to Lemma 3.3.4 we can write, for any $n \in \mathbb{N}$,

$$\rho v(\eta_n) = \langle \eta_n, A^* \nabla v(\eta_n) \rangle_H + f(\eta_n) v_{\eta_0}(\eta_n) + U_2(\eta_0^n) + \mathcal{H}(v_{\eta_0}(\eta_n)),$$

$$\rho v(\tilde{\eta}_n) = \langle \tilde{\eta}_n, A^* \nabla v(\tilde{\eta}_n) \rangle_H + f(\tilde{\eta}_n) v_{\eta_0}(\tilde{\eta}_n) + U_2(\eta_0^n) + \mathcal{H}(v_{\eta_0}(\tilde{\eta}_n)).$$

Passing to the limit we get

$$\langle \eta, A^* \zeta \rangle_H + f(\eta)\zeta_0 + U_2(\eta_0) + \mathcal{H}(\zeta_0) = \rho v(\eta) = \langle \eta, A^* \xi \rangle_H + f(\eta)\xi_0 + U_2(\eta_0) + \mathcal{H}(\xi_0). \tag{3.21}$$

On the other hand $\lambda\zeta + (1-\lambda)\xi \in D^+ v(\bar{\eta})$, for any $\lambda \in (0,1)$, so that we have the subsolution inequality

$$\begin{aligned}\rho v(\eta) \le \langle \eta, A^*[\lambda\zeta + (1-\lambda)\xi] \rangle_H + f(\eta)[\lambda\zeta_0 + (1-\lambda)\xi_0] + U_2(\eta_0) \\ + \mathcal{H}(\lambda\zeta_0 + (1-\lambda)\xi_0), \quad \forall \lambda \in (0,1).\end{aligned} \tag{3.22}$$

Combining (3.21) and (3.22) we get

$$\mathcal{H}(\lambda\zeta_0 + (1-\lambda)\xi_0) \geq \lambda\mathcal{H}(\zeta_0) + (1-\lambda)\mathcal{H}(\xi_0);$$

since $\mathcal{H}$ is strictly convex, the previous inequality implies $\zeta_0 = \xi_0$. This means that the projection of $D^*v(\eta)$ onto $\hat{n}$ is a singleton. Thanks to (3.15) this implies also that the projection of $D^+v(\eta)$ onto $\hat{n}$ is a singleton and therefore that v is differentiable in the direction $\hat{n}$ at η.

We prove now that the map $\eta \mapsto v_{\eta_0}(\eta)$ is continuous on $\mathcal{D}(V)$. To this aim take $\eta \in \mathcal{D}(V)$ and let (η^n) be a sequence such that $\eta^n \to \eta$. We have to show that $v_{\eta_0}(\eta^n) \to v_{\eta_0}(\eta)$. Of course for any $n \in \mathbb{N}$ there exists $p_1^n \in L^2_{-T}$ such that $(v_{\eta_0}(\eta^n), p_1^n) \in D^+v(\eta^n)$. Since v is concave, it is also locally Lipschitz continuous so that the super-differential is locally bounded. Therefore, from any subsequence $(v_{\eta_0}(\eta^{n_k}))$, we can extract a sub-subsequence $(v_{\eta_0}(\eta^{n_{k_h}}))$ such that $\left(v_{\eta_0}(\eta^{n_{k_h}}), p_1^{n_{k_h}}\right)$ is weakly convergent towards some limit point. Due to the concavity of v this limit point must live in the set $D^+v(\eta)$. In particular the limit point of $(v_{\eta_0}(\eta^{n_{k_h}}))$ must coincide with $v_{\eta_0}(\eta)$. This holds true for any subsequence $(v_{\eta_0}(\eta^{n_k}))$, so that the claim follows by the usual argument on subsequences. □

Remark 3.3.6. Notice that in the assumptions of Theorem 3.3.5 we do not require that v is the value function, but only that it is a concave $\|\cdot\|_{-1}$-continuous viscosity solution of (3.16). □

3.3.3 Verification theorem and optimal feedback strategies

Thanks to the regularity result of the previous subsection we can define, at least formally, the "candidate" optimal feedback map on $\mathcal{D}(V)$, which is given by

$$C(\eta) := \operatorname{argmax}_{c\geq 0}\left(U_1(c) - cV_{\eta_0}(\eta)\right), \quad \eta \in \mathcal{D}(V). \tag{3.23}$$

Note that this map is well-defined since V is concave and, by Proposition 3.1.14, strictly increasing, so that we have $V_{\eta_0}(\eta) \in (0, +\infty)$ for all $\eta \in \mathcal{D}(V)$. Existence and uniqueness of the argmax follow from the assumptions on U_1. Moreover, since V_{η_0} is continuous on $\mathcal{D}(V)$, also C is continuous on $\mathcal{D}(V)$. The closed-loop delay state equation associated with this map is, for $\eta \in \mathcal{D}(V)$,

$$\begin{cases} x'(t) = rx(t) + f_0\left(x(t), \int_{-T}^0 a(\xi)x(t+\xi)d\xi\right) - C\left((x(t), x(t+\xi)|_{\xi\in[-T,0]})\right), \\ x(0) = \eta_0, \ x(s) = \eta_1(s), \ s \in [-T, 0). \end{cases} \tag{3.24}$$

Now we want to prove a Verification Theorem: if the closed loop equation (3.24) has a strictly positive solution $x^*(\cdot)$ (so that $\left(x^*(t), x^*(t+\xi)|_{\xi\in[-T,0]}\right) \in$

$\mathcal{D}(V)$ for all $t \geq 0$ and the term $C\left((x^*(t), x^*(t+\xi)|_{\xi\in[-T,0]})\right)$ is well-defined for every $t \geq 0$), then the feedback strategy

$$c^*(t) := C\left((x^*(t), x^*(t+\xi)|_{\xi\in[-T,0]})\right) \tag{3.25}$$

is optimal. Notice that, by definition of $c^*(\cdot)$, if $x^*(\cdot)$ is a strictly positive solution of (3.24), then $c^*(\cdot)$ is admissible and, setting $X^*(t) := X(t;\eta, c^*(\cdot))$, we have

$$X^*(t) = \left(x^*(t), x^*(t+\xi)|_{\xi\in[-T,0]}\right) \in \mathcal{D}(V), \quad \forall t \geq 0.$$

The proof of the Verification Theorem, in the classical case, is done by computing the derivative

$$t \mapsto \frac{d}{dt}\left[e^{-\rho t}V(X^*(t))\right]. \tag{3.26}$$

and then using the HJB equation and integrating the resulting equality.

We want to get exactly the classical statement but we cannot proceed with the classical proof since we cannot compute the derivative (3.26). So we proceed using the fact that V is a viscosity solution (as e.g. in [Yong, Zhou; 1999], Theorem 3.9, Chapter 5, and in [Li, Yong; 1995], Theorems 5.4, 5.5, Chapter 6). But two main difficulties arise (strongly connected each other):

- The function
$$t \mapsto e^{-\rho t}V(X^*(t)) \tag{3.27}$$
is not Lipschitz continuous so, it may not have a.e. derivative. Indeed we do not require the initial datum η belonging to $\mathcal{D}(A)$ and the operator A works as a shift operator on the infinite-dimensional component so we do not have the condition $X^*(t) \in \mathcal{D}(A)$ for almost every $t \geq 0$ that would give the required Lipschitz regularity for the function (3.27): only continuity is ensured. Without this Lipschitz regularity we cannot apply the Fundamental Theorem of Calculus as done e.g. in [Li, Yong; 1995], Theorems 5.4, 5.5, Chapter 6.

- Consequently we have to deal with the concept of Dini derivatives of the function (3.27) and, since we want to integrate them, we need something like a Fundamental Theorem of Calculus in inequality form relating the function and the integral of its Dini derivative. Such a result in the context of stochastic verification theorems for viscosity solutions is given in [Yong, Zhou; 1999], Lemma 5.2, Chapter 5. Unfortunately we discovered that such result is not true as it is stated (we give a counterexample in Remark 3.3.10), so we have to use a more refined result, the so called Saks Theorem, that needs stronger assumptions and that is based on the theory of Dini derivatives.

Recall first that, if g is a continuous function on some interval $[\alpha, \beta] \subset \mathbb{R}$, the right Dini derivatives of g are defined by

$$D^+g(t) = \limsup_{h \downarrow 0} \frac{g(t+h) - g(t)}{h}, \ D_+g(t) = \liminf_{h \downarrow 0} \frac{g(t+h) - g(t)}{h}, \ \ t \in [\alpha, \beta),$$

and the left Dini derivatives by

$$D^-g(t) = \limsup_{h \uparrow 0} \frac{g(t+h) - g(t)}{h}, \ D_-g(t) = \liminf_{h \uparrow 0} \frac{g(t+h) - g(t)}{h}, \ \ t \in (\alpha, \beta].$$

Proposition 3.3.7. *If g is a continuous real function on $[\alpha, \beta]$, then the bounds of each Dini's derivative are equal to the bounds of the set of the difference quotients*

$$\left\{ \frac{g(t) - g(s)}{t - s} \;\middle|\; t, s \in [\alpha, \beta] \right\}.$$

□

Proof. See [Bruckner; 1978], Theorem 1.2, Chapter 4. □

An immediate consequence of Proposition 3.3.7 above is the following.

Proposition 3.3.8 (Monotonicity result)**.** *Let $g \in C([\alpha, \beta]; \mathbb{R})$ be such that*

$$D^+g(t) \geq 0, \quad \forall t \in [\alpha, \beta).$$

Then g is nondecreasing on $[\alpha, \beta]$. □

The following Lemma is a special case of the Saks Theorem (see, e.g., Chapter VI, Theorem 7.3, of [Saks; 1964]). We give the proof in a special case using the Monotonicity result above.

Lemma 3.3.9. *Let $g \in C([0, +\infty); \mathbb{R})$. Suppose that there exists $\mu \in L^1([0, +\infty); \mathbb{R})$ such that*

$$D_-g(t) \geq \mu(t), \quad for \ a.e. \ t \in (0, +\infty). \tag{3.28}$$

and that

$$D_-g(t) > -\infty \quad \forall t \in (0, +\infty) \tag{3.29}$$

except at most for those of a countable set. Then, for every $0 \leq \alpha \leq \beta < +\infty$,

$$g(\beta) - g(\alpha) \geq \int_\alpha^\beta \mu(t) dt. \tag{3.30}$$

Proof. We give the proof in the special case when μ is continuous and (3.28) holds *for every* $t \in (0,+\infty)$[1]. Since $D_- g(t) \geq \mu(t)$ for every $t \in (0,+\infty)$, we have

$$D_- \left[g(t) - \int_0^t \mu(s)ds \right] \geq 0, \;\; \forall t \in (0,+\infty).$$

Thanks to Proposition 3.3.7 we have also

$$D^+ \left[g(t) - \int_0^t \mu'(s)ds \right] \geq 0, \quad \forall t \in [0,+\infty).$$

Therefore, due to Proposition 3.3.8, the function

$$t \mapsto g(t) - \int_0^t \mu(s)ds$$

is nondecreasing, getting the claim. □

Remark 3.3.10. We give some remarks on Lemma 3.3.9.

- If μ is continuous and (3.28) holds for all $t > 0$ (as we assume in the proof of Lemma 3.3.9), then (3.29) is verified.

- In general we cannot avoid to assume (3.29): without it, then (3.30) is no longer true. For example, if $g = -f$ on $[0,1]$, where f is the Cantor function and $\mu \equiv 0$, we have

  $$\mu(t) = 0 = g'(t) = D_- g(t) \quad \text{for a.e. } t \in (0,1].$$

  Therefore, taking $\alpha = 0, \; \beta = 1$, the left handside of (3.30) is -1, while the right handside is 0. Indeed in this case $D_- g = -\infty$ on the Cantor set. So Lemma 5.2, Chapter 5, of [Yong, Zhou; 1999] is not correct. Indeed the condition required therein is not sufficient to apply Fatou's Lemma: it is assumed that only the limsup of difference quotients is estimated from above by an integrable function while also all difference quotients should be estimated from above by the same integrable function (and this is not true in the case of our counterexample). Therefore, one could substitute the assumption (3.29) with the following: there exists $\mu \in L^1([0,+\infty);\mathbb{R})$ such that for some $h_0 > 0$ and for almost every $t > 0$ we have

  $$\frac{g(t+h) - g(t)}{h} \geq \mu(t), \;\; -h_0 \leq h < 0.$$

□

[1] In the use of this Lemma that we will do in the next Verification Theorem 3.3.11, actually these conditions are satisfied.

Theorem 3.3.11 (Verification)**.** *Let $\eta \in H_+$ and let $x^*(\cdot)$ be a solution of* (3.24) *such that $x^*(\cdot) > 0$; let $c^*(\cdot)$ be the strategy defined by* (3.25)*. Then $c^*(\cdot)$ is admissible and optimal for the problem.*

Proof. As said above the fact that $c^*(\cdot)$ is admissible is a direct consequence of the assumption $x^*(\cdot) > 0$ and of the definition of $c^*(\cdot)$.

Set $X^*(\cdot) := X(\cdot\,;\eta, c^*(\cdot))$ and let $s > 0$. Let $p_1(s) \in L^2_{-T}$ be such that

$$(V_{\eta_0}(X^*(s)), p_1(s)) \in D^+V(X^*(s))$$

and let

$$\varphi(\zeta) := V(X^*(s)) + \langle (V_{\eta_0}(X^*(s)), p_1(s))\,, \zeta - X^*(s)\rangle_H, \quad \zeta \in H,$$

so that

$$\varphi(X^*(s)) = V(X^*(s)), \quad \varphi(\zeta) \geq V(\zeta),\ \ \zeta \in H.$$

From Proposition 3.2.12 we know that $\varphi \in \tau$, so that

$$\begin{aligned}
\liminf_{h\uparrow 0} & \frac{e^{-\rho(s+h)}V(X^*(s+h)) - e^{-\rho s}V(X^*(s))}{h} \\
&\geq \liminf_{h\uparrow 0} \frac{e^{-\rho(s+h)}\varphi(X^*(s+h)) - e^{-\rho s}\varphi(X^*(s))}{h} \\
= e^{-\rho s}\left[\mathcal{L}^{c^*(s)}\varphi\right](X^*(s)) &= e^{-\rho s}\Big[-\rho V(X^*(s)) + \langle X(s), A^*\left(V_{\eta_0}(X^*(s)), p_1(s)\right)\rangle_H \\
&\quad + f(X^*(s))V_{\eta_0}(X^*(s)) - c^*(s)V_{\eta_0}(X^*(s))\Big].
\end{aligned}$$

By definition of $c^*(\cdot)$ we get

$$\begin{aligned}
\liminf_{h\uparrow 0} & \frac{e^{-\rho(s+h)}V(X^*(s+h)) - e^{-\rho s}V(X^*(s))}{h} + e^{-\rho s}[U_1(c^*(s)) + U_2(X_0^*(s))] \\
&\geq e^{-\rho s}\Big[-\rho V(X^*(s)) + \langle X^*(s), A^*\left(V_{\eta_0}(X^*(s)), p_1(s)\right)\rangle_H \\
&\quad + f(X^*(s))V_{\eta_0}(X^*(s)) + \mathcal{H}(X^*(s)) + U_2(X_0^*(s))\Big].
\end{aligned}$$

Due to the subsolution property of V we get

$$\liminf_{h\uparrow 0} \frac{e^{-\rho(s+h)}V(X^*(s+h)) - e^{-\rho s}V(X^*(s))}{h} + e^{-\rho s}[U_1(c^*(s)) + U_2(X_0^*(s))] \geq 0.$$

The function $s \mapsto e^{-\rho s}V(X^*(s))$ and the function $s \mapsto e^{-\rho s}[U_1(c^*(s)) + U_2(X_0^*(s))]$ are continuous; therefore we can apply Lemma 3.3.9 on $[0, M]$, $M > 0$, getting

$$e^{-\rho M}V(X^*(M)) + \int_0^M e^{-\rho s}[U_1(c^*(s)) + U_2(X_0^*(s))]ds \geq V(\eta).$$

Since V, U_1, U_2 are bounded from above, taking the limsup for $M \to +\infty$ we get by Fatou's Lemma

$$\int_0^{+\infty} e^{-\rho s}[U_1(c^*(s)) + U_2(X_0^*(s))]ds \geq V(\eta),$$

which gives the claim. □

Remark 3.3.12. We have given in Theorem 3.3.11 a sufficient condition of optimality: indeed, we have proved that if the feedback map defines an admissible strategy then such a strategy is optimal. Of course, a natural question arising is whether, at least with a special choice of data, such a condition is also necessary for the optimality, i.e. if, given an optimal strategy, it can be written as feedback of the associated optimal state. From the viscosity point of view the answer to this question relies in requiring that the value function is a *bilateral* viscosity subsolution of (3.16) along the optimal state trajectories, i.e. requiring that the value function satisfies the property of Definition 3.3.2-(i) also with the reverted inequality along such trajectories.

Such a property of the value function is related to the so-called *backward dynamic programming principle* which is, in turn, related to the backward study of the state equation (see [Bardi, Capuzzo-Dolcetta; 1997], Chapter III, Section 2.3). Differently from the finite-dimensional case, this topic is not standard in infinite-dimension unless the operator A is the generator of a strongly continuous *group*, which is not our case.

However, in our case we can use the delay original setting of the state equation to approach this topic. Then the problem reduces to find, at least for sufficient regular data, a *backward continuation* of the solution. This problem is faced, e.g., in [Hale, Verduyn; 1993], Chapter 2, Section 5. Unfortunately our equation does not fit the main assumption required therein, which in our setting basically corresponds to require that the function $a(\cdot)$, seen as measure, has an atom at $-T$. Investigation on this is left for future research. □

Up to now we did not make any further assumption on the functions a and U_2 beyond Hypotheses 3.1.2 and 3.1.5; in particular it could be $U_2 \equiv 0$. However without any further assumption we have no information on the behaviour of V_{η_0} when we approach the boundary of $\mathcal{D}(V)$ and therefore we are not able to say anything about the existence of solutions of the closed loop equation and whether they satisfy or not the state constraint. So basically we cannot say whether the hypothesis of Theorem 3.3.11 is satisfied or not. In order to give sufficient conditions for that, we need to do some further assumptions.

Hypothesis 3.3.13. We will make use of the following assumptions

$$(i)\ U_2 \text{ is not integrable at } 0^+, \quad (ii)\ \int_{-\varepsilon}^{0} a(\xi)d\xi > 0, \quad \forall \varepsilon > 0. \tag{3.31}$$

□

Lemma 3.3.14.

1. *The following holds*
$$\partial_{\|\cdot\|_H}\mathcal{D}(V) = \partial_{\|\cdot\|_{-1}}\mathcal{D}(V).$$
Thanks to the previous equality we write without ambiguity $\partial\mathcal{D}(V)$ for denoting the boundary of $\mathcal{D}(V)$ referred to $\|\cdot\|_H$ or $\|\cdot\|_{-1}$ indifferentely.

2. *Suppose that* (3.31)-*(i) holds; then*
$$\lim_{\eta\to\bar\eta} V_{\eta_0}(\eta) = +\infty, \quad \forall \bar\eta \in \partial\mathcal{D}(V),$$
where the limit is taken with respect to $\|\cdot\|_H$.

Proof. We work with the original one-dimensional state equation with delay.

1. First of all note that, thanks to Proposition 3.1.14 and Proposition 3.2.8, the set $\mathcal{D}(V)$ has the following structure

$$\mathcal{D}(V) = \bigcup_{\eta_1\in L^2_{-T}} \left((\eta_0^{\eta_1}, +\infty)\times\{\eta_1\}\right), \tag{3.32}$$

where, for $\eta_1 \in L^2_{-T}$, we set $\eta_0^{\eta_1} = \inf\{\eta_0 > 0 \mid (\eta_0, \eta_1(\cdot)) \in \mathcal{D}(V)\}$. For any $\eta \in H$ set $x^\eta(\cdot) := x(\cdot;\eta,0)$ and consider the function $g : H \to \mathbb{R}$ defined by

$$g(\eta_0, \eta_1(\cdot)) := \inf_{t\in[0,T]} x^\eta(t).$$

Thanks to Lemma 3.2.7 this function is continuous (with respect to both the norms $\|\cdot\|_H$ and $\|\cdot\|_{-1}$), so we have the following representation of $\mathcal{D}(V)$ in terms of g:

$$\mathcal{D}(V) = \{g > 0\}.$$

Lemma 3.1.8 shows that g is increasing with respect to the first variable. Actually g is strictly increasing with respect to the first variable. Let us show this fact. Let $\eta_1 \in L^2_{-T}$ and take $\eta_0, \bar\eta_0 \in \mathbb{R}$ such that $\eta_0 > \bar\eta_0$. Define $y(\cdot) := x(\cdot;(\eta_0,\eta_1(\cdot)),0)$, $x(\cdot) := x(\cdot;(\bar\eta_0,\eta_1(\cdot)),0)$ and let $z(\cdot), \bar z(\cdot)$ be respectively the solutions on $[0,T]$ of the differential problems *without delay*

$$\begin{cases} z'(t) = rz(t) + f_0\left(z(t), \int_{-T}^{0} a(\xi)x(t+\xi)d\xi\right), \\ z(0) = \eta_0, \end{cases}$$

$$\begin{cases} \bar{z}'(t) = r\bar{z}(t) + f_0\left(\bar{z}(t), \int_{-T}^{0} a(\xi)x(t+\xi)d\xi\right), \\ \bar{z}(0) = \bar{\eta}_0, \end{cases}$$

Then we have, on the interval $[0,T]$, $\bar{z}(\cdot) \equiv x(\cdot)$ and, by comparison criterion, $y(\cdot) \geq z(\cdot)$; moreover we can apply the classic Cauchy-Lipschitz Theorem for ODEs getting uniqueness for the solutions of the above ODEs, which yields $z(\cdot) > \bar{z}(\cdot)$ on $[0,T]$. Thus $y(\cdot) > x(\cdot)$ on $[0,T]$, proving that g is strictly increasing with rescpect to the first variable.

The continuity (with respect to both the norms) of g, (3.32) and the fact that g is strictly increasing with respect to the first variable lead to have

$$\partial_{\|\cdot\|_H}\mathcal{D}(V) = \partial_{\|\cdot\|_{-1}}\mathcal{D}(V) = \{g=0\} = \bigcup_{\eta_1 \in L^2_{-T}} \left(\{\eta_0^{\eta_1}\} \times \{\eta_1\}\right).$$

2. First we prove that

$$\lim_{\eta\to\bar{\eta}} V(\eta) = -\infty, \quad \forall \bar{\eta} \in \partial\mathcal{D}(V).$$

Let $\bar{\eta} \in \partial\mathcal{D}(V)$ and let $(\eta^n) \subset \mathcal{D}(V)$ be a sequence such that $\eta^n \to \bar{\eta}$. We can suppose without loss of generality that $(\eta^n) \subset B(\bar{\eta}, 1)$. Set

$$x^n(\cdot) := x(\cdot;\eta^n,0), \quad p^n := \sup_{\xi\in[0,2T]} x^n(\xi).$$

Thanks to Lemma 3.2.7 there exists $K>0$ such that $p^n \leq K$ for any $n \in \mathbb{N}$. So, since $f_0(x,y) \leq C_0(1+|x|+|y|)$ for some $C_0 > 0$, we have for the dynamics of $x^n(\cdot)$ in the interval $[0,2T]$

$$\frac{d}{dt}\, x^n(t) \leq rx^n(t) + R,$$

where

$$R := C_0\left(1 + K + \|a\|_{L^2_{-T}}(\|\bar{\eta}_1\|_{L^2_{-T}}+1) + \|a\|_{L^2_{-T}}T^{1/2}K\right).$$

Therefore there exists $C>0$ such that, for any $s \in [0,T)$, $n \in \mathbb{N}$,

$$x^n(t) \leq x^n(s)\, e^{r(t-s)} + \frac{R}{r}\,(e^{r(t-s)}-1) \leq x^n(s)(1+C(t-s)) + C(t-s), \quad t \in [s,2T]. \tag{3.33}$$

By continuity of g we have $\lim_{n\to\infty} g(\eta_0^n, \eta_1^n(\cdot)) = 0$. Thus for any $\varepsilon > 0$ we can find $n_0 \in \mathbb{N}$ such that, for $n \geq n_0$, there exists $s_n \in [0,T)$ such that

$$x^n(s_n) \leq \varepsilon. \tag{3.34}$$

We want to show that

$$\int_0^{+\infty} e^{-\rho t} U_2(x^n(t))dt \longrightarrow -\infty, \quad n \to \infty. \tag{3.35}$$

For this purpose, since U_2 is bounded from above, it is clear that we can assume without loss of generality $U_2(\cdot) \leq 0$. We have for $n \geq n_0$, taking into account (3.33) and (3.34),

$$\begin{aligned}
\int_0^{+\infty} e^{-\rho t} U_2(x^n(t))dt &\leq e^{-2\rho T} \int_{s_n}^{2T} U_2\big(x^n(s_n)(1 + C(t - s_n)) + C(t - s_n)\big)dt \\
&\leq e^{-2\rho T} \int_{s_n}^{2T} U_2(\varepsilon(1 + C(t - s_n)) + C(t - s_n))dt \\
&\leq \frac{e^{-2\rho T}}{C(\varepsilon + 1)} \int_{\varepsilon}^{CT} U_2(x)dt.
\end{aligned}$$

Therefore, by the arbitrariness of ε and since U_2 is not integrable at 0^+, we get (3.35). This is enough to conclude that $J(\eta^n; 0) \to -\infty$, as $n \to \infty$. Of course we have $x^n(\cdot) \geq x(\cdot; \eta^n, c(\cdot))$ for any $c(\cdot) \in \mathcal{C}(\eta^n)$. Since U_1 is bounded from above this is enough to say that also $V(\eta^n) \to -\infty$, as $n \to \infty$.

Now we prove the claim. Let $\bar{\eta} \in \partial\mathcal{D}(V)$ and $(\eta^n) \subset \mathcal{D}(V)$ be such that $\eta^n \to \bar{\eta}$, suppose without loss of generality $(\eta_n) \subset B(\bar{\eta}, 1)$ and set $x^n(\cdot) := x(\cdot\,; (\eta_0^n + 1, \eta_1^n), 0) > 0$. Since f_0 is Lipschitz continuous and nondecreasing on the second variable, there exists $C > 0$ such that

$$f_0\left(x(t), \int_{-T}^0 a(\xi)x(t+\xi)d\xi\right) \geq -C\left(1 + x(t) + \|a\|_{L^2_{-T}}(\|\bar{\eta}_1\|_{L^2_{-T}} + 1)\right) =: -\tilde{R}.$$

Suppose $\tilde{R} \leq 0$. Then $\frac{d}{dt}x^n(t) \geq rx^n(t)$, so that, since $\eta_0^n > 0$, we have $x^n(t) \geq \eta_0^n + 1 \geq 1$. This leads to the estimate

$$V(\eta_0^n + 1, \eta_1^n(\cdot)) \geq K, \;\; n \in \mathbb{N},$$

for some $K > 0$. Due to the concavity of V we have the estimate

$$V_{\eta_0}(\eta^n) \geq V(\eta_0^n + 1, \eta_1^n(\cdot)) - V(\eta_0^n, \eta_1^n(\cdot)) \geq K - V(\eta_0^n, \eta_1^n(\cdot)) \to +\infty,$$

i.e. the claim.
Suppose then $\tilde{R} > 0$ and set $x^n(\cdot) := x(\cdot\,; (\eta_0^n + \tilde{R}/r, \eta_1^n), 0)$. Then

$$\frac{d}{dt}x^n(t) \geq rx^n(t) - \tilde{R},$$

so that, since $\eta_0^n > 0$, we have $x^n(t) \geq \tilde{R}/r > 0$. This leads to the estimate

$$V(\eta_0^n + \tilde{R}/r, \eta_1^n(\cdot)) \geq K, \;\; n \in \mathbb{N},$$

for some $K > 0$. Due to concavity of V we have the estimate

$$\begin{aligned}
V_{\eta_0}(\eta^n) &\geq \frac{r}{\tilde{R}}\left[V(\eta_0^n + \tilde{R}/r, \eta_1^n(\cdot)) - V(\eta_0^n, \eta_1^n(\cdot))\right] \\
&\geq \frac{r}{\tilde{R}}\left[K - V(\eta_0^n, \eta_1^n(\cdot))\right] \to +\infty,
\end{aligned}$$

i.e. the claim. □

Proposition 3.3.15. *Let* (3.31) *hold, let* $\eta \in H_{++}$ *and consider the closed-loop delay state equation* (3.24). *Then this equation admits a solution* $x^*(\cdot) \in C^1([0,+\infty);\mathbb{R})$. *Moreover, for all* $t \geq 0$,

$$x^*(t) > 0, \quad (x^*(t), x^*(t+\xi)|_{\xi\in[-T,0]}) \in \mathcal{D}(V).$$

In particular the feedback strategy defined in (3.25) *is admissible.*

Proof. Thanks to Lemma 3.3.14, if U_2 is not integrable at 0^+ we can extend the map C to a continuous map defined on the whole space $(H, \|\cdot\|_H)$ defining $C \equiv 0$ on $\mathcal{D}(V)^c$. We set

$$G(\eta) := r\eta_0 + f(\eta) - C(\eta), \quad \eta \in H,$$

and note that G is continuous.

