To my wonderful family

Acknowledgements

First and foremost I would like to thank my advisor, Dr. Salvatore Federico, for his continuous support, advice and patience during these years. He introduced me to the field of Control Theory and I owe him my personal discovery of this stimulating line of research. I am also grateful to him for our fruitful collaboration.

I would also like to thank my committee for their time and comments: Professor Fausto Gozzi closely followed all my research projects and Professor Marco Papi introduced me to the Malliavin Calculus.

I would like to express my gratitude towards Professor Francesco Russo for stimulating discussions, for having me introduced to the exciting field of Stochastic Analysis, and, mainly, for having first been my professor before becoming a personal friend.

I want to thank Professor Massimo Bernaschi for having been close to me during all these years: as my professor, my boss, my tutor. He always stood intellectually and financially by my side, allowing me to exploit research opportunities and to develop my skills as a researcher.

I would like to thank my family for their deep love and their patience: without their encouragement this thesis would not have been possible!

I want to thank Nicolas Serrano-Velarde for helping me patiently with English, and mainly, for giving me fondly his support in these years.

I also would like to thank my friends of the "Research Lab" in via della Marranella in Rome for their support and for welcoming me during my last year of the thesis.

Last but not least: in the course of my thesis I also had the opportunity to meet exceptional persons that I would like to thank for their love, friendship and support: Eva Tennina, Graham Mott, Cristina Di Girolami, Mme Serrano-Velarde and her family.

Abstract

The thesis is composed by two different parts, which are not related each other.

The first part is devoted to study a class of optimal control problems, where the state equation is an ordinary differential equation with delay in the control variable. This class of problems arises in economic applications, in particular in optimal advertising problems (see [36, 60, 66]). The control problem is embedded in a suitable Hilbert space and the associated Hamilton-Jacobi-Bellman (HJB) equation considered in this space. It is proved that the value function is continuous with respect to a weak norm and that it solves in the viscosity sense the associated HJB equation. The main result is the proof of a directional C^1-regularity result for the value function. This result represents the starting point to define a feedback map in classical sense going towards a verification theorem yielding optimal feedback controls for the problem.

In the second part of the thesis, the techniques of the Malliavin Calculus are applied to a stochastic differential equation whose coefficients depend on a control process, in particular in the special case of Markovian controls. It is calculated the stochastic derivative of the stochastic differential equation and it is proved that the Malliavin matrix is strictly positive, assuring the results of existence and regularity of densities for the controlled process.

Introduction

The first part of the thesis is devoted to study of a class of optimal control problems with distributed delay in the control variable. The study of problems with delays is motivated by economic applications (see, e.g., [1, 2, 3, 36, 60, 66, 41, 42, 12, 40]) and engineering applications (see, e.g., [51]). In this work we focus on the economic issues of our model.

There is a wide variety of models with memory structures considered by the economic literature. We refer, for instance, to models where the memory structure arises in the state variable, as growth models with time to build in the production (see [1, 2, 3]) and as vintage capital model (see [12, 40]); to models where the memory structure arises in the control variable, as in advertising models (see [36, 60, 66, 41, 42]) or even in growth models with time to build in the investment (see [52, 69]). Our model covers a class of optimal advertising problems.

The analysis of advertising policies has always been occupying a front-and-center place in the marketing research since the seminal paper by Nerlove and Arrow [56]. Indeed, their model has paved the way for the development of a number of models dealing with the optimal distribution of advertising expenditure over time in both monopolistic and competitive settings. However, apart from the fact that the optimal advertising policy in the Nerlove- Arrow model is of the bang-bang type, their model has another rather unattractive feature. Indeed, the dynamics of the goodwill assumes that there is no time lag between advertising expenditures and the goodwill's growth. Therefore, attempts have been made to incorporate different distributions on the lifetime of each unit of goodwill into the dynamics of advertising capital. Indeed, it has been advocated in the literature (see the survey [36] and references therein) that a realistic dynamic model for the goodwill should allow for lags in the effect of advertisement. First of all, it is natural to assume that there will be a time lag between advertising expenditure and the corresponding effect on the goodwill level. Moreover, the literature has also considered lag structures allowing for a distribution of the forgetting time. The model by Pawels [60] introduced a time lag between the rate of advertising and its effect on the rate of sales. This leads to a control problem where the dynamics is given by a differential equation with delay in the control variable. In [60], Pawels approaches the optimal control problem by means of a maximum principle for this kind of problems provided by Sethi [66]. Instead, our approach to the problem is based on the dynamic programming techniques in infinite dimension.

From a mathematical point of view, our aim is to study the optimal control of the differential equation

$$\begin{cases} y'(t) = ay(t) + b_0 u(t) + \int_{-r}^{0} b_1(\xi) u(t+\xi) d\xi; \\ y(0) = y_0; \quad u(s) = \delta(s), \ s \in [-r, 0); \end{cases} \tag{0.1}$$

subject to the state constraint $y(\cdot) > 0$ and to the control constraint $u(\cdot) \in U \subset \mathbb{R}$. The objective is to maximize a functional of the form

$$\int_0^{+\infty} e^{-\rho t} \Big(g_0(y(t)) - h_0(u(t)) \Big) dt,$$

where $\rho > 0$ is a discount factor and $g_0 : \mathbb{R}^+ \to \mathbb{R}$ and $h_0 : U \to \mathbb{R}$ are respectively a utility and a cost function.

The presence of the delay in the state equation (0.1) renders not possible to apply the dynamic programming techniques to the problem in its current form. A general way to tackle this problem consists in representing the controlled system in a suitable infinite-dimensional space (see [11, Part II, Chapter 1]). In this way the delay is absorbed by the infinite-dimensional state. On the other hand, the price to pay is that the resulting system is infinite-dimensional. By the way, once the delay has been absorbed by the infinite-dimensional state, in principle we can apply the technique of Dynamic Programming in infinite dimension. In some cases this method can lead to fully or partially satisfactory answers to the problem. Indeed sometimes explicit solutions are available (see e.g. [3, 40]), or, alternatively one can hope to give at least a partial characterization of the optimal controls (see e.g. [34, 35]). The core of the approach, as usual dealing with the Dynamic Programming, is the study of the Hamilton-Jacobi-Bellman (briefly, HJB) equation, which is in this case an infinite-dimensional partial differential equation. Obviously, due to the infinite dimension, the study of the HJB becomes much more problematic. However, the aspects to investigate are the same as in the finite-dimensional case. In particular it is crucial to try to show that the value function of the control problem solves in some sense the HJB equation; possibly that it is the unique solution in this sense of the HJB equation; that some regularity holds for the solutions of the HJB equation. The last point is strictly connected to the possibility of giving a satisfactory answer to the control problem. This is due to the fact that, in order to obtain an optimal strategy in feedback form, one 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. [9, 71]), but this is usually not satisfactory in applied problems since the closed loop equation becomes very hard to treat in such cases.

The C^1 regularity of solutions to HJB equations is particularly important in infinite dimension since in this case verification theorems in the framework of viscosity solutions are rather weak (see e.g [29, 53]). To the best of our knowledge, C^1 regularity for first order HJB equation was proved by method of convex regularization introduced by Barbu and Da Prato [4] and then developed by various authors (see e.g. [5, 6, 7, 8, 25, 26, 30, 38, 39]). All these results do not cover problems originated by control problems with delays. In the papers [16, 17, 32] a class of state constraints problems is treated using the method of convex regularization, but the C^1 regularity is not proved.

We follow an approach similar to the one used in [34], adapting it to our specific case. Indeed, while in [34] the delay is in the state variable, in our case the delay is in the control variable. This fact requires more care in the representation in infinite dimensional representation. In particular, if we want to get a directional regularity result for the solutions of the HJB equation similar to the one obtained in [34], we need to embed the problem in a more regular infinite-dimensional space. While in [34] the product space $\mathbb{R} \times L^2$ is used to represent the delay system, we need to use the product space $\mathbb{R} \times W^{1,2}$. We observe that the theory of the infinite-dimensional representation of delay systems has been developed mainly in spaces of continuous function or in product space of type $\mathbb{R} \times L^2$ (see [11]). Therefore we need to put more care in our infinite-dimensional representation in $\mathbb{R} \times W^{1,2}$ (even if at the end it looks like in $\mathbb{R} \times L^2$). We prove that the value function is continuous in the interior of its domain with respect to a weak norm (Proposition 3.4), that it solves in the viscosity sense the associated HJB equation (Theorem 3.3) and that it has continuous classical derivative along a suitable direction in the space H (Theorem 3.4). Exactly as in [34, 35], this regularity result just allows to define the formal optimal feedback strategy in classical sense. So, it represents the starting point to construct optimal feedbacks for the problem as in [35].

In the second part of the thesis we wanted to study an alternative approach to control theory that takes into account probability densities.
Stochastic dynamic programming has been recognized as a very important

tool for dealing with stochastic dynamic optimization problems. One of the most important areas of application of this technique is related to portfolio optimization. Various modeling approaches have been developed within this framework both in the discrete and continuous-time setting.
One of the main approach relies on the concept of expected utility, where the investor's risk bearing attitude under uncertainty is modeled as a von Neumann-Morgenstern (1947) utility function. The user needs to determine the utility function and risk aversion attitude that represent his preferences best. Following this direction, there has been considerable development of multi-period stochastic models and, pioneered by the famous work by Merton [55], continuous-time models for portfolio management. Merton used dynamic programming and partial differential equation theory to derive and analyze the relevant Hamilton-Jacobi-Bellman (HJB) equation. From then on, many variations of this problem have been investigated by this approach. The book by Karatzas and Shreve [49] summarizes much of this continuous-time portfolio management theory.

With regard to the object of the optimization, the literature has usually considered the cited Von Neumann-Morgestern criterion. It is clear that such kind of optimization takes into account only very little information about the law structure of the terminal wealth. On the other hand, the investor might want to optimize also with respect to some other features, thus to take into account also the structure of the law of the terminal wealth. In other terms, the investor might want to also look at the density of the terminal wealth (supposing that it exists). This is where the Malliavin Calculus comes into play.
Malliavin Calculus is a very powerful tool to prove existence and regularity of densities for solutions of SDE's (e.g., see [58]), also providing a very interesting link with the analytic theory of hypoelliptic operators (e.g., see [47]).

In this work, we have applied the techniques of the Malliavin Calculus to a stochastic differential equation whose coefficients depend on a control process. Consider the following example of a stochastic differential equation in the canonical probability space $(\Omega, \mathcal{F}, \mathbb{P})$:

$$X_t = x_0 + \int_0^t b(X_s, u_s)\, ds + \sum_{j=1}^{d} \int_0^t \sigma_j(X_s, u_s)\, dW_s^j, \tag{0.2}$$

where $x_0 \in \mathbb{R}^m$ is a random variable $\mathcal{F}_0$-adapted, $b : \mathbb{R}_+ \times \mathbb{R}^m \times U \longrightarrow \mathbb{R}^m$ and $\sigma : \mathbb{R}_+ \times \mathbb{R}^m \times U \longrightarrow \mathbb{R}^{m\times d}$ are measurable functions satisfying globally Lipschitz and boundness conditions, u is the control process belongs to $\mathcal{U} \subset \mathbb{R}^m$ such that $u \in \mathbb{L}^{1,\infty}$ and W^d is a d-dimensional Brownian motion

$W = \{W_t^j, t \in [0,T], 1 \le j \le d\}$ on a finite interval $[0,T]$.
In Chapter 5, we calculate the stochastic derivative of stochastic differential equation (0.2). In Chapter 6 we focus our attention on the particular case of Markovian controls, and we write the Malliavin matrix γ_t of the process X_t in the following form

$$\gamma_t = e^{A_t} \left(\int_0^t \left(e^{A_s}\right)^{-1} \sigma(X_s, u_s) \left(\left(e^{A_s}\right)^{-1} \sigma(X_s, u_s) \right)^* ds \right) \left(e^{A_t}\right)^*,$$

where A_t is a $m \times m$ matrix whose elements are stochastic differential equations depending on the coefficients of (0.2).
Our objective for future research is to prove strict positiveness of the Malliavin matrix γ_t, ensuring existence and regularity of densities for the controlled process.
Once we obtain these results for controlled systems, we intend to reformulate the portfolio allocation problem accordingly. In both cases, this new formulation should pass through another step, which is the characterization of these densities.

Contents

Part I

On optimal control problems with DDE'S and delay in the control

Chapter 1

Preliminary

1.1 Notations

For the first part of the work, we will use the following notations.
We denote by $\mathcal{L}(X)$ the space of the bounded linear operator from a Banach space X to itself and by $\|L\|_{\mathcal{L}(X)}$ the norm of an operator $L \in \mathcal{L}(X)$, i.e,

$$\|L\|_{\mathcal{L}(X)} := \sup_{x \in X} \|Lx\|_X.$$

Throughout paper we consider the Lebesgue space

$$L_r^2 := L^2([-r, 0]; \mathbb{R}),$$

endowed with inner product

$$\langle f, g \rangle_{L_r^2} := \int_{-r}^{0} f(\xi) g(\xi) d\xi,$$

which renders it a Hilbert space.
Also we consider the Sobolev spaces

$$W_r^{k,2} := W^{k,2}([-r, 0]; \mathbb{R}), \quad k = 1, 2, \ldots$$

endowed with inner products

$$\langle f, g \rangle_{W_r^{k,2}} := \sum_{i=0}^{k} \int_{-r}^{0} f^{(i)}(\xi) g^{(i)}(\xi) d\xi, \quad k = 1, 2, \ldots, \tag{1.1}$$

which render them Hilbert spaces. The well-known Sobolev's inclusions (see [54]) imply that

$$W^{k,2}([-r, 0]; \mathbb{R}) \hookrightarrow C^{k-1}([-r, 0]; \mathbb{R}), \quad k = 1, 2, \ldots$$

with continuous embedding. Throughout the paper we will confuse the elements of $W_r^{k,2}$, which are classes of equivalence of functions, with their (unique) representatives in $C^{k-1}([-r,0];\mathbb{R})$, which are pointwise well defined functions. Given that, we define the spaces

$$W_{r,0}^{k,2} := \{f \in W_r^{k,2} \mid f^{(i)}(-r) = 0,\ \forall i = 0,1,...,k-1\} \subset W_r^{k,2}, \quad k = 1,2,\ldots$$

We notice that in our definition of $W_0^{k,2}$ the boundary condition is only required at $-r$. The spaces $W_0^{k,2}$ are Hilbert spaces as closed subsets of the Hilbert spaces $W^{k,2}$. However, on these spaces we may consider the inner products

$$\langle f, g\rangle_{W_0^{k,2}} := \int_{-r}^{0} f^{(k)}(\xi) g^{(k)}(\xi) d\xi, \quad k = 1, 2, \ldots \tag{1.2}$$

It is easy to see that, due to the boundary condition in the definition of the subspaces $W_0^{k,2}$, the inner products $\langle \cdot,\cdot\rangle_{W_{r,0}^{k,2}}$ are equivalent to the original inner products $\langle \cdot,\cdot\rangle_{W^{k,2}}$ on $W_0^{k,2}$, in the sense that they induce equivalent norms. Due to that, dealing with topological concepts, we will consider the simpler inner products (1.2) on the spaces $W_0^{k,2}$.

We consider the Banach space

$$X := \mathbb{R} \times L_r^2.$$

This is a Hilbert space when endowed with the inner product

$$\langle \eta, \zeta\rangle := \eta_0\zeta_0 + \langle \eta_1, \zeta_1\rangle_{L_r^2},$$

where $\eta = (\eta_0, \eta_1(\cdot))$ is the generic element of X. The norm on this space is defined as

$$\|\eta\|_X^2 = |\eta_0|^2 + \|\eta_1\|_{L_r^2}^2.$$

We consider the space $H \subset X$ defined as

$$H := \mathbb{R} \times W_{r,0}^{1,2}.$$

This is a Hilbert space when endowed with the inner product

$$\langle \eta, \zeta\rangle := \eta_0\zeta_0 + \langle \eta_1, \zeta_1\rangle_{W_0^{1,2}}.$$

which induces the norm

$$\|\eta\|^2 = |\eta_0|^2 + \int_{-r}^{0} |\eta_1'(\xi)|^2 d\xi.$$

This will be the Hilbert spaces where our infinite-dimensional system will live.

Chapter 2

Setup of the Control Problem and infinite dimensional differential equation

In this chapter we give the precise formulation of the optimal control problem that we are going to study.

Given $y_0 \in (0, +\infty)$ and $\delta \in L^2_r$, we consider the optimal control of the following differential equation with delay in the control variable

$$\begin{cases} y'(t) = ay(t) + b_0 u(t) + \int_{-r}^{0} b_1(\xi) u(t+\xi) d\xi; \\ y(0) = y_0; \quad u(s) = \delta(s), \ s \in [-r, 0); \end{cases} \tag{2.1}$$

with state constraint $y(\cdot) > 0$ and with control constraint $u(\cdot) \in U \subset \mathbb{R}$. The value $y_0 \in (0, +\infty)$ in the state equation (2.1) represents the initial state of the system, while the function δ represents the past of the control, which is considered as given.

About U we assume the following

Hypothesis 2.1. $U = [0, \bar{u}]$, *where* $\bar{u} \in [0, +\infty]$. *When* $\bar{u} = +\infty$, *the set* U *is intended as* $U = [\underline{u}, +\infty)$. [1]

Moreover, with regard to the parameters appearing in (2.1) we assume the following, that will be standing assumptions throughout the paper

[1] Actually the assumption that $0 \in U$ is not strictly necessary. However it is quite natural in the application and makes simpler some proofs and the statements of the assumptions on the parameters. The case $0 \notin U$ can be treated as well, modifying accordingly the assumptions on the other parameters and some proofs.

Hypothesis 2.2.

(i) $a, b_0 \in \mathbb{R}$;

(ii) $b_1 \in W^{1,2}_{r,0}$, *and* $b_1 \neq 0$.