Local existence. Let $\bar{\eta} \in H$ the initial datum for the equation. We have to show the local existence of a solution of

$$\begin{cases} x'(t) = G\left((x(t), x(t+\xi)|_{\xi\in[-T,0]})\right), \\ x(0) = \bar{\eta}_0, \; x(s) = \bar{\eta}_1(s), \; s \in [-T, 0), \end{cases}$$

Since G is continuous, there exists $b > 0$ such that

$$m := \sup_{\|\eta-\bar{\eta}\|^2 \leq b} |G(\eta)| < +\infty.$$

By continuity of translations in $L^2(\mathbb{R};\mathbb{R})$ we can find $a \in [0,T]$ such that

$$\int_{-T}^{-t} |\bar{\eta}_1(t+\xi) - \bar{\eta}_1(\xi)|^2 d\xi \leq b/4, \quad \forall t \in [0,a];$$

moreover, without loss of generality, we can suppose that

$$\int_{-a}^{0} |\bar{\eta}_1(\xi)|^2 d\xi \leq b/16.$$

Set

$$\alpha := \min\left\{a, \frac{b}{2m}, \frac{b}{16}\left(b + 2|\bar{\eta}_0|^2\right)^{-1}\right\}.$$

Define

$$M := \left\{x(\cdot) \in C([0,\alpha];\mathbb{R}) \;\middle|\; |x(\cdot) - \bar{\eta}_0|^2 \leq b/2\right\};$$

M is a convex closed subset of the Banach space $C([0,\alpha];\mathbb{R})$ endowed with the sup-norm. Define

$$x(t+\xi) := \bar{\eta}_1(t+\xi), \quad \text{if } t+\xi \leq 0,$$

and observe that, for $t \in [0,\alpha]$, $x(\cdot) \in M$,

$$\begin{aligned}\int_{-t}^{0} |x(t+\xi) - \bar\eta_1(\xi)|^2 d\xi &\leq \int_{-t}^{0} \left(2|x(t+\xi)|^2 + 2|\bar\eta_1(\xi)|^2\right) d\xi \\ &\leq 2\left[\int_{-t}^{0}\left(2\left(|x(t+\xi)-\bar\eta_0|\right)^2 + 2|\bar\eta_0|^2\right) d\xi + \int_{-t}^{0} |\bar\eta_1(\xi)|^2 d\xi\right] \\ &\leq 2\left[2t\left(\frac{b}{2} + |\bar\eta_0|^2\right) + \frac{b}{16}\right] \leq b/4.\end{aligned}$$

So, for $t \in [0,\alpha]$, $x(\cdot) \in M$, we have

$$\begin{aligned}\left\|\left(x(t), x(t+\xi)|_{\xi\in[-T,0]}\right) - \bar\eta\right\|_H^2 &\leq |x(t)-\bar\eta_0|^2 + \int_{-t}^{0} |x(t+\xi)-\bar\eta_1(\xi)|^2 d\xi \\ &\quad + \int_{-T}^{-t} |\bar\eta_1(t+\xi) - \eta_1(\xi)|^2 d\xi \\ &\leq b/2 + b/4 + b/4 = b.\end{aligned}$$

Define, for $t \in [0,\alpha]$, $x(\cdot) \in M$,

$$[\mathcal{J}x](t) := \bar\eta_0 + \int_0^t G\left(x(s), x(s+\xi)|_{\xi\in[-T,0]}\right) ds, \quad t \in [0,\alpha].$$

We have

$$\begin{aligned}\left|\,[\mathcal{J}x](t) - \eta_0\right| &\leq \int_0^t \left|\, G\left(x(s), x(s+\xi)|_{\xi\in[-T,0]}\right)\right| ds \\ &\leq tm \leq b/2.\end{aligned}$$

Therefore we have proved that $\mathcal{J}$ maps the closed and convex set M in itself. We want to prove that $\mathcal{J}$ admits a fixed point, i.e., by definition of $\mathcal{J}$, the solution we are looking for. By Schauder's Theorem it is enough to prove that $\mathcal{J}$ is completely continuous, i.e. that $\overline{\mathcal{J}(M)}$ is compact. For any $x(\cdot) \in M$, we have the estimate

$$\left|\,[\mathcal{J}x](t) - [\mathcal{J}x](\bar t)\right| \leq \int_{t\wedge\bar t}^{t\vee\bar t} \left|\, G\left(x(s), x(s+\xi)|_{\xi\in[-T,0]}\right)\right| ds \leq m|t-\bar t|, \quad t,\bar t \in [0,\alpha].$$

Therefore $\mathcal{J}(M)$ is a uniformly bounded and equicontinuous family in the space $C([0,\alpha];\mathbb{R})$. Thus, by Ascoli-ArzelàTheorem, $\overline{\mathcal{J}(M)}$ is compact.

Global existence. Let $\eta \in H_{++}$ and let $x^*(\cdot)$ be the solution of equation (3.24) defined on an interval $[0,\beta)$, $\beta > 0$. Note that, by continuity of f_0, C, we have $x^*(\cdot) \in C^1([0,\beta);\mathbb{R})$.
Since $C(\cdot) \geq 0$, we have $x^*(\cdot) \leq x(\cdot;\eta,0)$; therefore $x^*(\cdot)$ is dominated from above on $[0,\beta)$ by

$$\max_{t\in[0,\beta]} x(\cdot;\eta,0).$$

We want to show that it is also dominated from below in order to apply the extension argument. Let us suppose that $x^*(\bar{t}) = 0$ for some $\bar{t} \in [0, \beta)$. We want to show that this leads to a contradiction, so that, without loss of generality, we can suppose that

$$\bar{t} = \min\{t \in (0, \beta) \mid x^*(t) = 0\}.$$

Therefore $x^*(\cdot) > 0$ in a left neighborhood of $\bar{t}$. Since f_0 satisfies (3.2) and thanks to (3.31)-(ii), we must have $\frac{d}{dt}x^*(\bar{t}) > 0$, which contradicts $x^*(\cdot) > 0$ in a left neighborhood of $\bar{t}$. Therefore we can say that $x^*(\cdot) > 0$ on $[0, \beta)$, so that in particular $x^*(\cdot)$ is bounded from below by 0 on $[0, \beta)$. Therefore, arguing as in the classical extension theorems for ODE, we could show that we can extend $x^*(\cdot)$ to a solution defined on $[0, +\infty)$ and, again by the same argument above, it will be $x^*(\cdot) > 0$ on $[0, +\infty)$. □

3.4 Approximation results

In this section we obtain some approximation results which may be used in order to produce ε-optimal controls for a wider class of problems. Herein we assume that

$$rx + f_0(x, 0) \geq 0, \quad \forall x \geq 0, \tag{3.36}$$

which implies in particular (3.4). In fact, all the results given below hold under (3.4) as well. We assume (3.36) only to simplify the proofs. Moreover we incorporate the term rx in the state equation within the term f_0, so that consistently with (3.36) we assume that

$$f_0(x, 0) \geq 0, \quad \forall x \geq 0. \tag{3.37}$$

3.4.1 The case without utility on the state

In the previous section we introduced an assumption of no integrability of the utility function U_2. This was necessary in order to ensure the existence of solutions for the closed loop equation and the admissibility of the feedback strategy. This fact is quite uncomfortable, because usually in consumption problems the objective functional is given by a utility depending only on the consumption variable, i.e. the case $U_2 \equiv 0$ should be considered. Of course we could take a U_2 heavily negative in a right neighborhood of 0 and equal to 0 out of this neighborhood, considering this as a forcing on the state constraint (states too near to 0 must be avoided). However we want to give here an approximation procedure to partly treat also the case $U_2 \equiv 0$, giving a way to construct at least ε-optimal strategies in this case.

So, let us consider a sequence of real functions (U_2^n) such that

$$U_2^n \uparrow 0, \quad U_2^n \text{ not integrable at } 0^+, \quad U_2^n \equiv 0 \text{ on } [1/n, +\infty). \tag{3.38}$$

Let us denote by J^n and V^n respectively the objective functionals and the value functions of the problems where the utility on the state is given by U_2^n and by J^0 and V^0 respectively the objective functional and the value function of the problem where the utility on the state disappears, i.e. $U_2 \equiv 0$. It is immediate to see that monotonicity implies

$$V^n \uparrow g \leq V^0. \tag{3.39}$$

Thanks to the previous section, for any problem V^n, $n \in \mathbb{N}$, we have an optimal feedback strategy $c_n^*(\cdot)$.

Lemma 3.4.1. *Let $\eta \in \mathcal{D}(V^0) \subset H_+$. Then, for any $\varepsilon > 0$, there exists an ε-optimal strategy $c^\varepsilon(\cdot) \in \mathcal{C}(\eta)$ for $V^0(\eta)$ such that*

$$\inf_{t \in [0,+\infty)} x(t; \eta, c^\varepsilon(\cdot)) > 0.$$

Proof. Let $\varepsilon > 0$ and take an $\varepsilon/2$-optimal control $c^{\varepsilon/2}(\cdot) \in \mathcal{C}(\eta)$ for $V^0(\eta)$. Let $M > T$ be such that

$$\frac{1}{\rho} e^{-\rho M} (\bar{U}_1 - U_1(0)) < \varepsilon/2. \tag{3.40}$$

Define the control

$$c^\varepsilon(t) := \begin{cases} c^{\varepsilon/2}(t), & \text{for } t \in [0, M], \\ 0, & \text{for } t > M. \end{cases}$$

By Lemma 3.1.8 we have

$$x(\cdot; \eta, c^\varepsilon(\cdot)) \geq x(\cdot; \eta, c^{\varepsilon/2}(\cdot))$$

and, by the assumption (3.37) and since $c^\varepsilon(t) = 0$ for $t \geq M$, it is not difficult to see that

$$x(t; \eta, c^\varepsilon(\cdot)) \geq x(M; \eta, c^\varepsilon(\cdot)), \qquad \text{for } t \geq M,$$

so that

$$\inf_{t \in [0,+\infty)} x(t; \eta, c^\varepsilon(\cdot)) = \inf_{t \in [0,M]} x(t; \eta, c^\varepsilon(\cdot)) > 0.$$

We claim that $c^\varepsilon(\cdot)$ is ε-optimal for $V^0(\eta)$, which yields the claim. Since $c^{\varepsilon/2}(\cdot)$ is $\varepsilon/2$-optimal for $V^0(\eta)$, taking also into account (3.40), we get

$$V^0(\eta) - \int_0^{+\infty} e^{-\rho t} U_1(c^\varepsilon(t))dt = V^0(\eta) - \int_0^{+\infty} e^{-\rho t} U_1(c^{\varepsilon/2}(t))dt + \int_M^{+\infty} e^{-\rho t}(U_1(c^{\varepsilon/2}(t)) - U_1(0))dt < \varepsilon/2 + \varepsilon/2 = \varepsilon.$$

□

Proposition 3.4.2. *Let $\eta \in \mathcal{D}(V^0)$ and $\varepsilon > 0$. Then $V^n(\eta) \to V^0(\eta)$ and, when n is large enough, $c_n^*(\cdot)$ is ε-optimal for $V^0(\eta)$.*

Proof. Let $\varepsilon > 0$ and take an ε-optimal control $c^\varepsilon(\cdot) \in \mathcal{C}(\eta)$ for $V^0(\eta)$ such that (Lemma 3.4.1)

$$m := \inf_{t\in[0,+\infty)} x(t;\eta;c^\varepsilon(\cdot)) > 0.$$

Take $n \in \mathbb{N}$ such that $1/n < m$. Since $U_2^n \equiv 0$ on $[m, +\infty)$, we have

$$\begin{aligned} V^0(\eta) - \varepsilon \le J(\eta;c^\varepsilon(\cdot)) &= \int_0^{+\infty} e^{-\rho t} U_1(c^\varepsilon(t))dt \\ &= \int_0^{+\infty} e^{-\rho t}\left[U_1(c^\varepsilon(t)) + U_2^n(x(t;\eta,c^\varepsilon(\cdot)))\right]dt \\ &= J^n(\eta,c^\varepsilon(\cdot)) \le V^n(\eta) = J^n(\eta,c_n^*(\cdot)) \le J^0(\eta,c_n^*(\cdot)). \end{aligned}$$

The latter inequality, toghether with (3.39), proves both the claims. □

3.4.2 The case with pointwise delay in the state equation

In this subsection we want to show that our problem is a good approximation for growth models with time to build and concentrated lag and discuss why our approach cannot work directly when the delay is concentrated in a point. In this case the state equation is

$$\begin{cases} y'(t) = f_0\left(y(t), y\left(t - \frac{T}{2}\right)\right) - c(t), \\ y(0) = \eta_0,\ y(s) = \eta_1(s),\ s \in [-T, 0). \end{cases} \tag{3.41}$$

It is possible to prove, as done in Theorem 3.1.4, that this equation admits, for every $\eta \in H_+$, and for every $c(\cdot) \in L^1_{loc}([0,+\infty);\mathbb{R})$, a unique absolutely continuous solution. We denote this solution by $y(\cdot;\eta,c(\cdot))$. The aim is to maximize, over the set

$$\mathcal{C}^0_{ad}(\eta) := \{c(\cdot) \in L^1_{loc}([0,+\infty);\mathbb{R}) \mid y(\cdot;\eta,c(\cdot)) > 0\}, \tag{3.42}$$

the functional

$$J_0(\eta, c(\cdot)) = \int_0^{+\infty} e^{-\rho t}\left[U_1(c(\cdot)) + U_2(y(t;\eta,c(\cdot)))\right] dt.$$

Denote by V_0 the associated value function. By monotonicity of f_0 we straightly get $H_{++} \subset \mathcal{D}(V_0)$.

Let us take a sequence $(a_k)_{k\in\mathbb{N}} \subset W^{1,2}_{-T}$ such that

$$\begin{cases} a_k(-T) = 0, \\ \|a_k\|_{L^2_{-T}} \le 1, \\ \text{(3.31)-(ii) holds true } \forall a_k, \\ a_k \stackrel{*}{\rightharpoonup} \delta_{-T/2} \text{ in } \left(C([-T,0];\mathbb{R})\right)^*, \end{cases} \tag{3.43}$$

where $\delta_{-T/2}$ is the Dirac measure concentrated at $-T/2$. We denote by $x_k(\cdot;\eta,c(\cdot))$ the unique solution of (3.1) where $a(\cdot)$ is replaced by $a_k(\cdot)$.

Proposition 3.4.3. *Let $\eta \in H_+$, $c(\cdot) \in L^1_{loc}([0,+\infty);\mathbb{R})$ and set $y(\cdot) := y(\cdot;\eta,c(\cdot))$, $x_k(\cdot) := x_k(\cdot;\eta,c(\cdot))$. Then there exists a continuous and increasing function h such that $h(0) = 0$ and*

$$\sup_{s\in[0,t]} |x_k(s) - y(s)| \le h(t)u_k(t), \quad t \in [0,+\infty), \tag{3.44}$$

where $u_k(t) \to 0$, as $k \to \infty$, uniformly on bounded sets.

Proof. Note that

$$\begin{aligned} \|a_k\|_{(C([-T,0];\mathbb{R}))^*} &= \sup_{\|f\|_\infty = 1} \left| \int_{-T}^0 a_k(\xi) f(\xi) d\xi \right| \\ &\le \int_{-T}^0 |a_k(\xi)| d\xi \le \|a_k\|_{L^2_{-T}} \cdot T^{1/2} \le T^{1/2}. \end{aligned} \tag{3.45}$$

Let $t \ge 0$; we have, for any $\zeta \in [0,t]$,

$$\begin{aligned} &|x_k(\zeta) - y(\zeta)| \\ &= \int_0^\zeta \left[f_0\left(x_k(s), \int_{-T}^0 a_k(\xi) x_k(s+\xi) d\xi \right) - f_0\left(y(s), y\left(s - \frac{T}{2}\right)\right)\right] ds \\ &\le C_{f_0} \left[\int_0^t |x_k(s) - y(s)| ds + \int_0^t \left| \int_{-T}^0 a_k(\xi)(x_k(s+\xi) - y(s+\xi)) d\xi \right| ds \right. \\ &\qquad \left. + \int_0^t \left| \int_{-T}^0 a_k(\xi) y(s+\xi) d\xi - y\left(s - \frac{T}{2}\right) \right| ds \right]. \end{aligned} \tag{3.46}$$

Call

$$g_k(t) := \sup_{s\in[-T,t]} |x_k(s) - y(s)| = \sup_{s\in[0,t]} |x_k(s) - y(s)|,$$

and set, for $s \in [0,t]$,

$$u_k(s) := C_{f_0} \int_0^s \left| \int_{-T}^0 a_k(\xi) y(r+\xi) d\xi - y\left(r - \frac{T}{2}\right) \right| dr.$$

Note that, for every $s \in [0,t]$, the function $[-T,0] \ni \xi \mapsto x_k(s+\xi) - y(s+\xi)$ is continuous, therefore thanks to (3.45) we can write from (3.46)

$$g_k(t) \leq C_{f_0} \left[\int_0^t g_k(s) ds + T^{1/2} \int_0^t g_k(s) ds + u_k(t) \right].$$

Therefore, setting $K := C_{f_0}(1 + T^{1/2})$, we get by Gronwall's Lemma

$$g_k(t) \leq u_k(t) + Kte^{Kt} u_k(t) =: h(t) u_k(t). \tag{3.47}$$

Note that, since $a_k \overset{*}{\rightharpoonup} \delta_{-T/2}$ in $(C([-T,0];\mathbb{R}))^*$, we have the pointwise convergence

$$\int_{-T}^0 a_k(\xi) y(s+\xi) d\xi \longrightarrow y\left(s - \frac{T}{2}\right), \quad s \in [0,t];$$

moreover

$$\left| \int_{-T}^0 a_k(\xi) y(s+\xi) d\xi \right| \leq \|a_k\|_{L^2_{-T}} \cdot \left\| y(s+\xi)|_{\xi \in [-T,0]} \right\|_{L^2_{-T}}$$
$$\leq C_{\eta, c(\cdot)} < +\infty, \quad \forall s \in [0,t],$$

where the last inequality follows from the fact that the function $[0,t] \to L^2_{-T}$, $s \mapsto y(s+\xi)|_{\xi \in [-T,0]}$ is continuous. Therefore we have by dominated convergence $u_k(t) \to 0$. By (3.47) we get (3.44). $\square$

For $k \in \mathbb{N}$, $\eta \in H_+$, let

$$\mathcal{C}^k_{ad}(\eta) := \{c(\cdot) \in L^1_{loc}([0,+\infty);\mathbb{R}) \mid x_k(\cdot;\eta,c(\cdot)) > 0\}, \tag{3.48}$$

Consider the problem of maximizing over $\mathcal{C}^k_{ad}(\eta)$ the functional

$$J_k(\eta, c(\cdot)) := \int_0^{+\infty} e^{-\rho t} \left[U_1(c(t)) + U_2(x_k(t;\eta,c(\cdot))) \right] dt$$

and denote by V_k the associated value function. Note that, since we have assumed (3.37), straightly we get $H_{++} \subset \mathcal{D}(V_k)$ for every $k \in \mathbb{N}$. Thanks to the previous section we have a sequence of optimal feedback strategies for the sequence of problems $(V_k(\eta))_{k \in \mathbb{N}}$, in the sense that we have a sequence $(c^*_k(\cdot))_{k \in \mathbb{N}}$ of feedback controls such that $c^*_k(\cdot) \in \mathcal{C}^k_{ad}(\eta)$ for every $k \in \mathbb{N}$ and

$$J_k(\eta; c^*_k(\cdot)) = \sup_{c(\cdot) \in \mathcal{C}^k_{ad}(\eta)} J_k(\eta; c(\cdot)) =: V_k(\eta), \qquad \forall k \in \mathbb{N}.$$

Lemma 3.4.4. *Let $\eta \in H_{++}$.*

- *For any $\varepsilon > 0$ there exists an ε-optimal strategy $c^\varepsilon(\cdot) \in \mathcal{C}^0_{ad}(\eta)$ for the problem $V_0(\eta)$ such that*
$$\inf_{t\in[0,+\infty)} y(t;\eta,c^\varepsilon(\cdot)) > 0.$$
- *Assume that*
$$\lim_{x\to 0^+} [x\, U_2(x)] = -\infty. \tag{3.49}$$
Then, for any $\varepsilon > 0$ there exists $\nu > 0$ such that for any $k \in \mathbb{N}$ there exists an ε-optimal control $c^\varepsilon_k(\cdot) \in \mathcal{C}^k_{ad}(\eta)$ for the problem $V_k(\eta)$ such that
$$\inf_{t\in[0,+\infty)} x_k(t;\eta,c^\varepsilon_k(\cdot)) \geq \nu. \tag{3.50}$$

Proof. (i) Let $\varepsilon > 0$ and take an $\varepsilon/2$-optimal control $c^{\varepsilon/2}(\cdot) \in \mathcal{C}^0_{ad}(\eta)$ for the problem $V_0(\eta)$. Take $M > 0$ large enough to satisfy
$$\frac{1}{\rho} e^{-\rho M}(\bar{U}_1 - U_1(0)) < \varepsilon/2. \tag{3.51}$$
Define the control
$$c^\varepsilon(t) := \begin{cases} c^{\varepsilon/2}(t), & \text{for } t \in [0, M], \\ 0, & \text{for } t > M. \end{cases}$$
A comparison criterion like the one proved in Lemma 3.1.8 can be proved also for equation (3.41). Therefore we have
$$y(\cdot;\eta,c^\varepsilon(\cdot)) \geq y(\cdot;\eta,c^{\varepsilon/2}(\cdot)) \tag{3.52}$$
and, since we have assumed (3.37),
$$\inf_{t\in[0,+\infty)} y(t;\eta,c^\varepsilon(\cdot)) = \inf_{t\in[0,M]} y(t;\eta,c^\varepsilon(\cdot)).$$
We claim that $y(\cdot;\eta,c^\varepsilon(\cdot))$ is ε-optimal for V_0, which yields the claim. Since $c^{\varepsilon/2}(\cdot)$ is $\varepsilon/2$-optimal for $V_0(\eta)$, taking also into account (3.51) and (3.52),
$$\begin{aligned} V_0(\eta) - \int_0^{+\infty} e^{-\rho t}\left[U_1(c^\varepsilon(t)) + U_2(y(t;\eta,c^\varepsilon(\cdot)))\right] dt \\ = V_0(\eta) - \int_0^{+\infty} e^{-\rho t}(U_1(c^{\varepsilon/2}(t)) + U_2(y(t;\eta,c^{\varepsilon/2}(\cdot)))dt \\ + \int_M^{+\infty} e^{-\rho t}(U_1(c^{\varepsilon/2}(t)) - U_1(0))dt \\ + \int_M^{+\infty} e^{-\rho t}\left[U_2(y(t;\eta,c^{\varepsilon/2}(\cdot)) - U_2(y(t;\eta,c^\varepsilon(t)))\right] dt < \varepsilon/2 + \varepsilon/2 = \varepsilon. \end{aligned}$$

(ii) Due to (3.37) we have $x_k(\cdot;\eta,0) \geq \eta_0$ for every $k \in \mathbb{N}$. Let

$$j_0 := \frac{U_1(0) + U_2(\eta_0)}{\rho}.$$

Then we have $V_k(\eta) \geq J_k(\eta, 0) \geq j_0$ for every $k \in \mathbb{N}$. Take $M > 0$ large enough to satisfy

$$\frac{1}{\rho} e^{-\rho M} (\bar{U}_1 - U_1(0)) < \varepsilon. \tag{3.53}$$

Arguing as done to get (3.33) and taking into account the comparison criterion, we can find $C_M > 0$ such that, for every $k \in \mathbb{N}$ and for every $c(\cdot) \in \mathcal{C}^k_{ad}(\eta)$, we have for all $s \in [0, M]$ and for all $t \in [s, M+1]$,

$$x_k(t;\eta,c(\cdot)) \leq x_k(s;\eta,c(\cdot))(1 + C_M(t-s)) + C_M(t-s). \tag{3.54}$$

Now take $\nu > 0$ small enough to have

$$\begin{cases} \text{(i)} \quad \nu < 1, \\ \text{(ii)} \quad \frac{\nu}{2C_M} < 1, \\ \text{(iii)} \quad \frac{\nu}{2C_M} U_2(2\nu) e^{-\rho(M+1)} < j_0 - \frac{\bar{U}_1 + \bar{U}_2}{\rho} - 1 < 0. \end{cases} \tag{3.55}$$

For $k \in \mathbb{N}$, thanks to the previous section we have optimal strategies in feedback form $c^*_k(\cdot) \in \mathcal{C}^k_{ad}(\eta)$ for V_k; we claim that $x_k(t;\eta,c^*_k(\cdot)) > \nu$ for $t \in [0, M]$ for every $k \in \mathbb{N}$. Indeed suppose by contradiction that for some $t_0 \in [0, M]$ we have $x_k(t_0;\eta,c^*_k(\cdot)) = \nu$; then by (3.54) and (3.55)-(i),(ii) we get that

$$x_k(t;\eta,c^*_k(\cdot)) \leq 2\nu, \quad \text{for } t \in \left[t_0, t_0 + \frac{\nu}{2C_M}\right].$$

Therefore, by (3.55)-(iii),

$$\int_{t_0}^{t_0 + \frac{\nu}{2C_M}} e^{-\rho t} U_2^n(x_k(t;\eta,c^*_k(\cdot)))dt \leq j_0 - \frac{\bar{U}_1 + \bar{U}_2}{\rho} - 1.$$

This shows that

$$J_k(\eta, c^*_k(\cdot)) \leq j_0 - 1 \leq V_k(\eta) - 1.$$

This fact contradicts the optimality of $c^*_k(\cdot)$. Therefore we have proved that for the choice of ν given by (3.55) we have

$$x_k(t;\eta,c^*_k(\cdot)) > \nu, \quad \text{for } t \in [0, M].$$

We can continue the strategy $c^*_k(\cdot)$ after M taking the null strategy, i.e. defining the strategy

$$c^\varepsilon_k(\cdot) := \begin{cases} c^*_k(t), & \text{for } t \in [0, M], \\ 0, & \text{for } t > M. \end{cases} \tag{3.56}$$

Then by (3.37) we have $x_k(\cdot;\eta,c^\varepsilon_k(\cdot)) > \nu$ for every $k \in \mathbb{N}$. We claim that $c^\varepsilon_k(\cdot)$ is ε-optimal for $V_k(\eta)$ for every $k \in \mathbb{N}$, which proves the claim. Indeed, taking into account the comparison criterion and (3.53) for the inequality in the following,

$$
\begin{aligned}
V_k(\eta) - \int_0^{+\infty} & e^{-\rho t}(U_1(c^\varepsilon_k(t)) + U_2(x_k(t;\eta,c^\varepsilon_k(\cdot)))dt \\
&= V_k(\eta) - \int_0^{+\infty} e^{-\rho t}(U_1(c^*_k(t)) + U_2(x_k(t;\eta,c^*_k(\cdot)))dt \\
&\qquad + \int_M^{+\infty} e^{-\rho t}(U_1(c^*_k(t)) - U_1(0))dt \\
&\qquad + \int_M^{+\infty} e^{-\rho t}(U_2(x_k(t;\eta,c^*_k(\cdot)) - U_2(x_k(t;\eta,c^\varepsilon_k(t))))dt < \varepsilon.
\end{aligned}
$$

□

Proposition 3.4.5. *Let $\eta \in H_{++}$ and suppose that* (3.49) *holds true. We have $V_k(\eta) \to V_0(\eta)$, as $k \to \infty$. Moreover for every $\varepsilon > 0$ we can find a constant M_ε and a k_ε such that the strategy (c^*_k is the optimal feedback strategy for the problem of V_k)*

$$
c_{k_\varepsilon,M_\varepsilon}(t) := \begin{cases} c^*_{k_\varepsilon}(t), & \text{for } t \in [0, M_\varepsilon], \\ 0, & \text{for } t > M_\varepsilon. \end{cases} \tag{3.57}
$$

is ε-optimal strategy for the problem $V_0(\eta)$.

Proof. (i) Here we show that

$$
\liminf_{k\to\infty} V_k(\eta) \geq V_0(\eta), \tag{3.58}
$$

Let $\varepsilon > 0$ and let $c^\varepsilon(\cdot) \in \mathcal{C}^0_{ad}(\eta)$ be an ε-optimal strategy for the problem $V_0(\eta)$. Thanks to Lemma 3.4.4-(i) we can suppose without loss of generality $2\nu_1 := \inf_{t\in[0,+\infty)} y(t;\eta,c^\varepsilon(\cdot)) > 0$. The function U_1 is uniformly continuous on $[0,+\infty)$ and the function U_2 is uniformly continuous on $[\nu_1,+\infty)$. Let ω_{ν_1} be a modulus of uniform continuity for both these functions. Take $M > 0$ such that

$$
-\frac{1}{\rho}e^{-\rho M}(\bar{U}_1 + \bar{U}_2) - \frac{1-e^{-\rho M}}{\rho}\omega_{\nu_1}(\nu_1) + \frac{1}{\rho}e^{-\rho M}(U_1(0) + U_2(\nu_1)) \geq -\varepsilon, \tag{3.59}
$$

Define

$$
c^\varepsilon_M(t) := \begin{cases} c^\varepsilon(t), & \text{for } t \in [0, M], \\ 0, & \text{for } t > M. \end{cases}
$$

Let k_M be such that

$$
h(M)u_k(M) < \nu_1, \quad \forall k \geq k_M, \tag{3.60}
$$

where u_k and h are the functions appearing in (3.44). Then, thanks to Proposition 3.4.3 and to the monotonicity property of f_0, it is straightforward to see that $x_k(t;\eta,c^\varepsilon_M(\cdot)) \geq \nu_1 > 0$, so that in particular $c^\varepsilon_M(\cdot) \in \mathcal{C}^k_{ad}(\eta)$ for all $k \geq k_M$. For all $k \geq k_M$, we have, thanks to Proposition 3.4.3 and by definition of k_M,

$$\begin{aligned}
&\int_0^{+\infty} e^{-\rho t}[U_1(c^\varepsilon_M(t)) + U_2(x_k(t;\eta,c^\varepsilon_M(\cdot)))]dt \\
&\qquad = \int_0^M e^{-\rho t}[U_1(c^\varepsilon_M(t)) + U_2(x_k(t;\eta,c^\varepsilon_M(\cdot)))]dt \\
&\qquad + \int_M^{+\infty} e^{-\rho t}[U_1(c^\varepsilon_M(t)) + U_2(x_k(t;\eta,c^\varepsilon_M(\cdot)))]dt \\
&\qquad \geq \int_0^M e^{-\rho t}[U_1(c^\varepsilon(t)) + U_2(y(t;\eta,c^\varepsilon(\cdot)))]dt \\
&\qquad - \frac{1-e^{-\rho M}}{\rho}\omega_{\nu_1}(\nu_1) + \frac{1}{\rho}e^{-\rho M}(U_1(0)+U_2(\nu_1)) \\
&\qquad \geq \int_0^{+\infty} e^{-\rho t}[U_1(c^\varepsilon(t)) + U_2(y(t;\eta,c^\varepsilon(\cdot)))]dt \\
&\qquad - \frac{1}{\rho}e^{-\rho M}(\bar{U}_1+\bar{U}_2) - \frac{1-e^{-\rho M}}{\rho}\omega_{\nu_1}(\nu_1) + \frac{1}{\rho}e^{-\rho M}(U_1(0)+U_2(\nu_1)).
\end{aligned}$$

so that by (3.59)

$$V_k(\eta) \geq V_0(\eta) - 2\varepsilon, \tag{3.61}$$

which shows (3.58).

(ii) Now we show that

$$\limsup_{k\to\infty} V_k(\eta) \leq V_0(\eta). \tag{3.62}$$

Let $\varepsilon > 0$; thanks to Lemma 3.4.4-(ii) we can construct a sequence $(c^\varepsilon_k(\cdot))_{k\in\mathbb{N}}$, $c^\varepsilon_k(\cdot) \in \mathcal{C}^k_{ad}(\eta)$ for every $k \in \mathbb{N}$, of ε-optimal controls for the sequence of problems $(V_k(\eta))_{k\in\mathbb{N}}$ such that

$$2\nu_2 := \inf_{k\in\mathbb{N}} \inf_{t\in[0,+\infty)} x_k(t;\eta,c^\varepsilon_k(\cdot)) > 0.$$

Let ω_{ν_2} be a modulus of uniform continuity for U_1 on $[0,+\infty)$ and for U_2 on $[\nu_2,+\infty)$. Take $\tilde{M} > 0$ such that

$$-\frac{1}{\rho}e^{-\rho\tilde{M}}(\bar{U}_1+\bar{U}_2) - \frac{1-e^{-\rho\tilde{M}}}{\rho}\omega_{\nu_2}(\nu_2) + \frac{1}{\rho}e^{-\rho\tilde{M}}(U_1(0)+U_2(\nu_2)) > -\varepsilon \tag{3.63}$$

and define the controls

$$c^\varepsilon_{k,\tilde{M}}(t) := \begin{cases} c^\varepsilon_k(t), & \text{for } t \in [0,\tilde{M}], \\ 0, & \text{for } t > \tilde{M}. \end{cases} \tag{3.64}$$

As before we can find $k_{\tilde{M}}$ such that we have

$$h(\tilde{M})u_k(\tilde{M}) < \nu_2, \quad \forall k \geq k_{\tilde{M}}, \tag{3.65}$$

In this case we have $y(\cdot;\eta, c_k^\varepsilon(\cdot)) \geq \nu_2$ and

$$\begin{aligned}
&\int_0^{+\infty} e^{-\rho t}[U_1(c^\varepsilon_{k,\tilde{M}}(t)) + U_2(y(t;\eta, c^\varepsilon_{k,\tilde{M}}(\cdot)))]dt \\
&\qquad = \int_0^{\tilde{M}} e^{-\rho t}[U_1(c^\varepsilon_{k,\tilde{M}}(t)) + U_2(y(t;\eta, c^\varepsilon_{k,\tilde{M}}(\cdot)))]dt \\
&\qquad + \int_{\tilde{M}}^{+\infty} e^{-\rho t}[U_1(c^\varepsilon_{k,\tilde{M}}(t)) + U_2(y(t;\eta, c^\varepsilon_{k,\tilde{M}}(\cdot)))]dt \\
&\qquad \geq \int_0^{\tilde{M}} e^{-\rho t}[U_1(c^\varepsilon_{k}(t)) + U_2(x_k(t;\eta, c^\varepsilon_{k}(\cdot)))]dt \\
&\qquad - \frac{1-e^{-\rho\tilde{M}}}{\rho}\omega_{\nu_2}(\nu_2) + \frac{1}{\rho}e^{-\rho\tilde{M}}(U_1(0)+U_2(\nu_2)) \\
&\qquad \geq \int_0^{+\infty} e^{-\rho t}[U_1(c^\varepsilon_{k}(t)) + U_2(x_k(t;\eta, c^\varepsilon_{k}(\cdot)))]dt \\
&- \frac{1}{\rho}e^{-\rho\tilde{M}}(\bar{U}_1+\bar{U}_2) - \frac{1-e^{-\rho\tilde{M}}}{\rho}\omega_{\nu_2}(\nu_2) + \frac{1}{\rho}e^{-\rho\tilde{M}}(U_1(0)+U_2(\nu_2)).
\end{aligned} \tag{3.66}$$

By (3.63), we get, for $k \geq k_{\tilde{M}}$,

$$V_0(\eta) \geq V_k(\eta) - 2\varepsilon,$$

which proves (3.62).

(iii) The procedure of construction of $c^\varepsilon_{k,M}$ in (ii) yields ε-optimal controls for the limit problem $V_0(\eta)$. Indeed, starting from $\varepsilon > 0$, we can compute $\nu_1, \nu_2, M, \tilde{M}$ depending on ε such that (3.59) and (3.63) hold true. Then, if $(a_k)_{k\in\mathbb{N}}$ is chosen in a clever way, for example if $(a_k)_{k\in\mathbb{N}}$ is a sequence of gaussian densities, we can compute $k_M, k_{\tilde{M}}$ such that (3.60),(3.65) hold true. Thanks to (3.66) and (3.61), for every $k \geq k_M \vee k_{\tilde{M}}$ the controls $c^\varepsilon_{k,\tilde{M}}(\cdot)$ defined in (3.64) are 4ε-optimal for the limit problem $V^0(\eta)$. Replacing ε with $\varepsilon/4$ we get the controls in (3.57). $\square$

Remark 3.4.6. When the delay is concentrated in a point in a linear way, we could tempted to insert the delay term in the infinitesimal generator A and try to work as done in Section 3.2. Unfortunately this is not possible. Indeed consider this simple case:

$$\begin{cases} y'(t) = ry(t) + y(t-T), \\ y(0) = \eta_0, \; y(s) = \eta_1(s), \; s \in [-T, 0), \end{cases}$$

In this case we can define

$$A : \mathcal{D}(A) \subset H \longrightarrow H, \qquad (\eta_0, \eta_1(\cdot)) \longmapsto (r\eta_0 + \eta_1(-T), \eta_1'(\cdot)).$$

where again

$$\mathcal{D}(A) := \{\eta \in H \mid \eta_1(\cdot) \in W^{1,2}([-T,0];\mathbb{R}),\ \eta_1(0) = \eta_0\}.$$

The inverse of A is the operator

$$A^{-1} : (H, \|\cdot\|_H) \longrightarrow (\mathcal{D}(A), \|\cdot\|_H) \qquad (\eta_0, \eta_1(\cdot)) \longmapsto \left(\frac{\eta_0 - c}{r},\ c + \int_{-T}^{\cdot} \eta_1(\xi)d\xi\right),$$

where

$$c = \frac{1}{r+1}\eta_0 - \frac{r}{r+1}\int_{-T}^{0} \eta_1(\xi)d\xi.$$

In this case we would have the first part of Lemma 3.2.7, but not the second part, because it is not possible to control $|\eta_0|$ by $\|\eta\|_{-1}$. Indeed take for example r such that $\frac{1-r}{1+r} = \frac{1}{2}$, and $(\eta^n)_{n\in\mathbb{N}} \subset H$ such that

$$\eta_0^n = 1/2, \quad \int_{-T}^{0} \eta_1^n(\xi)d\xi = 1, \quad n \in \mathbb{N}.$$

We would have $c = 1/2$, so that $\left|\frac{\eta_0^n - c}{r}\right| = 0$. Moreover we can choose η_1^n such that, when $n \to \infty$,

$$\int_{-T}^{0} \left| \frac{1}{2} + \int_{-T}^{s} \eta_1^n(\xi)d\xi \right|^2 ds \longrightarrow 0.$$

Therefore we would have $|\eta_0^n| = 1/2$ and $\|\eta^n\|_{-1} \to 0$. This shows that the second part of Lemma 3.2.7 does not hold. Once this part does not hold, then everything in the following argument breaks down. □

3.4.3 The case with pointwise delay in the state equation and without utility on the state

Now we want to approximate the problem of optimizing, for $\eta \in H_{++}$,

$$J_0^0(\eta, c(\cdot)) := \int_0^{+\infty} e^{-\rho t} U_1(c(t))dt,$$

over the set (3.42), where $y(\cdot;\eta, c(\cdot))$ follows the dynamics given by (3.41). Let us denote by V_0^0 the corresponding value function and let us take a sequence of real functions (U_2^n) as in (3.38), but with the assumption of no integrability at 0^+ replaced by the stronger assumption

$$\lim_{x\to 0^+} x\, U_2^n(x) = -\infty, \qquad \forall n \in \mathbb{N}.$$

Fix $n \in \mathbb{N}$ and consider the sequence of functions $(a_k)_{k\in\mathbb{N}}$ defined in (3.43). For $k \in \mathbb{N}$, consider the problem of maximizing over the set $\mathcal{C}^k_{ad}(\eta)$ defined in (3.48) the functional

$$J^n_k(\eta, c(\cdot)) := \int_0^{+\infty} e^{-\rho t}(U_1(c(t)) + U^n_2(x_k(t;\eta, c(\cdot)))dt,$$

where $x_k(\cdot;\eta, c(\cdot))$ follows the dynamics given by (3.1) when $a(\cdot)$ is replaced by $a_k(\cdot)$, and denote by V^n_k the associated value function.

Moreover, for $k \in \mathbb{N}$, consider the problem of maximizing

$$J^0_k(\eta, c(\cdot)) := \int_0^{+\infty} e^{-\rho t}U_1(c(t))dt,$$

over $\mathcal{C}^k_{ad}(\eta)$ and denote by V^0_k the associated value function.

Finally consider the problem of maximizing over the set $\mathcal{C}^0_{ad}(\eta)$ the functional

$$J^n_0(\eta, c(\cdot)) := \int_0^{+\infty} e^{-\rho t}(U_1(c(t)) + U^n_2(y(t;\eta, c(\cdot)))dt,$$

and denote by V^n_0 the associated value function.

For fixed $n \in \mathbb{N}$, the problems V^n_k approximate, when $k \to \infty$, the problem V^n_0 in the sense of Proposition 3.4.5, i.e. we are able to produce $k_{\varepsilon,n}, M_{\varepsilon,n}$ large enough to make the strategy $c_{k_{\varepsilon,n},M_{\varepsilon,n}}(\cdot)$ defined as in (3.57) admissible and ε-optimal for the problem $V^n_0(\eta)$.

Proposition 3.4.7. *Let $\eta \in H_{++}$, let $k_{\varepsilon,n}, M_{\varepsilon,n}$, $c_{k_{\varepsilon,n},M_{\varepsilon,n}}(\cdot)$ as above. For every $\varepsilon > 0$ we can find n_ε such that*

$$\lim_{\varepsilon\downarrow 0} V^{n_\varepsilon}_{k_{\varepsilon,n_\varepsilon}}(\eta) = V^0_0(\eta). \tag{3.67}$$

Moreover the controls $c_{k_{\varepsilon,n_\varepsilon},M_{\varepsilon,n_\varepsilon}}(\cdot)$ defined as in (3.57) *are admissible and 3ε-optimal for the problem $V^0_0(\eta)$.*

Proof. Let $\varepsilon > 0$ and consider the strategies $c^\varepsilon(\cdot)$ and $c^\varepsilon_M(\cdot)$ defined as in the part (i) of the proof of Proposition 3.4.5. Notice that actually $M = M(\varepsilon, n) =: M^\varepsilon_n$. Notice also that by definition of $k_{\varepsilon,n}$, M^ε_n we have $x_{k_{\varepsilon,n}}(\cdot;\eta, c^\varepsilon_{M^\varepsilon_n}(\cdot)) \geq \nu_1$ and that (3.51) in particular implies

$$\frac{1}{\rho}e^{-\rho M^\varepsilon_n}(\bar{U}_1 - U_1(0)) \leq \varepsilon.$$

Take $n_\varepsilon \in \mathbb{N}$ such that $1/n_\varepsilon < \nu_1$ (notice that ν_1 depends on ε and does not depend on n). Then, since $U_2^{n_\varepsilon} \equiv 0$ on $[\nu_1, +\infty)$, we can write

$$
\begin{aligned}
V_0^0(\eta) - \varepsilon \le J_0^0(\eta; c^\varepsilon(\cdot)) &= \int_0^{+\infty} e^{-\rho t} U_1(c^\varepsilon(t))dt \\
&= \int_0^{+\infty} e^{-\rho t} \left[U_1(c^\varepsilon(t)) + U_2^{n_\varepsilon}(x_{k_{\varepsilon,n_\varepsilon}}(t; \eta, c^\varepsilon_{M^\varepsilon_{n_\varepsilon}}(\cdot))) \right] dt \\
&\le J^{n_\varepsilon}_{k_{\varepsilon,n_\varepsilon}}(\eta, c^\varepsilon_{M^\varepsilon_{n_\varepsilon}}(\cdot)) + \frac{1}{\rho} e^{-\rho M^\varepsilon_{n_\varepsilon}} (\bar{U}_1 - U_1(0)) \le V^{n_\varepsilon}_{k_{\varepsilon,n_\varepsilon}}(\eta) + \varepsilon, \quad (3.68)
\end{aligned}
$$

so that

$$
\liminf_{\varepsilon \downarrow 0} V^{n_\varepsilon}_{k_{\varepsilon,n_\varepsilon}}(\eta) \ge V_0^0(\eta). \tag{3.69}
$$

On the other hand the strategies $c_{k_{\varepsilon,n}, M_{\varepsilon,n}}(\cdot)$ defined in (3.57) are admissible for the problem V_0^0 (since the state equation related to V_0^n and to V_0^0 is the same) and $c_{k_{\varepsilon,n}, M_{\varepsilon,n}}(\cdot)$ is ε-optimal for $V^n_{k_{\varepsilon,n}}(\eta)$ for every $n \in \mathbb{N}$. Therefore

$$
\begin{aligned}
V^{n_\varepsilon}_{k_{\varepsilon,n_\varepsilon}}(\eta) - \varepsilon \le J^{n_\varepsilon}_{k_{\varepsilon,n_\varepsilon}}(\eta; c_{k_{\varepsilon,n_\varepsilon}, M_{\varepsilon,n_\varepsilon}}(\cdot)) &\le J^0_{k_{\varepsilon,n_\varepsilon}}(\eta; c_{k_{\varepsilon,n_\varepsilon}, M_{\varepsilon,n_\varepsilon}}(\cdot)) \\
&= J_0^0(\eta; c_{k_{\varepsilon,n_\varepsilon}, M_{\varepsilon,n_\varepsilon}}(\cdot)) \le V_0^0(\eta), \quad (3.70)
\end{aligned}
$$

which shows

$$
\limsup_{\varepsilon \downarrow 0} V^n_{k_{\varepsilon,n}}(\eta) \le V_0^0(\eta). \tag{3.71}
$$

Combining (3.69) and (3.71) we get (3.67). Combining (3.68) and (3.70) we get

$$
V_0^0(\eta) \le J_0^0(\eta; c_{k_{\varepsilon,n_\varepsilon}, M_{\varepsilon,n_\varepsilon}}(\cdot)) + 3\varepsilon,
$$

i.e. the last claim. □

Chapter 4

Constrained choices in the decumulation phase of a pension plan

In this chapter we investigate a different problem in the pension funds' field, i.e. the problem of managing the pension of a single representative partecipant after retirement. The main reference for this chapter is the working paper [Di Giacinto, Federico, Gozzi, Vigna; WP].

In countries where immediate annuitization is the only option available to retiring members of defined contribution pension schemes, members who retire at a time of low bond yield rates have to accept a pension lower than the one available with higher bond yields. In UK and US the retiree is allowed to defer annuitization at some time after retirement, withdraw periodic income from the fund and invest the rest of it in the period between retirement and annuitization. This allows the retiree to postpone the decision to purchase an annuity until a more propitious time. In UK there are limits imposed on both the consumption (which must be between 35% and 100% of the annuity which would have been purchasable immediately on retirement) and on how long the annuity purchase can be deferred (the fund must be used to purchase an annuity at age 75, if this has not been done earlier). On the other hand, there is virtually unlimited freedom to invest the fund in a broad range of assets.

The three degrees of freedom of the retiree (amount of consumption, investment allocation and time of annuitization), together with the important issue of ruin possibility, have been investigated in the actuarial and financial literature in many papers. Among others we mention

- [Albrecht, Maurer; 2002],
- [Cairns, Blake, Dowd; 2000],

- [Gerrard, Haberman, Vigna; 2004],
- [Gerrard, Haberman, Vigna; 2006],
- [Milevsky, 2001],
- [Milevsky, Moore, Young; 2006],
- [Milevsky, Young; 2007].

Here we consider the position of a representative participant to a defined contribution pension fund who retires and compulsorily has to purchase an annuity within a certain period of time after retirement. In the interim the accumulated capital is dynamically allocated while the pensioner withdraws periodic amounts of money to provide for daily life in accordance with restrictions imposed by the scheme's rules or by legislation. in particular we assume that an individual who retires at time $t = 0$ acquires control of a fund of size x_0 which is invested in a market that consists of a risky and a riskless asset. The value of the risky asset is assumed to follow a geometric Brownian motion model. At age T the entire fund must be invested in an annuity. The retiree is given only one degree of freedom, namely the investment allocation. The income withdrawn from the fund in the unit time is assumed to be fixed and equal to b_0 and the retiree is obliged to annuitize at future time T.

Due to the difficulty that arises by the inclusion of constraints on the variables, we do not consider here the first and the third degrees of freedom of the pensioner, namely the optimal consumption strategy and the optimal annuitization time, treated – without restrictions on the variables – e.g. in the papers [Gerrard, Haberman, Vigna; 2006] and [Gerrard, Hojgaard, Vigna; 2008]. On the other hand, differently from the previous literature, we analyze the problem in the presence of short selling constraints, extending the work done by [Gerrard, Haberman, Vigna; 2004].

In this chapter we will solve explicitely the problem when the constraint is only on the strategy. When the constraint is also on the state we will characterize the value function as unique viscosity solution of the HJB equation and, in the special case without running cost in the objective functional, we will solve the problem explicitely (in this case the solution is much less trivial).

4.1 The model

In our model we consider the position of an individual who chooses the drawdown option at retirement, i.e. withdraws a certain income until he achieves the age at which the purchase of the annuity is compulsory. As in Chapters 1 & 2, the fund is invested in two assets, a riskless asset, with constant instantaneous rate of return $r \geq 0$, and a risky asset, whose price follows a geometric

Brownian motion with constant volatility $\sigma > 0$ and drift $\mu := r + \sigma\lambda$, where $\lambda \geq 0$ is the risk premium. The pensioner withdraws an amount b_0 in the unit of time. Therefore, according to [Merton; 1969] the state equation that describes the dynamics of the fund wealth is the following

$$\begin{cases} dX(s) = [rX(s) + \sigma\lambda\pi(s) - b_0]\, ds + \sigma\pi(s)dB(s), \\ X(0) = x_0, \end{cases}$$

where x_0 is the fund wealth at the retirement date $t = 0$, $B(\cdot)$ is a standard Brownian motion on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ and $\pi(\cdot)$ is the progressively measurable process (with respect to the filtration generated by $B(\cdot)$) representing the amount of money invested in the risky asset at time $t \in [0, T]$. We impose the constraint $\pi(\cdot) \geq 0$, i.e. short selling of risky asset is not allowed.

We introduce the loss function

$$L(t, x) = (F(t) - x)^2\,, \tag{4.1}$$

where $F(\cdot)$ is the target function, i.e. the target that the agent wishes to track at any time $t \in [0, T]$.

In [Højgaard, Vigna; 2007] it is proved that the quadratic loss function applied to the defined contribution pension schemes is a particular case of mean variance portfolio optimization approach. In other words, the optimal portfolio found via the quadratic loss function (4.1) is an efficient portfolio in the mean-variance setting. Namely, there is no other portfolio that provides a higher expected return with the same variance, and no other portfolio that provides a lower variance with the same mean.

According to [Gerrard, Haberman, Vigna; 2004], we choose

$$F(t) = \frac{b_0}{r} + \left(\bar{F} - \frac{b_0}{r}\right) e^{-r(T-t)}, \tag{4.2}$$

where $\bar{F} = F(T) \in (0, b_0/r]$ can be chosen arbitrarily. The interpretation of the target function is straightforward. Should the fund hit $F(t)$ at time $t \leq T$, the pensioner would be able, by investing the whole portfolio in the riskless asset, to consume b_0 from t to T and receive the desired target $\bar{F}$ at the time T of compulsory annuitization.

As it is shown in [Gerrard, Haberman, Vigna; 2004], with this choice of the target function the fund never reaches the target, provided that at initial time $t = 0$ the fund x_0 is lower than the target $F(0)$.

The optimization problem consists in minimizing, over the set of the admissible strategies, the functional

$$\mathbb{E}\left[\int_0^T \kappa e^{-\rho t} L(t, X(t))dt + e^{-\rho T} L(T, X(T))\right],$$

where $\rho \geq 0$ is the individual discount factor, $\kappa \geq 0$ is a weighting constant which measures the importance the running cost for deviation experienced before T.

4.2 Constraint on the strategies

In this section we study the optimization problem considering only the constraint on the strategies. We suppose that short selling of risky asset is not allowed, i.e. it must be $\pi(\cdot) \geq 0$. So, the set of admissible strategy is

$$\Pi_{ad} = \big\{\pi(\cdot) \text{ progressively measurable with respect to } (\mathcal{F}_t^B)_{t\in[0,T]}, \\ \pi(\cdot) \in L^2(\Omega \times [0,T]; [0,+\infty))\big\}.$$

In this case the set of admissible strategies can be taken independent of the initial (s,x). Given the initial data $(s,x) \in [0,T] \times \mathbb{R}$, the state equation is

$$\begin{cases} dX(t) = [rX(t) + \sigma\lambda\pi(t) - b_0]\, ds + \sigma\pi(t)dB(t), \\ X(s) = x. \end{cases}$$

This equation admits, for given $\pi(\cdot) \in \Pi_{ad}$, a unique strong solution on $(\Omega, \mathcal{F}, \mathbb{P})$ (see, e.g., Theorem 6.16, Chapter 1, of [Yong, Zhou; 1999] or Section 5.6.C of [Karatzas, Shreve; 1991]) and we denote it by $X(\cdot; t, x, \pi(\cdot))$. The objective functional is given by

$$J(s,x;\pi(\cdot)) := \mathbb{E}\left[\int_s^T \kappa e^{-\rho t}(F(t) - X(t;s,x,\pi(\cdot)))^2 dt + e^{-\rho T}(\bar{F} - X(T;s,x,\pi(\cdot)))^2\right].$$

The value function is defined by

$$V(s,x) := \inf_{\pi(\cdot)\in\Pi_{ad}} J(s,x;\pi(\cdot)).$$

4.2.1 Properties of the value function

In this subsection we prove some properties of the value function. We start with a lemma that analyzes the behaviour of the state trajectory under the null control.

Lemma 4.2.1. *Let $s \in [0,T]$, $x = F(s)$. Then $X(t;s,x,0) = F(t)$ for all $t \in [s,T]$.*

Proof. Let $s \in [0,T]$, $x \in \mathbb{R}$, and set $X(\cdot) := X(\cdot; s, x, 0)$. The dynamics of $X(\cdot)$ is given by

$$\begin{cases} dX(t) = (rX(t) - b_0)\, dt, \\ X(s) = x. \end{cases}$$

The "dynamics" of the target $F(\cdot)$ after s is given by

$$\begin{cases} dF(t) = (rF(t) - b_0)\, dt, \\ F(s) = F(s). \end{cases}$$

Therefore $X(\cdot)$ and $F(\cdot)$ solve the same ordinary differential equation. If $x = F(s)$ they have also the same initial condition, so they coincide. $\square$

Lemma 4.2.1 shows that the null strategy $\pi(\cdot) \equiv 0$ is optimal for the initial datum $(s, F(s))$, since we have $J(s, F(s); 0) = 0$ and, on the other hand, $V(\cdot, \cdot) \geq 0$. In particular $V(s, F(s)) = 0$ for each $s \in [0, T]$. Note also that, if $x \neq F(s)$, then it has to be $V(s, x) > 0$. This suggests that the graph of $F(\cdot)$ works as a barrier for the problem, so that we are led to separate the space $[0, T] \times \mathbb{R}$ in two regions

$$U_1 := \{(s, x) \mid s \in [0, T],\ x \leq F(s)\}, \quad U_2 := \{(s, x) \mid t \in [0, T],\ x \geq F(s)\}.$$

Notice that

$$U_1 \cup U_2 = [0, T] \times \mathbb{R}, \quad U_1 \cap U_2 = \{(s, F(s)) \mid s \in [0, T]\}.$$

Remark 4.2.2. The financial problem makes sense for $x_0 \leq F(0)$ and actually in applying the dynamic programming we can restrict to study the problem in the region U_1 (thanks to Lemma 4.2.3 below). However, by sake of completeness we treat also the problem in the region U_2. $\square$

Lemma 4.2.3. *Let* $(s, x) \in [0, T] \times \mathbb{R}$, $\pi(\cdot) \geq 0$ *a strategy; set* $X(\cdot) := X(\cdot; s, x, \pi(\cdot))$ *and define, with the convention* $\inf \emptyset = T$*, the stopping time*

$$\tau := \inf\{t \geq s \mid X(t) = F(t)\};$$

define the strategy

$$\pi^\tau(t) := \begin{cases} \pi(t), & \text{if } t < \tau, \\ 0, & \text{if } t \geq \tau. \end{cases}$$

Then $J(s, x; \pi^\tau(\cdot)) \leq J(s, x; \pi(\cdot))$.