Given $u(\cdot) \in L^2_{loc}([0,+\infty);\mathbb{R})$, there is a unique locally absolutely continuous solution of (2.1), provided explicitly by the variation of constants formula

$$y(t) = y_0 e^{at} + \int_0^t e^{a(t-s)} f(s) ds, \tag{2.2}$$

where

$$f(s) = b_0 u(s) + \int_{-r}^0 b_1(\xi) u(s+\xi) d\xi; \qquad u(s) = \delta(s), \; s \in [-r, 0].$$

We notice that f is well defined, as b_1 is bounded and $u(\cdot) \in L^2_{loc}([0,+\infty),\mathbb{R})$. We denote by $y(t; y_0, \delta(\cdot), u(\cdot))$ the solution to (2.2) with initial datum $(y_0, \delta(\cdot))$ and under the control $u(\cdot) \in L^2_{loc}$. We notice that $y(t; y_0, \delta(\cdot), u(\cdot))$ solves the delay differential equation (2.1) only for almost every $t \geq 0$.
We define the class of the admissible controls for the problem with state constraint $y(\cdot) > 0$ as

$$\mathcal{U}(y_0, \delta(\cdot)) := \{u(\cdot) \in L^2_{loc}([0,+\infty); U) \mid y(\cdot, y_0, \delta; u(\cdot)) > 0\}.$$

Setting $y(t) := y(t; y_0, \delta(\cdot), u(\cdot))$, we define the objective functional

$$J_0(y_0, \delta(\cdot); u(\cdot)) = \int_0^{+\infty} e^{-\rho t} \Big(g_0(y(t)) - h_0(u(t)) \Big) dt, \tag{2.3}$$

where $\rho > 0$ and $g_0 : (0,+\infty) \to \mathbb{R}$, $h_0 : U \to \mathbb{R}$ are measurable functions satisfying

Hypothesis 2.3.

(i) *The function* $g_0 : (0,+\infty) \to \mathbb{R}$ *is continuous, concave, nondecreasing and bounded from above.*

(ii) *The cost function* $h_0 \in C^1(U)$, *is convex and bounded from below. Without loss of generality we assume* $h_0(0) = 0$. *Moreover*

$$\lim_{u \downarrow \underline{u}} h_0'(u) \geq 0; \quad \lim_{u \uparrow \underline{u}} h_0'(u) = +\infty. \tag{2.4}$$

Finally in the case $\bar{u} = +\infty$ *we assume*

$$\exists \, \alpha > 0 : \; \liminf_{u \to \infty} \frac{h_0(u)}{u^{1+\alpha}} > 0. \tag{2.5}$$

The optimization problem consists in the maximization of the objective functional J_0 over the set of all admissible strategies $\mathcal{U}(y_0, \delta(\cdot))$, i.e.

$$\max_{u(\cdot)\in\mathcal{U}(\eta)} J_0(y_0, \delta(\cdot); u(\cdot)). \tag{2.6}$$

Remark 2.1. *We comment on the modeling features.*

(i) *We consider the optimal control problem imposing the strict constraint $y(\cdot) > 0$ on the state variable. The case of large state constraint $y(\cdot) \geq 0$ can be treated as well.*

(ii) *The assumption that g_0 is bounded from above (Hypothesis 2.3-(i)) is quite unpleasant, if we think about the applications. However we stress that this assumption is taken here just for convenience and can be replaced with a suitable assumption on the growth of g_0, relating it to the requirement of a large enough discount factor ρ.*

(iii) *The conditions required for the cost function h_0 ensure that the Legendre transform h_0 is strictly convex $(0, +\infty)$, see Lemma . This is easy in the case $\bar{u} < +\infty$, less standard in the case $\bar{u} = +\infty$.*

(iv) *We consider a delay r belonging to $(-\infty, 0]$. However one can obtain the same results even allowing $r = -\infty$ as in [60], suitably redefining the boundary conditions as limits. In the definition of the Sobolev spaces $W^{k,2}_{r,0}$, the boundary conditions required become*

$$W^{k,2}_{r,0} := \left\{ f \in W^{k,2} \mid \lim_{r\to-\infty} f^{(i)}(r) = 0, \ \forall i = 0, 1, ..., k-1 \right\} \subset W^{k,2}_r.$$

$\square$

2.1 Representation in infinite dimension

In this section we restate the delay differential equation (2.1) as an abstract evolution equation in infinite dimension. The infinite-dimensional setting will be represented by the Hilbert space $H = \mathbb{R} \times W^{1,2}_{r,0}$. We use the Semigroups Theory, for which we refer to [28]. The following argument is just a suitable rewriting in H of the method illustrated in [11] in the framework of the product space $\mathbb{R} \times L^2$.

Let A be the unbounded linear operator on $\mathcal{D}(A) \subset H$, defined as

$$A : \mathcal{D}(A) \subset H \to H, \quad (\eta_0, \eta_1(\cdot)) \mapsto (a\eta_0 + \eta_1(0), -\eta_1'(\cdot)), \tag{2.7}$$

where a is the constant appearing in (2.1), defined on

$$\mathcal{D}(A) = \mathbb{R} \times W^{2,2}_{r,0}.$$

It is possible to show by direct computations that A is a (densely defined) closed operator and generates a C_0-semigroup $(S_A(t)_t)_{t\geq 0}$ in H. However, we provide the proof of that in the following subsection 2.1.1 by using some known facts from the Semigroups Theory.

The explicit expression of $S_A(t)$ is

$$\begin{aligned} S_A(t)\eta = & \left(\eta_0 e^{at} + \int_{(-t)\vee(-r)}^{0} \eta_1(\xi) e^{a(\xi+t)} d\xi,\ \eta_1(\cdot - t)\mathbf{1}_{[-r,0]}(\cdot - t)\right), \\ & \eta = (\eta_0, \eta_1(\cdot)) \in H. \end{aligned} \tag{2.8}$$

By Chapter 2, Proposition 4.7 in [53], there exist $M > 0$, $\omega \in \mathbb{R}$ such that

$$\|S_A(t)\| \leq Me^{\omega t}, \qquad t \geq 0, \tag{2.9}$$

where M and ω depend on a and r. For the exactly calculations of (2.8) and (2.9), we refer the reader to the Appendix 2.1.1.

In the space H we set

$$b := (b_0, b_1(\cdot)) \quad \text{and} \quad \hat{b} := b/\|b\|$$

and consider the bounded linear operator B defined as

$$B : U \to H, \quad u \mapsto bu. \tag{2.10}$$

Often we will identify the operator B with b.

Given $u(\cdot) \in L^2_{loc}([0, +\infty), \mathbb{R})$ and $\eta \in H$, we consider the abstract problem in the Hilbert space H,

$$\begin{cases} Y'(t) = AY(t) + Bu(t), \\ Y(0) = \eta. \end{cases} \tag{2.11}$$

We will use two concepts of solution to (2.11), that in our case coincide each other. For details we refer to [53, Chapter 2, Section 5].

In the definitions below the integral in dt is intended in Bochner sense in the Hilbert space H.

Definition 2.1.

(i) *We call mild solution of* (2.11) *the function* $Y \in C([0,+\infty), H)$ *defined as*

$$Y(t) = S_A(t)\eta + \int_0^t S_A(t-\tau)Bu(\tau)d\tau, \quad t \geq 0. \tag{2.12}$$

(ii) *We call weak solution of* (2.11) *a function* $Y \in C([0,+\infty), H)$ *such that, for any* $\phi \in \mathcal{D}(A^*)$,

$$\langle Y(t), \phi\rangle = \langle \eta, \phi\rangle + \int_0^t \langle Y(\tau), A^*\phi\rangle d\tau + \int_0^t \langle Bu(\tau), \phi\rangle d\tau, \quad \forall t \geq 0. \tag{2.13}$$

From now on we denote by $Y(\cdot;\eta,u(\cdot))$ the mild solution of (2.11). We notice that the definition of mild solution is the infinite-dimensional version of the variation of constants formula. By a well-known result (see [53, Chapter 2, Proposition 5.2]), the mild solution is also the (unique) weak solution.

2.1.1 The semigroup $S_A(t)$ on H

Hereafter, given $f \in L^2$, with a slight abuse of notation we shall intend it extended on $[-r,+\infty)$ setting $f \equiv 0$ on $(0,+\infty)$.
Consider the space $X = \mathbb{R} \times L^2$ endowed with the inner product

$$\langle \cdot,\cdot\rangle_X = \langle \cdot,\cdot\rangle_{\mathbb{R}} + \langle \cdot,\cdot\rangle_{L^2},$$

which makes it a Hilbert space. On this space consider the unbounded operator

$$\bar{A}^* : \mathcal{D}(\bar{A}^*) \subset X \longrightarrow X, \quad (\eta_0, \eta_1(\cdot)) \longmapsto (a\eta_0, \eta_1'(\cdot)) \tag{2.14}$$

defined on the domain

$$\mathcal{D}(\bar{A}^*) = \left\{\eta = (\eta_0, \eta_1(\cdot)) \,|\, \eta_1 \in W^{1,2},\ \eta_1(0) = \eta_0\right\}.$$

It is well known (see [28]) that $\bar{A}^*$ is a closed operator which generates a C_0-semigroup $(S_{\bar{A}^*}(t))_{t\geq 0}$ on X. More precisely the explicit expression of $S_{\bar{A}^*}(t)$ acting on $\psi = (\psi_0, \psi_1(\cdot)) \in X$ is

$$S_{\bar{A}^*}(t)\psi = \left(e^{at}\psi_0, \mathbf{1}_{[-r,0]}(t+\xi)\psi_1(t+\xi) + \mathbf{1}_{[0,+\infty)}(t+\xi)e^{a(t+\xi)}\psi_0\Big|_{\xi\in[-r,0]}\right). \tag{2.15}$$

On the other hand it is possible to show (see e.g. [34]) that $\bar{A}^*$ is the adjoint of the $\bar{A}$ on X defined by

$$\begin{array}{rlcc} \bar{A}: & \mathcal{D}(\bar{A}) \subset X & \longrightarrow & X \\ & (\eta_0, \eta_1(\cdot)) & \longmapsto & (a\eta_0 + \eta_1(0), -\eta_1'(\cdot)), \end{array}$$

where

$$\mathcal{D}(\bar{A}) = \mathbb{R} \times W_0^{1,2} = H.$$

It follows (see [28]) that $\bar{A}$ generates on X the C_0-semigroup $(S_{\bar{A}}(t))_{t\geq 0}$ where

$$S_{\bar{A}}(t) = S_{\bar{A}^*}(t)^*, \quad \forall t \geq 0.$$

We can compute the explicit expression of the semigroup $S_{\bar{A}}(t)$ through the relation

$$\langle S_{\bar{A}}(t)\phi, \psi\rangle = \langle \phi, S_{\bar{A}^*}(t)\psi\rangle, \quad \forall\, \phi = (\phi_0, \phi_1(\cdot)) \in X,\ \forall\, \psi = (\psi_0, \psi_1(\cdot)) \in X.$$

By (2.15), we calculate

$$\begin{aligned} \langle \bar{S}_{\bar{A}}(t)\phi, \psi\rangle = {} & \phi_0 e^{at}\psi + \int_{-r}^{(-t)\vee(-r)} \phi_1(\xi)\psi_1(t+\xi)d\xi \\ & + \int_{-(t)\vee(-r)}^{0} \phi_1(\xi)\psi_0 e^{a(t+\xi)}d\xi = \phi_0 e^{at}\psi_0 + \int_{(-r+t)\wedge 0}^{0} \phi_1(\xi - t)\psi_1(\xi)d\xi \\ & + \int_{(-t)\vee(-r)}^{0} \phi_1(\xi)e^{a(\xi+t)}\psi_0 d\xi. \end{aligned} \tag{2.16}$$

So we can write the explicit form of the semigroup $\bar{S}(t)$ as

$$S_{\bar{A}}(t)\phi = \left(\phi_0 e^{at} + \int_{(-t)\vee(-r)}^{0} \phi_1(\xi)e^{a(\xi+t)}d\xi,\ T(t)\phi_1\right), \quad \phi = (\phi_0, \phi_1(\cdot)) \in X, \tag{2.17}$$

where $(T(t))_{t\geq 0}$ is the semigroup of truncated right shifts on L^2 defined as

$$[T(t)f](\xi) = \begin{cases} f(\xi - t), & -r \leq \xi - t, \\ 0, & \text{otherwise}, \end{cases} \tag{2.18}$$

for $f \in L^2$. So, we may rewrite the above expression as

$$\begin{aligned} S_{\bar{A}}(t)\phi = {} & \left(\phi_0 e^{at} + \int_{(-t)\vee(-r)}^{0} \phi_1(\xi)e^{a(\xi+t)}d\xi, \phi_1(\cdot - t)\mathbf{1}_{[-r,0]}(\cdot - t)\right), \\ & \phi = (\phi_0, \phi_1(\cdot)) \in X. \end{aligned} \tag{2.19}$$

Equation (2.19) defines the explicit form of the semigroup $(\bar{S}(t))_{t\geq 0}$.

We have defined the semigroup $S_{\bar{A}}(t)$ and its infinitesimal generator $(\bar{A}, \mathcal{D}(\bar{A}))$ in the space X. Therefore, by well-known results (see [28, chapter II, pag 124]), we get that $\bar{A}|_{\mathcal{D}(\bar{A}^2)}$ is the generator of a C_0-semigroup on $(\mathcal{D}(\bar{A}), \|\cdot\|_{\mathcal{D}(\bar{A})})$, which is nothing but the restriction of $S_{\bar{A}}$ to this subspace. Now we notice that

$$\mathcal{D}(\bar{A}) = H, \;\; \|\cdot\|_{\mathcal{D}(\bar{A})} \sim \|\cdot\|, \;\; \mathcal{D}(\bar{A}^2) = W_0^{2,2} = \mathcal{D}(A), \;\; \bar{A}|_{W_0^{2,2}} = A,$$

where A is the operator defined in (2.7). Hence, we conclude that A generates a C_0-semigroup on H, whose expression is the same given in (2.17). We denote such semigroup by $S_A(t)$.
Moreover we recall that if $S(t)$ is a C_0 semigroup on a Banach space H, then there exist constants $M \geq 1$ and $\omega \in \mathbb{R}$, such that

$$\|S(t)\| \leq Me^{\omega t}, \qquad t \geq 0. \tag{2.20}$$

[53, Chapter 2, Proposition 4.7]. Then, in this case, using Holder's inequality and taking into account that $\phi_1(-r) = 0$ we compute for every $t \geq 0$

$$\begin{aligned}
\left|\phi_0 e^{at} + \int_{(-t)\vee(-r)}^0 \phi_1(\xi)e^{a(\xi+t)}d\xi\right|^2 &\leq 2e^{2at}|\phi_0|^2 + 2e^{2at}\left(\int_{-r}^0 |\phi_1(\xi)|d\xi\right)^2 \\
&\leq 2e^{-at}|\phi_0|^2 + 2e^{2at}r\left(\int_{-r}^0 |\phi_1(\xi)|^2 d\xi\right) \\
&\leq 2e^{2at}|\phi_0|^2 + 2e^{2at}r\left(\int_{-r}^0 \left|\int_{-r}^{\xi} \phi_1'(s)ds\right|^2 d\xi\right) \\
&\leq 2e^{2at}|\phi_0|^2 \\
&\quad + 2e^{2at}r\left(\int_{-r}^0 (r+\xi)\left(\int_{-r}^{\xi} |\phi_1'(s)|^2 ds\right)d\xi\right) \\
&\leq 2e^{2at}|\phi_0|^2 + e^{2at}r^3\|\phi_1\|^2_{W^{1,2}_{r,0}}.
\end{aligned}$$

Moreover

$$\|T(t)\|_{\mathcal{L}(W^{1,2}_{r,0})} \leq 1, \;\forall t \in [0,r]; \quad \|T(t)\|_{\mathcal{L}(W^{1,2}_{r,0})} = 0, \;\forall t > r.$$

The computations above show that

$$\|e^{tA}\|_{\mathcal{L}(H)} \leq (2+r^3)^{1/2}e^{at}, \quad \forall t \geq 0. \tag{2.21}$$

So, setting

$$\omega = a \quad \text{and} \quad M = (2+r^3)^{1/2}, \tag{2.22}$$

(2.20) is verified.

2.1.2 Equivalence with the original problem

In order to state equivalence results between the DDE (2.1) and the abstract evolution equation (2.11) when A and B are defined by (2.10) and (2.7), respectively, we need to link the canonical $\mathbb{R}$-valued integration with the $W^{1,2}_{r,0}$-valued integration. This is provided by the following lemma, whose proof for L^2_r-valued processes can be found in [33, Chapter 2]. A similar argument holds for $W^{1,2}_{r,0}$-valued processes (this is even easier, as the functions of $W^{1,2}_{r,0}$ are pointwise well defined), and the proof is provided in the Appendix (see B).

Lemma 2.1. *Let* $0 \leq a < b$ *and* $f \in L^2\left([a,b]; W^{1,2}_0\right)$. *Then*

$$\left(\int_a^b f(t)dt\right)(\xi) = \int_a^b f(t)(\xi)dt, \qquad \forall \xi \in [-r,0],$$

where the integral in dt *in the left handside is intended as Bochner integral in the space* $W^{1,2}_0$. □

Also we need to study the adjoint operator A^* in order to use the concept of weak solution of (2.11).

Proposition 2.1. *We have*

$$\mathcal{D}(A^*) = \left\{\phi = (\phi_0, \phi_1(\cdot)) \in H \mid \phi_1 \in W^{2,2}_r, \ \ \phi_1'(0) = 0\right\}$$

and

$$A^*\phi = \left(a\phi_0, \ \xi \mapsto \phi_1'(\xi) + \phi_0(\xi + r) + \int_{-r}^0 \phi_1''(s)ds\right), \quad \phi \in \mathcal{D}(A^*). \quad (2.23)$$

Proof. Let

$$\mathcal{D} := \left\{\phi = (\phi_0, \phi_1(\cdot)) \in H \mid \phi_1 \in W^{2,2}_r, \ \ \phi_1'(0) = 0\right\}.$$

First of all we notice that, defining $A^*\phi$ on $\mathcal{D}$ as in (2.23), we have $A^*\phi \in H$. Now notice that

$$\psi_1'(-r) = 0, \ \ \psi_1(0) = \int_{-r}^0 \psi_1'(\xi)d\xi, \qquad \forall \psi \in \mathcal{D}(A). \qquad (2.24)$$

Therefore, taking into account (2.24), we have for every $\psi \in \mathcal{D}(A)$ and every

$\phi \in \mathcal{D}$

$$
\begin{aligned}
\langle A\psi, \phi\rangle &= a\psi_0\phi_0 + \psi_1(0)\phi_0 - \int_{-r}^{0} \psi_1''(\xi)\phi_1'(\xi)d\xi \\
&= a\psi_0\phi_0 + \left(\int_{-r}^{0} \psi_1'(\xi)d\xi\right)\phi_0 - \psi_1'(0)\phi_1'(0) + \psi_1'(-r)\phi_1'(-r) \\
&+ \int_{-r}^{0} \psi_1'(\xi)\phi_1''(\xi)d\xi \\
&= a\psi_0\phi_0 + \int_{-r}^{0} \psi_1'(\xi)\left(\phi_0 + \phi_1''(\xi)\right)d\xi.
\end{aligned}
\tag{2.25}
$$

The equality above shows that $\mathcal{D} \subset \mathcal{D}(A^*)$ and that A^* acts as claimed in (2.23) on the elements of $\mathcal{D}$.

Now we have to show that $\mathcal{D} = \mathcal{D}(A^*)$. For sake of brevity we only sketch the proof of this fact here[2], as a complete proof would require a study of the adjoint semigroup e^{tA^*}. We observe that $\mathcal{D}$ is dense in H. Moreover an explicit computation of the adjoint semigroup e^{tA^*} would show that $e^{tA^*}\mathcal{D} \subset \mathcal{D}$ for any $t \geq 0$. Hence, by [24, Th.1.9, p.8], $\mathcal{D}$ is dense in $\mathcal{D}(A^*)$ endowed with the graph norm. Finally, using (2.23) it is easy to show that $\mathcal{D}$ is closed in the graph norm of A^* and therefore $\mathcal{D}(A^*) = \mathcal{D}$. $\square$

Let $v \in L^2_r$ and consider the function

$$(v * b_1)(\xi) = \int_{-r}^{\xi} b_1(\tau)v(\tau - \xi)d\tau, \quad \xi \in [-r, 0].$$

First of all we notice that $(v * b_1)(-r) = 0$. Extend b_1 to a $W^{1,2}_r(\mathbb{R})$ function on $\mathbb{R}$ equal to 0 in $(-\infty, -r)$ (recall that $b_1(-r) = 0$) and extend v to an $L^2(\mathbb{R};\mathbb{R})$ function simply defining it equal to 0 out of $[-r, 0]$. Then the function above can be rewritten as

$$(v * b_1)(\xi) = \int_{\mathbb{R}} b_1(\tau)v(\tau - \xi)\mathbf{1}_{(-\infty,0]}(\tau - \xi)d\tau, \quad \xi \in [-r, 0].$$

Since $v\mathbf{1}_{(-\infty,0]} \in L^2(\mathbb{R};\mathbb{R})$ and $b_1 \in W^{1,2}(\mathbb{R};\mathbb{R})$, [14, Lemma VIII.4] yields $v * b_1 \in W^{1,2}_{r,0}$ and

$$(v * b_1)'(\xi) = \int_{-r}^{\xi} b_1'(\tau)v(\tau - \xi)d\tau. \tag{2.26}$$

[2] To this regard we observe that we will use in the following only the fact $\mathcal{D} \subset \mathcal{D}(A^*)$ and that (2.23) holds true on the elements of $\mathcal{D}$, which has been proven rigorously. More precisely we will use the fact that $(1, 0) \in \mathcal{D} \subset \mathcal{D}(A^*)$ in the proof of Theorem 2.4.

Consider still v extended to 0 out of $[-r,0]$ and set $v_\xi(\tau) := v(\tau-\xi)$, $\tau \in [-r,0]$ for $\xi \in [-r,0]$. Of course $v_\xi \in L^2_r$ and $\|v_\xi\|_{L^2_r} \le \|v\|_{L^2_r}$ for every $\xi \in [-r,0]$. Then, due to (2.26) and by Holder's inequality we have

$$\|v * b_1\|^2_{W^{1,2}_{r,0}} = \int_{-r}^0 \left| \int_{-r}^{\xi} b_1'(\tau) v(\tau-\xi) d\tau \right|^2 d\xi = \int_{-r}^0 \left| \int_{-r}^{\xi} b_1'(\tau) v_\xi(\tau) d\tau \right|^2 d\xi$$
$$\le \int_{-r}^0 \left(\int_{-r}^0 |b_1'(\tau) v_\xi(\tau)| d\tau \right)^2 d\xi \ \le \int_{-r}^0 \left(\|b_1'\|^2_{L^2_r} \|v_\xi\|^2_{L^2_r} \right) d\xi \le r \|b_1'\|^2_{L^2_r} \|v\|^2_{L^2_r}. \tag{2.27}$$

Let us introduce the continuous linear map M:

$$\begin{array}{llll} M: & \mathbb{R} \times L^2([-r,0];\mathbb{R}) & \longrightarrow & H \\ & (z,v) & \longmapsto & (z, v*b_1) = \left(z, \int_{-r}^{\cdot} b_1(\tau) v(\tau - \cdot) d\tau\right). \end{array} \tag{2.28}$$

Due to (2.27), M is bounded. Call

$$\mathcal{M} := \mathrm{Im}(M). \tag{2.29}$$

Remark 2.2. *Of course $\mathcal{M}$ is a linear subspace of H. It should be possible using [10] that is not closed.* □

Theorem 2.4. *Let $y_0 \in \mathbb{R}$, $\delta \in L^2_r$, $u(\cdot) \in L^2([0,+\infty),\mathbb{R})$. Set*

$$\eta := M(y_0, \delta(\cdot)) \in \mathcal{M}; \quad u(s) := \delta(s), \ \ s \in [-r,0); \quad Y(t) := Y(t;\eta,u(\cdot)), \ \ t \ge 0. \tag{2.30}$$

Then

$$Y(t) = M(Y_0(t), u(t+\cdot)), \qquad \forall t \ge 0. \tag{2.31}$$

Moreover, let $y(\cdot) := y(\cdot; y_0, \delta, u(\cdot))$ be the unique solution to (2.1). Then

$$y(t) = Y_0(t), \qquad \forall t \ge 0. \tag{2.32}$$

Proof. Let Y be the mild solution defined by (2.12) with initial condition η given by (2.30). On the second component it reads

$$\begin{aligned} Y_1(t) &= T(t)\eta_1 + \int_0^t [T(t-s)b_1] u(s) ds \\ &= \mathbf{1}_{[-r,0]}(\cdot - t)\eta_1(\cdot - t) + \int_0^t \mathbf{1}_{[-r,0]}(\cdot - t + s) b_1(\cdot - t + s) u(s) ds \end{aligned} \tag{2.33}$$

where $(T(t))_{t\geq 0}$ is the semigroup of truncated right shifts on $W_0^{1,2}$ just defined, that is

$$[T(t)\phi](\xi) = \mathbf{1}_{[-r,0]}(\xi - t)\phi(\xi - t), \quad \xi \in [-r, 0].$$

We recall that by hypothesis $\eta = M(y_0, \delta(\cdot))$, so we write the second component of the initial datum

$$\eta_1(\xi) = \int_{-r}^{\xi} b_1(\alpha)\delta(\alpha - \xi)d\alpha.$$

Then, by (2.33) and due to Lemma B.3, the second component evaluated at ξ is

$$Y_1(t)(\xi) = \mathbf{1}_{[-r,0]}(\xi-t)\int_{-r}^{\xi-t} b_1(\alpha)u(\alpha-\xi+t)d\alpha + \int_0^t \mathbf{1}_{[-r,0]}(\xi-t+s)b_1(\xi-t+s)u(s)ds. \tag{2.34}$$

Taking into account that $0 \leq s \leq t$, we have $\xi - t \leq \xi - t + s \leq \xi$, so that, setting $\alpha = \xi - t + s$ in the second part of the right handside of (2.34), it becomes

$$\begin{aligned} Y_1(t)(\xi) &= \mathbf{1}_{[-r,0]}(\xi - t)\int_{-r}^{\xi-t} b_1(\alpha)u(\alpha - \xi + t)d\alpha + \int_{\xi-t}^{\xi} \mathbf{1}_{[-r,0]}(\alpha)b_1(\alpha)u(\alpha - \xi + t)d\alpha \\ &= \int_{-r}^{(\xi-t)\vee(-r)} b_1(\alpha)u(\alpha - \xi + t)d\alpha + \int_{(\xi-t)\vee(-r)}^{\xi} b_1(\alpha)u(\alpha - \xi + t)d\alpha. \\ &= \int_{-r}^{\xi} b_1(\alpha)u(\alpha + t - \xi)d\alpha. \end{aligned} \tag{2.35}$$

Therefore, due to the (2.28), (2.31) is proved.
It follows immediately, setting $\xi = 0$ in (2.35), that

$$Y_1(t)(0) = \int_{-r}^{0} b_1(\alpha)u(t + \alpha)d\alpha. \tag{2.36}$$

Let us show (2.32). We use the fact that $Y(\cdot)$ is also a weak solution of (2.11).
From Proposition 2.1 we know that

$$(1, 0) \in \mathcal{D}(A^*), \quad A^*(1, 0) = (a,\ \xi \mapsto \xi + r)\,. \tag{2.37}$$

Therefore taking into account (2.37) and (2.36) and Definition 2.1-(i), we have for every $t \geq 0$

$$
\begin{aligned}
Y_0'(t) = \frac{d}{dt}\langle Y(t),(1,0)\rangle &= \langle Y(t), A^*(1,0)\rangle + \langle Bu(t),(1,0)\rangle \\
&= aY_0(t) + \int_{-r}^{0} Y_1(t)'(\xi)d\xi + b_0u(t) \\
&= aY_0(t) + Y_1(t)(0) - Y_1(t)(-r) + b_0u(t) \\
&= aY_0(t) + \int_{-r}^{0} b_1(\xi)u(t+\xi)d\xi + b_0u(t).
\end{aligned}
\tag{2.38}
$$

Therefore $Y_0(t)$ solves (2.1), with initial data $(y_0, \delta(\cdot))$, so it must coincide with $y(t)$. $\square$

We use the above result to reformulate the optimization problem (2.3) in H. Let $\eta \in H$ and define the (possibly empty) set

$$\mathcal{U}(\eta) := \{u(\cdot) \in L^2([0,+\infty);U) \mid Y_0(\cdot;\eta,u(\cdot)) > 0, \quad \forall\, t \geq 0\}.$$

Given $u(\cdot) \in \mathcal{U}(\eta)$ define

$$J(\eta;u(\cdot)) = \int_0^{+\infty} e^{-\rho t}\Big(g(Y(t;\eta,u(\cdot))) + h(u(t))\Big)dt. \tag{2.39}$$

where

$$h : U \to \mathbb{R}, \quad h := -h_0; \qquad g : H \to \mathbb{R}, \quad g(\eta) := g_0(\eta_0). \tag{2.40}$$

Due to (2.32), if $\eta = M(y_0, \delta(\cdot))$ then

$$\mathcal{U}(\eta) = \mathcal{U}(y_0, \delta(\cdot))$$

and

$$J(\eta;u(\cdot)) = J_0(y_0,\delta(\cdot);u(\cdot)), \tag{2.41}$$

where J_0 is the objective functional defined in (2.3). Therefore, we have reduced the original problem (2.6) to

$$\max_{u(\cdot)\in\mathcal{U}(y_0,\delta(\cdot))} J(\eta;u(\cdot)), \quad \eta = M(y_0,\delta(\cdot)) \in \mathcal{M}. \tag{2.42}$$

Although we are interested to solve the problem for initial data $\eta \in \mathcal{M}$, as these are the initial data coming from the real original problem, we enlarge the problem to data $\eta \in H$ and consider the functional (2.39) defined also for these data. So the problem is

$$\max_{u(\cdot)\in\mathcal{U}(\eta)} J(\eta;u(\cdot)), \quad \eta \in H. \tag{2.43}$$

2.1.3 Continuous dependence on initial data

In this subsection we introduce a weaker norm and we study the continuous dependence on initial data of the mild solution (2.12) with respect to this norm. We introduce the following assumption

Hypothesis 2.5. $a \neq 0$.

From now on we assume Hypothesis 2.5.

Remark 2.3. *First of all we notice that Hypothesis 2.5 is not restrictive for the applications, as the $a = 0$ can be treated translating the problem as follows. Take $a = 0$. The state equation in infinite dimension is* (2.11) *with*

$$A : (\phi_0, \phi_1(\cdot)) \mapsto \Big(\phi_1(0), -\phi_1'(\cdot)\Big).$$

However, we can rewrite it as

$$Y'(t) = \tilde{A}Y(t) - P_0 Y(t) + Bu(t),$$

where

$$P_0 : H \mapsto H, \quad P_0\phi = (\phi_0, 0); \qquad \tilde{A} = A + P_0.$$

Then everything we will do can be suitably replaced dealing with this translated equation. □

Due to Hypothesis 2.5, the inverse operator of A is well defined. This is a bounded operator $H \to D(A)$ whose explicit expression is

$$\begin{array}{rcl} A^{-1} : (H, \|\cdot\|_H) & \longrightarrow & (\mathcal{D}(A), \|\cdot\|_H), \\ \eta & \mapsto & \left(\dfrac{\eta_0 + \int_{-r}^0 \eta_1(s)ds}{-a}, -\displaystyle\int_{-r}^{\xi} \eta_1(s)ds \right). \end{array}$$

We define on H the norm $\|\cdot\|_{-1}$ as

$$\|\eta\|_{-1} := \|A^{-1}\eta\|, \tag{2.44}$$

so

$$\|\eta\|_{-1}^2 = \left| \frac{\eta_0 + \int_{-r}^0 \eta_1(s)ds}{a} \right|^2 + \int_{-r}^0 |\eta_1(s)|^2 ds. \tag{2.45}$$

Lemma 2.2. *The norms $\|\cdot\|_{-1}$ and $\|\cdot\|_X$ are equivalent in H.*

Proof. Let $\eta = (\eta_0, \eta_1) \in H$. Taking into account (2.45) and by Hölder's inequality, we have

$$
\begin{aligned}
\|\eta\|_X^2 &= |\eta_0|^2 + \int_{-r}^0 |\eta_1(\xi)|^2 d\xi \\
&= \left|\eta_0 + \int_{-r}^0 \eta_1(\xi)d\xi - \int_{-r}^0 \eta_1(\xi)d\xi\right|^2 + \int_{-r}^0 |\eta_1(\xi)|^2 d\xi \\
&\leq 2\left|\eta_0 + \int_{-r}^0 \eta_1(\xi)d\xi\right|^2 + 2\left|\int_{-r}^0 \eta_1(\xi)d\xi\right|^2 + \int_{-r}^0 |\eta_1(\xi)|^2 d\xi \\
&\leq 2\left|\eta_0 + \int_{-r}^0 \eta_1(\xi)d\xi\right|^2 + 2\left(\int_{-r}^0 |\eta_1(\xi)|d\xi\right)^2 + \int_{-r}^0 |\eta_1(\xi)|^2 d\xi \\
&\leq 2a^2\left|\frac{\eta_0 + \int_{-r}^0 \eta_1(\xi)d\xi}{a}\right|^2 + 2r^2\int_{-r}^0 |\eta_1(\xi)|^2 d\xi + \int_{-r}^0 |\eta_1(\xi)|^2 d\xi \\
&\leq C^2\|\eta\|_{-1}^2,
\end{aligned}
$$

where $C^2 = \max\{2a^2, 2r^2+1\}$.
On the other hand, still using (2.45) and Hölder's inequality, we have

$$
\begin{aligned}
\|\eta\|_{-1}^2 &= \left|\frac{\eta_0 + \int_{-r}^0 \eta_1(s)ds}{a}\right|^2 + \int_{-r}^0 |\eta_1(s)|^2 ds \\
&\leq \frac{2}{a^2}|\eta_0|^2 + \int_{-r}^0 |\eta_1(s)|^2 ds \\
&\leq N^2\|\eta\|_X^2,
\end{aligned}
$$

where $N^2 = \max\left\{\frac{2}{a^2}, 1\right\}$. The claim is proved. □

From Lemma 2.2 we get the following

Corollary 2.1. *There exists a constant $C_{a,r} > 0$ such that*

$$
|\eta_0| \leq C_{a,r}\|\eta\|_{-1}, \quad \forall \eta \in H. \tag{2.46}
$$

Proof. We have, taking into account the Hölder inequality,

$$\begin{aligned}
|\eta_0|^2 &= \left|\eta_0 + \int_{-r}^0 \eta_1(\xi)d\xi - \int_{-r}^0 \eta_1(\xi)d\xi\right|^2 \\
&\leq 2\left|\eta_0 + \int_{-r}^0 \eta_1(\xi)d\xi\right|^2 + 2\left|\int_{-r}^0 \eta_1(\xi)d\xi\right|^2 \\
&\leq 2\left|\eta_0 + \int_{-r}^0 \eta_1(\xi)d\xi\right|^2 + 2\left(\int_{-r}^0 |\eta_1(\xi)|d\xi\right)^2 \\
&\leq 2a^2\left|\frac{\eta_0 + \int_{-r}^0 \eta_1(\xi)d\xi}{a}\right|^2 + 2r^2\int_{-r}^0 |\eta_1(\xi)|^2 d\xi \\
&\leq (2a^2 + 2r^2)\|\eta\|_{-1}^2. \qquad \square
\end{aligned}$$