Proof. It follows straightly from Lemma 4.2.1. $\square$

Definition 4.2.4. Let $(s, x) \in [0, T] \times \mathbb{R}$, $\delta > 0$; a strategy $\pi^\delta(\cdot) \geq 0$ is called δ-optimal if

$$J(s, x; \pi^\delta(\cdot)) \leq V(s, x) + \delta.$$

$\square$

Proposition 4.2.5. *Let* $s \in [0, T]$*. The function* $\mathbb{R} \to \mathbb{R}$, $x \mapsto V(s, x)$ *is convex.*

Proof. *Step 1.* In this step we will prove that $x \mapsto V(s,x)$ is convex on $(-\infty, F(s)]$. Let us suppose $x, y \le F(s)$. Let $\delta > 0$ and let $\pi_x^\delta(\cdot), \pi_y^\delta(\cdot)$ two controls δ-optimal for x, y respectively, i.e.

$$J(s,x;\pi_x^\delta(\cdot)) \le V(s,x) + \delta, \qquad J(s,y;\pi_y^\delta(\cdot)) \le V(s,y) + \delta.$$

Set $X(t) := X(t;s,x,\pi_x^\delta(\cdot))$, $Y(t) := X(t;s,y,\pi_y^\delta(\cdot))$. Without loss of generality, thanks to Lemma 4.2.3, we can suppose $X(t), Y(t) \le F(t)$ for all $t \in [s,T]$. We want to prove that, for all $\gamma \in [0,1]$,

$$V(s, \gamma x + (1-\gamma)y) \le \gamma V(s,x) + (1-\gamma)V(s,y).$$

Fix $\gamma \in [0,1]$ and set $Z(t) := \gamma X(t) + (1-\gamma)Y(t)$; of course $Z(t) \le F(t)$, for all $t \in [s,T]$. We have

$$\begin{aligned}\gamma V(s,x) + (1-\gamma)V(s,y) + \delta &\ge \gamma J(s,x;\pi_x^\delta(\cdot)) + (1-\gamma)J(s,y;\pi_y^\delta(\cdot)) \\ &= \gamma \mathbb{E}\left[\int_s^T \kappa e^{-\rho t}(F(t)-X(t))^2 dt + e^{-\rho T}(\bar{F} - X(T))^2\right] \\ &+ (1-\gamma)\mathbb{E}\left[\int_s^T \kappa e^{-\rho t}(F(t)-Y(t))^2 dt + e^{-\rho T}(\bar{F} - Y(T))^2\right] \\ &\ge \mathbb{E}\left[\int_s^T \kappa e^{-\rho t}(F(t)-Z(t))^2 dt + e^{-\rho T}(\bar{F} - Z(T))^2\right], \end{aligned} \tag{4.3}$$

where the last inequality follows by convexity of $\xi \mapsto (F(t)-\xi)^2$. Let us write the dynamics for $Z(\cdot)$:

$$\begin{aligned} dZ(t) &= \gamma dX(t) + (1-\gamma)dY(t) \\ &= \gamma \left[rX(t) + \sigma\lambda\pi_x^\delta(t) - b_0\right] dt + (1-\gamma)\left[rY(t) + \sigma\lambda\pi_y^\delta(t) - b_0\right] dt \\ &\quad + \gamma\sigma\pi_x^\delta(t)dB(t) + (1-\gamma)\sigma\pi_y^\delta(t)dB(t) \\ &= \left[rZ(t) + \sigma\lambda(\gamma\pi_x^\delta(t) + (1-\gamma)\pi_y^\delta(t)) - b_0\right] dt \\ &\quad + \sigma\left(\gamma\pi_x^\delta(t) + (1-\gamma)\pi_y^\delta(t)\right) dB(t). \end{aligned}$$

Thus, if we define $\pi_z(\cdot) := \gamma\pi_x^\delta(\cdot) + (1-\gamma)\pi_y^\delta(\cdot) \ge 0$, we get

$$Z(t) = X(t;s,\gamma x + (1-\gamma)y, \pi_z(\cdot)).$$

Therefore,

$$\mathbb{E}\left[\int_s^T \kappa e^{-\rho t}(F(t)-Z(t))^2 dt + e^{-\rho T}(F(T) - Z(T))^2\right] \ge V(s, \gamma x + (1-\gamma)y) \tag{4.4}$$

and comparing (4.3) with (4.4) we get the claim in this case by the arbitrariness of δ.

Step 2. We can argue exactly as in the step 1 and conclude that $x \mapsto V(s,x)$ is convex on $[F(s), +\infty)$.

Step 3. We can notice that $V(s,\cdot)$ is nonnegative and that, thanks to Lemma 4.2.1, $V(s,F(s)) = 0$, so that $F(s)$ is a minimum for $V(s,\cdot)$. Thus the global convexity of $V(s,\cdot)$ follows from the convexity on the two half lines $(-\infty, F(s)]$, $[F(s), +\infty)$ and from the fact that it has a minimum in $F(s)$. $\square$

Since, for $s \in [0,T]$, $V(s,\cdot)$ is convex and admits a minimum at $x = F(s)$, we have directly the following result. However we give a proof of the statement independent on Proposition 4.2.5.

Proposition 4.2.6. *Let $s \in [0,T]$; the function $x \mapsto V(s,x)$ is decreasing on $(-\infty, F(s)]$ and increasing on $[F(s), +\infty)$.*

Proof. We prove the statement on $(-\infty, F(s)]$; the other one follows as well. So, let $x \le y \le F(s)$, let $\delta > 0$ and $\pi^\delta(\cdot) \ge 0$ a δ-optimal strategy for x, so that

$$V(s,x) + \delta \ge J(s,x;\pi^\delta(\cdot)). \tag{4.5}$$

Set $X(t) := X(t;s,x,\pi^\delta(\cdot))$ and $Y(t) := X(t;s,y,\pi^\delta(\cdot))$. Again, thanks to Lemma 4.2.3, we can suppose without loss of generality that $X(t) \le F(t)$ for all $t \in [s,T]$. By comparison criterion (see, e.g., [Karatzas, Shreve; 1991]) we have $X(t) \le Y(t)$ for all $t \in [s,T]$. Let us define the strategy

$$\tilde{\pi}(t) := \begin{cases} \pi^\delta(s), & \text{if } Y(t) < F(t), \\ 0, & \text{if } Y(t) = F(t); \end{cases}$$

of course $\tilde{\pi}(\cdot) \ge 0$ and, if we set $\tilde{Y}(t) := X(t;s,y,\tilde{\pi}(\cdot))$, again thanks to Lemma 4.2.1 we get $X(t) \le \tilde{Y}(t) \le F(t)$, for all $t \in [s,T]$. Thus, by monotonicity of $L(t,\cdot)$,

$$J(s,x;\pi^\delta(\cdot)) \ge J(s,y;\tilde{\pi}(\cdot)) \tag{4.6}$$

and of course

$$J(s,y;\tilde{\pi}(\cdot)) \ge V(s,y). \tag{4.7}$$

Comparing (4.5), (4.6) and (4.7) we get the claim by the arbitrariness of δ. $\square$

4.2.2 The HJB equation

Since we are in the context of optimal control problems with finite horizon, the HJB equation associated with the value function V is a nonlinear parabolic

PDE with terminal boundary condition. We are going to define this equation. To this aim let us define the Hamiltonian current-value

$$\begin{array}{rccc} \mathcal{H}_{cv}: & \mathbb{R}^2 \times [0,+\infty) & \longrightarrow & \mathbb{R}, \\ & (p,P;\pi) & \longmapsto & \frac{1}{2}\sigma^2 P\pi^2 + \sigma\lambda p\pi. \end{array}$$

and the Hamiltonian

$$\begin{array}{rccc} \mathcal{H}: & \mathbb{R}^2 & \longrightarrow & \mathbb{R}\cup\{-\infty\}, \\ & (p,P) & \longmapsto & \inf_{\pi\geq 0} \mathcal{H}_{cv}(p,P,\pi). \end{array}$$

Given $(p,P) \in \mathbb{R}\times(0,+\infty)$, the function $\pi \mapsto \mathcal{H}_{cv}(p,P;\pi)$ has a unique minimum point on $[0,+\infty)$ given by

$$\pi^* = -\frac{\lambda\, p}{\sigma\, P} \vee 0,$$

so in this case the Hamiltonian can be written as

$$\mathcal{H}(p,P) = \begin{cases} -\frac{\lambda^2 p^2}{2P}, & \text{if } p<0, \\ 0, & \text{if } p \geq 0. \end{cases} \tag{4.8}$$

The HJB equation is

$$\begin{cases} v_s(s,x) + (rx - b_0)v_x(s,x) + \kappa e^{-\rho s}(F(s)-x)^2 + \mathcal{H}\left(v_x(s,x), v_{xx}(s,x)\right) = 0, \\ \qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad \text{on } [0,T]\times\mathbb{R}, \\ v(T,x) = e^{-\rho T}(\bar{F}-T)^2, \quad x\in\mathbb{R}. \end{cases} \tag{4.9}$$

Recall that we have set

$$U_1 := \{(s,x) \mid s\in[0,T],\ x \leq F(s)\}, \quad U_2 := \{(s,x) \mid s\in[0,T],\ x\geq F(s)\}.$$

If we suppose that the value function is smooth on U_1 and on U_2, then, inspired by the previous subsection that gives information on the signs of V_x, V_{xx} on the regions U_1 and U_2, we can split the HJB equation in these two regions. Supposing $V_{xx}>0$, we get that V should satisfy the equation

$$\begin{cases} \kappa e^{-\rho s}(F(s)-x)^2 + v_s(s,x) + (rx-b_0)v_x(s,x) - \dfrac{\lambda^2 v_x^2(s,x)}{2v_{xx}(s,x)} = 0, \\ \qquad\qquad\qquad\qquad\qquad\qquad \text{on } U_1\backslash\{(s,F(s)) \mid s\in[0,T]\}, \\ \kappa e^{-\rho s}(F(s)-x)^2 + v_s(s,x) + (rx-b_0)v_x(s,x) = 0, \\ \qquad\qquad\qquad\qquad\qquad\qquad \text{on } U_2\backslash\{(s,F(s)) \mid s\in[0,T]\}, \end{cases} \tag{4.10}$$

with boundary conditions

$$\begin{cases} v_x(s,F(s)) = 0, & s\in[0,T], \\ v(T,x) = e^{-\rho T}(\bar{F}-x)^2, & x\in\mathbb{R}. \end{cases} \tag{4.11}$$

Definition 4.2.7. A function v is called a classical solution to (4.10)-(4.11) if

- $v \in C^{1,1}([0,T] \times \mathbb{R};\mathbb{R}) \cap C^{1,2}\big(([0,T] \times \mathbb{R})\backslash\{(s,F(s)) \,|\, s \in [0,T]\};\mathbb{R}\big)$,
- v satisfies pointwise in classical sense (4.10) (the derivatives with respect to the time variable at $s = 0$ and $s = T$ have to be intended respectively as right and left derivative)
- v satisfies the boundary conditions (4.11) $\square$

We look for an explicit classical solution to (4.10)-(4.11),

Lemma 4.2.8. *1. Let $v_1(s,x) = e^{-\rho s}A_1(s)(F(s)-x)^2$, where $A_1(\cdot)$ is the unique solution of the ordinary differential equation*

$$\begin{cases} A_1'(s) = \big(\rho + \lambda^2 - 2r\big)\, A_1(s) - \kappa, \\ A_1(T) = 1. \end{cases}$$

i.e., setting $a_1 := \rho + \lambda^2 - 2r$,

$$A_1(s) = \begin{cases} \left(1 - \dfrac{\kappa}{a_1}\right) e^{-a_1(T-s)} + \dfrac{\kappa}{a_1}, & \textit{if } a_1 \neq 0, \\ \kappa(T-s) + 1, & \textit{if } a_1 = 0. \end{cases}$$

Then

(a) $v_{1_x} \leq 0$ on U_1;

(b) $v_{1_{xx}} > 0$ on U_1;

(c) v_1 solves

$$\begin{cases} \kappa e^{-\rho s}(F(s)-x)^2 + v_s(s,x) + (rx - b_0)v_x(s,x) - \dfrac{\lambda^2 v_x^2(s,x)}{2v_{xx}(s,x)} = 0, \\ \hfill \textit{on } [0,T] \times \mathbb{R}, \\ v(T,x) = e^{-\rho T}(\bar{F} - x)^2, \quad x \in \mathbb{R}. \end{cases} \tag{4.12}$$

2. Let $v_2(s,x) = e^{-\rho s}A_2(s)(F(s)-x)^2$, where $A_2(\cdot)$ is the unique solution of the ordinary differential equation

$$\begin{cases} A_2'(s) = (\rho - 2r)\, A_2(s) - \kappa, \\ A_2(T) = 1, \end{cases}$$

i.e., setting $a_2 := \rho - 2r$.

$$A_2(s) = \begin{cases} \left(1 - \dfrac{\kappa}{a_2}\right) e^{-a_2(T-s)} + \dfrac{\kappa}{a_2}, & \textit{if } a_2 \neq 0, \\ \kappa(T-s) + 1, & \textit{if } a_2 = 0. \end{cases}$$

Then

(a) $v_{2_x} \geq 0$ *on* U_2;

(b) $v_{2_{xx}} > 0$ *on* U_2;

(c) v_2 *solves*

$$\begin{cases} \kappa e^{-\rho s}(F(s)-x)^2 + v_s(s,x) + (rx-b_0)v_x(s,x) = 0, & \text{on } [0,T]\times\mathbb{R}, \\ v(T,x) = e^{-\rho T}(\bar{F}-x)^2, & x\in\mathbb{R}. \end{cases} \tag{4.13}$$

3. *For* $s \in [0,T]$, *we have*

(a) $v_1(s,F(s)) = v_2(s,F(s)) = 0$;

(b) $v_{1_t}(s,F(s)) = v_{2_s}(s,F(s)) = 0$;

(c) $v_{1_x}(s,F(s)) = v_{2_x}(s,F(s)) = 0$.

Moreover

- *if* $\lambda = 0$, *then* $v_{1_{xx}}(s,F(s)) = v_{2_{xx}}(s,F(s))$ *for* $s \in [0,T]$;
- *if* $\lambda > 0$, *then* $v_{1_{xx}}(s,F(s)) \neq v_{2_{xx}}(s,F(s))$ *for* $s \in [0,T)$.

Proof. Let us consider, for $A(\cdot) \in C^1([0,T];\mathbb{R})$, the function

$$v(s,x) = e^{-\rho s}A(s)(F(s)-x)^2.$$

We have

$$v_t(s,x) = -\rho e^{-\rho s}A(s)(F(s)-x)^2 + e^{-\rho s}A'(s)(F(s)-x)^2 + 2e^{-\rho s}A(s)(F(s)-x)F'(s),$$

$$v_x(s,x) = -2e^{-\rho s}A(s)(F(s)-x),$$

$$v_{xx}(s,x) = 2e^{-\rho s}A(s).$$

Notice also that

$$F'(s) = rF(s) - b_0.$$

Finally it is immediate to check, splitting the proofs in the cases $a_1 < 0$, $a_1 = 0$, $a_1 > 0$ and $a_2 < 0$, $a_2 = 0$, $a_2 > 0$, that the functions $A_1(\cdot), A_2(\cdot)$ are strictly positive on $[0,T]$. Therefore all the statements follow by simple computations taking $A(\cdot) = A_1(\cdot)$ and $A(\cdot) = A_2(\cdot)$. □

Proposition 4.2.9. *Define on* $[0,T]\times\mathbb{R}$ *the function*

$$v(s,x) := \begin{cases} v_1(s,x), & \text{if } (s,x)\in U_1, \\ v_2(s,x), & \text{if } (s,x)\in U_2. \end{cases} \tag{4.14}$$

Then v *is a classical solution to* (4.10)-(4.11) *in the sense of Definition 4.2.7.*

Proof. It follows from Lemma 4.2.8 □

4.2.3 The Verification Theorem and the optimal feedback strategy

The aim of this subsection is to prove a Verification Theorem stating that the function v defined in (4.14) is actually the value function and moreover giving an optimal feedback strategy for the problem.

Lemma 4.2.10 (Fundamental identity)**.**

1. *Let $(s,x) \in U_1$, let v_1 be the function defined in Lemma 4.2.8-(1) and finally let $\pi(\cdot) \geq 0$ be a strategy such that $X(\cdot;s,x,\pi(\cdot)) \in U_1$. Then*

$$v_1(s,x) = J(s,x;\pi(\cdot))$$
$$+\mathbb{E}\left[\int_s^T \Big(\mathcal{H}\big(v_{1_x}(t,X(t)), v_{1_{xx}}(t,X(t))\big) - \mathcal{H}_{cv}\big(v_{1_x}(t,X(t)), v_{1_{xx}}(t,X(t));\pi(t)\big)\Big)\, dt\right].$$

2. *Let $(s,x) \in U_2$, let v_2 be the function defined in Lemma 4.2.8-(2) and finally let $\pi(\cdot) \geq 0$ be a strategy such that $X(\cdot;s,x,\pi(\cdot)) \in U_2$. Then*

$$v_2(s,x) = J(s,x;\pi(\cdot))$$
$$+\mathbb{E}\left[\int_s^T \Big(\mathcal{H}\big(v_{2_x}(t,X(t)), v_{2_{xx}}(t,X(t))\big) - \mathcal{H}_{cv}\big(v_{2_x}(t,X(t)), v_{2_{xx}}(t,X(t));\pi(t)\big)\Big)\, dt\right].$$

Proof. (1) Let v_1 be the function defined in Lemma 4.2.8-(1); by the same lemma v_1 solves (4.12) on U_1. Since $v_{1_x} \leq 0, v_{1_{xx}} > 0$ on U_1, we get

$$\mathcal{H}(v_{1_x}(s,x), v_{1_{xx}}(s,x)) = -\frac{\lambda^2 v_x^2(s,x)}{2v_{xx}(s,x)}, \quad (s,x) \in U_1,$$

so that v_1 solves the HJB equation (4.9) on U_1. Let us take $\pi(\cdot)$ such that the corresponding state trajectory $X(\cdot) := X(\cdot;s,x,\pi(\cdot))$ remains in U_1 and apply the Dynkin formula to $X(\cdot)$ with the function v_1; we get

$$\mathbb{E}\left[v_1(T,X(T)) - v_1(s,x)\right] = \mathbb{E}\Bigg[\int_s^T \Big(v_{1_s}(t,X(t)) + (rX(t)-b_0)\, v_{1_x}(t,X(t))$$
$$+ \mathcal{H}_{cv}\big(v_{1_x}(t,X(t)), v_{1_{xx}}(t,X(t));\pi(t)\big)\Big)\, dt\Bigg],$$

i.e.

$$v_1(s,x) = \mathbb{E}\Bigg[e^{-\rho T}(\bar{F} - X(T))^2 - \int_s^T \Big(v_{1_s}(t,X(t)) + (rX(t)-b_0)\, v_{1_x}(t,X(t))$$
$$+ \mathcal{H}_{cv}\big(v_{1_x}(t,X(t)), v_{1_{xx}}(t,X(t));\pi(t)\big)\Big)\, dt\Bigg].$$

Taking into account the assumption on $\pi(\cdot)$ and the fact that, as shown, v_1 solves the originary HJB equation on U_1, we can write

$$\begin{aligned}
v_1(s,x) = \mathbb{E}\Bigg[& e^{-\rho T}(\bar{F}-X(T))^2 + \int_s^T \kappa e^{-\rho t}(F(t)-X(t))^2\,ds \\
+ \int_s^T & \Big(\mathcal{H}\big(v_{1_x}(t,X(t)), v_{1_{xx}}(t,X(t))\big) - \mathcal{H}_{cv}\big(v_{1_x}(t,X(t)), v_{1_{xx}}(t,X(t));\pi(t)\big)\Big)dt\Bigg] \\
& = J(s,x;\pi(\cdot)) + \mathbb{E}\Bigg[\int_s^T \Big(\mathcal{H}\big(v_{1_x}(t,X(t)), v_{1_{xx}}(t,X(t))\big) \\
& \qquad - \mathcal{H}_{cv}\big(v_{1_x}(t,X(t)), v_{1_{xx}}(t,X(t));\pi(t)\big)\Big)dt\Bigg].
\end{aligned}$$

(2) Let v_2 be the function defined in Lemma 4.2.8-(2); by the same lemma v_2 solves (4.13) on U_2. Since $v_{2_x} \geq 0, v_{2_{xx}} > 0$ on U_1, we get

$$\mathcal{H}(v_{2_x}(s,x), v_{2_{xx}}(s,x)) = 0, \quad (s,x) \in U_2,$$

so v_2 solves the HJB equation (4.9) on U_2. Now the proof follows the same line of the proof of the previous statement. □

Lemma 4.2.11. .

1. *Let $(s,x) \in U_1$. There exists a unique process $X(\cdot)$ solution of the equation*

$$\begin{cases} dX(t) = \left[rX(t) + \lambda^2(F(t)-X(t)) - b_0\right]dt + \lambda(F(t)-X(t))dB(t), \\ X(s) = x. \end{cases}$$

   Moreover the process $X(\cdot)$ is such that $(t, X(t))$, $t \in [s,T]$, lives in U_1.

2. *Let $(s,x) \in U_2$. The (deterministic) process $X(\cdot) := X(\cdot; s, x, 0)$ is such that $(t, X(t))$, $t \in [s,T]$, lives in U_2.*

Proof. **(1)** The proof of the existence and uniqueness of X is standard (see, e.g., [Karatzas, Shreve; 1991], Chapter 5, Theorem 2.9). About the second part of the statement, notice that, if write the dynamics of $Z(\cdot) := F(\cdot) - X(\cdot)$ in the interval $[s,T]$, we get

$$\begin{cases} dZ(t) = \left(r-\lambda^2\right)Z(t)dt - \lambda Z(t)dB(t), \\ Z(s) = F(s) - x. \end{cases}$$

Therefore $Z(\cdot)$ is a geometric Brownian motion with positive starting point, so that it has to be positive and this claim is proved.

(2) Let $(s,x) \in U_2$; the explicit expression of $X(t) := X(t; s, x, 0)$ is

$$X(t) = \frac{b_0}{r} + \left(x - \frac{b_0}{r}\right) e^{r(t-s)},$$

Comparing this expression with (4.2) we get the claim. $\square$

Let us define the feedback map

$$(s,y) \mapsto G(s,y) := \begin{cases} \frac{\lambda}{\sigma}\,(F(s) - y), & \text{if } (s,y) \in U_1, \\ 0, & \text{if } (s,y) \in U_2. \end{cases} \tag{4.15}$$

Notice that, thanks to Lemma 4.2.8-(1a), we have

$$F(s) - y = -\frac{v_{1_x}(s,y)}{v_{1_{xx}}(s,y)}. \tag{4.16}$$

Moreover, by definition of G and due to Lemma 4.2.11, the strategy $\pi_G^{s,x}(\cdot)$ defined by the feedback map (4.15), i.e.

$$\pi_G^{s,x}(t) := G(t, X_G(t; s, x)), \quad t \in [s,T], \tag{4.17}$$

where $X_G(\cdot; s, x)$ is the solution of the closed loop equation

$$\begin{cases} dX(t) = [rX(t) + \sigma\lambda G(X(t)) - b_0]\, dt + \sigma G(X(t)) dB(t), \\ X(s) = x, \end{cases}$$

is admissible.

Theorem 4.2.12 (Verification). *Let $(s,x) \in [0,T] \times \mathbb{R}$ let v be the function defined in* (4.14). *Then $V(s,x) = v(s,x)$. Moreover $\pi(\cdot) \in \Pi_{ad}$ is optimal for the initial (s,x) if and only if*

$$\pi(t) = G(t, X(t; s, x, \pi(\cdot)), \quad \mathbb{P} - a.s., \; \forall t \in [s,T]. \tag{4.18}$$

In particular the feedback strategy $\pi_G^{s,x}(\cdot)$ defined in (4.17) *is the unique optimal strategy starting from the initial (s,x).*

Proof. Let $(s,x) \in U_1$ and $\pi(\cdot) \in \Pi_{ad}$ and set $X(\cdot) := X(\cdot; s, x, \pi(\cdot))$. Let us suppose $X(\cdot) \in U_1$. Thus we can apply Lemma 4.2.10 to $X(\cdot)$ with v_1 getting

$$\begin{aligned} &v_1(t,x) = J(t,x;\pi(\cdot)) \\ &+ \mathbb{E}\left[\int_s^T \Big(\mathcal{H}\big(v_{1_x}(t,X(t)), v_{1_{xx}}(t,X(t))\big) - \mathcal{H}_{cv}\big(v_{1_x}(t,X(t)), v_{1_{xx}}(t,X(t)); \pi(t)\big)\Big)\, dt\right] \\ &\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad \leq J(s,x;\pi(\cdot)). \end{aligned} \tag{4.19}$$

Taking into account Lemma 4.2.3, this shows that $v_1(s,x) \le V(s,x)$.

Now consider $X(\cdot; s, x, \pi_G^{s,x}(\cdot)) = X_G(\cdot; s, x) =: X_G(\cdot)$. We see from Lemma 4.2.11-(1) that also $X_G(\cdot) \in U_1$, so that we can apply the fundamental identity also to $X_G(\cdot)$ with v_1. Taking into account (4.16) we see that, by (4.8) and by Lemma 4.2.8-(1b), the feedback map minimizes at any time $t \in [s,T]$ the Hamiltonian current value. Thus we get in this case $v_1(s,x) = J(s,x;\pi_G^{s,x}(\cdot))$, which shows that $v_1(s,x) = V(s,x) = J(s,x;\pi_G^{s,x}(\cdot))$.

The fact that an optimal strategy must satisfy (4.18) is consequence of the first claim and of (4.19).

Finally the uniqueness of the optimal strategy is consequence of the characterization (4.18) and of the uniqeness of solutions to the closed loop equation stated in Lemma 4.2.11-(1).

If $(s,x) \in U_2$ we can argue exactly in the same way with v_2 getting the claim also in this case. □

Remark 4.2.13. Theorem 4.2.12 and Lemma 4.2.8-(3) say that

- if $\lambda = 0$, then the value function is C^2 with respect to the state variable;
- if $\lambda > 0$, then the value function is C^1 with respect to the state variable, but not C^2. □

4.3 Constraints on the state and on the control

In this section we study the problem with the additional capital requirement (state constraint) that $X(T) \ge \bar{S}$ almost surely, where $0 \le \bar{S} < \bar{F}$. It will turn out that this constraint implies the "no ruin" constraint, i.e. $X(t) \ge 0$ almost surely for every $t \in [0,T]$.

The setting is very similar to the one of Section 1.2, so that we will refer to that section for the proofs of some results.

For $s \in [0,T]$, $x \in \mathbb{R}$, let $\pi(\cdot)$ be progressively measurable with respect to $\mathcal{F}^{B^s}$, where $B^s(t) = B(t) - B(s)$, $t \in [s,T]$, $\pi(\cdot) \ge 0$ and let $X(\cdot; s, x, \pi(\cdot))$ be the unique solution to

$$\begin{cases} dX(t) = [rX(t) + \sigma\lambda\pi(t) - b_0]\, dt + \sigma\pi(t) dB^s(t), \\ X(s) = x. \end{cases}$$

The set of admissible strategies, in this case depending on the initial (s,x), is

$$\Pi^0_{ad}(s,x) = \big\{\pi(\cdot) \text{ progressively measurable with respect to } \mathcal{F}^{B^s}, \\ \pi(\cdot) \in L^2(\Omega \times [s,T]; [0,+\infty)) \;\big|\; X(T; s, x, \pi(\cdot)) \ge \bar{S}\big\}.$$

The objective functional is again, for $s \in [0,T],\ x \in \mathbb{R}$,

$$J(s,x;\pi(\cdot)) := \mathbb{E}\left[\int_s^T \kappa e^{-\rho t}(F(t)-X(t;s,x,\pi(\cdot)))^2 dt + e^{-\rho T}(\bar{F}-X(T;s,x,\pi(\cdot)))^2\right].$$

We denote the value function by W, so that

$$W(s,x) = \sup_{\pi(\cdot)\in\Pi^0_{ad}(s,x)} J(s,x;\pi(\cdot)).$$

4.3.1 The set of admissible strategies

In this subsection we analyze some properties of the set of admissible strategies and rewrite the problem in a more convenient form in the region that is meaningful from the financial point of view.

Let us set

$$S(t) := \frac{b_0}{r} - \left(\frac{b_0}{r} - \bar{S}\right) e^{-r(T-t)}. \tag{4.20}$$

Proposition 4.3.1. *Let $s \in [0,T]$, $x \in \mathbb{R}$. We have*

1. $\Pi^0_{ad}(s,x) \neq \emptyset$ *if and only if* $0 \in \Pi_{ad}(s,x)$.
2. $0 \in \Pi^0_{ad}(s,x)$ *if and only if* $x \geq S(s)$.
3. *If* $x = S(s)$, *then* $\Pi^0_{ad}(s,x) = \{0\}$.
4. *If* $x > S(s)$, *then* $\Pi^0_{ad}(s,x) \supsetneq \{0\}$.
5. *The state constraint* $X(T) \geq \bar{S}$ *is equivalent to*

$$X(t) \geq S(t),\ \ \mathbb{P} - a.s.\ \forall t \in [0,T].$$

□

Proof. 1. Clearly, if $0 \in \Pi^0_{ad}(s,x)$, then $\Pi^0_{ad}(s,x) \neq \emptyset$. Conversely suppose that $\Pi^0_{ad}(s,x) \neq \emptyset$ and let $\pi(\cdot) \in \Pi^0_{ad}(s,x)$. This means that $X(T;s,x,\pi(\cdot)) \geq \bar{S}$ almost surely, therefore $\tilde{\mathbb{E}}[X(T;s,x,\pi(\cdot))] \geq \bar{S}$, where $\tilde{\mathbb{E}}$ denotes the expectation under the probability $\tilde{\mathbb{P}} = e^{-\lambda B(T)-\frac{\lambda^2}{2}} \cdot \mathbb{P}$ given by the Girsanov transformation (see Theorem A.1.1). As, e.g., in Proposition 1.12, we have $X(T;s,x,0) = \tilde{\mathbb{E}}[X(T;s,x,\pi(\cdot)] \geq \bar{S}$, so that $0 \in \Pi^0_{ad}(s,x)$.