Remark 2.4. *Corollary 2.1 represents a crucial issue and motivates our choice of working in the product space $\mathbb{R} \times W_0^{1,2}$ in place of the more usual product space $\mathbb{R} \times L^2$. Indeed, embedding the problem in $\mathbb{R} \times L^2$ and defining everything in the same way in this bigger space, we would not be able to control $|\eta_0|$ by $\|\eta\|_{-1}$. But this estimate is necessary to prove the continuity of the value function with respect to $\|\cdot\|_{-1}$, since in this way g is Lipschitz continuous in $(H, \|\cdot\|_{-1})$ (see Proposition 3.4). And, on the other hand, the continuity of V with respect to $\|\cdot\|_{-1}$ is necessary to have a suitable property for the superdifferential of V (see Proposition 3.5), allowing to handle the unbounded linear term in the HJB equation.* $\square$

Lemma 2.3. *We denote*

$$C_t := (2 + r^3)^{\frac{1}{2}} e^{at} \quad \textit{and} \quad \tilde{C}_t := C_{a,r} e^{at}, \tag{2.47}$$

where $C_{a,r}$ is the constant appearing in (2.46).
Let $Y(\cdot)$, $\overline{Y}(\cdot)$ be the mild solutions to (2.11) *starting respectively from η, $\bar{\eta} \in H$. Then*

$$\|Y(t) - \overline{Y}(t)\|_{-1} \leq C_t\|\eta - \bar{\eta}\|_{-1}, \quad \forall t \geq 0. \tag{2.48}$$

In particular, by (2.46),

$$|Y_0(t) - \overline{Y}_0(t)| \leq \tilde{C}_t\|\eta - \bar{\eta}\|_{-1}, \quad \forall t \geq 0. \tag{2.49}$$

Proof. By Definition 2.1-(i), for all $t \geq 0$, we can write

$$Y(t) - \overline{Y}(t) = S_A(t)(\eta - \bar{\eta}).$$

Then we have

$$A^{-1}\left(Y(t) - \overline{Y}(t)\right) = S_A(t)A^{-1}(\eta - \bar{\eta})$$

and by (2.44), (2.47) and (2.21), it holds

$$\|Y(t) - \overline{Y}(t)\|_{-1} \leq C_t\|\eta - \bar{\eta}\|_{-1}$$

whit C_t is above defined and it depends on semigroup $S_A(\cdot)$ through (2.21). By (2.44) and (2.32) we have

$$|Y_0(t) - \overline{Y}_0(t)| = |y(t) - \bar{y}(t)| = |e^{at}(\eta_0 - \bar{\eta}_0)| \leq \tilde{C}_t\|\eta - \bar{\eta}\|_{-1}$$

and the second claim is proved. □

Chapter 3

The value function in the space H

In this section we study some qualitative properties of the value function V associated to the optimization problem (2.43) in the space H. Then we use such properties to investigate the nature of the superdifferential.

For $\eta \in H$ the value function of our problem is the function $V : H \longrightarrow \mathbb{R}$

$$V(\eta) := \sup_{u(\cdot)\in\mathcal{U}(\eta)} J(\eta, u(\cdot)) \tag{3.1}$$

with the convention

$$\sup \emptyset = -\infty.$$

We notice that the value function is bounded from above due to the Hypotheses 2.3 on the functions h and g.

The domain of the value function V is defined as

$$\mathcal{D}(V) := \{\eta \in H \mid V(\eta) > -\infty\}.$$

Since g and h are bounded from below, we have

$$\mathcal{D}(V) := \{\eta \in H \mid \mathcal{U}(\eta) \neq \emptyset\} \tag{3.2}$$

and V is bounded from below in $\mathcal{D}(V)$.

We define the space

$$H_+ := (0, +\infty) \times W^{1,2}_{r,0}.$$

Of course if $\eta \notin H_+$, we have $\mathcal{U}(\eta) = \emptyset$, so that $\eta \notin \mathcal{D}(V)$. This means that

$$\mathcal{D}(V) \subset H_+.$$

.

Definition 3.1. *Let $\eta \in \mathcal{D}(V)$.*

(i) An admissible control $u^*(\cdot) \in \mathcal{U}(\eta)$ is said to be optimal for the initial state η if $J(\eta; u^*(\cdot)) = V(\eta)$.

(ii) Let $\varepsilon > 0$. An admissible control $u^\varepsilon(\cdot) \in \mathcal{U}(\eta)$ is said ε-optimal for the initial state η if $J(\eta; u^\varepsilon(\cdot)) > V(\eta) - \varepsilon$.

Proposition 3.1. *The value function V is finite from below.*

Proof. The boundness from below is a direct consequence of the boundness of g. □

Proposition 3.2. *The set $\mathcal{D}(V)$ is convex and the value function V is concave on $\mathcal{D}(V)$.*

Proof. Let η, $\bar{\eta} \in \mathcal{D}(V)$ and set, for $\lambda \in [0,1]$, $\eta_\lambda := \lambda\eta + (1-\lambda)\bar{\eta}$. For $\varepsilon > 0$, let $u^\varepsilon(\cdot) \in \mathcal{U}(\eta)$ and $\bar{u}^\varepsilon(\cdot) \in \mathcal{U}(\bar{\eta})$ be two controls ε-optimal for the initial states $\eta, \bar{\eta}$ respectively. Set

$$y(\cdot) := y(\cdot; \eta, u^\varepsilon(\cdot)),\ \bar{y}(\cdot) := \bar{y}(\cdot; \bar{\eta}, \bar{u}^\varepsilon(\cdot)),\ u^\lambda(\cdot) := \lambda u^\varepsilon(\cdot) + (1-\lambda)\bar{u}^\varepsilon(\cdot).$$

Finally set $y_\lambda(\cdot) := \lambda y(\cdot) + (1-\lambda)\bar{y}(\cdot)$.
The function h is concave so one has

$$h(u^\lambda(t)) \geq \lambda h(u^\varepsilon(t)) + (1-\lambda)h(\bar{u}^\varepsilon(t)),\ t \geq 0. \tag{3.3}$$

Moreover, by linearity of the state equation, we have

$$Y(t; \eta_\lambda, u_\lambda(\cdot)) := \lambda Y(y; \eta, u^\varepsilon(\cdot)) + (1-\lambda)Y(t; \bar{\eta}, \bar{u}^\varepsilon(\cdot)).$$

Hence, by concavity of g we have

$$g(Y(t; \eta_\lambda, u_\lambda(\cdot))) \geq \lambda g(Y(t; \eta, u^\varepsilon(\cdot))) + (1-\lambda)g(Y(t; \bar{\eta}, \bar{u}^\varepsilon(\cdot))),\ t \geq 0. \tag{3.4}$$

So, we have

$$\begin{aligned} V(\eta_\lambda) \geq J(\eta_\lambda, u^\lambda(\cdot)) &= \int_0^{+\infty} e^{-\rho t} \left(g(Y(t; \eta_\lambda, u_\lambda)) + h(u^\lambda(t)) \right) dt \\ &\geq \int_0^{+\infty} e^{-\rho t} \Bigg(\lambda g(Y(t; \eta, u^\varepsilon)) + (1-\lambda) g(Y(t; \bar{\eta}, \bar{u}^\varepsilon)) + \lambda h(u^\varepsilon(t)) \\ &+ (1-\lambda) h(\bar{u}^\varepsilon(t)) \Bigg) dt \\ &= \lambda J(\eta, u^\varepsilon) + (1-\lambda) J(\bar{\eta}, \bar{u}^\varepsilon) > \lambda V(\eta) + (1-\lambda) V(\bar{\eta}) - \varepsilon \end{aligned}$$

Since ε is arbitrary, this shows both the claims.

We introduce the following assumptions.

Hypothesis 3.1.

(i) g_0 *is strictly increasing.*

(ii) $b_0 \geq 0$, $b_1(\cdot) \geq 0$ *almost everywhere.*

Proposition 3.3.

(i) *Let Hypothesis 3.1-(ii) hold. Then the value function V is nondecreasing along the direction $\hat{b}$ in $\mathcal{D}(V)$.*

(ii) *Let Hypotheses 3.1-(i)-(ii) hold. Then the value function V is strictly increasing along the direction $\hat{b}$ in $\mathcal{D}(V)$.*

Proof. (i) Given η, $\zeta \in H$ we say that $\eta \geq \zeta$ if

$$\eta_0 \geq \zeta_0 \quad \text{and} \quad \eta_1(\xi) \geq \zeta_1(\xi), \quad \forall\, \xi \in [-r, 0].$$

Let $\eta \in \mathcal{D}(V)$ and α_1, $\alpha_2 \in \mathbb{R}$ such that $\alpha_1 \leq \alpha_2$. We notice that $S_A(t)$ is positive preserving, which means that

$$\eta \geq 0 \implies S_A(t)\eta \geq 0.$$

Let $u(\cdot) \in \mathcal{U}(\eta + \alpha_1 \hat{b})$ and consider $Y(\cdot; \eta + \alpha_1 \hat{b}, u(\cdot))$. We have

$$Y(\cdot; \eta + \alpha_2 \hat{b}, u(\cdot)) - Y(\cdot; \eta + \alpha_1 \hat{b}, u(\cdot)) = S_A(t) \left((\alpha_2 - \alpha_1)\hat{b} \right) \geq 0. \tag{3.5}$$

Therefore in particular

$$Y_0(t; \eta + \alpha_2 b, u(\cdot)) \geq Y_0(t; \eta + \alpha_1 b, u(\cdot)). \tag{3.6}$$

This shows that $u(\cdot) \in \mathcal{U}(\eta + \alpha_2 \hat{b})$. Hypothesis 3.1-(i) implies that g is nondecreasing with respect to the order relation defined above.
Set

$$\beta(t) := \int_0^t S_A(t - \tau) B u(\tau) d\tau.$$

Then, also taking into account (3.5)

$$\begin{aligned}
&J(\eta + \alpha_2 \hat{b}; u(\cdot)) - J(\eta + \alpha_1 \hat{b}; u(\cdot)) \\
&= \int_0^{+\infty} e^{-\rho t} \left(g(Y(t; \eta + \alpha_2 \hat{b}, u(\cdot)) - g(Y(t; \eta + \alpha_1 \hat{b}, u(\cdot))) \right) dt \\
&= \int_0^{+\infty} e^{-\rho t} \left(g(S_A(t)(\eta + \alpha_2 \hat{b}) + \beta(t)) - g(S_A(t)(\eta + \alpha_1 \hat{b}) + \beta(t)) \right) dt.
\end{aligned}$$

So also

$$V(\eta + \alpha_2\hat{b}) \geq V(\eta + \alpha_1\hat{b})$$

and the claim is proved.

(ii) We consider $\eta \in \mathcal{D}(V)$ and $\alpha_1, \alpha_2 \in \mathbb{R}$ such that $\alpha_1 \leq \alpha_2$. By item (i) follows the monotonicity of the value function V. Then assuming by contradiction that it is not strictly increasing on $\mathcal{D}(V)$, it follows that it must be constant $\bar{\alpha_1}$ such that $V(\eta + \alpha_1\hat{b})$ is constant on the half line $[\bar{\alpha_1}, +\infty)$. In this case we have

$$V(\eta + \alpha_2 b) = V(\eta + \alpha_1 b). \tag{3.7}$$

Let $u^\varepsilon(\cdot) \in \mathcal{U}(\eta + \alpha_1\hat{b})$ a ε-optimal control for initial state $\eta + \alpha_1\hat{b}$ such that

$$J(\eta + \alpha_1\hat{b}) = V(\eta + \alpha_1\hat{b}) - \varepsilon. \tag{3.8}$$

By (3.6), $u^\varepsilon(\cdot) \in \mathcal{U}(\eta + \alpha_1\hat{b})$ and it is an admissible control for initial state $\eta + \alpha_2\hat{b}$.
Taking into account Hypotheses 2.40, Hypothesis 3.1-(i) we can write

$$J(\eta + \alpha_2\hat{b}, u^\epsilon(\cdot)) > J(\eta + \alpha_1\hat{b}, u^\epsilon(\cdot)),$$

so we have

$$J(\eta+\alpha_2\hat{b}, u^\varepsilon(\cdot)) > J(\eta+\alpha_1\hat{b}, u^\varepsilon(\cdot))+\varepsilon = V(\eta+\alpha_1\hat{b}) = V(\eta+\alpha_2\hat{b}) \geq J(\eta+\alpha_2\hat{b}).$$

□

3.0.4 $\|\cdot\|_{-1}$-continuity of the value function

In this subsection we prove a continuity property of the value function which is the key to treat the unbounded term in the HJB equation. From now on we assume that Hypothesis 3.1 holds true. In order to proceed we introduce the sets

$$\mathcal{F} := \left\{ \eta \in H_+ \,\middle|\, \int_{-\xi}^{0} \eta_1(s) e^{as} ds \geq 0 \quad \forall \xi \in [-r, 0] \right\}$$

and

$$H_{++} := (0, +\infty) \times \{\eta_1(\cdot) \in W^{1,2}_{r,0} \mid \eta_1(\cdot) > 0 \ a.e.\}.$$

We note that

$$H_{++} \subset \mathcal{F} \subset H_+.$$

Proposition 3.4.

(i) *$\forall \eta \in \mathcal{F}$, we have $0 \in \mathcal{U}(\eta)$. In particular, by* (3.2) *$\mathcal{F} \subset \mathcal{D}(V)$.*

(ii) *Let Hypothesis 3.1-(ii) hold. Then the set $\mathcal{D}(V)$ is open in the space $(H, \|\cdot\|_{-1})$.*

(iii) *Let Hypothesis 3.1-(ii) hold. Then the value function V is continuous with respect to $\|\cdot\|_{-1}$ in $\mathcal{D}(V)$.*

Proof. (i) Let $\eta \in \mathcal{F}$ and set $Y(\cdot) := Y(\cdot;\eta,0)$. Due to Definition 2.1-(i) and to the definition of the set $\mathcal{F}$,

$$\begin{aligned} Y_0(t) = [S_A(t)\eta]_0 &= \eta_0 e^{at} + \int_{(-t)\vee(-r)}^0 e^{a(t+\xi)}\eta_1(\xi)d\xi, \\ &= e^{at}\left(\eta_0 + \int_{(-t)\vee(-r)}^0 e^{a\xi}\eta_1(\xi)d\xi\right) > 0, \ \forall t \geq 0. \end{aligned}$$

(ii) Let $\bar\eta \in \mathcal{D}(V)$. Then in particular $\mathcal{U}(\bar\eta) \neq \emptyset$, so there exists a control $u(\cdot) \in \mathcal{U}(\bar\eta)$ such that $Y_0(t;\bar\eta,u(\cdot)) > 0$ for every $t \geq 0$. Given $\varepsilon > 0$ define

$$B_{-1}(\bar\eta,\varepsilon) := \{\eta \in H_+ \mid \|\eta - \bar\eta\|_{-1} < \varepsilon\}.$$

Due to (3.2) we have to prove that

$$\exists\, \varepsilon > 0, \text{ s.t. } \quad \mathcal{U}(\eta) \neq \emptyset \ \forall \eta \in B_{-1}(\bar\eta,\varepsilon). \tag{3.9}$$

Let $\eta \in B_{-1}(\bar\eta,\varepsilon)$. By Lemma 2.3, we know that

$$|Y_0(t;\eta,u(\cdot)) - Y_0(t;\bar\eta,u(\cdot))| \leq aC_t\|\eta - \bar\eta\|_{-1}, \ \forall t \in [0,r].$$

Since $Y_0(t;\bar\eta,u(\cdot)) > 0$ in $[0,r]$,

$$\exists\, \varepsilon > 0 \text{ s.t. } Y(t;\eta,u(\cdot)) > 0, \ \forall\, t \in [0,r]. \tag{3.10}$$

Define the control

$$\tilde u(\cdot) = \begin{cases} u(\cdot), & \text{if } t \in [0,s], \\ 0, & \text{if } t > s. \end{cases} \tag{3.11}$$

We have

$$Y_0(t;\eta,\tilde u(\cdot)) = Y_0(t;\eta,u(\cdot)) > 0, \quad \forall t \in [0,r].$$

By the semigroup properties of the mild solution of (2.11) (see [28, Chapter II]), setting $\tilde{t} = t - r$ we have

$$\tilde{\eta} := Y(t;\eta,\tilde{u}(\cdot)) = Y(t-r; Y(s;\eta,u(\cdot)),\tilde{u}(\cdot+r)) = Y(\tilde{t},\tilde{\eta},0),\ \forall\, t > r. \tag{3.12}$$

By (2.31) we have $Y_1(r,\eta,u(\cdot))(\xi) = \int_{-r}^{\xi} b_1(\alpha)u(r+\alpha-\xi)d\alpha$ that is nonnegative, since by Hypothesis 2.1 $u(\cdot) \geq 0$ and by Hypotheses 3.1-(ii) $b_1(\cdot) \geq 0$. This means that $\tilde{\eta} \in H_{++} \subset \mathcal{F}$, so that by (i)

$$Y(\tilde{t},\tilde{\eta},0)) > 0 \quad \forall\, \tilde{t} \geq 0.$$

Thus, by (4.5), we have that $Y(t;\eta,\tilde{u}(\cdot)) > 0$ for every $t \geq 0$, so (3.9) is proved with $\varepsilon > 0$ realizing (4.6).
(iii) The function V is concave and bounded from below in the $\|\cdot\|_{-1}$ open set $\mathcal{D}(V)$. Therefore the claim follows by a result of Convex Analysis ([27, Chapter 1, Corollary 2.4]) □

Remark 3.1. *$\mathcal{D}(V)$ is open also with respect to $\|\cdot\|$.* □

3.0.5 Superdifferential of concave $\|\cdot\|_{-1}$-continuous function

Motivated by Proposition 3.2 and Proposition 3.4, in this subsection we focus on the properties of the superdifferential of concave and $\|\cdot\|_{-1}$-continuous functions. This will be useful to prove the regularity result in the next section. We recall first some definitions and basic results from non-smooth analysis concerning the generalized differentials. For rigorous details we remind the reader to [53].

Let v be a concave continuous function defined on some open set $\mathcal{A}$ of H. We have the following

Definition 3.2. *For each $\eta \in \mathcal{A}$ the set*

$$D^-v(\eta) := \left\{ p \in H \mid \liminf_{\zeta \to \eta} \frac{v(\zeta) - v(\eta) - \langle \zeta - \eta, p\rangle_H}{\|\zeta - \eta\|} \geq 0 \right\},$$

$$D^+v(\eta) := \left\{ p \in H \mid \limsup_{\zeta \to \eta} \frac{v(\zeta) - v(\eta) - \langle \zeta - \eta, p\rangle_H}{\|\zeta - \eta\|} \leq 0 \right\}.$$

are called, respectively the (Fréchet) *subdifferential and superdifferential of v at η. They are convex and closed sets.*

Remark 3.2. *If $D^+v(\eta)\bigcap D^-v(\eta)\neq\emptyset$, then $D^+v(\eta)\bigcap D^-v(\eta)=\{p\}$, v is differentiable at η and $\nabla v(\eta)=p$.*

The set of the *reachable gradients* at $\eta\in\mathcal{A}$ is defined as

$$D^*v(\eta):=\{p\in H\mid\exists\,\eta_n\longrightarrow\eta,\,\eta_n\in\mathcal{A},\text{ such that }\exists\,\nabla v(\eta_n)\,\text{and}\,\nabla v(\eta_n)\longrightarrow p\}\,.$$

We suppose that the function v is concave and continuous; then, in this case, due to the [61, Chapter 1, Proposition 1.11], we can assert that D^+v is non empty at every point of $\mathcal{A}$, bounded, weakly closed and it holds

$$D^+v(\eta)=\{p\in H\mid v(\zeta)-v(\eta)\leq\langle\zeta-\eta,p\rangle_H,\quad\forall\,\zeta\in\mathcal{A}\}\,.$$

As known (see [61, Chapter 1, Proposition 1.11]) $D^*v(\eta)$ is a closed convex not empty subset of H. Moreover the set-valued map $\mathcal{A}\longrightarrow\mathcal{P}(H)$, $\eta\mapsto D^*v(\eta)$ is locally bounded (see again [61, Chapter 1, Proposition 1.11]). Also we have the representation (see [62, pp. 319-320])

$$D^+v(\eta)=\overline{co}(D^*v(\eta)),\quad\eta\in\mathcal{A}.\tag{3.13}$$

Given $p\in H$ we denote

$$p_{\hat b}:=\langle p,\hat b\rangle.\tag{3.14}$$

We introduce the directional superdifferential of v at η along the direction $\hat b$

$$D^+_{\hat b}v(\eta):=\Big\{\alpha\in\mathbb{R}\mid v(\eta+\gamma\hat b)-v(\eta)\leq\gamma\alpha,\ \forall\,\gamma\in\mathbb{R}\Big\}.\tag{3.15}$$

We have that this set is a nonempty closed and bounded interval $[a,c]\subset\mathbb{R}$. More precisely, since $v(\eta)$ is concave, we have

$$a=v^+_{\hat b}(\eta),\quad c=v^-_{\hat b}(\eta),$$

where $v^+_{\hat b}(\eta),v^-_{\hat b}(\eta)$ denote respectively the right and the left derivatives of the function $v(\eta)$ at the point $\hat b$. By definition of $D^+v(\eta)$, the projection of $D^+v(\eta)$ on $\hat b$ must be contained in $D^+_{\hat b}v(\eta)$, that is

$$D^+_{\hat b}v(\eta)\supset\left\{p_{\hat b}\mid p\in D^+v(\eta)\right\}.\tag{3.16}$$

On the other hand, Proposition 2.24 in [61], Chapter 1, states that

$$a=\inf\{\langle p,\hat b\rangle\mid p\in D^+v(\eta)\}\quad c=\sup\{\langle q,\hat b\rangle,\mid q\in D^+v(\eta)\},$$

and the sup and inf above are attained. This means that there exist $p, q \in D^+v(\eta)$ such that

$$a = \langle p, \hat{b} \rangle, \quad c = \langle q, \hat{b} \rangle.$$

Since $D^+v(\eta)$ is convex, we see that also the converse inclusion of (3.16) is true. Therefore

$$D^+_{\hat{b}} v(\eta) = \left\{ p_{\hat{b}} \mid p \in D^+v(\eta) \right\}. \tag{3.17}$$

Lemma 3.1. *The following statements hold:*

1. $A^{-1}(\mathcal{D}(V))$ *is a convex open set of* $(\mathcal{D}(A), \|\cdot\|_H)$.
2. $\mathcal{O} := Int_{(H,\|\cdot\|_H)}(Clos_{(H,\|\cdot\|_H)}(A^{-1}(\mathcal{D}(V))))$ *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. For the proof see [33, Chapter 3, Lemma 3.2.11]. □

Proposition 3.5. *Let* $v : \mathcal{D}(V) \longrightarrow \mathbb{R}$ *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* $p \in D^*v(\eta)$, *then there exists a sequence* $\eta_n \to \eta$ *such that*

$$\exists \nabla v(\eta_n), \quad \forall\, n \in \mathbb{N}, \quad \textit{and} \quad \nabla v(\eta_n) \rightharpoonup p, \; A^*\nabla v(\eta_n) \rightharpoonup A^*p.$$

Proof. Observe first that, since A^{-1} is one-to-one, there is a one-to-one correspondence between the elements $\eta \in \mathcal{D}(V)$ and $p \in A^{-1}(\mathcal{D}(V))$.

1. Let us define the function $u_0 : A^{-1}(\mathcal{D}(V)) \to \mathbb{R}$ by

$$u_0(p) := v(Ap).$$

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 the Lemma 3.1-3., $A^{-1}(\mathcal{D}(V))$ is $\|\cdot\|_H$ dense in $\mathcal{O}$. Since v is concave and locally Lipschitz continuous, u_0 can be extended to a concave $\|\cdot\|_H$ continuous function u defined on $\mathcal{O}$. This function satisfies the claim by construction.

2. Let $\eta \in \mathcal{D}(V)$, $\alpha \in D^+v(\eta)$. Then

$$v(\zeta) - v(\eta) \leq \langle \zeta - \eta, \alpha\rangle, \quad \forall \eta \in \mathcal{D}(V).$$

So, setting $p = A^{-1}\eta$, $q = A^{-1}\zeta$,

$$u(q) - u(p) \leq \langle A(q-p), \alpha\rangle \leq \quad \forall q \in A^{-1}\mathcal{D}(V).$$

Hence, the function $(\mathcal{D}(A), \|\cdot\|) \to, \mathbb{R}$, $q \mapsto \langle Aq, \alpha\rangle$ is lower semicontinuous at p. It is also linear and therefore it is continuous on $(\mathcal{D}(A), \|\cdot\|)$. So, we conclude that $\alpha \in \mathcal{D}(A^*)$.

3. Again we consider let $\eta \in \mathcal{D}(V)$, $\alpha \in D^+v(\eta)$. Then

$$v(\zeta) - v(\eta) \leq \langle \zeta - \eta, \alpha\rangle, \quad \forall \eta \in \mathcal{D}(V).$$

So, setting $p = A^{-1}\eta$, $q = A^{-1}\zeta$,

$$u(q) - u(p) \leq \langle A(q-p), \alpha\rangle = \langle q - p, A^*\alpha\rangle, \quad \forall q \in A^{-1}\mathcal{D}(V).$$

So, $A^*\alpha \in D^+u(p)$. This proves the inclusion $D^+u(A^{-1}\eta) \supset A^*D^+v(\eta)$. Conversely let $p \in A^{-1}(\mathcal{D}(V))$ and $\omega \in D^+u(p)$. Then

$$u(q) - u(p) \leq \langle q - p, \omega\rangle, \quad \forall q \in A^{-1}\mathcal{D}(V).$$

Thus, setting $\eta = Ap$, $\zeta = Aq$,

$$v(\zeta) - v(\eta) \leq \langle A^{-1}(\zeta - \eta), \omega\rangle = \langle (\zeta - \eta), (A^{-1})^*\omega\rangle \quad \forall \zeta \in \mathcal{D}(V).$$

Since $A(^{-1})^* = (A^*)^{-1}$, we get $(A^*)^{-1}\omega \in D^+v(\bar{\eta})$. This proves the inclusion $D^+u(A^{-1}\eta) \subset A^*D^+v(\eta)$.

4. Let $\eta \in \mathcal{D}(V)$ and $\alpha \in D^*v(\eta)$. By definition of $D^*v(\eta)$, we can find a sequence $(\eta_n) \subset \mathcal{D}(V)$ such that $\eta_n \to \eta$, $\nabla v(\eta_n)$ exists for any $n \in \mathbb{N}$ and $\nabla v(\eta_n) \rightharpoonup \alpha$. Setting $p_n = A^{-1}\eta_n$, thanks to claim 3 also $\nabla u(p_n)$ exists and $\nabla u(p_n) = A^*\nabla v(\eta_n)$. Since u is concave, the set-valued map $p \mapsto D^+u(p)$ is locally bounded. Hence the sequence $\nabla u(p_n)$ is bounded. Therefore from any subsequence we can extract a sub-subsequence converging to some element $j \in H$. The operator A^* is closed, so that the graph of A^* is closed in $(H \times H, \|\cdot\|_H \times \|\cdot\|_H)$. Therefore we can say that $\alpha \in \mathcal{D}(A^*)$ and $j = A^*\alpha$. Since this holds for any subsequence, we conclude that $A^\nabla v(\eta_n) = \nabla u(p_n) \rightharpoonup A^*\alpha$. $\square$

3.1 Dynamic Programming

Theorem 3.2. ***Dynamic Programming Principle***
For any $\eta \in H$ and for any $\tau \geq 0$,

$$V(\eta) = \sup_{u(\cdot)\in\mathcal{U}} \left[\int_0^\tau e^{-\rho t} \left(g\left(Y^{\eta,u(\cdot)}(t)\right) \; + \; h(u(t)) \right) dt + e^{-\rho\tau} V\left(Y^{\eta,u(\cdot)}(\tau)\right) \right],$$

where $Y^{\eta,u(\cdot)}(\cdot) := Y(\cdot)$.

Proof. See e.g. [53], Theorem 1.1 in Chapter 6. □

The differential version of the Dynamic Programming Principle is the HJB equation, which formally in our case reads as

$$\rho v(\eta) = \langle A\eta, \nabla v(\eta)\rangle \; + \; g(\eta) \; + \; \sup_{u\in U} \left\{ \langle Bu, \nabla v(\eta)\rangle + h(u) \right\}, \qquad \eta \in \mathcal{D}(V). \tag{3.18}$$

We have

$$\sup_{u\in U} \left\{ \langle Bu, p\rangle \; + \; h(u) \right\} = \sup_{u\in U} \left\{ \langle u, B^* p\rangle \; + \; h(u) \right\}.$$

Therefore, defining the Legendre transform of h

$$h^*(r) := \sup_{u\in U} \{ur + h(u)\},$$

and taking into account that

$$B^* p = \langle \hat{b}, p\rangle,$$

(3.18) can be rewritten as

$$\rho v(\eta) = \langle \eta, A^* \nabla v(\eta)\rangle \; + \; g(\eta) \; + \; h^*(\langle \nabla v(\eta), \hat{b}\rangle), \qquad \eta \in \mathcal{D}(V). \tag{3.19}$$

We note that the nonlinear term in (3.19) can be defined without requiring the full regularity of v, but only the C^1-smoothness of v with respect to the direction $\hat{b}$. Indeed, denoting coherently with (3.17) by $v_{\hat{b}}$ the directional derivative of v with respect to $\hat{b}$, we can intend the nonlinear term in (3.19) as

$$h^*(\langle \nabla v(\eta), \hat{b}\rangle) = \mathcal{H}(v_{\hat{b}}(\eta)),$$

where

$$\mathcal{H}(r) = \; h^*(\|b\| r), \quad r \in \; (0, +\infty).$$

So we write (3.19) as

$$\rho v(\eta) = \langle \eta, A^* \nabla v(\eta)\rangle \; + \; g(\eta) \; + \; \mathcal{H}(v_{\hat{b}}(\eta)), \qquad \eta \in \mathcal{D}(V). \tag{3.20}$$

Due to Hypothesis 2.3-(ii), the function $\mathcal{H}$ is finite on $(0, +\infty)$ and strictly convex and increasing therein (see [64, Corollary 26.4.1]).

3.1.1 The HJB equation: viscosity solutions

In this subsection we are going to prove that the value function V is a viscosity solution of the HJB equation (3.20). We will make use of the Yoshida approximations A_n of the unbounded operator A; for further details on that we refer to [11, section 2.4].

Next we give with an approximation result (Lemma 3.2) that we need to prove a chain's rule in infinite dimension for suitable regular functions (provided in Lemma 3.3 below).

Lemma 3.2.

(i) *Let A_n, $n \in \mathbb{N}$, be the (bounded) Yosida approximations of the (unbounded) operator A and let S_{A_n}, $n \in \mathbb{N}$ be the associated uniformly continuous semigroups. Let*

$$Y_n(t) = S_{A_n}(t)\eta + \int_0^t S_{A_n}(t-\tau)Bu(\tau)d\tau. \tag{3.21}$$

Then, for each $n \in \mathbb{N}$, we have $Y_n \in C^1([0,+\infty); H)$ and moreover Y_n solves the differential equation

$$\begin{cases} Y_n'(t) = \ A_nY_n(t)dt \ + \ Bu(t)dt, \\ \\ Y_n(0) = \ \eta, \end{cases} \tag{3.22}$$

(ii) *Let Y be the mild solution to* (2.11) *as in Definition 2.1-(i) and let Y_n, $n \in \mathbb{N}$, be the functions defined by* (3.21).

Then, for each $T \geq 0$,

$$Y_n(t) \to \ Y(t), \quad \textit{in } L^2([0,T]; H). \tag{3.23}$$

Proof. For details of the proof we remind the reader to [53, Chapter 2, Proposition 5.4].
Let S_{A_n} be the uniformly continuous semigroups on H generated by the bounded operators A_n and S_A the C_0-semigroup generated by the unbounded operator A. We know that there exist $\omega, M, \tilde{M}$ $\mathbb{R}$-valued constants such that

$$\|S_{A_n}(t)\| \leq \ Me^{\omega t}, \tag{3.24}$$

and

$$\|S_A(t)\| \leq \ \tilde{M}e^{\omega t}, \tag{3.25}$$

for each $t \geq 0$, (see [53, Chapter 2, Proposition 4.7]). So, for any $t \geq 0$ and $\eta \in H$ we have the convergence, for $n \to \infty$,

$$\|S_{A_n}(t)\eta - S_A(t)\eta\|_H \to 0. \tag{3.26}$$

The proof of the existence and uniqueness for the solution to (3.22) follows [11, Part II, Section 3.1] and by a consequence of the boundness of A_n.
As for equation (2.11), we write the mild solution to (3.22)

$$Y_n(t) = S_{A_n}(t)\eta + \int_0^t S_{A_n}(t-\tau)Bu(\tau)d\tau. \tag{3.27}$$

Therefore we can write

$$\|Y_n(t)-Y(t)\|_H \leq \|S_{A_n}(t)\eta - S_A(t)\eta\|_H + \int_0^t \|\left(S_{A_n}(t-\tau) - S_A(t-\tau)\right)Bu(\tau)\|_H d\tau. \tag{3.28}$$

The first term of the right-handside of (3.28) can be dominated in $L^2([0,T];\mathbb{R})$ thanks to (3.24) and (3.25). The second one can be dominated thanks to (3.24), (3.25) and by Hölder's inequality.
Moreover, the right-handside of (3.28) converges pointwise to 0, when $n \to \infty$, thanks to (3.26). Therefore (3.23) follows by dominated convergence from (3.28), letting $n \to \infty$. □

Now we need to define a suitable set of regular test functions. This is the set This is the set

$$\mathcal{T} := \left\{\varphi \in C^1(H) \mid \nabla\varphi(\cdot) \in \mathcal{D}(A^*),\ A^*\nabla\varphi : H \to H \text{ is continuous}\right\}. \tag{3.29}$$

Let us define, for $u \geq 0$, the operator $\mathcal{L}^u$ on $\mathcal{T}$ by

$$[\mathcal{L}^u\varphi](\eta) := -\rho\varphi(\eta) + \langle \eta, A^*\nabla\varphi(\eta)\rangle_H + u\langle\nabla\varphi(\eta), \hat{b}\rangle.$$

Lemma 3.3. *Let $\eta \in H$, $\varphi \in \mathcal{T}$, $u(\cdot) \in L^2_{loc}([0,+\infty);\mathbb{R})$ and set $Y(t) := Y(t;\eta,u(\cdot))$. Then the following chain's rule holds*

$$e^{-\rho t}\varphi(Y(t)) - \varphi(\eta) = \int_0^t e^{-\rho s}[\mathcal{L}^{u(s)}\varphi](Y(s))ds \quad \forall\, t \geq 0.$$

Proof. For details of the proof we remind the reader to [53, Chapter 2, Proposition 5.5].