2. The state equation yields

$$X(t;s,x,0) = \frac{b_0}{r} - \left(\frac{b_0}{r} - x\right) e^{r(t-s)},$$

so that from the expression of $S(\cdot)$ in (4.20) we get the claim.

3. If $x = S(s)$, we have $X(t; s, x, 0) = S(t)$ on $[s, T]$; therefore $0 \in \Pi^0_{ad}(s, x)$. On the other hand, arguing as in Proposition 1.2.6, one can see that

$$\pi(\cdot) \in \Pi^0_{ad}(s, x) \Longrightarrow \pi(\cdot) \equiv 0,$$

so the claim.

4. Let $x > S(s)$. Define the stopping time

$$\tau := \inf\{t \geq \in [s, T] \mid X(t; s, x, 1) = S(t)\} \wedge T.$$

Define the strategy

$$\pi^\tau(t) := \begin{cases} \pi(t), & \text{if } t \in [s, \tau], \\ 0, & \text{if } t \in [\tau, T]. \end{cases}$$

Then $\pi^\tau(\cdot) \in \Pi^0_{ad}(s, x)$ and $\pi(\cdot)$ is not identically null, so the claim.

5. The claim reduces to show that, for every $\pi(\cdot) \in \Pi^0_{ad}(s, x)$, we have $X(t) \geq S(t)$ almost surely for every $t \in [s, T]$. This follows arguing by contradiction and using the previous item. □

Due to Proposition 4.3.1, the value function is defined on the set

$$\mathcal{D} = \{(s, x) \in [0, T] \times \mathbb{R} \mid x \geq S(s)\}.$$

Consider the set

$$\mathcal{C} := \{(s, x) \in [0, T] \times \mathbb{R} \mid S(s) \leq x \leq F(s)\} \subset \mathcal{D}.$$

Since the optimal strategy $\pi^*(\cdot) \equiv 0$ of the (state) unconstrained problem starting from a point of the set $(s, x) \in \mathcal{D} \setminus \mathcal{C}$ satisfies $X(t; s, x, 0) \geq F(t)$, we have $W = V$ on this set, where V is the value function of the (state) unconstrained problem studied in the previous section. In other words the state constrained problem is already solved on the region $\mathcal{C}^c$ keeping the strategy $\pi^*(\cdot) \equiv 0$.

For the points belonging to $\mathcal{C}$ we have the following representation for the value function W.

Proposition 4.3.2. *Let $(s, x) \in \mathcal{C}$ and consider the set of admissible strategies*

$$\begin{aligned} \Pi_{ad}(s, x) = \big\{ & \pi(\cdot) \textit{ progressively measurable with respect to } \left(\mathcal{F}_t^{B^s}\right)_{t \in [s,T]}, \\ & \pi(\cdot) \in L^2(\Omega \times [s, T]; [0, +\infty)), \\ & S(t) \leq X(t; s, x, \pi(\cdot)) \leq F(t),\ t \in [s, T] \big\} \subset \Pi^0_d(s, x). \end{aligned}$$

Then we have

$$W(s, x) = \sup_{\pi(\cdot) \in \Pi_{ad}(s,x)} J(s, x; \pi(\cdot)).$$

Proof. The argument of Lemma 4.2.3 yields the claim. □

Proposition 4.3.2 says that on the set $\mathcal{C}$ the original problem is equivalent to the problem with state constraint

$$S(t) \leq X(t) \leq F(t), \quad t \in [s, T].$$

The analogue of Proposition 4.3.1 is the following.

Proposition 4.3.3. *Let $(s, x) \in \mathcal{C}$. We have*

- $0 \in \Pi_{ad}(s, x)$.
- *If $x = S(s)$ or $x = F(s)$, then $\Pi_{ad}(s, x) = \{0\}$.*
- *If $S(s) < x < F(s)$, then $\Pi_{ad}(s, x) \supsetneq \{0\}$.*

Proof. The claims can be obtained as in the proof of Proposition 4.3.1. □

Notice that, rephrasing the problem in these new terms, both the lateral boundaries

$$\partial^*_F \mathcal{C} := \{(s, x) \in [0, T] \times \mathbb{R} \mid x = F(s)\}, \quad \partial^*_S \mathcal{C} := \{(s, x) \in [0, T] \times \mathbb{R} \mid x = S(s)\},$$

are absorbing for the problem, in the sense that, if $x = S(s)$ (respectively $x = F(s)$), then the only admissible strategy is $\pi(\cdot) \equiv 0$ and $X(t; s, x, 0) = S(t)$ for $t \in [s, T]$ (respectively $X(t; s, x, 0) = F(t)$ for $t \in [s, T]$).

4.3.2 Properties of the value function

As in Section 1.2 we have the following properties for the value function W.

Proposition 4.3.4. *Let $s \in [0, T]$. The function $[S(s), F(s)] \to \mathbb{R}$, $x \mapsto W(s, x)$ is convex.*

Proof. It follows the line of the proof of Proposition 1.2.10. □

Proposition 4.3.5. *Let $s \in [0, T]$. The function $[S(s), F(s)] \to \mathbb{R}$, $x \mapsto W(s, x)$ is decreasing.*

Proof. It follows the line of Proposition 1.2.11 □

Proposition 4.3.6. *The function W is continuous on $\mathcal{C}$.*

Proof. It follows the line the case $\beta = r$ of Subsection 1.2.4 □.

4.3.3 The HJB equation

We want to associate to the control problem an HJB equation. Thanks to Subsection 4.3.2 this equation is the same as the one of Section 4.2 and differs from it because of the presence of Dirichlet lateral boundary conditions coming from the state constraints. Calling

$$\mathrm{Int}^*(\mathcal{C}) := \{(s,x) \in \mathcal{C} \mid s \in [0,T),\ S(s) < x < F(s)\}$$

the HJB equation reads as

$$w_s(s,x) + (rx - b_0)w_x(s,x) + \kappa e^{-\rho s}(F(s) - x)^2 + \mathcal{H}(w_x(s,x), w_{xx}(s,x)) = 0, \\ (s,x) \in \mathrm{Int}^*(\mathcal{C}), \quad (4.21)$$

with boundary conditions

$$\begin{cases} w(s,F(s)) = 0,\ s \in [0,T]; \\ w(s,S(s)) = g(s) + e^{-\rho T}(\bar{F} - \bar{S})^2,\ s \in [0,T]; \\ w(T,x) = e^{-\rho T}(\bar{F} - x)^2,\ x \in [\bar{S}, \bar{F}], \end{cases} \quad (4.22)$$

where

$$g(s) = \kappa(\bar{F} - \bar{S})^2 \int_s^T e^{-(\rho+2r)(T-t)} dt. \quad (4.23)$$

Next we give the definitions of classical and viscosity solution for (4.21)-(4.22).

Definition 4.3.7. A function w is called a classical solution to (4.21)-(4.22) if

- $w \in C(\mathcal{C};\mathbb{R}) \cap C^{1,2}(\mathrm{Int}^*(\mathcal{C});\mathbb{R})$,
- w satisfies pointwise in classical sense (4.21) (the derivative with respect to the time variable at $s = 0$ has to be intended as right derivative)
- w satisfies the boundary Dirichlet conditions (4.22). □

Definition 4.3.8. A function w is called a viscosity solution to (4.21)-(4.22) if

- $w \in C(\mathcal{C};\mathbb{R})$;
- w is a viscosity subsolution of (4.21) on $\mathrm{Int}^*(\mathcal{C})$, i.e. for every $(s_M, x_M) \in \mathrm{Int}^*(\mathcal{C})$ and $\varphi \in C^{1,2}(\mathrm{Int}^*(\mathcal{C});\mathbb{R})$ such that $w - \varphi$ has a local maximum at (s_M, x_M) we have

$$- \varphi_s(s_M,x_M) - (rx_M - b_0)\varphi_x(s_M,x_M) - \kappa e^{-\rho s_M}(F(s_M) - x_M)^2 \\ - \mathcal{H}(\varphi_x(s_M,x_M), \varphi_{xx}(s_M,x_M)) \leq 0; \quad (4.24)$$

- w is a viscosity supersolution of (4.21) on $\mathrm{Int}^*(\mathcal{C})$, i.e. for every $(s_m, x_m) \in \mathrm{Int}^*(\mathcal{C})$ and $\varphi \in C^{1,2}(\mathrm{Int}^*(\mathcal{C}); \mathbb{R})$ such that $w - \varphi$ has a local minimum at (s_m, x_m) we have

$$\begin{aligned}- \varphi_s(s_m, x_m) - (rx_m - b_0)\varphi_x(s_m, x_m) - \kappa e^{-\rho s_m}(F(s_m) - x_m)^2 \\ - \mathcal{H}(\varphi_x(s_m, x_m), \varphi_{xx}(s_m, x_m)) \geq 0;\end{aligned}$$

- w satisfies the Dirichlet boundary conditions (4.22). □

Now we operate a transformation to get a nicer domain. Let us consider the diffeomorphism

$$\begin{array}{cccc}\mathcal{L}: & [0,T] \times [\bar{S}, \bar{F}] & \longrightarrow & \mathcal{C}, \\ & (s,z) & \longmapsto & (s,x) = \mathcal{L}(s,z) := \left(s, ze^{-r(T-s)} + \frac{b_0}{r}\left(1 - e^{-r(T-s)}\right)\right).\end{array}$$

Equation (4.21)-(4.22) transformed via the diffeomorphism $\mathcal{L}$ leads to consider the equation

$$\begin{aligned}h_s(s,z) + \kappa e^{-\rho s}(F(s) - \mathcal{L}(s,z))^2 + \mathcal{H}(h_z(s,z), h_{zz}(s,z)) = 0, \\ (s,z) \in [0,T) \times (\bar{S}, \bar{F}), \quad (4.25)\end{aligned}$$

with boundary conditions

$$\begin{cases} h(s, \bar{F}) = 0, \ s \in [0,T]; \\ h(s, \bar{S}) = g(s) + e^{-\rho T}(\bar{F} - \bar{S})^2, \ s \in [0,T]; \\ h(T,z) = e^{-\rho T}(\bar{F} - z)^2, \ z \in [\bar{S}, \bar{F}], \end{cases} \tag{4.26}$$

where $g(s)$ was defined in (4.23). We note that (4.25)-(4.26) is associated with the control problem

$$\sup_{\pi(\cdot) \in \tilde{\Pi}_{ad}(s,z)} E\left[\int_s^T \kappa e^{-\rho t}(F(s) - \mathcal{L}(t, Z(t; s, x, \pi(\cdot))))^2 dt + (\bar{F} - Z(T; s, z, \pi(\cdot)))^2\right],$$

where

$$\begin{cases} dZ(t) = \sigma\lambda\pi(t)dt + \sigma\pi(t)dB(t), \\ Z(s) = z, \end{cases}$$

and

$$\begin{aligned}\tilde{\Pi}_{ad}(s,z) = \big\{\pi(\cdot) \text{ progressively measurable with respect to } \left(\mathcal{F}_t^{B^s}\right)_{t \in [s,T]}, \\ \pi(\cdot) \in L^2(\Omega \times [s,T]; [0, +\infty)), \\ \bar{S} \leq Z(t; s, z, \pi(\cdot)) \leq \bar{F}, \ t \in [s,T]\big\}.\end{aligned}$$

Remark 4.3.9. Although a priori the stochastic optimal control problem associated with (4.25)-(4.26) can be defined on any probability space, we see it defined on the originary probability space $(\Omega, \mathcal{F}, \mathbb{P})$ and taking the same Brownian motion B. In this way we have the relationship

$$Z(t; s, z, \pi(\cdot)) = [\mathcal{L}(t, \cdot)]^{-1} \left(X(t; s, x, \pi(\cdot))\right),$$

where $z = [\mathcal{L}(s, \cdot)]^{-1}(x)$, $x \in [S(s), F(s)]$. □

Let us give also the definition of viscosity solution to (4.25)-(4.26).

Definition 4.3.10. A function h is called a viscosity solution to (4.25)-(4.26) if

- $h \in C([0, T] \times [\bar{S}, \bar{F}]; \mathbb{R})$;
- h is a viscosity subsolution of (4.25) on $[0, T) \times (\bar{S}, \bar{F})$, i.e. for every $(s_M, z_M) \in [0, T) \times (\bar{S}, \bar{F})$ and $\varphi \in C^{1,2}([0, T) \times (\bar{S}, \bar{F}); \mathbb{R})$ such that $h - \varphi$ has a local maximum at (s_M, z_M) we have

$$\begin{aligned} &- \varphi_s(s_M, z_M) - \kappa e^{-\rho s_M} (F(s_M) - \mathcal{L}(s_M, z_M))^2 \\ &\qquad\qquad - \mathcal{H}(\varphi_z(s_M, z_M), \varphi_{zz}(s_M, z_M)) \leq 0; \end{aligned}$$

- h is a viscosity supersolution of (4.25) on $[0, T) \times (\bar{S}, \bar{F})$, i.e. for every $(s_m, z_m) \in [0, T) \times (\bar{S}, \bar{F})$ and $\varphi \in C^{1,2}([0, T) \times (\bar{S}, \bar{F}); \mathbb{R})$ such that $h - \varphi$ has a local minimum at (s_m, z_m) we have

$$\begin{aligned} &- \varphi_s(s_m, z_m) - \kappa e^{-\rho s_m} (F(s_m) - \mathcal{L}(s_m, z_m))^2 \\ &\qquad\qquad - \mathcal{H}(\varphi_z(s_m, z_m), \varphi_{zz}(s_m, z_m)) \geq 0; \end{aligned}$$

- h satisfies the Dirichlet boundary conditions (4.26). □

Let us make more precise the relationship between (4.21)-(4.22) and (4.25)-(4.26). We have the following.

Proposition 4.3.11.

1. *Equation* (4.21)-(4.22) *admits a unique viscosity solution if and only if equation* (4.25)-(4.26) *admits a unique viscosity solution. Moreover, in this case, h is the unique viscosity solution to* (4.25)-(4.26) *if and only*

$$w(s, x) := h(s, [\mathcal{L}(s, \cdot)]^{-1}(x)), \quad (s, x) \in \mathcal{C}, \tag{4.27}$$

is the unique viscosity solution to (4.21)-(4.22).

2. *Uniqueness of viscosity solutions holds for equation* (4.25)-(4.26) *(so, by the previous item, also for equation* (4.21)-(4.22)*).*

3. *The value function W is a viscosity solution of* (4.21)-(4.22)*. Therefore, in particular, existence holds for equation* (4.21)-(4.22) *(so, by item (1), also for equation* (4.25)-(4.26)*).*

Proof. 1. This claim is consequence of straightforward computations transforming equation (4.21)-(4.22) into equation (4.25)-(4.26) via the diffeomorphism $\mathcal{L}$.

2. This claim follows from Theorem V.8.1 of [Fleming, Soner; 1993]. Indeed, condition V-(7.1) of [Fleming, Soner; 1993] is easily verified in our case.

3. The proof that the value function W is a viscosity solution to HJB (4.21)-(4.22) basically follows the line of Theorem 1.2.23. We do not give the proof here, but we point out that we must only take care of the fact that the space of controls is unbounded here. □

Thanks to Proposition 4.3.11 we have obtained the following characterization of the value function W.

Theorem 4.3.12. *The value function W is the unique viscosity solution to* (4.21)-(4.22)*.* □

4.3.4 An example with explicit solution

This subsection is devoted to the study of a case allowing an explicit solution for the problem (when the running cost is null, i.e. $\kappa = 0$). Such a solution is not trivial and presents an interesting link with the Black-Scholes formula for the price of the European put option (see Remark 4.3.15).

Hereafter we assume that the risk premium λ is strictly positive. The case $\lambda = 0$, less meaningful from a financial point of view, can be treated apart: it could be seen quite easily that in this case the optimality consists simply in keeping the null strategy, i.e. in investing the whole wealth in the riskless asset.

We make a further transformation to get a linear equation. This method can be successfully used in the case of HJB equation coming from optimal portfolio allocation problems, for which the nonlinearity in the equation takes the form v_x^2/v_{xx}. We refer, e.g., to the papers [Elie, Touzi; 2008], [Gao; 2008], [Højgaard, Vigna; 2007], [Milevsky, Moore, Young; 2006], [Milevsky, Young; 2007].

So, let $\kappa = 0$; in this case we can suppress the term $e^{-\rho T}$ appearing in the bequest functional, i.e. considering without loss of generality $\rho = 0$. We will do it for sake of simplicity. In this case (4.25)-(4.26) becomes

$$h_s(s,z) + \mathcal{H}(h_z(s,z), h_{zz}(s,z)) = 0, \qquad (s,z) \in [0,T) \times (\bar{S}, \bar{F}), \tag{4.28}$$

with boundary conditions

$$\begin{cases} h(s,\bar{F}) = 0, \ s \in [0,T]; \\ h(s,\bar{S}) = (\bar{F}-\bar{S})^2, \ s \in [0,T]; \\ h(T,z) = (\bar{F}-z)^2, \ z \in [\bar{S},\bar{F}], \end{cases} \tag{4.29}$$

Let us give the definition of classical solution to (4.28)-(4.29).

Definition 4.3.13. A function h is called a classical solution to (4.28)-(4.29) if

- $h \in C([0,T] \times [\bar{S},\bar{F}];\mathbb{R}) \cap C^{1,2}([0,T) \times (\bar{S},\bar{F});\mathbb{R})$,
- h satisfies pointwise in classical sense (4.28) (the derivative with respect to the time variable at $s=0$ has to be intended as right derivative),
- h satisfies the boundary Dirichlet conditions (4.29). □

Suppose that the unique viscosity solution h to (4.28)-(4.29) is such that $h \in C^{1,2}([0,T) \times (\bar{S},\bar{F}))$. Suppose also that for every $s \in [0,T]$

$$h_z(s,z) < 0, \ \forall z \in (\bar{S},\bar{F}), \quad h_{zz}(s,z) > 0, \ \forall z \in (\bar{S},\bar{F}). \tag{4.30}$$

i.e. in particular $h(s,\cdot)$ is strictly decreasing and strictly convex. Moreover suppose that

$$\lim_{z \downarrow \bar{S}} h_z(s,z) = -\infty, \quad \forall s \in [0,T). \tag{4.31}$$

Due to (4.30), taking into account the structure (4.8) for the Hamiltonian, we see that h satisfies in classical sense

$$h_s(s,z) - \frac{\lambda^2 h_z^2(s,z)}{2\, h_{zz}(s,z)} = 0, \qquad (s,z) \in [0,T) \times (\bar{S},\bar{F}), \tag{4.32}$$

and is the unique classical solution to (4.28)-(4.29). Given such a solution h, for every $(s,y) \in [0,T] \times [0,+\infty)$, there exists a unique minimizer $g(s,y) \in [\bar{S},\bar{F}]$ of the function $[\bar{S},\bar{F}] \to \mathbb{R}^+$, $z \mapsto h(s,z) + zy$. Since h is continuous, also g is continuous.

Let us look at the behaviour of g for fixed $s \in [0,T)$. First of all we note that, since $h(s,\cdot)$ is decreasing, also $g(s,\cdot)$ is decreasing. Moreover, for every $y \in (0,+\infty)$ the minimizer $g(s,y)$ of $z \mapsto h(s,z) + zy$ belongs to $(\bar{S},\bar{F})$ and is characterized by the relationship

$$h_z(s, g(s,y)) = -y. \tag{4.33}$$

Finally we have

$$g(s,0) = \bar{F}; \quad \lim_{y\to+\infty} g(s,y) = \bar{S}.$$

When $s = T$, the unique minimizer $g(T,y) \in [\bar{S}, \bar{F}]$ of

$$[\bar{S}, \bar{F}] \ni z \mapsto h(T,z)$$

is explicitely computable, since $h(T,\cdot)$ is known. Indeed

$$g(T,y) = \left(\bar{F} - \frac{y}{2}\right) \vee \bar{S}.$$

Proposition 4.3.14. *Suppose that the unique viscosity solution h to* (4.28)-(4.29) *is of class $C^{1,3}([0,T) \times (\bar{S}, \bar{F}); \mathbb{R})$ and satisfies* (4.30)-(4.31) *and let g be defined as above. Then g is a classical solution (in a sense analogous to Definitions 4.3.7-4.3.13) of*

$$g_s(s,y) + \lambda^2 y g_y(s,y) + \frac{\lambda^2}{2} y^2 g_{yy}(s,y) = 0 \ \text{ on } [0,T) \times (0,+\infty), \tag{4.34}$$

with boundary conditions

$$\begin{cases} g(s,0) = \bar{F}, \quad s \in [0,T]; \\ g(T,y) = \left(\bar{F} - \frac{y}{2}\right) \vee \bar{S}, \quad y \in (0,+\infty); \\ \lim_{y\to+\infty} g(s,y) = \bar{S}, \ \ s \in [0,T). \end{cases} \tag{4.35}$$

Conversely, if $g \in C([0,T] \times [0,+\infty); \mathbb{R}) \cap C^{1,2}([0,T) \times (0,+\infty); \mathbb{R})$ is a classical solution to (4.34)-(4.35) *satisfying*

$$g(s,\cdot) \in [\bar{S}, \bar{F}], \ \ g_y(s,\cdot) < 0, \quad \forall s \in [0,T), \tag{4.36}$$

then $h \in C^{1,3}([0,T) \times (\bar{S}, \bar{F}); \mathbb{R})$ defined by

$$\begin{cases} h(s,z) := \bar{S} - \displaystyle\int_{\bar{S}}^{z} [g(s,\cdot)]^{-1}(\xi) d\xi, & (s,z) \in [0,T) \times [\bar{S}, \bar{F}], \\ h(T,z) = (\bar{F} - z)^2, & z \in [\bar{S}, \bar{F}], \end{cases} \tag{4.37}$$

is a classical solution (actually the unique classical solution) to (4.28)-(4.29) *and satisfies* (4.30)-(4.31).

Proof. Let h be the unique viscosity solution to (4.28)-(4.29) and suppose that $h \in C^{1,3}([0,T) \times (\bar{S}, \bar{F}); \mathbb{R})$ and that it satisfies (4.30)-(4.31). Let g be defined as above. Due to (4.30), we know that h satisfies (4.32). Deriving this equation with respect to z we get

$$h_{sz}(s,z) = \frac{\lambda^2}{2} \frac{2h_z(s,z)h_{zz}(s,z)^2 - h_z^2 h_{zzz}(s,z)}{h_{zz}(s,z)^2}, \quad (s,z) \in [0,T) \times (\bar{S}, \bar{F}). \tag{4.38}$$

Deriving (4.33) with respect to s, y and twice with respect to z yields

$$h_{sz}(s,g(s,y)) + h_{zz}(s,g(s,y))g_s(s,y) = 0, \tag{4.39}$$

$$h_{zz}(s,g(s,y))g_y(s,y) = -1, \tag{4.40}$$

$$h_{zzz}(s,g(s,y))g_y^2(s,y) + h_{zz}(s,g(s,y))g_{yy}(s,y) = 0. \tag{4.41}$$

Plugging (4.39)–(4.41) into (4.38) yields

$$\frac{g_s(s,y)}{g_y(s,y)} = \frac{\lambda^2}{2}\left[-2y - y^2\frac{g_{yy}(s,y)}{g_y(s,y)}\right], \quad (s,y)\in[0,T)\times(0,+\infty).$$

Multiplying by g_y we get (4.34). The boundary conditions (4.35) for g are easily verified starting from the boundary conditions (4.29) for h.

Conversely, let $g \in C([0,T]\times[0,+\infty);\mathbb{R}) \cap C^{1,2}([0,T)\times(0,+\infty);\mathbb{R})$ be a classical solution to (4.34)-(4.35) satisfying (4.36) and let h be defined by (4.37). The structure of g (we have to consider also the probabilistic representation of *the* solution g given below and Remark 4.3.15 for an idea about the shape of g and its derivatives) yields (4.30)-(4.31) for h.

Using backward the argument above we get that h solves (4.38). Integrating (4.38) with respect to s we see that, for some continuus function $C(s)$,

$$h_s(s,z) - \frac{\lambda^2 h_z^2(s,z)}{2\,h_{zz}(s,z)} = C(s), \qquad (s,z)\in[0,T)\times(\bar{S},\bar{F}), \tag{4.42}$$

Deriving h with respect to s, z and twice with respect to z and computing these derivatives at $(0,\bar{F}^-)$ yields (again we have to consider the probabilistic representation of g for an idea about the shape of g and about the behaviour at $y=0^+$ of g and g_y)

$$h_s(s,\bar{F}^-) = 0,\; h_z(s,\bar{F}^-) = 0,\; h_{zz}(s,\bar{F^-}) = 1/2;\;\; \forall s\in[0,T). \tag{4.43}$$

Plugging (4.43) into (4.42) yields $C(s)\equiv 0$ therein. So, taking also into account (4.30), we see that h solves (4.28). The boundary conditions (4.29) for h are verified starting from the boundary conditions (4.35) for g. □

As known, the classical solution to (4.34)-(4.35) exists, is unique and satisfies (4.36)[1]. Indeed it is given by the Kolmogorov probabilistic representation

$$g(s,y) = \mathbb{E}\left[g(T,Y(T;s,y))\right], \quad (s,y)\in[0,T]\times[0,+\infty),$$

[1] The uniqueness of such a solution g is consistent with the fact that (4.25)-(4.26) admits a unique viscosity solution, so that the function h defined in (4.37) must be the *unique* viscosity (actually classical) solution to (4.32)-(4.29). Indeed two different solutions g_1, g_2 to (4.34)-(4.35) and satisfying (4.36) would give rise to two different solutions h_1, h_2 to (4.32)-(4.29), which is not possible due to Proposition 4.3.11.

where $Y(\cdot;s,y)$ is the solution of

$$\begin{cases} dY(t) = \lambda^2 Y(t)dt + \lambda Y(t)d\tilde{B}(t), \\ Y(s) = y, \end{cases} \tag{4.44}$$

where $\tilde{B}$ is a standard Brownian motion defined on some probability space $(\tilde{\Omega}, \tilde{\mathcal{F}}, \tilde{\mathbb{P}})$. Since the law of $Y(T;s,y)$ is known (it is the log-normal law), we can explicitely compute g. It is given by

$$\begin{cases} g(s,y) = (\bar{F}-\bar{S})\Phi(k(s,y)) - \dfrac{y}{2}e^{\lambda^2(T-s)}\Phi(k(s,y)-\lambda\sqrt{T-s}) + \bar{S}, \\ \qquad\qquad\qquad\qquad\qquad\qquad\qquad (s,y) \in [0,T) \times [0,+\infty), \\ g(T,y) = \left(\bar{F} - \frac{y}{2}\right) \vee \bar{S}, \end{cases} \tag{4.45}$$

where

$$k(s,y) = \frac{\log\left(\frac{2(\bar{F}-\bar{S})}{y}\right) - \frac{\lambda^2}{2}(T-s)}{\lambda\sqrt{T-s}} \tag{4.46}$$

and where $\Phi(\cdot)$ is the cumulative distribution function of a standard normal random variable, i.e.

$$\Phi(x) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{x} e^{-\frac{\xi^2}{2}}d\xi.$$

Moreover, g satisfies (4.35)-(4.36), thus, by Proposition 4.3.14, h defined in (4.37) is the unique classical solution to (4.32)-(4.29).

Remark 4.3.15. Despite of some constants, we see that the expression of g is related to the price of a European put option in an appropriate Black-Scholes market. □

4.3.5 The feedback map

Due to the previous subsection we can say that the value function W is the unique classical solution of the HJB equation

$$w_s(s,x) + (rx - b_0)w_x(s,x) - \frac{\lambda^2 w_x^2(s,x)}{2\,w_{xx}(s,x)} = 0, \ (s,x) \in \text{Int}^*(\mathcal{C}), \tag{4.47}$$

with boundary conditions

$$\begin{cases} w(s,F(s)) = 0, \ s \in [0,T]; \\ w(s,S(s)) = g(s) + e^{-\rho T}(\bar{F}-\bar{S})^2, \ s \in [0,T]; \\ w(T,x) = e^{-\rho T}(\bar{F}-x)^2, \ x \in [\bar{S},\bar{F}], \end{cases} \tag{4.48}$$

(the structure (4.8) for the Hamiltonian is provided by the fact that W_{xx}, as well as h_{zz}, is strictly positive). Moreover W is explicitely computable starting from g given in (4.45).