We consider the functions $Y_n(t)$ of Lemma 3.2-(i) and calculate the derivatives of the functions $e^{-\rho t}\varphi(Y_n(t))$. Due to (3.22), we have

$$\begin{aligned}\frac{d}{dt}\left[e^{-\rho t}\varphi(Y_n(t))\right] &= -\rho e^{\rho t}\varphi(Y_n(t)) + e^{-\rho t}\langle\nabla\varphi(Y_n(t)), Y_n'(t)\rangle \\ &= e^{-\rho t}\left[-\rho\varphi(Y_n(t) + \langle\nabla\varphi(Y_n(t)), A_nY_n(t) + Bu(t)\rangle\right] \\ &= e^{-\rho t}\big[-\rho\varphi(Y_n(t)) + \langle A_n^*\nabla\varphi(Y_n(t)), Y_n(t)\rangle \\ &+ \langle\nabla\varphi(Y_n(t)), Bu(t)\rangle\big].\end{aligned} \tag{3.30}$$

Therefore

$$\begin{aligned}&e^{-\rho t}\varphi(Y_n(t)) - \varphi(\eta) \\ &= \int_0^t e^{-\rho s}\left[-\rho\varphi(Y_n(s)) + \langle A_n^*\nabla\varphi(Y_n(t)), Y_n(t)\rangle + \langle\nabla\varphi(Y_n(t)), Bu(t)\rangle\right] ds.\end{aligned} \tag{3.31}$$

We want to get the claim letting $n \to \infty$ and taking into account Lemma 3.2 and the continuity properties of φ and its derivatives.
Due to (3.23) we have $Y_n \to Y$ in $L^2([0,T];H)$ for each $t \geq 0$ such that, for the left-handside of the above equation we can write the following

$$e^{-\rho t}\varphi(Y_n(t)) - \varphi(\eta) \longrightarrow e^{-\rho t}\varphi(Y(t)) - \varphi(\eta), \tag{3.32}$$

letting $n \to \infty$.
Consider now the right handside of (3.31). We recall that the operator $A^*\nabla\varphi : H \longrightarrow H$ is continuous, so that, taking into account the regularity properties of φ, of its derivatives, and the hypotheses of B and u, we can state that, letting $n \to \infty$ we have

$$\begin{aligned}&-\rho\varphi(Y_n(s)) + \langle A_n^*\nabla\varphi(Y_n(t)),Y_n(t)\rangle + \langle\nabla\varphi(Y_n(t)), Bu(t)\rangle \longrightarrow \\ &\qquad - \rho\varphi(Y(s)) + \langle A^*\varphi'(Y(s)), Y(s)\rangle + u\langle\varphi'(Y(s)), \hat{b}\rangle.\end{aligned} \tag{3.33}$$

By (3.32) and (3.33) we can conclude that it holds

$$\begin{aligned}&e^{-\rho t}\varphi(Y(t)) - \varphi(\eta) \\ &= \int_0^t e^{-\rho s}\Big(-\rho\varphi(Y(s)) + \langle A^*\varphi'(Y(s)), Y(s)\rangle + u\langle\varphi'(Y(s)), \hat{b}\rangle\Big)\, ds \\ &= \int_0^t e^{-\rho s}[\mathcal{L}^{u(s)}\varphi](Y(s))ds\end{aligned}$$

which proves the claim. $\square$

Next we give the definition of viscosity solution to (3.20).

Definition 3.3.

(i) *A continuous function $v : \mathcal{D}(V) \to \mathbb{R}$ is called a viscosity subsolution of* (3.20) *if, for each couple $(\eta_M, \varphi) \in \mathcal{D}(V) \times \mathcal{T}$ such that $v - \varphi$ has a local maximum at η_M, we have*

$$\rho v(\eta_M) \leq \langle \eta_M, A^* \nabla \varphi(\eta_M) \rangle + g(\eta_M) + \mathcal{H}\left(\varphi_{\hat{b}}(\eta_M)\right).$$

(ii) *A continuous function $v : \mathcal{D}(V) \to \mathbb{R}$ is called a viscosity supersolution of* (3.20) *if, for each couple $(\eta_m, \varphi) \in \mathcal{D}(V) \times \mathcal{T}$ such that $v - \varphi$ has a local minimum at η_m, we have*

$$\rho v(\eta_m) \geq \langle \eta_m, A^* \nabla \varphi(\eta_m) \rangle + g(\eta_m) + \mathcal{H}\left(\varphi_{\hat{b}}(\eta_M)\right).$$

(iii) *A continuous function $v : \mathcal{D}(V) \to \mathbb{R}$ is called a viscosity solution of* (3.20) *if it is both a viscosity sub and supersolution of* (3.20).

We need the following technical Lemma 3.4 to prove that V is a viscosity solution.

Lemma 3.4. *For every 1-optimal control $u(\cdot) \in \mathcal{U}(\eta)$ we have*

$$\int_0^1 |u(s)|^{1+\alpha} ds \leq M$$

where α is the constant appearing in (2.5).

Proof. The proof of the lemma follows directly by Hypotheses 2.3. □

Theorem 3.3. *The value function V is a viscosity solution of* (3.20).

Proof. *Subsolution property.* Let $(\eta_M, \varphi) \in \mathcal{D}(V) \times \mathcal{T}$ 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

$$2\nu \leq \rho V(\eta_M) - \left(\langle \eta_M, A^* \nabla \varphi(\eta_M) \rangle + g(\eta_M) + \mathcal{H}\left(\varphi_{\hat{b}}(\eta_M)\right)\right).$$

Let us define the function

$$\tilde{\varphi}(\eta) := V(\eta_M) + \langle \nabla \varphi(\eta_M), \eta - \eta_M \rangle, \quad \eta \in H. \tag{3.34}$$

We have
$$\nabla\tilde{\varphi}(\eta) = \nabla\varphi(\eta_M), \quad \forall\, \eta \in H.$$
Thus $\tilde{\varphi}$ is also test function and we have as well
$$2\nu \leq \rho V(\eta_M) - \left(\langle \eta_M, A^*\nabla\tilde{\varphi}(\eta_M)\rangle + g(\eta) + \mathcal{H}\left(\tilde{\varphi}_{\hat{b}}(\eta_M)\right)\right).$$
Now, we know that V is a concave function and that $V - \varphi$ has a local maximum at η_M, so that
$$V(\eta) \leq V(\eta_M) + \langle \nabla\varphi(\eta_M), \eta - \eta_M\rangle. \tag{3.35}$$
Thus, by (3.34) and (3.35)
$$\tilde{\varphi}(\eta_M) = V(\eta_M) \quad \tilde{\varphi}(\eta) \geq V(\eta), \quad \forall\, \eta \in \mathcal{D}(V). \tag{3.36}$$
Let $B_\varepsilon := B_{(H,\|\cdot\|)}(\eta_M, \varepsilon)$. Due to the properties of the functions belonging to $\mathcal{T}$, we can find $\varepsilon > 0$ such that
$$\nu \leq \rho V(\eta) - \left(\langle \eta, A^*\nabla\tilde{\varphi}(\eta)\rangle + g(\eta) + \mathcal{H}\left(\tilde{\varphi}_{\hat{b}}(\eta_M)\right)\right), \quad \forall \eta \in B_\varepsilon. \tag{3.37}$$
Take a sequence $\delta_n > 0$ such that $\delta_n \to 0$. For each $n \in \mathbb{N}$, take a δ_n-optimal control $u_n(\cdot) \in \mathcal{U}(\eta)$ and set $Y^n(\cdot) := Y(\cdot; \eta_M, u_n(\cdot))$. Define
$$t_n := \inf\{t \geq 0 \mid \|Y^n(t) - \eta_M\| = \varepsilon\} \wedge 1 \tag{3.38}$$
with the agrement that $\inf \emptyset = +\infty$. Of course t_n is well defined and belongs to $(0,1]$. Moreover, by continuity of $t \mapsto Y^n(t)$, we have $Y^n(t) \in B_\varepsilon$, for $t \in [0, t_n)$.
By definition of δ_n-optimal control, we have as consequence of the Dynamic Programming Principle
$$\delta_n \geq -\int_0^{t_n} e^{-\rho t}\left[g(Y^n(t)) + h(u_n(t))\right] dt - \left(e^{-\rho t_n} V(Y(t_n)) - V(\eta_M)\right). \tag{3.39}$$
Therefore, by (3.36) and (3.39),
$$\begin{aligned}
\delta_n &\geq -\int_0^{t_n} \left[g(Y^n(t)) + h(u_n(t))\right] dt - \left(e^{-\rho t_n}\left(\tilde{\varphi}(Y^n(t_n))\right) - \tilde{\varphi}(\eta_M)\right) \\
&= \int_0^{t_n} e^{-\rho t}\left[g(Y^n(t)) + h(u_n(t)) + [\mathcal{L}^{u_n(t)}\tilde{\varphi}](Y^n(t))\right] dt \\
&\geq -\int_0^{t_n} e^{-\rho t}\Big[g(Y^n(t)) - \rho\tilde{\varphi}(Y^n(t)) + \langle A^*\nabla\tilde{\varphi}(Y^n(t)), Y^n(t)\rangle \\
&+ \mathcal{H}(\tilde{\varphi}_{\hat{b}}(Y^n(t)))\Big] dt \\
&\geq t_n \nu.
\end{aligned}$$

Since $\delta_n \longrightarrow 0$ we also have $t_n \longrightarrow 0$. Now we want to prove that the following convergence holds true

$$\|Y^n(t_n) - \eta_M\| \longrightarrow 0. \tag{3.40}$$

We use the definition of mild solution (2.1) of $Y^n(t_n)$, so we have

$$\begin{aligned}
\|Y^n(t_n) - \eta_M\| &= \left\| S_A(t_n)\eta_M + \int_0^{t_n} S_A(t_n - \tau)Bu_n(\tau)d\tau \; - \; \eta_M \right\| \\
&\le \|(S_A(t_n) - I)\,\eta_M\| + \left\| \int_0^{t_n} S_A(t_n - \tau)Bu_n(\tau)d\tau \right\| \\
&\le \|(S_A(t_n) - I)\,\eta_M\| + \int_0^{t_n} \|S_A(t_n - \tau)\|_{\mathcal{L}(H)}\|B\||u_n(\tau)|d\tau.
\end{aligned}$$

By properties of semigroup $S_A(\cdot)$, to prove that the right side of above inequality converges to 0, it suffices to prove that

$$\int_0^{t_n} |u_n(s)|ds \to 0. \tag{3.41}$$

Set $\beta > 1$ and $1/\beta + 1/\alpha = 1$, so that $\alpha = \beta/(\beta - 1)$. By Hölder's inequality

$$\int_0^{t_n} |u_n(s)|ds \le \left(\int_0^{t_n} |u_n(\tau)|^\beta d\tau \right)^{\frac{1}{\beta}} t_n^{\frac{\beta-1}{\beta}}.$$

Since by Lemma 3.4 we know that $\left(\int_0^{t_n} |u_n(\tau)|^\beta d\tau \right)^{\frac{1}{\beta}}$ is bounded and since $t_n \to 0$, we have (3.41) and the convergence (3.40) is true. But (3.40) contradicts the definition of t_n, so the claim is proved.

Supersolution property. The proof that V is a viscosity supersolution is more standard. We refer to [53, Chapter 6, Theorem 3.2]. □

3.1.2 Smoothness of viscosity solutions

In this subsection we are going to show a C^1-directional regularity result for the value function. To this aim we need the following.

Lemma 3.5. *Let $v : \mathcal{D}(V) \longrightarrow \mathbb{R}$ be a concave $\|\cdot\|_{-1}$-continuous function and suppose that $\eta \in \mathcal{D}(V)$ is a differentiability point for v and that $\nabla v(\eta) = \alpha$. Then*

1. *There exists a test function φ such that $v - \varphi$ has a local maximum at η and $\nabla\varphi(\eta) = \alpha$.*

2. *There exists a test function φ such that $v - \varphi$ has a local minimum at η and $\nabla\varphi(\eta) = \alpha$.*

Proof. Thanks to Proposition 3.5-(2) and to the concavity of v, the first statement is clearly satisfied by the function $\langle \cdot, \alpha\rangle$. We prove now the second statement, which is more delicate.
Let u be defined as in Proposition 3.5-(1). The first and third claim of Proposition 3.12 yield that $\alpha \in \mathcal{D}(A^*)$, u is differentiable at $p := A^{-1}\eta$ and $\nabla u(p) = A^*\alpha$. This yields

$$u(q) - u(p) - \langle q - p, A^*\alpha\rangle \geq -\|q-p\| \cdot \epsilon\left(\|q-p\|\right),\ \forall q \in A^{-1}\mathcal{D}(V),$$

for some $\epsilon : [0,+\infty) \longrightarrow [0,+\infty)$ increasing and such that $\epsilon(r) \longrightarrow 0$, when $r \longrightarrow 0$. The previous inequality can be rewritten as

$$u(q) - u(p) - \langle A\left(q-p\right), \alpha\rangle \geq -\|q-p\| \cdot \epsilon\left(\|q-p\|\right),\ \forall q \in A^{-1}\mathcal{D}(V).$$

Therefore, defining $\zeta = Aq$ for $q \in A^{-1}\mathcal{D}(V)$ and recalling that A is one-to-one from $A^{-1}\mathcal{D}(V)$ to $\mathcal{D}(V)$,

$$v(\zeta) - v(\eta) - \langle \zeta - \eta, \alpha\rangle \geq -\|\zeta-\eta\|_{-1}\epsilon\left(\|\zeta-\eta\|_{-1}\right), \forall\, \zeta \in \mathcal{D}(V). \quad (3.42)$$

We look for a test function of this form:

$$\varphi(\zeta) := v(\eta) + \langle \zeta - \eta, \alpha\rangle - g(\|\zeta-\eta\|_{-1}),\ \zeta \in \mathcal{D}(V),$$

where $g : [0,\infty) \longrightarrow [0,+\infty)$ is a suitable increasing C^1 function such that $g(0) = g'(0) = 0$. Notice that, since $g(0) = 0$, we have $\varphi(\eta) = v(\eta)$. So, in order to prove that $v - \varphi$ has a local minimum at η, we have to prove that $\varphi \leq v$ in a neighborhood of η. Let us define the function

$$g(r) := \int_0^{2r} \epsilon(s)ds.$$

We see that $g(0) = g'(0) = 0$ and

$$g(r) \geq \int_r^{2r} \epsilon(s)ds \geq r\epsilon(r).$$

By (3.42),

$$\begin{aligned}\varphi(\zeta) &= v(\eta) + \langle \zeta-\eta, \alpha\rangle - g(\|\zeta-\eta\|_{-1})\\ &= v(\eta) + \langle \zeta-\eta, \alpha\rangle - \|\zeta-\eta\|_{-1} \cdot \epsilon(\|\zeta-\eta\|_{-1}) \leq v(\zeta), \quad \forall\, \zeta \in \mathcal{D}(V).\end{aligned}$$

Moreover, recalling that $(A^{-1})^* = (A^*)^{-1}$,

$$\nabla\varphi(\zeta) = \begin{cases} \alpha - (A^*)^{-1} g'(\|\zeta - \eta\|_{-1}) \dfrac{A^{-1}(\zeta - \eta)}{\|A^{-1}(\zeta - \eta)\|}, & \text{if } \zeta \neq \eta; \\ \\ \alpha, & \text{if } \zeta = \eta. \end{cases}$$

This expression of $\nabla\varphi$ shows that $\zeta \mapsto A^*\nabla\varphi(\zeta)$ is continuous. Therefore, φ is a test function. Finally, $\nabla\varphi(\eta) = \alpha$ and the proof is complete. □

Now we can state and prove the main result.

Theorem 3.4. *Let v be a concave $\|\cdot\|_{-1}$-continuous viscosity solution of* (3.20) *on $\mathcal{D}(V)$, which is strictly increasing along the direction $\hat{b}$. Then v is differentiable along $\hat{b}$ at any point $\eta \in \mathcal{D}(V)$ and the function $\eta \mapsto v_{\hat{b}}(\eta)$ is continuous in $\mathcal{D}(V)$.*

Proof. Let $\eta \in \mathcal{D}(V)$ and $p,\, q \in D^*v(\eta)$. Thanks to Proposition 3.5, there exist sequences (η_n), $(\tilde{\eta}_n) \subset \mathcal{D}(V)$ such that:

- $\eta_n \to \eta$, $\tilde{\eta}_n \to \eta$;
- $\nabla v(\eta_n)$ and $\nabla v(\tilde{\eta}_n)$ exist for all $n \in \mathbb{N}$ and $\nabla v(\eta_n) \rightharpoonup p$, $\nabla v(\tilde{\eta}_n) \rightharpoonup q$;
- $A^*\nabla v(\eta_n) \rightharpoonup A^*p$ and $A^*\nabla v(\tilde{\eta}_n) \rightharpoonup A^*q$.

Recall that, given $\eta \in H$, we have defined

$$\eta_{\hat{b}} := \langle \eta, \hat{b} \rangle.$$

Thanks to Lemma 3.5 and Theorem 3.3 we can write, for any $n \in \mathbb{N}$,

$$\rho v(\eta_n) = \langle \eta_n, A^*\nabla v(\eta_n) \rangle_H + g(\eta_n) + \mathcal{H}(v_{\hat{b}}(\eta_n))$$

$$\rho v(\tilde{\eta}_n) = \langle \eta_n, A^*\nabla v(\tilde{\eta}_n) \rangle + g(\tilde{\eta}_n) + \mathcal{H}(v_{\hat{b}}(\tilde{\eta}_n)).$$

So, passing to the limit, we get

$$\langle \eta, A^*p \rangle + g(\eta) + \mathcal{H}(p_{\hat{b}}) = \rho v(\eta) = \langle \eta, A^*q \rangle + g(\eta) + \mathcal{H}(q_{\hat{b}}). \tag{3.43}$$

On the other hand $\lambda p + (1-\lambda)q \in D^+v(\eta)$ for any $\lambda \in (0,1)$, so that we have the subsolution inequality

$$\rho v(\eta) \leq \langle \eta, A^*[\lambda p + (1-\lambda)q] \rangle + g(\eta) + \mathcal{H}(\lambda p_{\hat{b}} + (1-\lambda)q_{\hat{b}}),\ \forall\, \lambda \in (0,1). \tag{3.44}$$

Combining (3.43) and (3.44) we get

$$\mathcal{H}(\lambda p_{\hat{b}} + (1-\lambda) q_{\hat{b}}) \geq \lambda \mathcal{H}(p_{\hat{b}}) + (1-\lambda)\mathcal{H}(q_{\hat{b}}). \tag{3.45}$$

Notice that, since $p, q \in D^* v(\eta)$, we have also $p, q \in D^+ v(\eta)$. Since v is strictly increasing along $\hat{b}$ we must have $p_{\hat{b}}, q_{\hat{b}} \in (0, +\infty)$. Since $\mathcal{H}$ is strictly convex on $(0, +\infty)$, (3.45) yields $p_{\hat{b}} = q_{\hat{b}}$. Due to (3.17) we have that $p_{\hat{b}}$, $q_{\hat{b}} \in D^+_{\hat{b}} v(\eta)$. With this argument we have shown that the projection of $D^* v(\eta)$ onto $\hat{b}$ is a singleton. Due to (3.13), this implies that also the projection of $D^+ v(\eta)$ onto $\hat{b}$ is a singleton. Due to (3.17) we have that $D^+_{\hat{b}} v(\eta)$ is a singleton too. Since v is concave, this is enough to conclude that it is differentiable along the direction $\hat{b}$ at η.

Now we prove the second claim of the Theorem, that is that the map $\eta \mapsto v_b(\eta)$ is continuous in $\mathcal{D}(V)$. To this aim we take $\eta \in \mathcal{D}(V)$ and a sequence $(\eta^n) \subset \mathcal{D}(V)$ such that $\eta^n \to \eta$. We have to prove that $v_{\hat{b}}(\eta^n) \to v_{\hat{b}}(\eta)$.
Being v concave, by definition of superdifferential (3.17) for every $n \in \mathbb{N}$, there exists $p_n \in D^+ v(\eta_n)$ such that $\langle p_n, \hat{b} \rangle = v_{\hat{b}}(\eta_n) \in D^+_{\hat{b}}(\eta_n)$.
Since v is concave, it is also locally Lipschitz continuous, so that the superdifferential is a locally bounded multi-function (see [61, Chapter 1, Proposition 2.5]). Therefore, from each subsequence (p_{n_k}) we can extract a sub-subsequence $(p_{n_{k_h}})$ such that

$$p_{n_{k_h}} \longrightarrow p \in \mathcal{D}(V)$$

for some limit point p. Due to concavity of v, this limit point must belong to $D^+ v(\eta)$. We have shown in the first part of the proof that the projection of $D^+ v(\eta)$ onto $\hat{b}$ is the singleton $v_{\hat{b}}(\eta)$, so that it must be

$$\langle p, \hat{b} \rangle = v_{\hat{b}}(\eta).$$

With this argument we have shown that, from each subsequence $(v_{\hat{b}}(\eta^{n_k}))$, we can extract a sub-subsequence $(v_{\hat{b}}(\eta^{n_{k_h}}))$ such that

$$v_{\hat{b}}(\eta^{n_{k_h}}) = \langle p_{n_{k_h}}, \hat{b} \rangle \to \langle p, \hat{b} \rangle = v_{\hat{b}}(\eta).$$

The claim follows by the usual argument on subsequences. □

Remark 3.3. *Notice that in the assumption of Theorem 3.4 we do not require that v is the value function, but only that it is a concave $\|\cdot\|_{-1}$-continuous viscosity solution of* (3.20) *strictly increasing along $\hat{b}$.*

Chapter 4

Verification Theorem

In this chapter we draw a simple consequence on the infinite horizon optimal problem from the results on the associated Hamilton-Jacobi-Bellman equation of the previous section, that is, we want to prove the Verification Theorem. To this aim we will recall some definitions and propositions which will be needed in the proof of the theorem. For details we remind the reader to [9, Chapter III, Section 2].

The characterization of optimal *closed loop controls* provides a method constructing an optimal pair control-trajectory for every initial condition.
The first step is to find a map $G : \mathbb{R}^n \longrightarrow A$, with the property that

$$G(z) \in \operatorname{argmax}_{u \in A} \{-f(z,u)Dv(z) - l(z,u)\},$$

for $z \in \mathbb{R}^n$; if v is known, this is a static, finite dimensional, mathematical programming problem. Such a map G is called an *optimal feedback map.*
The second step is solving

$$\begin{cases} y' = f(y, G(y)), & t > 0, \\ y(0) = x \end{cases} \tag{4.1}$$

and a solution $y^*(t)$ generates a control $\alpha^*(t) := G(y^*(t))$, which is optimal for the initial state x.
Finally the use of viscosity subsolutions as verification functions gives verification theorems as a very easy consequence of the comparison principle.
The applicability of this method requires the regularity of the value function for the characterization of optimal controls and some regularity of the feedback map G for the solvability of (4.1).

In our case, to prove the Verification Theorem, we will follow the procedure used by Federico (see [33, Chapter 3, Section 3.3.3].

Thanks to the regularity result of the previous section (Theorem 3.4), we can define a feedback map on $\mathcal{D}(V)$ for our problem in classical form, that is:

$$G(\eta) := \operatorname{argmax}_{u \in U} \{h(u) + uV_{\hat{b}}(\eta)\|b\|\}, \quad \eta \in \mathcal{D}(V).$$

This map is well-defined: since V is concave and strictly increasing along $\hat{b}$, so that by Theorem 3.4

$$\exists\, V_{\hat{b}}(\eta) \in (0, +\infty), \quad \forall \eta \in \mathcal{D}(V); \tag{4.2}$$

therefore existence and uniqueness of the argmax follow from Hypothesis 2.3-(ii).
Moreover, since $V_{\hat{b}}$ is continuous on $\mathcal{D}(V)$, then also G is continuous on $\mathcal{D}(V)$.
Thus, for $\eta \in \mathcal{D}(V)$, the closed-loop delay state equation associated with this map is defined as

$$\begin{cases} Y'(t) = AY(t) + BG(Y(t)) \\ \\ Y(0) = \eta. \end{cases} \tag{4.3}$$

We recall that we are considering the infinite dimensional case, so this means that

$$G(Y(t)) = G\left(Y_0(t), Y_1(t)(\xi)\,|_{\xi\in[-r,0]}\right) = G\left(y(t), \int_{-r}^{\xi} b_1(\alpha)u(\alpha + t - \xi)\,|_{\xi\in[-r,0]}\, d\alpha\right).$$

Given $u(\cdot) \in U$ and recalling the Hypothesis 2.1 on U, we write the explicit form of the mild solution $Y \in C([0, +\infty), H)$ of (4.3) defined as

$$Y(t) = S_A(t)\eta + \int_0^t S_A(t-\tau)BG(Y(\tau))d\tau, \quad t \geq 0. \tag{4.4}$$

With Verification theorem, we want to prove that if the closed-loop equation (4.3) has a strictly positive solution $Y^*(\cdot)$ defined by (4.4), (so that $(Y_0^*(t), Y_1^*(\cdot)(\xi)\,|_{\xi\in[-r,0]})$ belongs to $\mathcal{D}(V)$ for all $t \geq 0$ and the term $G\left((Y_0^*(t), Y_1^*(\cdot)(\xi)\,|_{\xi\in[-r,0]})\right)$ is well-defined for all $t \geq 0$), then the feedback strategy defined as

$$u^*(t) := G(Y^*(t)) = G\left(Y_0^*(t), Y_1^*(t)(\xi)\,|_{\xi\in[-r,0]}\right) \tag{4.5}$$

is optimal. We remark that, if $u^*(\cdot)$ is admissible and it is defined as in (4.5), then, setting $Y^*(t) := Y(t; \eta, u^*(\cdot))$, we have

$$Y^*(t) = (Y_0^*(t), Y_1(t)(\xi)|_{\xi\in[-r,0]}) \in \mathcal{D}(V), \quad \forall t \geq 0.$$

Remark 4.1. *We have to remark that in this case, the feedback strategy* $u^*(\cdot)$ *above defined, depends on the* $Y_0^*(\cdot)$ *defined in* (4.4)*, and on the* $Y_1^*(\cdot)(\xi)$ *for* $\xi \in [-r, 0)$ *that represents the past-value of the control* $u(\cdot + \xi)$*, for* $\xi \in [-r, 0)$. □

The proof of the Verification Theorem, in the classical case, is done by computing the derivative

$$t \mapsto \frac{d}{dt}[e^{-\rho t} V(Y^*(t))] \tag{4.6}$$

and then using the HJB equation and integrating the resulting equality.
We cannot proceed with the classical proof since we cannot compute the derivative (4.6). So we proceed considering the fact that V is a viscosity solution (as in [71, Chapter 5, Theorem 3.9] and in [53, Chapter 6, Theorem 5.4, 5.5]). But we have to overcome two main difficulties (see [35, pag 40]).

- The function
$$t \mapsto e^{-\rho t} V(Y^*(t)) \tag{4.7}$$
is not Lipschitz continuous so it may not have almost everywhere derivative. Indeed we do not require the initial datum $\eta \in \mathcal{D}(A)$ and the operator A works as a shift operator on the infinite-dimensional component. Then we do not have the condition $Y^*(t) \in \mathcal{D}(A)$ for almost $t \geq 0$ that would give the required Lipschitz regularity for the function (4.7): only continuity is insuring so we cannot apply the Fundamental Theorem of Calculus (see [53, Chapter 6, Theorem 5.4, 5.5]).

- Consequently we had to deal with the concept of Dini derivatives of the function (4.7), 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 derivatives. We will use the so called SaKs Theorem, that need 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

$$D^-g(t) = \limsup_{h \uparrow 0} \frac{g(t+h) - g(t)}{h}, \; D_-g(t) = \liminf_{h \uparrow 0} \frac{g(t+h) - g(t)}{h}, \; t \in [\alpha, \beta).$$

For details we refer the reader to [18].

Proposition 4.1. *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*

$$\left\{ \frac{g(t) - g(s)}{t - s} \,\Big|\, t, s \in [\alpha, \beta] \right\}.$$

Proof. See [15, Chapter 4, Theorem 1.2]. $\square$

An immediate consequence of the above Proposition is the following

Proposition 4.2. (Monotonicity result). *Let $g \in C([\alpha, \beta]; \mathbb{R})$ be such that*

$$D^+ g(t) \geq 0, \;\; \forall\, t \in [\alpha, \beta).$$

Then g is nondecreasing on $[\alpha, \beta]$.

The following Lemma is a special case of the Saks Theorem (see [65, Chapter 6, Theorem 7.3]). We give the proof in a special case using the Monotonicity result above.

Lemma 4.1. *Let $g \in C([0, +\infty); \mathbb{R})$. Suppose that there exists $\mu \in L^1([0, +\infty); \mathbb{R})$ such that*

$$D_- g(t) \geq \mu(t), \;\; \textit{for a.e. } t \in (0, +\infty) \tag{4.8}$$

and

$$D_- g(t) > -\infty \;\; \forall\, t \in (0, +\infty) \tag{4.9}$$

except at most for those of a countable set.
Then, for every $0 \leq \alpha \leq \beta < +\infty$,

$$g(\beta) - g(\alpha) \geq \int_\alpha^\beta \mu(t) dt. \tag{4.10}$$

Proof. We give the proof in the special case when μ is continuous and (4.8) holds for every $t \in (0, +\infty)$[1]. Since $D_- g(t) \geq \mu(t)$ for every $t \in (0, +\infty)$, we have

$$D_- \left[g(t) - \int_0^t \mu(s) ds \right] \geq 0, \;\; \forall\, t \in (0, +\infty).$$

[1]We will use this Lemma in the Verification Theorem, and these conditions will be satisfied

Thanks to Proposition 4.1 we have also

$$D^+\left[g(t)-\int_0^t \mu'(s)ds\right] \geq 0,\ \forall\, t\, \in\, [0,+\infty).$$

Therefore, due to Proposition 4.2, the function

$$t \mapsto\, g(t)-\int_0^t \mu(s)ds$$

is nondecreasing, getting the claim. □

Remark 4.2. *We give some remarks on Lemma 4.1.*

- *If μ is continuous and condition* (4.8) *holds for all $t > 0$, then* (4.9) *is verified and so the claim of Lemma 4.1 holds without assuming it.*
- *We cannot avoid to assume* (4.8)*: without it, then* (4.9) *is no longer true. One could substitute the assumption* (4.8) *with the following: there exists $\mu \in L^1([0,+\infty);\mathbb{R})$ such that, for some $h_0 > 0$, we have $\dfrac{g(t+h)-g(t)}{h} \geq \mu(t)$, for $h_0 < h \leq 0$, for a.e. $t > 0$ (see* [72, Lemma 2.3]*). However this assumption is more difficult to check in our case than the one of our Lemma 4.1.*

□

Theorem 4.1. Verification. *Let $\eta \in H_+$ and let $y^*(\cdot)$ be a solution of* (4.3) *such that $y^*(\cdot) > 0$. Let $u^*(\cdot)$ be the strategy defined by* (4.5)*. Then $u^*(\cdot)$ is admissible and optimal for the problem.*

Proof. The fact that $u^*(\cdot)$ is admissible is a direct consequence of the assumption $y^*(\cdot) > 0$ and of the definition of $u^*(\cdot)$.
Set $Y^*(\cdot) := Y(\cdot;\eta,u^*(\cdot))$ and let $s \geq 0$. Let $\varsigma_1(s) \in W^{1,2}_{r,0}$ be such that, taking into account (4.2), it holds true that

$$(V_b(Y^*(s)),\varsigma_1(s))\, \in\, D^+V(Y^*(s)).$$

Let

$$\varphi(\zeta) := V(Y^*(s))\, +\, \langle (V_b(Y^*(s)),\varsigma_1(s))\,,\zeta - Y^*(s)\rangle_H, \quad \zeta \in H,$$

so that

$$\varphi(Y^*(s)) = V(Y^*(s)), \quad \varphi(\zeta) \geq V(\zeta),\ \zeta \in H.$$

From Proposition 3.5 we know that $\varphi \in \mathcal{T}$, so, thanks to Lemma 3.5, we have

$$\begin{aligned}
\liminf_{h\uparrow 0} & \frac{e^{-\rho(s+h)}V(Y^*(s+h)) - e^{-\rho s}V(Y^*(s))}{h} \\
&\geq \liminf_{h\uparrow 0} \frac{e^{-\rho(s+h)}\varphi(Y^*(s+h)) - e^{-\rho s}V(Y^*(s))}{h} \\
&= e^{-\rho s}[\mathcal{L}^{u^*(s)}\varphi](Y^*(s)) \\
&= e^{-\rho s}\Big[-\rho V(Y^*(s)) + \langle Y(s), A^*\left(V_b(Y^*(s)), \varsigma_1(s)\right)\rangle_H + u^*(s)\langle V(Y^*(s)), b\rangle \Big].
\end{aligned} \tag{4.11}$$

Recalling the definition of $u^*(\cdot)$, due to (D.2) and by above equation, we calculate

$$\begin{aligned}
&\liminf_{h\uparrow 0} \frac{e^{-\rho(s+h)}V(Y^*(s+h)) - e^{-\rho s}V(Y^*(s))}{h} + e^{-\rho s}\left(h(u^*(s)) + g(Y^*(s))\right) \\
&\geq e^{-\rho s}\Big[-\rho V(Y^*(s)) + \langle Y(s), A^*\left(V_b(Y^*(s)), \varsigma_1(s)\right)\rangle_H + u^*(s)\langle V(Y^*(s)), b\rangle \Big] \\
&\ + e^{-\rho s}\left(h(u^*(s)) + g(Y^*(s))\right) \\
&\geq e^{-\rho s}\Big[-\rho V(Y^*(s)) + \langle Y(s), A^*\left(V_b(Y^*(s)), \varsigma_1(s)\right)\rangle_H + \mathcal{H}(V_b(Y^*(s))) + g(Y^*(s)) \Big].
\end{aligned} \tag{4.12}$$

So, due to the subsolution property of V we get

$$\begin{aligned}
\liminf_{h\uparrow 0} & \frac{e^{-\rho(s+h)}V(Y^*(s+h)) - e^{-\rho s}V(Y^*(s))}{h} \\
& + e^{-\rho s}\left(h(u^*(s) + g(Y^*(s)))\right) \geq 0.
\end{aligned} \tag{4.13}$$

We know that the function $s \mapsto e^{-\rho s}V(Y^*(s))$ and the function $s \mapsto e^{-\rho s}\left(h(u^*(s) + g(Y^*(s)))\right)$ are continuous; therefore we can apply Lemma 4.1 on $[0, N]$, $N > 0$, getting

$$e^{-\rho N}V(Y^*(N)) + \int_0^N e^{-\rho s}\left(h(u^*(s) + g(Y^*(s)))\right) ds \geq V(\eta).$$

Now we take the $\limsup$ for $N \to +\infty$. Recalling that the functions V, h, g are bounded from above, we get by Fatou's Lemma

$$J(\eta; u^*(\cdot)) = \int_0^{+\infty} e^{-\rho s}\left(h(u^*(s) + g(Y^*(s)))\right) ds \geq V(\eta),$$

and the claim is proved. □

Remark 4.3. *We give some remarks on Theorem 4.1.*

- *Observe that no continuity or measurability property of $\varsigma_1(s)$ with respect to s is needed in the proof of the theorem above.*

- *In the Theorem 4.1 we have given a sufficient condition of optimality. A natural question arising is whether 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.*
  For what concerns the finite-dimensional case, it is possible to find an answer to this problem in the so called BACKWARD DYNAMIC PROGRAMMING PRINCIPLE, (see [9, Chapter 3, Proposition 2.25]). But this topic is not standard in the infinite-dimension case unless the operator A is the generator of a strongly continuous group which is not our case, so, the investigation on this left for future research.