The feedback map related to W is (extended for convenience to $[0,T) \times \mathbb{R}$)

$$G(s,x) := \begin{cases} -\dfrac{\lambda}{\sigma} \dfrac{W_x(s,x)}{W_{xx}(s,x)}, & s \in [0,T),\ S(s) < x < F(s), \\ 0, & s \in [0,T),\ x \geq F(s) \text{ or } x \leq S(s). \end{cases} \tag{4.49}$$

It can be expressed as well in terms of h, g, given by (4.37), (4.45) by means of the transformations done in the previous subsection. Indeed, for $s \in [0,T)$, $S(s) < x < F(s)$, we can write

$$\begin{aligned} G(s,x) &= -\frac{\lambda}{\sigma} \frac{W_x(s,x)}{W_{xx}(s,x)} = -\frac{\lambda}{\sigma} \frac{h_z(s,[\mathcal{L}(s,\cdot)]^{-1}(x))}{h_{zz}(s,[\mathcal{L}(s,\cdot)]^{-1}(x))}\, e^{-r(T-s)} \\ &= -\frac{\lambda}{\sigma} \left([g(s,\cdot)]^{-1}\left([\mathcal{L}(s,\cdot)]^{-1}(x)\right)\right) g_y\left(s,[g(s,\cdot)]^{-1}\left([\mathcal{L}(s,\cdot)]^{-1}(x)\right)\right) e^{-r(T-s)}. \end{aligned} \tag{4.50}$$

Proposition 4.3.16. *Let $s_0 \in [0,T)$. The feedback map G defined in* (4.49) *is continuous on $[0,s_0] \times \mathbb{R}$. Moreover it is α-Holder continuous with respect to x uniformly in $s \in [0,s_0]$ for every $\alpha \in (0,1)$.*

Proof. We prove the claim in several steps.

(i) First of all we notice that, since $\mathcal{L}$ is a diffeomorphism, the claim is equivalent to prove that the function

$$\tilde{G}(s,z) := \begin{cases} [g(s,\cdot)]^{-1}(z) \cdot g_y\left(s,[g(s,\cdot)]^{-1}(z)\right), & z \in (\bar{S},\bar{F}), \\ 0, & z \leq \bar{S} \text{ or } z \geq \bar{F}, \end{cases} \tag{4.51}$$

is continuous on $[0,s_0] \times \mathbb{R}$ and α-Holder continuous with respect to z uniformly in $s \in [0,s_0]$ for every $\alpha \in (0,1)$.

(ii) Since $\tilde{G} \in C^{1,2}([0,s_0] \times (\bar{S},\bar{F});\mathbb{R})$, the claim is true on every compact set contained in $[0,s_0] \times (\bar{S},\bar{F})$.

(iii) Here we prove $\tilde{G}(s,\cdot)$ is continuous on $[\bar{S},\bar{F}]$ (therefore on $\mathbb{R}$) for every $s \in [0,s_0]$. Thanks to (ii) and to the definition of $\tilde{G}(s,\cdot)$ on $(-\infty,\bar{S}] \cup [\bar{F},+\infty)$, we need to prove the claim only at the endpoint $\bar{S}$ from the right and at the endpoint $\bar{F}$ from the left. This is equivalent to prove that, for every $s \in [0,s_0]$,

$$\lim_{z \to \bar{S}^+} \tilde{G}(s,z) = 0, \qquad \lim_{z \to \bar{F}^-} \tilde{G}(s,z) = 0,$$

i.e.

$$\lim_{y \to +\infty} y g_y(s,y) = 0, \qquad \lim_{y \to 0^+} y g_y(s,y) = 0.$$

The limits above are true by straightforward computations.

(iv) Here we prove that $\tilde{G}(s, \cdot)$ is Lipschitz continuous on $[\bar{F} - \varepsilon, \bar{F}]$ for some $\varepsilon > 0$. uniformly with respect to $s \in [0, s_0]$. To this aim, it suffices to show that $\tilde{G}_z$ is bounded on $[0, s_0] \times [\bar{F} - \varepsilon, \bar{F}]$ for some $\varepsilon > 0$. We have

$$\tilde{G}_z(s, z) = 1 + \frac{[g(s, \cdot)]^{-1}(z) \cdot g_{yy}\left(s, [g(s, \cdot)]^{-1}(z)\right)}{g_y\left(s, [g(s, \cdot)]^{-1}(z)\right)}$$

Therefore we study the limit for $z \to \bar{F}^-$ of $\tilde{G}_z(s, \cdot)$, or equivalently

$$\lim_{y \to 0^+} \frac{y g_{yy}(s, y)}{g_y(s, y)}.$$

Straightforward computations show that

$$\lim_{y \to 0^+} \frac{y g_{yy}(s, y)}{g_y(s, y)} = 0$$

uniformly in $s \in [0, s_0]$, which is enough to get the claim of this step.

(v) This is the most difficult step. Here we prove that, for every $\alpha \in (0, 1)$, the map $\tilde{G}(s, \cdot)$ is α-Holder continuous on $[\bar{S}, \bar{S} + \varepsilon]$ for some $\varepsilon > 0$ uniformly with respect to $s \in [0, s_0]$. The argument below holds uniformly with respect to $s \in [0, s_0]$, so here we fix $s \in [0, s_0]$ without loss of generality for our goal. To this aim it suffices to show that $\tilde{G}_z(s, z) \cdot (z - \bar{S})^{1-\alpha}$ is bounded on $[0, s_0] \times [\bar{S}, \bar{S} + \varepsilon]$ for some $\varepsilon > 0$. Therefore, we are led to study the limit for $z \to \bar{S}$ of $\tilde{G}_z(s, z) \cdot (z - \bar{S})^{1-\alpha}$, or equivalently

$$\lim_{y \to +\infty} \frac{y g_{yy}(s, y)}{g_y(s, y)} (g(s, y) - \bar{S})^{1-\alpha}. \tag{4.52}$$

A direct study of the limit above is quite hard. However, we notice that we can simplify it applying de l'Hôpital's rule "backward". Indeed, consider

$$\lim_{y \to +\infty} \frac{y g_y(s, y) - (g(s, y) - \bar{S})}{\left(g(s, y) - \bar{S}\right)^{\alpha}}. \tag{4.53}$$

The limit above is in the form $\frac{0}{0}$ so that we can apply de l'Hôpital rule getting

$$\lim_{y \to +\infty} \frac{y g_y(s, y) - (g(s, y) - \bar{S})}{\left(g(s, y) - \bar{S}\right)^{\alpha}} = \frac{1}{\alpha} \cdot \lim_{y \to +\infty} \frac{y g_{yy}(s, y)}{g_y(s, y)} (g(s, y) - \bar{S})^{1-\alpha}. \tag{4.54}$$

Therefore, from (4.54) we see that (4.52) behaves as (4.53), i.e. as

$$\lim_{y \to +\infty} \frac{y g_y(s, y)}{(g(s, y) - \bar{S})^{\alpha}}. \tag{4.55}$$

We observe that

$$\frac{d}{dy}\left[y\left(g(s, y) - \bar{S}\right)^{1-\alpha}\right] = (1 - \alpha) \frac{y g_y(s, y)}{\left(g(s, y) - \bar{S}\right)^{\alpha}} + \left(g(s, y) - \bar{S}\right)^{1-\alpha},$$

so that (4.55) behaves as

$$\lim_{y\to+\infty}\left(\frac{d}{dy}\left[y\left(g(s,y)-\bar{S}\right)^{1-\alpha}\right]\right). \tag{4.56}$$

To study the limite above, we observe that, for sufficiently large $M > 0$ we have

$$\frac{d^2}{dy^2}\left[\frac{yg_y(s,y)}{(g(s,y)-\bar{S})^{\alpha}}\right] \geq 0, \quad y \geq M.$$

Therefeore, the function

$$y \mapsto \frac{yg_y(s,y)}{(g(s,y)-\bar{S})^{\alpha}} \tag{4.57}$$

is convex on $[M, +\infty)$. It is straight to check that

$$\lim_{y\to+\infty}\left(y\left(g(s,y)-\bar{S}\right)^{1-\alpha}\right) = 0. \tag{4.58}$$

By convexity of (4.57), we must have also

$$\lim_{y\to+\infty}\left(\frac{d}{dy}\left[y\left(g(s,y)-\bar{S}\right)^{1-\alpha}\right]\right) = 0$$

and the claim is proved.

(vi) Due to items (iii) and (iv) we can see that $\tilde{G}$ is continuous with respect to (s, z) up to the boundary $[0, s_0] \times \{\bar{F}\}$ on the left with respect to z, i.e. for $z \to \bar{F}^-$.

Due to items (iii) and (v) we can see that $\tilde{G}$ is continuous with respect to (s, z) up to the boundary $[0, s_0] \times \{\bar{S}\}$ on the right with respect to z, i.e. for $z \to \bar{S}^+$.

(vii) The claim of item (i) is obvious on the set $[0, s_0] \times ((-\infty, \bar{S}] \cup [\bar{F}, +\infty))$, so the general claim is definitively proved. □

Another important property of the feedback map G is that it is bounded.

Proposition 4.3.17. *The map G defined in* (4.50) *is bounded on* $[0, T) \times \mathbb{R}$.

Proof. It is clear that we can prove the statement for the map $\tilde{G}$ defined in (4.51) on the set $[0, T) \times (\bar{S}, \bar{F})$. This claim is equivalent to prove that the map

$$(s, y) \mapsto yg_y(s, y)$$

is bounded on the set $[0, T) \times (0, +\infty)$. It is also clear that, since as known[2] g_y is bounded on $[0, T) \times (0, +\infty)$, it suffices to prove the claim on the set $[0, T) \times [M, +\infty)$ for some $M > 0$.

[2]Recall that g has the same structure of the price of a European put option (see Remark 4.3.15).

We have

$$\begin{aligned} g_y(s,y) &= \frac{1}{\sqrt{2\pi}}(\bar{F}-\bar{S})e^{-\frac{k(s,y)^2}{2}}k_y(s,y) - \frac{1}{2}e^{\lambda^2(T-s)}\Phi\left(k(s,y)-\lambda\sqrt{T-s}\right) \\ & -\frac{1}{\sqrt{2\pi}}\frac{y}{2}e^{\lambda^2(T-s)}e^{-\frac{\left(k(s,y)-\lambda\sqrt{T-s}\right)^2}{2}}k_y(s,y) \end{aligned} \tag{4.59}$$

and

$$k_y(s,y) = -\frac{1}{\lambda\sqrt{T-s}}\cdot\frac{1}{y}.$$

We notice that $yg_y(s,y)$ is negative and that for $y \geq 2(\bar{F}-\bar{S})$ the first two terms in the right handside of (4.59) are negative, while the third one is positive. Therefore, for $y \geq 2(\bar{F}-\bar{S})$,

$$|yg_y(s,y)| \leq K_1\frac{e^{-\frac{k(s,y)^2}{2}}}{\lambda\sqrt{T-s}} + K_2\, y\, \Phi\left(k(s,y)-\lambda\sqrt{T-s}\right),$$

for some $K_1, K_2 > 0$. Set

$$p(s,y) := \frac{e^{-\frac{k(s,y)^2}{2}}}{\lambda\sqrt{T-s}}, \quad q(s,y) := y\Phi\left(k(s,y)-\lambda\sqrt{T-s}\right).$$

We claim that p, q are bounded on $[0,T)\times[M,+\infty)$ for some $M \geq 2(\bar{F}-\bar{S})$, that is enough to conclude. We have

$$p(s,y) = \frac{\left(e^{-\frac{1}{2}\left(\log\left(\frac{2(\bar{F}-\bar{S})}{y}\right)-\frac{\lambda^2}{2}(T-s)\right)^2}\right)^{\frac{1}{\lambda^2(T-s)}}}{\lambda\sqrt{T-s}}. \tag{4.60}$$

Notice that, if $\alpha(y) \in (0,\alpha_0]$, $\alpha_0 < 1$, there exists $C_{\alpha_0} > 0$ such that

$$\frac{1}{\sqrt{x}}\alpha(y)^{1/x} \leq C_{\alpha_0}, \qquad x > 0. \tag{4.61}$$

Take $M \geq 2(\bar{F}-\bar{S})$ large enough to have

$$e^{-\frac{1}{2}\left(\log\left(\frac{2(\bar{F}-\bar{S})}{M}\right)\right)^2} \leq \alpha_0 < 1.$$

Take $C_{\alpha_0} > 0$ such that (4.61) holds. From (4.60) we get on the set $[0,T)\times[M,+\infty)$

$$p(s,y) \leq C_{\alpha_0}.$$

Consider now q. We have on $[0,T)\times(0,+\infty)$

$$q(s,y) \leq y\Phi\left(\log\left(\frac{2(\bar{F}-\bar{S})}{y}\right)\right).$$

It is straightforward to check that

$$\lim_{y\to+\infty}\left[y\Phi\left(\log\left(\frac{2(\bar{F}-\bar{S})}{y}\right)\right)\right] = 0,$$

so q is bounded on $[0,T)\times(0,+\infty)$. Therefore we get the claim. □

4.3.6 The closed loop equation and the optimal feedback strategy

Thanks to Proposition 4.3.16 in the previous subsection, the classical theory on one dimensional SDEs allows to study the closed loop equation associated with G.

Proposition 4.3.18. *Let G be the map defined in* (4.50) *and let $s \in [0, T)$. For every $s_0 \in [s, T)$ the closed loop equation*

$$\begin{cases} dX(t) = [rX(t) + \sigma\lambda G(t, X(t)) - b_0]\, dt + \sigma G(t, X(t))dB(t), \\ X(s) = x \in [S(s), F(s)], \end{cases} \tag{4.62}$$

admits a unique strong solution $X_G^{s_0}(\cdot; s, x)$ on $(\Omega, \mathcal{F}, \mathbb{P})$ in the interval $[s, s_0]$. Moreover we have

$$S(t) \leq X_G^{s_0}(t; s, x) \leq F(t), \quad \mathbb{P} - a.s., \ \forall t \in [s, s_0]. \tag{4.63}$$

Proof. Since G is continuous on $[s, s_0] \times \mathbb{R}$, we see that Theorem 2.4, p. 163, of [Ikeda, Watanabe; 1981] provides the existence of a weak solution.

Since G is α-Holder continuous, $\alpha \geq 1/2$, with respect to y uniformly with respect to the time variable, the Yamabe-Watanabe Theorem (see, e.g.,

- [Yamada, Watanabe; 1971] or

- Theorem 3.5-(ii), p. 390, of [Revuz, Yor; 1999] or

- Proposition 2.13, p. 291, of [Karatzas, Shreve; 1991])

ensures pathwise uniqueness for the solution.

By Yamada-Watanabe theory, pathwise uniqueness and weak existence imply existence (and uniqueness) of a strong solution (see, e.g., Section 5.3.D, pp. 308–311, of [Karatzas, Shreve; 1991]).

The statement (4.63) follows by definition of G and due to the absorbing properties of $\partial_F^*\mathcal{C}$ and $\partial_S^*\mathcal{C}$ (see Proposition 4.3.3 and the considerations below). $\square$

By uniqueness straightly we see that

$$0 \leq s \leq s_0 \leq s_0' < T, \ x \in [S(s), F(s)] \implies X_G^{s_0'}(\cdot; s, x)|_{[s,s_0]} = X_G^{s_0}(\cdot; s, x).$$

So, we have a unique strong solution $X_G(\cdot; s, x)$ to (4.62) in the right-open interval $[s, T)$. Moreover

$$S(t) \leq X_G(t; s, x) \leq F(t), \quad \mathbb{P} - a.s., \ \forall t \in [s, T). \tag{4.64}$$

Therefore, for $s \in [0, T)$, $x \in [S(s), F(s)]$, consider the feedback strategy $\pi_G^{s,x}(\cdot)$ defined by

$$\pi_G^{s,x}(t) := \begin{cases} G(t, X_G(t; s, x)), & \text{if } t \in [s, T), \\ 0, & \text{if } t = T. \end{cases} \tag{4.65}$$

This is the candidate optimal strategy for the problem. First let us show that it is admissible.

Proposition 4.3.19. *For every $s \in [0,T)$, $x \in [S(s), F(s)]$, the feedback strategy $\pi_G^{s,x}(\cdot)$ defined in* (4.65) *is admissible, i.e.* $\pi_G^{s,x}(\cdot) \in \Pi_{ad}(s,x)$.

Proof. By Proposition 4.3.17 we know that G is bounded on $[0,T) \times \mathbb{R}$, so that $\pi_G^{s,x}(\cdot) \in L^\infty(\Omega \times [s,T];\mathbb{R}) \subset L^2([\Omega \times [s,T];\mathbb{R})$. Moreover clearly we have by definition of $\pi_G^{s,x}(\cdot)$

$$X(t; s, x, \pi_G^{s,x}(\cdot)) = X_G(t; s, x), \quad \mathbb{P} - a.s., \quad \forall t \in [s, T),$$

so that from (4.64) we see also that

$$S(t) \leq X(t; s, x, \pi_G^{s,x}(\cdot)) \leq F(t), \quad \mathbb{P} - a.s., \ \forall t \in [s, T).$$

By continuity of trajectories we have also

$$S(T) \leq X(T; s, x, \pi_G^{s,x}(\cdot)) \leq F(T), \quad \mathbb{P} - a.s.$$

and the claim is proved. □

Theorem 4.3.20 (Verification). *Let $(s,x) \in \mathcal{C}$ and let $\pi(\cdot) \in \Pi_{ad}(s,x)$. Then $\pi(\cdot)$ is optimal for the initial (s,x) if and only if*

$$\pi(t) = G(t, X(t; s, x, \pi(\cdot)), \quad \mathbb{P} - a.s., \ \forall t \geq s.$$

In particular the feedback strategy $\pi_G^{s,x}(\cdot)$ defined in (4.65) *is the unique optimal strategy starting from the initial (s,x).*

Proof. Since W is a classical solution of the HJB equation (4.47)-(4.48), the proof follows the line of the proof of Theorem 4.2.12. Here we have only to take care of the state constraint, as in the proof of Theorem 1.2.30. A similar argument works thanks to Proposition 4.3.19. □

Remark 4.3.21. Due to the fact that the feedback map is α-Holder continuous for every $\alpha \in (0,1)$ with respect to the space variable uniformly with respect to $s \in [0, s_0]$, it would be possible to prove that the boundaries $\partial_S^*\mathcal{C}, \partial_F^*\mathcal{C}$ are never reached by the optimal diffusion $(t, X_G(t; s, x))_{t \in [s, s_0]}$ when $(s,x) \in \mathrm{Int}^*(\mathcal{C})$. From the point of view of Feller's boundaries classification, this means that these boundaries are *natural* for the optimal diffusion in the time interval $[s,T)$. However, we cannot exclude that the optimal diffusion touches these boundaries at time T, i.e. that $X_G(T; s, x) = \bar{F}$ or $X_G(T; s, x) = \bar{S}$ with positive probability. □

Remark 4.3.22. The process Y defined by (4.44) is just an auxiliary process defined in order to give a computable representation of the solution g to (4.34)-(4.35). However, there is a relationship between the solution of (4.44) and the optimal process X_G, as we are going to show.

On the original probability space $(\Omega, \mathcal{F}, \mathbb{P})$ consider the SDE

$$\begin{cases} d\tilde{Y}(t) = \lambda \tilde{Y}(t) dB(t), \\ \tilde{Y}(s) = y, \end{cases} \tag{4.66}$$

where B is the originary Brownian motion defined on this space.

(Notice that this is equivalent to consider the equation

$$\begin{cases} dY(t) = \lambda^2 Y(t) dt + \lambda Y(t) d\tilde{B}(t), \\ Y(s) = y, \end{cases} \tag{4.67}$$

i.e. (4.44) on the probability space $(\Omega, \mathcal{F}, \tilde{\mathbb{P}})$, where

$$\tilde{\mathbb{P}} = e^{\lambda B(t) - \frac{\lambda^2}{2}} \cdot \mathbb{P}, \quad \tilde{B}(t) = B(t) - \lambda t, \quad t \in [0, T],$$

By Girsanov's Theorem $\tilde{B}$ is a Brownian motion under $\tilde{\mathbb{P}}$.)

Let g be the solution of (4.34)-(4.35) and consider the process

$$\tilde{Z}(t) := g(t, \tilde{Y}(t)),$$

where $\tilde{Y}$ is the solution to (4.66). The Itô formula, the fact that g solves (4.34) and the expression of $\tilde{G}$ in (4.51) yield

$$\begin{aligned} d\tilde{Z}(t) &= g_t(t, \tilde{Y}(t)) dt + g_y(t, \tilde{Y}(t)) d\tilde{Y}(t) + \frac{1}{2} \lambda^2 \tilde{Y}(t)^2 g_{yy}(t, \tilde{Y}(t)) dt \\ &= \left[g_t(t, \tilde{Y}(t)) + \lambda^2 \tilde{Y}(t) g_y(t, \tilde{Y}(t)) + \frac{1}{2} \lambda^2 \tilde{Y}(t)^2 g_{yy}(t, \tilde{Y}(t)) \right] dt \\ &\quad - \lambda^2 \tilde{Y}(t) g_y(t, \tilde{Y}(t)) dt + \lambda \tilde{Y}(t) g_y(t, \tilde{Y}(t)) dB(t) \\ &= -\lambda^2 \tilde{Y}(t) g_y(t, \tilde{Y}(t)) dt + \lambda \tilde{Y}(t) g_y(t, \tilde{Y}(t)) dB(t) \\ &= \sigma \lambda \tilde{G}(t, \tilde{Z}(t)) dt - \sigma \tilde{G}(t, \tilde{Z}(t)) dB(t). \end{aligned} \tag{4.68}$$

This shows that

$$\tilde{X}(t) := \mathcal{L}(t, \tilde{Z}(t)) = \mathcal{L}(t, g(t, \tilde{Y}(t))) \tag{4.69}$$

solves

$$\begin{cases} dX(t) = [rX(t) + \sigma \lambda G(t, X(t)) - b_0] \, dt - \sigma G(t, X(t)) dB(t), \\ X(s) = x = \mathcal{L}(t, g(t, \tilde{Y}(t))) \in [S(s), F(s)]. \end{cases} \tag{4.70}$$

Therefore $\tilde{X}(t)$ has the same law of $X_G(t; s, x)$. This can be used in the simulations, because simulating $\tilde{Y}$ is much easier than simulating X_G. Moreover, this argument can be used to study the closed loop equation through (4.66): indeed, we see that $\tilde{Y}$ solves (4.66) if and only if $\tilde{X}$ defined in (4.69) solves (4.70). Despite of a minus sign in the diffusion, (4.70) is the same as (4.62), so that, since (4.66) admits a unique strong solution on $(\Omega, \mathcal{F}, \mathbb{P})$, also (4.62) admits a unique solution on $(\Omega, \mathcal{F}, \mathbb{P})$, i.e. the statement of Proposition 4.3.18. $\square$

4.3.7 Numerical application

In this subsection we show a numerical application of the model presented in the present section. We consider the position of a male retiree aged 60 with initial wealth $x_0 = 100$. Consistently with the compulsory annuitization age of 75 holding in UK, we set $T = 15$. The market parameters are $r = 0.03, \mu = 0.08, \sigma = 0.15$, implying a Sharpe ratio equal to $\beta = 0.33$. The pension rate purchasable at retirement, using Italian projected mortality tables (RG48) is $b_0 = 6.22$. The choice of the final target F and the final guarantee S are evidently subjective and are determined by the member's risk aversion. High risk aversion will lead to a high guarantee and a low level of the target, while a high target and a low guarantee will be driven by low risk aversion. We have tested three levels of risk aversion. Thus, high risk aversion is associated to terminal safety level $S = \frac{2}{3} b_0 a_{75}$ and final target equal to $F = 1.5 b_0 a_{75}$, where a_{75} is the actuarial value of a unitary lifetime annuity issued to an individual aged 75; medium risk aversion is associated to terminal safety level $S = \frac{1}{2} b_0 a_{75}$ and final target equal to $F = 1.75 b_0 a_{75}$; low risk aversion is associated to terminal safety level[3] $S = 10^{-4} = 0^{+}$ and final target equal to $F = 2 b_0 a_{75}$. These values are reported in Table 1 below.

	S	F
High risk aversion	$\frac{2}{3} b_0 a_{75}$	$1.5 b_0 a_{75}$
Medium risk aversion	$\frac{1}{2} b_0 a_{75}$	$1.75 b_0 a_{75}$
Low risk aversion	0	$2 b_0 a_{75}$

Table 1. Terminal safety level S and final target F for different risk profiles.

The interpretation of these choices is immediate. With high and medium risk aversion, the minimum pension rate guaranteed is, respectively, two third and half of the annuity rate that was possible to have on immediate annuitization at retirement, b_0; the targeted wealth is sufficient to fund a final pension that

[3]According to Remark **??** we have to consider $S > 0$. Evidently, from a practical point of view, being owner of a fund of 10^{-4} EUR is equivalent to ruin.

amounts to, respectively, 1.5 and 1.75 of b_0. With low risk aversion, ruin is avoided but no money is left for annuitization at age 75; on the other hand, the targeted pension pursued is twice b_0. It is worth mentioning that even the more risk averse individual has some restrictions in choosing the minimum income guaranteed. Indeed, it is clear from the formulation of the problem that the value of $\bar{S}$ has to satisfy $\bar{S} \leq z_0 = x_0 e^{rT} - \frac{b_0}{r}(e^{rT} - 1)$. The most risk averse choice would be $S = z_0$, but in this case the only admissible strategy would be $\theta(\cdot) \equiv 0$, i.e. the whole fund wealth must be invested in the riskless asset (see Proposition (4.3.1)), and one would end up after 15 years with an annuity lower than that purchasable at retirement[4]. This choice makes little sense in a realistic framework, given that here the bequest motive is disregarded and the individual takes the income drawdown option only in the hope of being able to buy a better annuity than b_0. For this reason, we here consider only cases where $S < z_0$, which in this example means $S < 0.7 b_0 a_{75}$.

We have carried out 1000 Monte Carlo simulations for the behaviour of the risky asset, with discretization step equal to one week. In order to do so, we have simulated the process $\tilde{Y}$ given by equation (4.66) with starting point

$$\tilde{Y}(0) = y = [g(0, \cdot)]^{-1}(z_0),$$

inserting the corresponding values of S, F and z_0 as above. With each risk aversion we have generated the same 1000 scenarios, by applying in each case the same stream of pseudo random numbers.

For each risk aversion choice, we report the following results:

- evolution of the fund under optimal control during the 15 years time, by showing a graph with mean and standard deviation and a graph with some percentiles
- behaviour of the optimal investment strategy θ^* over the 15 years time, by showing a graph with some percentiles
- distribution of the final annuity that can be bought with the final fund at age 75, comparison with the annuity purchasable at retirement.

Figures 1–4 report results for high risk aversion, figures 5–8 those for medium risk aversion, figures 9–12 those for low risk aversion. In particular, figures 1,

[4]This is clear considering the mortality credits obtained when investing in an insurance product, that enhance the riskless rate.

5 and 9 report, over 15 years time, the mean and dispersion of the fund trajectories, while figures 2, 6 and 10 report their percentiles. Figures 3, 7 and 11 report some percentiles of the distribution of the optimal investment allocation θ^* over 15 years, and figures 4, 8 and 12 report the distribution of the final annuity upon annuitization at time T.

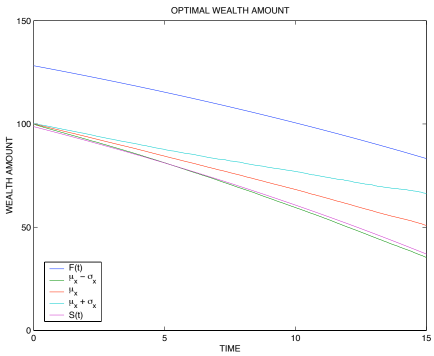


Figure 1. High risk aversion.

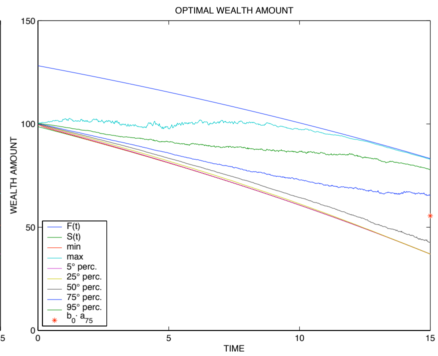


Figure 2. High risk aversion.

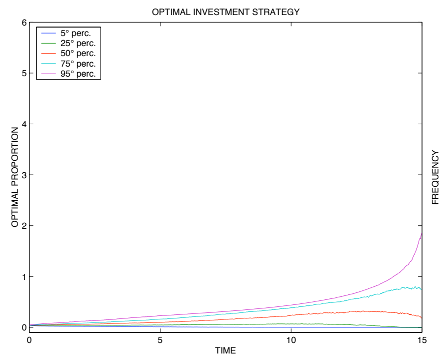


Figure 3. High risk aversion.

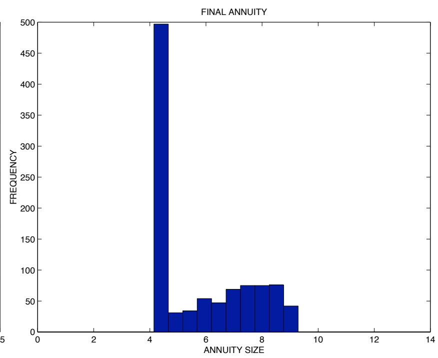


Figure 4. High risk aversion.

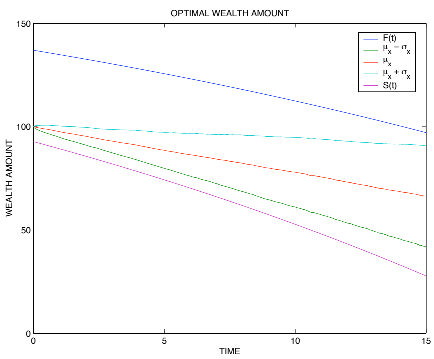


Figure 5. Medium risk aversion.

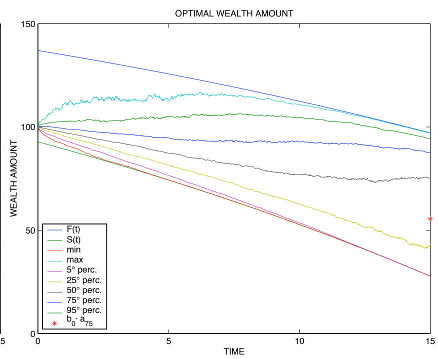


Figure 6. Medium risk aversion.

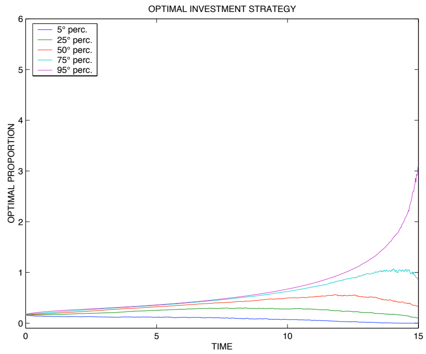


Figure 7. Medium risk aversion.

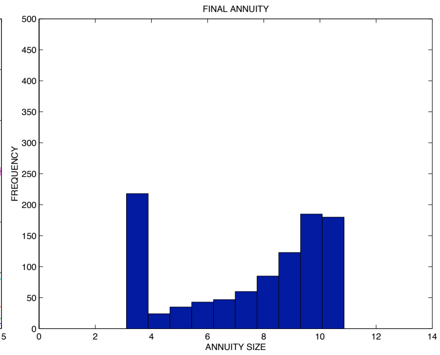


Figure 8. Medium risk aversion.

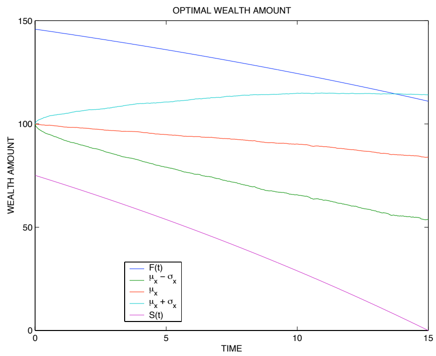


Figure 9. Low risk aversion.

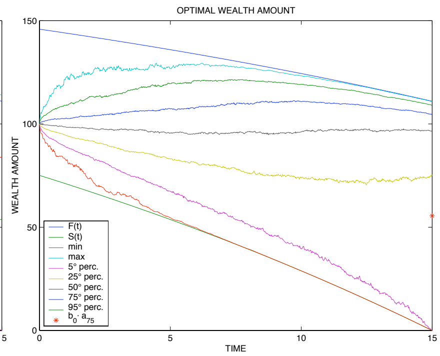


Figure 10. Low risk aversion.

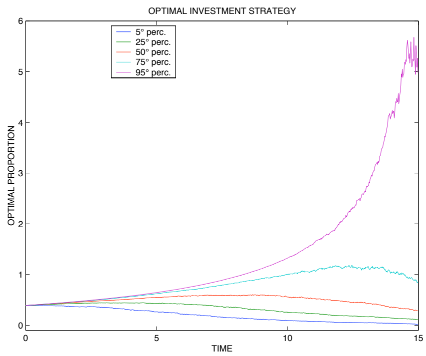


Figure 11. Low risk aversion.

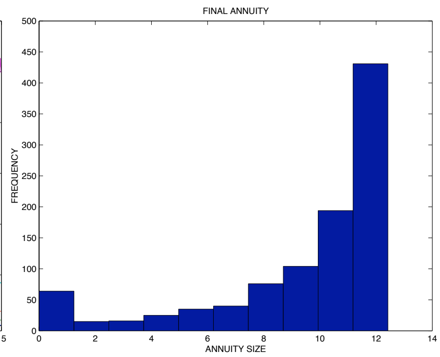


Figure 12. Low risk aversion.

From the graphs we can make the following comments:

- Obviously, the wealth trajectories always lie strictly between the two barriers $S(t)$ and $F(t)$ for $t < T$. In fact, the two bottom and upper absorbing barriers cannot be reached before time T.
- Looking at the graphs reporting the percentiles of the trajectories, Figg. 2, 6 and 10, however, it seems that in some cases the fund touches the bottom target $S(t)$. This is due to the approximation error made by the machine, that is unavoidable. In fact, for not too low values of η, $\Phi(\eta)$ is so close to 0 that it cannot be distinguished from it. The result is that in the practical applications for not too high values of y one has $\Phi(k(t,y)) = \Phi(k(t,y) - \beta\sqrt{T-t}) = 0$ and $g(t,y) = \bar{S}$, meaning that the fund is on the safety level $S(t)$, which is theoretically not true.
- When risk aversion decreases, the boundaries for the wealth process become larger. This is due to the obvious fact that strategies are more aggressive and the range of final outcomes increases, both in the positive as well as in the negative direction.
- Inspection of Figg. 3, 7 and 11 shows that when risk aversion decreases, optimal strategies become riskier. In fact, with high risk aversion the 95^{th} percentile of θ^* stays below 2 even immediately prior to time T, whereas with low risk aversion it lies between 5 and 6 close to T. On the other hand, clearly, all strategies are bounded away from 0.
- Comparing Figg. 4, 8 and 12 it is immediate to see that the distribution of the final annuity becomes more and more spread when risk aversion decreases. Moreover, with high risk aversion one can observe a considerable concentration around the guaranteed income $\frac{2}{3}b_0 = 4.15$. In fact, in almost 50% of the cases, the fund approaches $S(t)$ and stays close to it until T (this can be noticed also by thorough inspection of Fig. 2). On the contrary, the distribution of final annuity looks very favourable in the case of low risk aversion, where in most of the cases the annuity lies between 9 and 12, and unfavourable scenarios leading to final income equal to 0 happen in ca 5% of the cases.

One should not forget that the real goal of the pensioner who opts for phased withdrawals is to be better off than immediate annuitization when final annuitization takes place. Thus, it is of greatest interest to provide her with detailed information regarding the distribution of the final annuity achieved. To some extent, this has been already shown in Figg. 4, 8 and 12. However, the hystograms cannot report relevant information that are of immediate use for the member who has to choose a risk profile. In particular, for the member's decision making it is relevant the comparison between the final annuity achievable by taking income drawdown option and b_0, the pension rate purchasable at retirement. Table 2 reports useful statistics of the distribution of the final annuity achieved at age 75, for each risk aversion. The first nine line report mean, standard deviation, min, max and some percentiles of the distribution of the final annuity. Lines 10 and 11 report, respectively, the guaranteed income S/a_{75} and the targeted income F/a_{75} (as chosen in Table 1), while the last line reports the probability (i.e. the frequency over 1000 scenarios) that the final annuity is higher than b_0.

	HIGH RISK AVERSION	MEDIUM RISK AVERSION	LOW RISK AVERSION
mean	5.70	7.44	9.40
st.dev.	1.74	2.73	3.38
min	4.15	3.11	0.00
5th perc.	4.15	3.11	0.00
25th perc.	4.15	4.87	8.45
50th perc.	4.75	8.41	10.80
75th perc.	7.33	9.80	11.72
95th perc.	8.71	10.55	12.21
max	9.29	10.86	12.42
guaranteed income S/a_{75}	4.15	3.11	0
targeted income F/a_{75}	9.33	10.885	12.44
prob(final annuity $> b_0$)	39.20%	68.80%	84.10%

Table 2. Distribution of final annuity at age 75. Annuity on immediate annuitization $b_0 = 6.22$.

The following comments can be made:

- The mean of the final annuity is 5.70, 7.44, 9.40 with high, medium and low risk aversion, respectively. The probability of being able to afford a final annuity higher than $b_0 = 6.22$ is 39.20%, 68.80% and 84.10% with high, medium and low risk aversion, respectively.
- This shows that if risk aversion is too high, the price for having a high

guarantee on the final income[5] is that the chances of reaching the desired annuity reduce dramatically. In fact, in 60% of the cases the individual ends up with a final annuity lower than b_0 and, even worse, in almost 50% of the cases the individual receives exactly the guaranteed income, that is only two third of b_0. This is likely to be an undesirable result for the pensioner and it seems to indicate that if the member's risk aversion is too high, it is not convenient to take the income drawdown option. This result was already observed by [Gerrard, Haberman, Vigna; 2006].

- On the other hand, with medium and low risk aversion the chances of being better off with annuitization at time T are almost 70% and 85%, respectively. This is an encouraging result, given that from retirement to T the pensioner has withdrawn the prescribed rate of b_0 and that she was also guaranteed with a minimum lifetime income at retirement, or at worst against ruin.

- The low risk aversion profile could turn out to be particularly attractive to a member whose global post-retirement income was not heavily affected by the second pillar provision. In fact, 1) the chances of exceeding the immediate annuitization income b_0 are extremely high (84%), 2) in 75 cases out of 1000 the member ends up with an annuity higher than 8.45, that is well above $b_0 = 6.22$, 3) in about 5 cases out of 1000 the final annuity is null and 4) ruin never occurs.

- Clearly, the price to pay for having a favourable distribution of final income is to take more risk, which translates into more aggressive investment policies. This is highlighted by Fig.11, that reports the optimal investment strategies for low risk aversion. In more than 25% of the cases, the optimal strategy consists in borrowing considerable amounts of money to be invested in the risky asset. This kind of strategy is evidently not feasible in the presence of real world constraints.

[5] Observe in fact that the value of $S = 0.67 b_0 a_{75}$ is chosen to be very close to the upper boundary $z_0 = 0.70 b_0 a_{75}$

Appendix A

Here we provide some results of functional analysis and measure theory used in the thesis.

A.1 Some results in Probability and Measure Theory

Theorem A.1.1 (Girsanov). *Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space on which there is defined a Brownian motion $(B(t))_{t\geq 0}$. Let $(\mathcal{F}_t^B)_{t\geq 0}$ be the filtration generated by the Brownian motion. For an $(\mathcal{F}_t)_{t\geq 0}$-progressively measurable process γ consider the processes*

$$\tilde{B}(t) = B(t) + \int_0^t \gamma(s)ds,$$

$$L(t) = \exp\left(-\int_0^t \gamma(s)dB(s) - \frac{1}{2}\int_0^t \gamma^2(s)ds\right)$$

and assume that γ satisfies the Novikov condition

$$\mathbb{E}\left[\exp\left(\frac{1}{2}\int_0^t \gamma^2(s)ds\right)\right] < \infty, \qquad t \in [0, +\infty).$$

Then $(\tilde{B}(s))_{0\leq s\leq t}$ is a Brownian motion under the probability $\tilde{\mathbb{P}}_t := L(t)\cdot\mathbb{P} \ll \mathbb{P}|_{\mathcal{F}_t^B}$ on $\mathcal{F}_t^B$ defined by the Radon-Nikodym derivative $L(t)$. □

Proof. See, e.g., [Karatzas, Shreve; 1991], Chapter 3, Section 3.5, or also [Revuz, Yor; 1999], Chapter VIII. □

Lemma A.1.2. *Let $(M, \mathcal{M}, \mu)$ be a finite measure space. Let $p \geq 1$ and set $L^p := L^p(M, \mathcal{M}, \mu; \mathbb{R})$. Let $(f_n), f \subset L^p$ be such that $f_n \to f$ in L^p. Finally let $\psi : \mathbb{R} \to \mathbb{R}$ be uniformly continuous. Then $\psi(f_n), \psi(f) \in L^p$ and $\psi(f_n) \to \psi(f)$ in L^p.*

Proof. Let x be the generic element of M. Since ψ is uniformly continuous there exists a modulus of uniform continuity ω and it has sublinear growth. Therefore, for suitable $a, b > 0$,

$$|\psi(f(x))| \leq |\psi(f(x)) - \psi(0)| + |\psi(0)| \leq \omega(|f(x)|) + |\psi(0)| \leq a|f(x)| + b + |\psi(0)|.$$

Since μ is finite and $f \in L^p$, we get $\psi(f) \in L^p$. The same argument holds for $\psi(f_n)$, $n \in \mathbb{N}$.
Let us show that $\psi(f_n) \to \psi(f)$ in L^p. Let $\varepsilon > 0$; we have, for some $C > 0$,

$$\begin{aligned}
&\int_M |\psi(f) - \psi(f_n)|^p d\mu \\
&\quad = \int_{\{|f-f_n|\leq\varepsilon\}} |\psi(f) - \psi(f_n)|^p d\mu + \int_{\{|f-f_n|>\varepsilon\}} |\psi(f) - \psi(f_n)|^p d\mu \\
&\qquad \leq \mu(M)\omega(\varepsilon)^p + \mu\{|f - f_n| > \varepsilon\}\, C \left(\|f\|_{L^p}^p + \|f_n\|_{L^p}^p + 1\right).
\end{aligned}$$

Taking the limsup for $n \to \infty$, the last term of the rigth hand-side goes to 0, since $f_n \xrightarrow{\mu} f$ and $\|f_n\|_{L^p} \to \|f\|_{L^p}$. Therefore

$$\limsup_{n\to\infty} \int_M |\psi(f) - \psi(f_n)|^p d\mu \leq \mu(M)\omega(\varepsilon)^p.$$

Since $\omega(\varepsilon) \to 0$ when $\varepsilon \to 0$, by the arbitrariness of ε we get the claim. □

A.2 Some results in Functional Analysis

Theorem A.2.1 (Ascoli-Arzelà). *Let (S, d) be a compact metric space and let $C(S)$ be the Banach space of the real-valued functions $x : S \to \mathbb{R}$ endowed with the sup-norm $\|x\| = \sup_{s\in S} |x(s)|$. Let $(x_n)_{n\in\mathbb{N}} \subset C(X)$ be a sequence of equi-bounded and equi-uniformly continuous functions, i.e.*

- *there exists $M > 0$ such that $|x_n(s)| \leq M$ for every $n \in \mathbb{N}$,*
- *for every $\varepsilon > 0$ there exists $\delta > 0$ such that*

$$d(s, s') < \delta \Longrightarrow |x_n(s) - x - n(s')| < \varepsilon,\ \forall n \in \mathbb{N}.$$

Then $(x_n)_{n\in\mathbb{N}}$ is relatively compact in $C(S)$. In particular we can extract a subsequence $(x_{n_k})_{k\in\mathbb{N}}$ converging towards some $\bar{x} \in C(S)$.

Proof. See [Yosida; 1980], Chapter III, Section 3. □

Theorem A.2.2 (Schauder). *Let X be a Banach space and let $M \subset X$ be nonempty, closed, convex and bounded. Let $F : X \to X$ completely continuous (i.e. F is continuous and maps bounded sets into relatively compact sets). Then F admits a fixed point.*

Theorem A.2.3 (Banach-Steinhaus). *Let X, Y be Banach spaces and let $\{T_a\}_{a\in A}$ be a family of bounded linear operator from X to Y such that*

$$\sup_{a\in A} \|T_a x\|_Y < +\infty,\ \ \forall x \in X.$$

Then the family $\{T_a\}_{a\in A}$ is equibounded, i.e.

$$\sup_{a\in A} \|T_a\|_{\mathcal{L}(X,Y)} < +\infty.$$

Proof. See, in more generality, [Yosida; 1980], Chapter II, Section 1. $\square$

Lemma A.2.4. *Let $(X, \|\cdot\|)$ be a Banach space and $f : X \to \mathbb{R}$ uniformly continuous. Then f is bounded on the bounded sets of X.*

Proof. Fix $\varepsilon > 0$ and let $\delta > 0$ be such that

$$\|x - y\| < \delta \Longrightarrow |f(x) - f(y)| < \varepsilon.$$

Let us suppose by contradiction that

$$R := \sup\{r \geq 0 \mid f(B(0,r)) \text{ is bounded}\} < +\infty. \tag{A.1}$$

Of course, by continuity of f, it results $R > 0$; by (A.1), we can find a sequence

$$(x_n) \subset B(0, R + \delta/2) \backslash \overline{B(0,R)}$$

such that

$$|f(x_n)| \geq n, \quad n \in \mathbb{N}. \tag{A.2}$$

On the other hand, again for (A.1), if

$$y_n := \frac{R - (\delta/3 \wedge R/2)}{\|x_n\|} x_n,$$

then, by definition of R, $|f(y_n)| \leq K$, for some $K > 0$. Since $\|x_n - y_n\| < \delta$, by uniform continuity of f we have $|f(x_n) - f(y_n)| < \varepsilon$ for every $n \in \mathbb{N}$, which implies $|f(x_n)| \leq K + \varepsilon$ for every $n \in \mathbb{N}$. This contradicts (A.2), so that the claim is proved. $\square$

Bibliography

[Albrecht, Maurer; 2002] Albrecht P., Maurer R., *Self-annuitization, consumption shortfall in retirement and asset allocation: the annuity benchmark*, Journal of Pension Economics and Finance, **1**, pp. 269–288, 2002.

[Asea, Zak; 1999] Asea P.K., Zak P.J., *Time-to-build and cycles*, Journal of Economic Dynamics and Control, Vol. 23, No.8, pp. 1155–1175, 1999.

[Bambi; 2008] Bambi, M., *Endogenous growth and time to build: the AK case*, Journal of Economic Dynamics and Control, Vol. 32, pp. 1015–1040, 2008.

[Barbu, Da Prato; 1981] Barbu V., Da Prato G., *A direct method for studying the dynamic programming equation for controlled diffusion processes in Hilbert spaces*, Numerical Functional Analysis and Optimization, **4**, 1, 23–43,1981.

[Barbu, Da Prato; 1983a] Barbu V., Da Prato G., *Solution of Bellman equation associated with an infinite dimensional stochastic control problem and synthesis of oprtimal control*, SIAM Journal on Control and Optimization, **21**, 4, pp. 531–550, 1983.

[Barbu, Da Prato; 1983b] Barbu V., Da Prato G. *Hamilton-Jacobi equations in Hilbert spaces*, Research notes in Mathematics, Pitman, Boston, 1983.

[Barbu, Da Prato; 1985a] Barbu V., Da Prato G., *Hamilton-Jacobi equations in Hilbert spaces; variational and semigroup approach*, Annali di Matematica Pura e Applicata, **142**, 1, pp. 303–349, 1985.

[Barbu, Da Prato; 1985b] Barbu V., Da Prato G., *A note on a Hamilton-Jacobi equation in Hilbert space*, Nonlinear Analysis, 9, pp. 1337–1345, 1985.

[Barbu, Da Prato, Popa; 1983] Barbu V., Da Prato G., Popa C., *Existence and uniqueness of the dynamic programming equation in Hilbert spaces*, Nonlinear Anal., **7**, 3, pp. 283-299, 1983.

[Barbu, Precupanu; 1986] Barbu V., Precupanu Th., *Convexity and Optimization in Banach Spaces*, Editura Academiei, Bucharest, 1986.

[Bardi, Capuzzo-Dolcetta; 1997] Bardi M., Capuzzo-Dolcetta I., *Optimal control and viscosity solutions of Hamilton-Jacobi-Bellman equations*, Birkhauser, 1997.

[Barles, Souganidis; 1991] Barles G., Souganidis P.E., *Convergence of Approximation Schemes for Fully Nonlinear Second Order Equations*, Asymptotic Analysis, **4**, 271–283, 1991.

[Battocchio, Menoncin; 2004] Battocchio P., and Menoncin F., *Optimal pension management in a stochastic framework*, Insurance: Mathematics and Economics, **34**, 1, pp. 79–95, 2004.

[Battocchio, Menoncin, Scaillet; 2007] Battocchio P., Menoncin F., Scaillet O., *Optimal asset allocation for pension funds under mortality risk during the accumulation and decumulation phases*, Annals of Operations Research, **152**, 1, pp. 141–165, 2007.

[Bellman; 1957] Bellman R., *Dynamic programming*, Princeton University Press, New Jersey, 1957.

[Bensoussan, Da Prato, Delfour, Mitter; 2007] Bensoussan A., Da Prato G., Delfour M.C., Mitter S.K., *Representation and control of infinite dimensional systems. Second edition*, Birkhauser, 2007.

[Benton; 1977] Benton S.H., *The Hamilton-Jacobi equation: a global approach*, Academic Press, 1977.

[Blake, Cairns, Dowd; 2001] Blake D., Cairns A.J.G., Dowd K., *Pensionmetrics: stochastic pension plan design and value-at-risk during the accumulation phase*, Insurance: Mathematics and Economics, **29**, pp. 187–215, 2001.

[Blake, Cairns, Dowd; 2003] Blake D., Cairns A.J.G., Dowd K., *Pensionmetrics 2: stochastic pension plan design during the distribution phase*, Insurance: Mathematics and Economics, **33**, pp. 29–47, 2003.

[Blanchett-Scaillet, El Karoui, Jeanblanc, Martellini; 2008] Blanchett-Scalliet C., El Karoui N., Jeanblanc M., Martellini L., *Optimal investment decisions when time-horizon is random*, Journal of Mathematical Economics, **44**, 11, pp. 1100–1113, 2008.

[Booth; 1995] Booth P., *The management of Investment risk for defined contribution pension schemes*, Transaction of the 25th International Congress of Actuaries, **3**, International Actuarial Association, 1995.

[Booth, Yakoubov; 2000] Booth P., Yakoubov Y., *Investment policy for defined contribution pension scheme members close to retirement: an analysis of the lifestyle concept*, North American Actuarial Journal, **4**, 2, pp. 1–19, 2000.

[Bouchard, Pham; 2004] Bouchard B., Pham., *Wealth-path dependent utility maximization in incomplete markets*, Finance and Stochastics, **8**, pp. 579–603, 2004.

[Boulier, Huang, Taillard; 2001] Boulier J.F., Huang S.J., Taillard G., *Optimal management under stochastic interest rates: the case of a protected pension fund*, Insurance: Mathematics and Economics, **28**, 2, pp. 173–189, 2001.

[Boulier, Michel, Wisnia; 1996] Boulier J.F., Michel S., Wisnia V., *Optimizing investment and contribution policies of a defined benefit pension fund*, Proceedings of the 6th AFIR International Colloquim, **1**, pp. 593–607, 1996.

[Boulier, Trussant, Florens; 1995] Boulier J.F., Trussant E., Florence D., *A dynamic model for pension funds management*, Proceedings of the 5th AFIR International Colloquim, **1**, pp. 361–384, 1995.

[Bruckner; 1978] Bruckner A., *Differentiation of real functions*, Lecture Notes in Mathematics, **659**, Springer-Verlag, Berlin-Heidelberg-New York, 1978.

[Bjork, Svensson; 2001] Bjork T., Svensson L., *On the existence of finite dimensional realizations for nonlinear forward rate models*, mathematical Finance, **11**, 2001.

[Cabré, Caffarelli; 1995a] Cabré X., Caffarelli L., *Fully nonlinear elliptic equations*, American Mathematical Society, Providence, 1995.

[Cabré, Caffarelli; 1995b] Cabré X., Caffarelli L., *Regularity for viscosity solutions of fully nonlinear equations* $F(D^2u) = 0$, Topological Methods in Nonlinear Analysis, **6**, pp. 31–48, 1995.

[Cabré, Caffarelli; 2003] Cabré X., Caffarelli L., *Interior* $C^{2,\alpha}$ *regularity theory for a class of nonconvex fully nonlinear elliptic equations*, Journal de Mathématiques Pures et Appliquées, **82**, pp. 573–612, 2003.

[Cadenillas, Sethi; 1997] Cadenillas A., Sethi S., *Consumption-investment problem with subsistence consumption, bankruptcy, and random market coefficients*, Journal of Optimization Theory and Applications, **93**, 2, pp. 243–272, 1997.

[Caffarelli; 1989] Caffarelli L., *Interior a priori estimates for solutions of fully nonlinear equations*, Annals of Mathematics, **130**, pp. 189–213, 1989.

[Caffarelli, Crandall, Kocan, Swiech; 1996] Caffarelli L., Crandall M., Kocan M., Swiech A., *On viscosity solutions of fully nonlinear equations with measurable ingredients*, Communications on Pure and Applied Mathematics, **XLIX**, pp. 365–397, 1996.

[Caffarelli, Yuan; 2000] Caffarelli L., Yuan Y., *A priori estimates for solutions of fully nonlinear equations with convex level sets*, Indiana University Mathematics Journal, **49**, 2, pp. 681–695, 2000.

[Cairns; 1996] Cairns A., *Continuous-time stochastic pension fund modelling*, Proceedings of the 5th AFIR International Colloquim, **1**, pp. 609–624, 1996.

[Cairns; 2000] Cairns A., *Some notes on the dynamics and optimal control of stochstic pension fund models in continuous time*, Astin Bullettin, **30**, 1, pp.19–55, 2000.

[Cairns, Blake, Dowd; 2000] Cairns A.J G., Blake D., Dowd K., *Optimal dynamic asset allocation for defined-contribution pension plans*, Proceedings of the 10th AFIR International Colloquium, pp. 131–154, 2000.

[Cairns, Parker; 1997] Cairns A., Parker G., *Stochastic pension fund modelling*, Insurance: Mathematics and Economics, **21**, pp. 43–79, 1997.

[Cannarsa, Da Prato; 1991] Cannarsa P., Da Prato G., *Second order Hamilton-Jacobi equations in infinite dimensions*, SIAM Journal of Control and Optimization, **29**, 2, pp. 474–492, 1991.

[Cannarsa, Da Prato; 1992] Cannarsa P., Da Prato G., *Direct solution of a second order Hamilton-Jacobi equation in Hilbert spaces*, Stochastic partial differential equations and applications, G. Da Prato & L. Tubaro Editors, Pitman Research Notes in Mathematics **268**, pp. 72–85, 1992.

[Cannarsa, Di Blasio; 1995] Cannarsa P., Di Blasio G., *A direct approach to infinite dimensional Hamilton-Jacobi equations and applications to convex control with state constraints*, Differential Integral Equations, **8**, 2, pp. 225-246, 1995.

[Cannarsa, Di Blasio; 1993] Cannarsa P., Di Blasio G., *Dynamic programming for an abstract second order evolution equation with convex state constraints*, Control of Partial Differential Equations, IFIP WG 7.2 Conference, Villa Madruzzo, Trento, Italy, 1993.

[Cannarsa, Gozzi, Soner; 1991] Cannarsa P., Gozzi F., Soner H.M., *A boundary-value problem for Hamilton-Jacobi equations in Hilbert spaces*, Applied Mathematics and Optimization, **24**, 2, pp. 197-220, 1991.

[Cannarsa, Sinestrari; 2005] Cannarsa P., Sinestrari C., *Semiconcave functions, Hamilton-Jacobi equations and optimal control*, Springer, Progress in Nonlinear Equations and Their Applications, **58**, 2005.

[Cannarsa, Soner; 1989] Cannarsa P., Soner H.M. *Generalized one-sided estimates for solutions of Hamilton-Jacobi equations and applications*, Nonlinear Analysis, **13**, pp. 305–323, 1989.

[Capuzzo-Dolcetta, Lions; 1990] Capuzzo Dolcetta I., Lions P.L., *Viscosity solutions of Hamilton-Jacobi equations and state constraints*, Transactions of the American Mathematical Society, **318**, 2, pp. 643–683, 1990.

[Chang, Peng, Pemy; 2008] Chang M.H., Pang T., Pemy M., *Finite Difference Approximations for Stochastic Control Systems with Delay*, Stochastic Analysis and Applications, **26**, 3, 451-470, 2008.

[Charupat, Milevsky; 2002] Charupat N., Milevsky M.A., *Optimal asset allocation in life annuities: a note*, Insurance: Mathematics and Economics **30**, pp. 199–209, 2002.

[Chojnowska-Michalik, 1978] Chojnowska-Michalik A., *Representation theorem for general stochastic delay equations*, Bulletin de lAcadémie Polonaise des Sciences. Serie des Sciences Mathématiques, **26**, 7, pp. 635-642, 1978.

[Choulli, Taksar, Zhou; 2003] Choulli T., Taksar M., Zhou X.Y., *A diffusion model for optimal dividend distribution for a company with constraints on risk control*, SIAM Journal on Control and Optimization **41**, 6, pp. 1946–1979, 2003.

[Crandall, Ishii, Lions; 1992] Crandall M., Ishii H., Lions P.L., *User's guide to viscosity solutions of second order partial differential equations*, Bulletin of the American Mathematical Society (New Series), **27**, 1, pp. 1–67, 1992.

[Crandall, Kocan, Swiech; 2000] Crandall M., Kocan M., Swiech A., *L^p theory for fully nonlinear uniformly parabolic equations*, Communications in Partial Differential Equations, **25**, 11& 12, pp. 1997–2053, 2000.

[Crandall, Lions; 1990] Crandall M., Lions P.L., *Hamilton-Jacobi equations in infinite dimension (Part IV): unbounded linear term*, Journal of Functional Analysis, **90**, pp. 237—283, 1990.

[Crandall, Lions; 1991] Crandall M., Lions P.L., *Hamilton-Jacobi equations in infinite dimension (Part V): B-continuous solutions*, Journal of Functional Analysis, **97**, pp. 417—465, 1991.

[Da Prato, Zabczyk; 1992] Da Prato G., Zabczyk J., *Stochastic equations in infinite dimensions*, Encyclopedia of Mathematics and its Applications, Cambridge University Press, Cambridge (UK), 1992.

[Da Prato, Zabczyk; 1996] Da Prato G., Zabczyk J., *Ergodicity for infinite dimensional systems*, London mathematical society lecture note series **229**, Cambridge University Press, 1996.

[Da Prato, Zabczyk; 2002] Da Prato G., Zabczyk J., *Second order partial differential equations in Hilbert spaces*, London mathematical society lecture note series **293**, Cambridge University Press, 2002.

[Davies; 1980] Davies E. B., *One Parameter Semigroups*, Academic Press, 1980

[Davis; 1995] Davis E.P., *Pension funds: retirement-income security and capital markets: an international perspective*, Oxford University Press, Oxford, 1995.

[Deelstra, Grasselli, Koehl; 2003] Deelstra G., Grasselli M., Koehl P.F., *Optimal investment strategies in the presence of a minimum guarantee*, Insurance: Mathematics and Economics, **33**, 1, pp. 189–207, 2003.

[Deelstra, Grasselli, Koehl; 2004] Deelstra G., Grasselli M., Koehl P.F., *Optimal design of a guarantee for defined contribution funds*, Journal of Economic Dynamics and Control, **28**, 11, pp. 2239–2260, 2004.

[Devolder, Bosch Princep, Dominguez Fabian; 2003] Devolder P., Bosch Princep M., Dominguez Fabian I., *Stochastic optimal control of annuity contracts*, Insurance: Mathematics and Economics, **33**, pp. 227–238, 2003.

[Di Blasio; 1985] Di Blasio G., *Global solutions for a class of Hamilton-Jacobi equations in Hilbert spaces*, Numerical Functional Analysis and Optimization, Vol. 8, No. 3-4, pp. 261–300, 1985/86.

[Di Blasio; 1991] Di Blasio G., *Optimal control with infinite horizon for distributed parameter systems with constrained controls*, SIAM Journal on Control and Optimization, Vol. 29, No. 4, pp. 909–925, 1991.

[Di Giacinto, Federico, Gozzi; 2009] Di Giacinto M., Federico S., Gozzi F., *Pension funds with a minimum guarantee: a stochastic control approach*, Finance and Stochastics, accepted paper.

[Di Giacinto, Federico, Gozzi, Vigna; WP] Di Giacinto M., Federico S., Gozzi F., Vigna E., *Constrained portfolio choices in the decumulation phase of a pension plan*, Working Paper.

[Duffie, Fleming, Soner, Zariphopoulou; 1997] Duffie D., Fleming W., Soner H.M., Zariphopoulou T., *Hedging in incomplete markets with HARA utility*, Journal of Economic Dynamics and Control, **21**, 4–5, pp. 753–782, 1997.

[Dunford, Schwartz; 1958] Dunford N., Schwartz J., *Linear operators, Vol.1*, Interscience, 1958.

[Ekeland, Taflin; 2005] Ekeland I., Taflin E., *A Theory of Bond Portfolios*, The Annals of Applied Probability, **15**, 2, 1260–1305, 2005.

[Ekeland, Temam; 1976] Ekeland I., Temam R., *Convex analysis and variational problems*, North Holland Company, 1976.

[Elie, Touzi; 2008] Elie R., Touzi N., *Optimal lifetime consumption and investment under drawdown constraint*, Finance and Stochastics, **12**, 3, pp. 299–330, 2008.

[El Karoui, Jeanblanc, Lacoste; 2005] El Karoui N., Jeanblanc M., Lacoste V., *Optimal portfolio management with American capital guarantee*, Journal of Economic and Dynamic Control. **29**, 3, pp. 449–468, 2005.

[Elsanousi, Øksendal, Sulem; 2000] Elsanousi I., Øksendal B., Sulem A., *Some solvable stochastic control problems with delay*, Stochastics and Stochastics Reports, **71**, pp 69 - 89, 2000.

[Engel, Nagel; 2000] Engel K.J., Nagel R.,*One-paramater semigroups for linear evolution equations*, Springer-Verlag, New York, 2000.

[Escauriaza; 1993] Escauriaza L., $W^{2,n}$ *a priori estimates for solutions to fully nonlinear equations*, Indiana University Mathematics Journal, **42**, 2, pp. 413–423,1993.

[Evans; 1982] Evans L., *Classical solutions of fully nonlinear, convex, second-order elliptic equations*, Communications on Pure and Applied Mathematics, **XXV**, pp. 333–363, 1982.

[Evans; 1998] Evans L., *Partial Differential Equations*, American MAthematical Society, Graduate Studies in Mathematics, **19**, Providence, 1998.

[Fabbri, Gozzi, Swiech] Fabbri G., Gozzi F., Swiech A., *Verification Theorems and construction of ε-optimal controls*, Working paper.

[Faggian; 2005] Faggian S., *Regular solutions of Hamilton–Jacobi equations arising in Economics*, Applied Mathematics and Optimization, **51**, 2, pp. 123–162, 2005.

[Faggian; 2008] Faggian S., *Infinite dimensional Hamilton–Jacobi-Bellman equations and applications to boundary control with state constraints*, SIAM Journal on Control and Optimization, **47**, 4, pp.2157–2178, 2008.

[Federico; 2008] Federico S., *A pension fund in the accumulation phase: a stochastic control approach,* Banach Center Publications, **83**: Advances in Mathematics of Finance, 2008.

[Federico; 2007] Federico S., *A pension fund model with surplus: an infinite dimensional stochastic control approach*, Preprint n. 8, Pisa: Scuola Normale Superiore, 2007

[Federico; 2008] Federico S., *A stochastic control problem with delay arising in a pension fund model*, submitted to Finance and Stochastics, 2008.

[Federico; WP] Federico S., *A viscosity approach for a Hamilton-Jacobi-Bellman equation arising in a stochastic control problem with delay*, working paper, 2009.

[Federico, Goldys, Gozzi; 2009a] Federico S., Goldys B., Gozzi F., *HJB equations for the optimal control of differential equations with delays and state constraints, I: regularity and applications*, ArXiv preprint, submitted to SIAM Journal on Control and Optimization, 2009.

[Federico, Goldys, Gozzi; 2009b] Federico S., Goldys B., Gozzi F., *HJB equations for the optimal control of differential equations with delays and state constraints, II: Optimal Feedbacks and Approximations*, ArXiv preprint, 2009.

[Federico, Øksendal; 2009] Federico S., Øksendal B., *Optimal stopping of stochastic differential equations with delay driven by a Lévy noise*, Preprint of the University of Oslo, Submitted to Potential Analysis, 2009.

[Feller; 1952] Feller W., The parabolic differential equatios and the associated semigroup of transformations, Annals of Mathematics, **55**, 3, 1952.

[Feller; 1954] Feller W., Diffusion processes in one dimension, Transactions of the American Mathematical Society, **77**, 1, 1954.

[Filipovic; 2001] Filipovic D., *Consistency problems for Heath-Jarrow-Morton interest rate models*, Springer, 2001.

[Fischer, Nappo; 2008] Fischer M., Nappo G., *Time Discretisation and Rate of Convergence for the Optimal Control of Continuous-Time Stochastic Systems with Delay*, Applied Mathematics and Optimization, **57**, 2, 177–206, 2008.

[Fischer, Nappo; 2007] Fischer M., Reiss M., *Discretisation of Stochastic Control Problems for Continuous Time Dynamics with Delay*, Journal of Computational and Applied Mathematics, **205**, 969–981, 2007.

[Fleming, Soner; 1993] Fleming V.H., and Soner H.M., *Controlled Markov processes and viscosity solutions*, Springer-Verlag, New York, 1993.

[Fuhrman, Tessitore; 2004] Fuhrman M., Tessitore G., *Infinite horizon backward stochastic differential equations and elliptic equations in Hilbert spaces*, *Annals of Probability*, **32**, 1B, pp. 607–660, 2004.

[Gabay, Grasselli; 2009] Gabay D., Grasselli M., *Fair demographic risk sharing in defined contribution pension funds*, working paper ESILV.

[Gao; 2008] Gao J., *Stochastic optimal control of DC pension funds*, Insurance: Mathematics and Economics, **42**, pp.1159-1164, 2008.

[Gerrard, Hojgaard, Vigna; 2004] Gerrard R., Haberman S., Hojgaard B., Vigna E., *The income drawdown option: quadratic loss*, Actuarial Research No.155, Cass Business School, City University, London, 2004.

[Gerrard, Haberman, Vigna; 2004] Gerrard R., Haberman S., Vigna E., *Optimal investment choices post retirement in a defined contribution pension scheme*, Insurance: Mathematics and Economics, **35**, pp. 321–342, 2004.

[Gerrard, Haberman, Vigna; 2006] Gerrard R., Haberman S., Vigna E., *The management of decumulation risks in a defined contribution pension scheme*, North American Actuarial Journal, **10**, pp. 84–110, 2006.

[Gerrard, Hojgaard, Vigna; 2008] Gerrard R., Hojgaard B., Vigna E., *Choosing the optimal annuitization time post retirement*, 2008, mimeo.

[Gilbarg, Trudinger; 1983] Gilbarg D., Trudinger N.S., *Elliptic partial differential equations of second order*, Springer-Verlag, 1983.

[Goetzmann, Ingersoll, Ross; 2003] Goetzmann W.N., Ingersoll J.E., Ross S.A. *High-water marks and hedge fund management contracts*, *The Journal of Finance*, **58**, 4, pp. 1685–1718, 2003.

[Goldys, Gozzi; 2006] Goldys B., Gozzi F., *Second order Hamilton-Jacobi equations in Hilbert spaces and stochastic optimal control: L^2 approach*, *Stochastic Processes and their Applications*, **116**, 12, pp. 1932–1963, 2006.

[Goldys, Musiela, Sondermman; 2000] Goldys B., Musiela M., Sondermann D., *Lognormality of rates and term structure models*, Stochastic Analysis and Appllications, **18**, 375396, 2000.

[Gozzi; 1989] Gozzi F., *Some results for an infinite horizon control problem governed by a semilinear state equation*, Proceedings Vorau, July 10-16 1988; editors F.Kappel, K.Kunisch, W.Schappacher; International Series of Numerical Mathematics, **91**, Birkäuser - Verlag, Basel, pp.145-163, 1989.

[Gozzi; 1991] Gozzi F., *Some results for an optimal control problem with a semilinear state equation II*, SIAM Journal on Control and Optimization, **29**, 4, pp. 751-768, 1991.

[Gozzi; 1998] Gozzi F., *Second order Hamilton-Jacobi equations in Hilbert spaces and stochastic optimal control*, Tesi di Perfezionamento in Matematica, Scuola Normale Superiore, Pisa, 1998.

[Gozzi, Marinelli; 2004] Gozzi F., Marinelli C., *Stochastic optimal control of delay equations arising in advertising models*, Da Prato (ed.) et al., Stochastic partial differential equations and applications VII - Papers of the 7th meeting, Levico Terme, Italy, January 5–10, 2004. Boca Raton, FL: Chapman & Hall/CRC. Lecture Notes in Pure and Applied Mathematics, **245**, pp. 133–148, 2004.

[Gozzi, Marinelli, Savin] Gozzi F., Marinelli C., Savin S., *On controlled linear diffusions with delay in a model of optimal advertising under uncertainty with memory effects*, Journal of Optimization, Theory and Applications, to appear.

[Gozzi, Rouy, Swiech; 2000] Gozzi F., Rouy E., Swiech A., *Second order Hamilton-Jacobi equations in Hilbert spaces and Stochastic Boundary Control*, *Siam Journal on Control and Optimization*, **38**, 2, 400–430, 2000.

[Gozzi, Swiech; 2000] Gozzi F., Swiech A., *Hamilton-Jacobi-Bellman equations for the optimal control of the Duncan-Mortensen-Zakai equations*, Journal of Functional Analysis, **172**, pp. 466–510, 2000.

[Gozzi, Sritharan, Swiech; 2005] Gozzi F., Sritharan S.S., Swiech A., *Bellman equations associated to the optimal feedback control of stochastic Navier-Stokes equations*, Communications on Pure and Applied Mathematics, **LVIII**, pp. 671–700, 2005.

[Haberman, Sung; 1994] Haberman S., Sung J.H., *Dynamic approaches to pension funding*, Insurance: Mathematics and Economics, **15**, pp. 151–162, 1994.

[Haberman, Vigna; 2001] Haberman S., Vigna E., *Optimal investment strategy for defined contribution pension schemes*, Insurance: Mathematics and Economics, **28**, 2, pp. 233–262, 2001.

[Haberman, Vigna; 2002] Haberman S., Vigna E., *Optimal investment strategies and risk measures in defined contribution pension schemes*, Insurance: Mathematics and Economics, **31**, 1, pp. 35–69, 2002.

[Hagood, Thomson; 2006] Hagood J.W., Thomson B.S., *Recovering a function from a Dini derivative*, The American Mathematical Monthly, January 2006.

[Hale, Verduyn; 1993] Hale J.K., Verduyn Lunel S.M., *Introduction to Functional Differential Equations*, Springer-Verlag, Applied Mathematical Sciences, **99**, 1993.

[Heath, Jarrow, Morton; 1992] Heath D.C., Jarrow R.A., Morton A., *Bond pricing and term structures of interest rates: a new methodology for contingent claim valuation*, Econometrica, **60**, 77–105, 1992.

[Højgaard, Vigna; 2007] Højgaard B., Vigna E., *Mean-variance portfolio selection and efficient frontier for defined contribution pension schemes*, 2007.

[Ishii; 1993] Ishii H., *Viscosity solutions of nonlinear second-order partial differential equations in Hilbert spaces*, Communications in Partial Differential Equations, **18**, pp. 601–651, 1993.

[Ishii, Lions; 1990] Ishii H., Lions P.L., *Viscosity solutions of fully nonlinear second order partial differential equations*, Journal of Differential Equations, **83**, pp. 26–78, 1990.

[Ishii, Loreti; 2002] Ishii H., Loreti P., *A class of stochastic optimal control problems with state constraint*, Indiana University Mathematics Journal, **51**, 5, pp. 1167–1196, 2002.

[Ikeda, Watanabe; 1981] Ikeda N., Watanabe S., *Stochastic differential equations and diffusion processes*, North Holland, Amsterdam, 1981.

[Kallenberg; 1997] Kallenberg O., Fundations of Modern Probability, Springer, 1997.

[Kapur, Orszag; 1999] Kapur S., Orszag J.M., *A portfolio approach to investment and annuitization during retirement*, Proceedings of the Third International Congress of Insurance: Mathematics and Economics, 1999.

[Karatzas, Lehoczky, Sethi, Shreve; 1986] Karatzas I., Lehoczky J.P., Sethi S., Shreve S.E., *Explicit solution of a general consumption/investment problem*, Mathematics of Operations Research, **11**, 2, pp. 261–294, 1986.

[Karatzas, Shreve; 1991] Karatzas I., Shreve S.E., *Brownian motion and stochastic calculus*, Second Edition, Springer Verlag, New York, 1991.

[Katsoulakis; 1994] Katsoulakis M., *Viscosity solutions of second order fully nonlinear elliptic equations with state constraints*, Indiana University Mathematics Journal, **43**, 2, pp. 493–519, 1994.

[Kelome, Swiech; 2003] Kelome D., Swiech A., *Viscosity solution of an infinite-dimensional Black-Scholes-Barenblatt equation*, Applied Mathematics and Optimization, **47**, 3, pp. 253–278, 2003.

[Khorasanee; 1995] Khorasanee M.Z., *Simulation of investment returns for a money purchase fund*, Journal of Actuarial Practice, **3**, 1, pp. 93–115, 1995.

[Kohorasanee; 1996] Khorasanee M.Z., *Annuity choices for pensioners*, Journal of Actuarial Practice, **4**, pp. 229–235, 1996.

[Khorasanee; 1998] Khorasanee M.Z., *Deterministic modelling of defined-contributions pension funds*, North American Actuarial Journal, **1**, 4, pp. 83–103, 1998.

[Knox; 1993] Knox D.M., *A critique of defined contribution plans using a simulation approach*, Journal of Actuarial Practice, **1**, 2, pp. 49–66, 1993.

[Kocan, Soravia; 1998] Kozan M., Soravia P., *A viscosity approach to infinite-dimensional Hamilton-Jacobi equations arising in optimal control with state constraints*, SIAM Journal of Control and Optimization, **36**, 4, pp. 1348–1375, 1998.

[Kolmanovskii, Shaikhet; 1996] Kolmanovskii V.B., Shaikhet L.E., *Control of systems with aftereffect*, Providence: American Mathematical Society, Translations of mathematical monograph, **157**, 1996.

[Krylov; 1993] Krylov N.V., *Controlled diffusion processes*, Springer-Verlag, Berlin, New-York, 1993.

[Krylov; 1987] Krylov N.V., *Nonlinear elliptic and parabolic equations of the second order*, D. Reidel Publishing Company, 1987.

[Krylov; 1983] Krylov N.V., *Boundedly inhomogeneous elliptic and parabolic equations*, Izvestiya Rossiiskoi Akademii Nauk, **46**, 1982 (Russian). English translation: Mathematics of the USSR: Izvestija, **20**, pp. 459–492, 1983.

[Kushner; 1977] Kushner H.J., *Probability Methods for Approximations in Stochastic Control and for Elliptic Equations*, Academic Press, New York, 1977.

[Kushner; 2006] Kushner H.J., *Numerical Approximations for Stochastic Systems with Delays in the State and Control*, Stochastics: an International Journal of Probability and Stochastic Processes, **78**, 5, 343–376, 2006.

[Kydland, Prescott; 1982] Kydland F.E., Prescott E.C., *Time-to-build and aggregate fluctuations*, Econometrica, **50**, 6, pp. 1345–1370, 1982.

[Ladyzenskaja, 1968] Ladyzenskaja O.A., Solonnikov V.A., Uralceva N.N., *Linear and quasilinear equations of parabolic type*, American Mathematical Society, Providence, 1968.

[Larrsen; 2002] Larssen B., *Dynamic programming in stochastic control of systems with delay*, Stochastics and Stochastics Reports, **74**, 3-4, pp 651-673, 2002.

[Larrsen, Risebro; 2003] Larssen B., Risebro N.H., *When are HJB-equations for control problems with stochastic delay equations finite dimensional?*, Stochastic Analysis and Applications, **21**, 3, pp. 643-671, 2003.

[Lasry, Lions; 1989] Lasry J.M., Lions P.L., *Nonlinear elliptic equations with singular boundary conditions and state constraints. I. The model problem*, Mathematische Annalen, **283**-4, pp. 583–630, 1989.

[Li, Yong; 1995] Li X., Yong J., *Optimal control theory for infinite-dimensioal systems*, Birkhauser Boston, 1995.

[Lions; 1988] Lions P.L., *Viscosity solutions of fully nonlinear second-order equations and optimal stochastic control in infinite dimensions. Part I: The case of bounded stochastic evolution*, Acta Mathematica, **161**, pp. 243–278, 1988.

[Lions; 1989a] Lions P.L., *Viscosity solutions of fully nonlinear second-order equations and optimal stochastic control in infinite dimensions. Part II: Optimal Control of Zakais equation*, Lecture Notes in Mathematics, **1390**, eds. G. Da Prato & L.Tubaro, Springer-Verlag, Berlin, pp. 147–170, 1989.

[Lions; 1989b] Lions P.L., *Viscosity solutions of fully nonlinear second-order equations and optimal stochastic control in infinite dimensions. Part III: Uniqueness of viscosity solutions for general second order equations*, Journal of Functional Analysis, **86**, pp.1–18, 1989.

[Ludvik; 1994] Ludvik P., *Investment strategy for defined contribution plans*, 4th AFIR International Colloquium, **3**, pp. 1389–1400, 1994.

[Merton; 1969] Merton R.C., *Lifetime portfolio selection under uncertainty: the continuous-time case*, Review of Economics and Statistics, **51**, pp. 247–257, 1969.

[Merton; 1971] Merton R.C., *Optimum consumption and portfolio rules in a continuous-time model*, Journal of Economic Theory, **3**, pp. 373–413, 1971.

[Merton; 1992] Merton R.C., *Continuous-time finance*, Blackwell, United States of America, 1992.

[Milevsky, 2001] Milevsky M.A., *Optimal annuitization policies: analysis of the options*, North American Actuarial Journal, **5**, pp. 57–69, 2001.

[Milevsky, Moore, Young; 2006] Milevsky M.A., Moore K.S., Young V.R., *Optimal asset allocation and ruin-minimization annuitization strategies*, Mathematical Finance, **16**, pp. 647–671, 2006.

[Milevsky, Young; 2007] Milevsky M.A., Young V.R., *Annuitization and asset allocation*, Journal of Economic Dynamics and Control, **31**, pp. 3138–3177, 2007.

[Mohammed; 1984] Mohammed S.E., *Stochastic functional differential equations*, Research Notes in Mathematics, **99**, Pitman, London, 1984.

[Morimoto, 2008] Morimoto H., *Optimal consumption models in economic growth*, Journal of Mathematical Analysis and Applications, **337**, pp. 480–492, 2008.

[Musiela; 1993] Musiela M., *Stochastic PDEs and term structure models*, Journeés Internationales de Finance, IGR-AFFI, 1993.

[O'Brien; 1986] O' Brien T., *A stochastic dynamic approach to pension funding*, Insurance: Mathematics and Economics, **5**, pp. 141–146, 1986.

[Øksendal, Sulem; 2001] Øksendal B., Sulem A., *A maximum principle for optimal control of stochastic systems with delay with applications to finance*, Optimal Control and PDE, Essays in Honour of Alain Bensoussan, eds J.L. Menaldi, E. Rofman and A. Sulem, IOS Press, Amsterdam, pp. 64-79, 2001.

[Øksendal, Zhang; 2008] Øksendal B., Zhang H., *Optimal control with partial information for stochastic Volterra equations*, Preprint of University of Oslo, 2008.

[Pazy; 1983] Pazy A., *Semigroups of linear operators and applications to partial differential equations*, Springer-Verlag, 1983.

[Pontryagin, Boltyanski, Gamkerildze, Mishenko; 1962] Pontryagin L.S., Boltyanski V.G., Gamkelidze R.V., Mishenko E.S., *Mathematical theory of optimal processes*, Wiley, New York, 1962.

[Preisel, Jarner, Eliasen; 2008] Preisel M., Jarner S., Eliasen R., *Investing for retirement through a with-profits pension scheme: a client's perspective*, Scandinavian Actuarial Journal, pp. 1–21, 2008.

[Protter; 2003] Protter P.E., *Stochastic integration and differential equations*, 2nd edition, Springer-Verlag, Berlin-Heidelberg-New York, 2003.

[Revuz, Yor; 1999] Revuz D., Yor M., *Continuous martingales and Brownian motion*, 3rd edition, Springer, 1999.

[Rockafellar; 1970] Rockafellar R.T., *Convex analysis*, Princeton University Press, 1970.

[Saks; 1964] Saks S., *Theory of the integral*, Dover Publications, 1964.

[Sbaraglia, Papi, Briani, Bernaschi, Gozzi; 2003] Sbaraglia S., Papi M., Briani M., Bernaschi M., Gozzi F., A model for the optimal asset-liability management for insurance companies. *International Journal of Theoretical and Applied Finance* **6**, 3, pp. 277–299, 2003.

[Sethi, Taksar; 1992] Sethi S., Taksar M., *Infinite-horizon investment consumption model with a nonterminal bankruptcy*, Journal of Optimization Theory and Applications, **74**, 2, pp. 333–346, 1992.

[Sethi, Taksar, Presman; 1992] Sethi S., Taksar M., Presman E., *Explicit solution of a general consumption/portfolio problem with subsistence consumption and bankruptcy*, *Journal of Economic Dynamics and Control*, **16**, 3–4, pp. 747–768, 1992.

[Soner; 1986] Soner H.M., *Optimal control with state-space constraint I*, SIAM Journal on Control and Optimization, **24**, 3, pp. 552–561, 1986.

[Soner; 2004] Soner H.M., *Stochastic optimal control in finance*, Cattedra Galileiana, Pisa: Scuola Normale Superiore, 2004.

[Soner, Touzi; 2002] Soner M., Touzi N., *Stochastic Target Problems, Dynamic Programming and Viscosity Solutions*, SIAM Journal of Control and Optimization, **41**, 2, pp. 404–424, 2002.

[Stabile; 2006] Stabile G., *Optimal timing of the annuity purchase: combined stochastic control and optimal stopping problem*, International Journal of Theoretical and Applied Finance, **9**, pp. 151–170, 2006.

[Starks; 1987] Starks L. T., *Performance incentive fees: an agency theoretic approach*, The Journal of Financial and Quantitative Analysis, **22**, 1, pp. 17–32, 1987.

[Steffensen; 2004] Steffensen M., *On Merton's problem for life insurers*, Astin Bullettin, **34**, 1, pp. 5–25, 2004.

[Steffensen; 2006] Steffensen M., *Quadratic optimization of life and pension insurance payments*, Astin Bullettin, **36**, 1, pp. 245–267, 2006.

[Swiech; 1994] Swiech A., *Unbounded second order partial differential equations in infinite dimensional Hilbert spaces*, Communications in Partial Differential Equations **19**, pp. 1999–2036, 1994.

[Swiech; 1997] Swiech A., *$W^{1,p}$ interior estimates for solutions of fully nonlinear, uniformly elliptic equations*, Advances in Differential Equations, **2**, 6, pp. 1005–1027, 1997.

[Taylor I; 1996] Taylor M.E., *Partial differential equations I. Basic theory*, Springer, 1996.

[Taylor III; 1996] Taylor M.E., *Partial differential equations III. Nonlinear equations*, Springer, 1996.

[Tebaldi, Schwartz; 2006] Tebaldi C., Schwartz E.S., *Illiquid assets and optimal portfolio choice*, NBER Working Paper No. W12633, 2006.

[Vargiolu; 1999] Vargiolu T., *Invariant measures for the Musiela equation with deterministic diffusion term*, Finance and Stochastics, **3**, 4, 483492, 1999.

[Vinter, Kwong; 1981] Vinter R.B., Kwong R.H., *The infinite time quadratic control problem for linear systems with state and control delays: an evolution equation approach*, SIAM Journal of Control and Optimization, **19**, 1, pp. 139–153, 1981.

[Wang I; 1992] Wang L., *On the regularity of fully nonlinear parabolic equations. I*, Communications in Pure and Applied Mathematics, **45**, pp. 27–76, 1992.

[Wang II; 1992] Wang L., *On the regularity of fully nonlinear parabolic equations. II*, Communications in Pure and Applied Mathematics, **45**, pp. 141–178, 1992.

[Wang III; 1992] Wang L., *On the regularity of fully nonlinear parabolic equations. III*, Communications in Pure and Applied Mathematics, **45**, pp. 255–262, 1992.

[Xiao, Zhai, Qin; 2007] Xiao J., Zhai H., Qin C., *The constant elasticity of variance (CEV) model and the Legendre transform-dual solution for annuity contracts*, Insurance: Mathematics and Economics, **40**, pp. 302-310, 2007.

[Yaari; 1965] Yaari M.E., *Uncertain lifetime, life insurance and the theory of the consumer*, Review of Economic Studies, **32**, pp. 137–150, 1965.

[Yamada, Watanabe; 1971] Yamada T., Watanabe S., *On the uniqueness of solutions of stochastic differential equations*, Journal of Mathematics of Kyoto University, **11**, 1, 155–167, 1971.

[Yosida; 1980] Yosida K., *Functional analysis*, VI edition, Springer-Verlag, 1980.

[Yong, Zhou; 1999] Yong J., Zhou X.Y., *Stochastic controls: Hamiltonian systems and HJB equations*, Springer Verlag, New York, 1999.

[Zariphopoulou; 1994] Zariphopoulou T., *Consumption-investment models with constraints*, *SIAM Journal on Control and Optimization*, **32**, 1, pp. 59–85, 1994.

[Zhou, Yong, Li; 1997] Zhou X.Y., Yong J., Li X., *Stochastic Verification theorems within the framework of viscosity solutions*, SIAM Journal on Control and Optimization, **37**, pp. 243–253, 1997.