$\square$

Part II

Malliavin Calculus in the Control theory

4.1 Notations

In this second part of the work, we will use the classical notations of numbering definitions, theorems, formulas, etc...
We write a.s. for almost surely, a.e. for almost everywhere, and a.a for almost all.
$\mathbb{R}^n$ denotes the real euclidean n-dimensional space, with elements $x = (x_1, \ldots, x_n)$.
We write

$$\langle x, y \rangle = \sum_{i=1}^{n} x_i y_i.$$

$L^2([0,T])$ denotes the space of function f such that $\int_0^T |f(s)|^2 ds < \infty$. We write

$$\langle f, g \rangle_{L^2([0,T])} = \int_0^T f(s)\overline{g(s)} ds = \int_0^T f(s)g(s)ds.$$

The last equality holds if f and g are $\mathbb{R}$-valued.
Given a function $f : \mathbb{R}^m \times \mathbb{R}^\nu \longrightarrow \mathbb{R}^n$, we denote

$$\partial_k f^i_X(x,u) = \frac{\partial f^i}{\partial X_k}(x,u)$$
$$\partial_k f^i_u(x,u) = \frac{\partial f^i}{\partial u_k}(x,u)$$

for all $i = 1, ..., m$.
Let $\sigma^i_h : \mathbb{R}^m \times \mathbb{R}^\nu \longrightarrow \mathbb{R}^n$ $i = 1, ..., m$.
Let $\sigma^i_h(\cdot,\cdot)$, $i = 1, ..., m$, $h = 1, ..., d$ be the element of a matrix $(m \times d)$. We denote

$$\partial_k \sigma^i_{h,X} = \frac{\partial \sigma^i_h}{\partial X_k}$$
$$\partial_k \sigma^i_{h,u} = \frac{\partial \sigma^i_h}{\partial u_k}$$

for all $i = 1, ..., m$, $h = 1, ..., d$.
Let Y a metric space.
A function f is called polynomially growing if there exist constants K, m, ≥ 0 such that

$$|f(x)| \leq K(1 + |x|^m), \forall\, x \in Y.$$

For an open set $O \subset \mathbb{R}^n$, and a positive integer k,

$$\begin{aligned} C^k(O) &= \{\text{all k-times continuously differentiable functions on O}\} \\ C^k_b(O) &= \{f \in C^k(O) : \text{all partial derivatives of f of orders} \leq \text{k are} \\ &\quad \text{bounded}\} \\ C^k_p(O) &= \{f \in C^k(O) : \text{all partial derivatives of f of orders} \leq \text{k are} \\ &\quad \text{polynomially growing}\}. \end{aligned}$$

In this chapter we will give some basilar definitions about the well known stochastic control problems. Moreover we will recall some well-known results which we will use to prove our main aim.

4.2 Stochastic control problems

We will indicate with $(\Omega, \mathcal{F}, \mathbb{P})$ the probability space with filtration $\mathcal{F} = \{\mathcal{F}_t, t \geq 0\}$ satisfying the usual conditions, i.e.,

- all the set P-negligible belong to $\mathcal{F}_0$,
- $\mathcal{F}_t$ is right-continuous, i.e. $\mathcal{F}_{t^+} := \cap_{s>t}\mathcal{F}_s = \mathcal{F}_t$.

Let $W = \{W^1_t, W^2_t, \dots W^d_t, t \geq 0\}$ be a $(\mathcal{F}_t)$ d-dimensional standard brownian motion, defined on $(\Omega, \mathcal{F}, \mathbb{P})$.

Definition 4.1. *A processes $(X_t)_{t\geq 0}$ is said progressively measurable with respect to $\mathcal{F}_t$ if for all $t \geq 0$, the application*

$$(\omega, s) \mapsto X_s(\omega) \quad of \quad (\Omega \times [0, t], \mathcal{F} \otimes B([0, t]))$$

is measurable.

Let the control space U be a subset of $\mathbb{R}^n$. We denote by $\mathcal{U}_0$ the set of all progressively measurable processes $\nu = \{\nu_t, t \geq 0\}$ valued in U. The elements of $\mathcal{U}_0$ are called **control processes.**

We are interested to feedback controls, which will also be called **Markovian controls**, i.e., to processes $u \in \mathcal{U}$ which can written in the form $u_s = \tilde{u}(s, X_s)$ for some measurable map $\tilde{u}$ from $[0,T] \times \mathbb{R}^n$ into U. This means that the control u_s is chosen based on knowing not only time s but also the state X_s, where $s \in [0,T]$.

Let

$$b : (t, X_t, u_t) \in \mathbb{R}_+ \times \mathbb{R}^n \times U \longrightarrow b(t, X_t, u_t) \in \mathbb{R}^n$$

and

$\sigma : (t, X_t, u_t) \in \mathbb{R}_+ \times \mathbb{R}^n \times U \longrightarrow \sigma(t, X_t, u_t) \in \mathbb{R}^{n \times d}$ (set of matrix $n \times d$)

be two borelian functions such that $b(t, X_t, u_t) = (b^1(t, X_t, u_t), \dots b^n(t, X_t, u_t))$ and $\sigma(t, X_t, u_t) = (\sigma_{ij}(t, X_t, u_t))_{1 \leq i \leq n, 1 \leq j \leq d}$.
The vector function $b = (b_t^1, .., b_t^n)$ is called the *local drift coefficient* and the matrix-valued function $\sigma(t, X_t, u_t)$ the *local covariance matrix,* or *diffusion coefficient.*
For each control process $\nu \in U$ we consider the state stochastic differential equation

$$dX_t = b(t, X_t, \nu_t)dt + \sum_{j=1}^{d} \sigma_j(t, X_t, \nu_t)\, dW_t^j. \tag{4.14}$$

The differential form of the above equation is

$$X_t = Z + \int_0^t b\,(s, X_s, \nu_s), ds + \sum_{j=1}^{d} \int_0^t \sigma_j\,(s, X_s, \nu_s)\, dW_s^j$$

where $Z = (Z^1, Z^2, ..., Z^n)$ is a $\mathcal{F}_0$ random variable $\mathbb{R}^n$-valued.
If the (4.14) has a unique solution X, for a given initial data, then the process X is called the **controlled process**, as his dynamics is driven by the action of the control process ν.
Let $T > 0$ be some given time horizon. We shall denote by $\mathcal{U}$ the subset of all control processes $\nu \in \mathcal{U}_0$ which satisfy the additional requirement:

$$\mathbb{E} \int_0^T \left(|b(t, X, \nu_t)| + |\sigma(t, X, \nu_t)|^2 \right) dt < \infty \quad \text{for} \quad X \in \mathbb{R}^n. \tag{4.15}$$

This condition guarantees the existence of a controlled process for each given initial condition and control.
In the following we will state a theorem (Theorem 4.4)that guarantees the existence and the uniqueness of the solution of (4.14).
In order to proceed, we will give some definition about optimal control theory. The reason is to understand what must changed.
Let

$$f,\, k : [0, T] \times \mathbb{R}^n \times U \longrightarrow \mathbb{R} \quad \text{and} \quad g : \mathbb{R}^n \longrightarrow \mathbb{R}$$

be given continuous functions, f is called **running cost function**, g **terminal cost function** and k is a **discount factor**.
We assume that $\|k^-\|_\infty < \infty$ and f and g satisfy the quadratic growth condition:

$$|f(t, x, u)| + |g(x)| \leq C(1 + |x|^2) \quad \text{for some constant C independent of}\,(t, u).$$

We define the **cost function** J on $[0,T] \times \mathbb{R}^n \times \mathcal{U}$ by:

$$J(t,x;\nu) := \mathbb{E}_{t,x}\left[\int_t^T \beta(t,s) f(s,X_s,\nu_s)\,ds + \beta(t,T)g(X_T)\right]$$

with

$$\beta(t,s) := e^{-\int_t^s k(r,X_r,\nu_r)dr}.$$

The problem is to choose $\nu(\cdot)$ to minimize J.
Here $\mathbb{E}_{t,x}$ is the expectation operator conditional on $X_t = x$, and X_t is the solution of (4.14) with control ν and initial condition $X_t = x$.
The quadratic growth condition on f an g together with the bound on k^- ensure that $J(t,x;\nu)$ is well defined for all admissible control $\nu \in \mathcal{U}$ (Theorem 4.4).
We have to study the minimization problem

$$V(t,x) := \inf_{\nu\in\mathcal{U}} J(t,x;\nu) \quad \text{for} \quad (t,x) \in [0,T) \times \mathbb{R}^n.$$

This is called **the stochastic control problem** and V is called the value function.

4.3 Malliavin calculus

We first provide several useful results concerning Malliavin Calculus. In the following, the same results will be slightly modified at some points to obtain our main result.

4.3.1 The derivative operator

Suppose that H is a real separable Hilbert space with scalar product denoted by $\langle \cdot, \cdot \rangle_H$. The norm of an element $h \in H$ will be denoted by $\|h\|_H$.
Let $W = \{W(h),\, h \in H\}$ denote an isonormal Gaussian process associated with the Hilbert space H, by definition E.2.
We assume that W is defined on a complete probability space $(\Omega, \mathcal{F}, \mathbb{P})$, and $\mathcal{F}$ is generated by W.
We want to introduce the derivative DF of a square integrable random variable $F : \Omega \longrightarrow \mathbb{R}$. This means that we want to differentiate F with respect to the chance parameter $\omega \in \Omega$.
Denote by $C_p^\infty(\mathbb{R}^n)$ the set of all infinitely continuously differentiable functions $f : \mathbb{R}^n \longrightarrow \mathbb{R}$ such that f and all of its partial derivatives have polynomial growth.

Let $\mathcal{S}$ denote the class of smooth random variable such that $F \in \mathcal{S}$ has the form

$$F = f(W(h_1), W(h_2), ..., W(h_n)), \tag{4.16}$$

where f belongs to $C_p^\infty(\mathbb{R}^n)$, $h_1, h_2, ..., h_n$ are in H, and $n \geq 1$.

Definition 4.2. *The derivative of a smooth random variable F of the form* (4.16) *is the H-valued random variable given by*

$$DF = \sum_{i=1}^{n} \partial_i f(W(h_1), W(h_2), ..., W(h_n))h_i. \tag{4.17}$$

The operator D is closable from $L^p(\Omega)$ to $L^p(\Omega; H)$ for any $p \geq 1$.
For any $p \geq 1$ we will denote the domain of D in $L^p(\Omega)$ by $\mathbb{D}^{1,p}$, meaning that $\mathbb{D}^{1,p}$ is the closure of the class of smooth random variables $\mathcal{S}$ with respect to the norm

$$\| F \|_{1,p} = [\mathbb{E}(|F|^p) + \mathbb{E}(\| DF \|_H^p)]^{\frac{1}{p}}.$$

We can define the iteration of the operator D in such a way that for a smooth random variable F, the iterated derivative $D^k F$ is a random variable with values in $H^{\otimes k}$. Then for every $p \geq 1$ and any natural number $k \geq 1$ we introduce the seminorm on $\mathcal{S}$ defined by

$$\| F \|_{k,p} = \left[\mathbb{E}(|F|^p) + \sum_{j=1}^{k} \mathbb{E}(\| D^j F \|_{H^{\otimes j}}^p) \right]^{\frac{1}{p}}. \tag{4.18}$$

We will denote by $\mathbb{D}^{k,p}$ the completion of the family of smooth random variables $\mathcal{S}$ with respect to the norm $\| \cdot \|_{k,p}$.
In the following we state some propositions and Lemma that we will use to prove our results. For the proofs see in order [58, Chapter 1, Proposition 1.2.3, Lemma 1.2.3, Proposition 1.2.4].

Proposition 4.3. ***Chain Rule.*** *Let $\varphi : \mathbb{R}^m \longrightarrow \mathbb{R}$ be a continuously differentiable function with bounded partial derivatives, and fix $p \geq 1$. Suppose that $F = (F^1, F^2, ..., F^m)$ is a random vector whose components belong to the space $\mathbb{D}^{1,p}$. Then $\varphi(F) \in \mathbb{D}^{1,p}$, and*

$$D(\varphi(F)) = \sum_{i=1}^{m} \partial_i \varphi(F) DF^i. \tag{4.19}$$

Lemma 4.2. *Let $\{F_n, n \geq 1\}$ be a sequence of random variables in $\mathbb{D}^{1,2}$ that converges to F in $L^2(\Omega)$ and such that*

$$\sup_n \mathbb{E}\left(\|DF_n\|_H^2 \right) < +\infty.$$

Then F belongs to $\mathbb{D}^{1,2}$, and the sequence of derivatives $\{ DF_n, n \geq 1 \}$ converges to DF in the weak topology of $L^2(\Omega; H)$.

Proposition 4.4. *Let $\varphi : \mathbb{R}^m \longrightarrow \mathbb{R}$ be a function such that*

$$|\varphi(x) - \varphi(y)| \leq K|x - y|$$

for any x and $y \in \mathbb{R}^m$. Suppose that $F = (F^1, F^2, ..., F^m)$ is a random vector whose components belong to the space $\mathbb{D}^{1,2}$. Then $\varphi(F) \in \mathbb{D}^{1,2}$, and there exists a random vector $G = (G_1, G_2, ..., G_m)$ bounded by K such that

$$D(\varphi(F)) = \sum_{i=1}^{m} G_i \, DF^i. \tag{4.20}$$

For the proof see [58, pag 29].

Remark 4.4. *Proposition 4.4 and Lemma 4.2 still hold if we replace $\mathbb{D}^{1,2}$ by $\mathbb{D}^{1,p}$ for any $p > 1$. This follows from [58, Lemma 1.53 pag 79] and the duality relationship between D and δ.*

Now let $H = L^2([0,T], \mathcal{B}, \mu)$, where μ is a σ-finite atomless measure on a measurable space $([0,T], \mathcal{B})$.
The derivative of a random variable $F \in \mathbb{D}^{1,2}$ will be a stochastic process denoted by $\{ D_t F, t \in [0,T] \}$ due to the identification between the Hilbert spaces $L^2(\Omega; H)$ and $L^2([0,T] \times \Omega)$. Notice that $D_t F$ is defined almost everywhere with respect to the measure $\mu \times \mathbb{P}$. More generally, if $k \geq 2$, the derivative $D^k F = \{ D^k_{t_1,t_2,...,t_k} F, t_i \in [0,T] \}$, is a measurable function on the product space $[0,T]^k \times \Omega$, which is defined everywhere with respect to the measure $\mu^k \times \mathbb{P}$.

4.3.2 The divergence operator

The divergence operator is defined as the adjoint of the derivative operator. Let the underlying space H be of the form $L^2([0,T], \mathcal{B}, \mu.)$

Definition 4.3. *We denote by δ the adjoint of the operator D. That is, δ is an unbounded operator on $L^2(\Omega; H)$ with values in $L^2(\Omega)$ such that:*
(i) the domain of δ, denoted by Dom δ, is the set of H-valued square integrable random variables $u \in L^2(\Omega; H)$ such that

$$|\mathbb{E}(\langle DF, u \rangle_H)| \leq c\| F \|_2, \tag{4.21}$$

for all $F \in \mathbb{D}^{1,2}$, where c is some constant depending on u.
(ii) If u belongs to Dom δ, then $\delta(u)$ is the element of $L^2(\Omega)$ characterized by

$$\mathbb{E}(F\delta(u)) = \mathbb{E}(\langle DF, u \rangle_H) \tag{4.22}$$

for any $F \in \mathbb{D}^{1,2}$.

The operator δ is called the divergence operator and is closed as the adjoint of an unbounded and densely defined operator.
We denote by $\mathcal{S}_H$ the class of smooth elementary elements of the form

$$u = \sum_{j=1}^{n} F_j h_j, \tag{4.23}$$

where the F_j are smooth random variables, and h_j are elements of H.
The space $\mathbb{D}^{1,2}(H)$ is included in the domain of δ. In fact, if $u \in \mathbb{D}^{1,2}(H)$, there exists a sequence $u^n \in \mathcal{S}_H$ such that u^n converges to u in $L^2(\Omega)$ and Du^n converges to Du in $L^2(\Omega; H \otimes H)$. Therefore, $\delta(u^n)$ converges in $L^2(\Omega)$ and its limit is $\delta(u)$.

4.3.3 The Skorohod integral

Let $H = L^2([0,T], \mathcal{B}, \mu)$ where μ is the Lebesgue measure.
In this case the elements of Dom $\delta \subset L^2([0,T] \times \Omega)$ are square integrable processes, and the divergence $\delta(u)$ is called the *Skorohod stochastic integral* of the process u. We will use the following notation:

$$\delta(u) = \int_0^T u_t \, dW_t.$$

Definition 4.4. *The space $\mathbb{D}^{1,2}(L^2([0,T]))$, denoted by $\mathbb{L}^{1,2}$, coincides with the class of processes $u \in L^2([0,T] \times \Omega)$ such that $u_t \in \mathbb{D}^{1,2}$ for almost all t, and there exists a measurable version of the two parameter process $D_s u_t$ verifying $\mathbb{E} \int_0^T \int_0^T (D_s u_t)^2 \, ds dt < \infty$.*

The space $\mathbb{D}^{1,2}(L^2([0,T]))$ is included in Dom δ.
$\mathbb{L}^{1,2}$ is a Hilbert space with the norm

$$\| u \|^2_{1,2,L^2([0,T])} = \| u \|^2_{L^2([0,T] \times \Omega)} + \| Du \|^2_{L^2([0,T]^2 \times \Omega)}.$$

Note that $\mathbb{L}^{1,2}$ is isomorphic to $L^2([0,T]; \mathbb{D}^{1,2})$.

Definition 4.5. *Let* $\mathbb{L}^{1,p}$ *be the class of the processes* $u \in L^p([0,T] \times \Omega)$ *such that* $u_t \in \mathbb{D}^{1,p}$ *for almost all* t*, and there exists a measurable version of the two-parameter process* $D_s u_t$ *such that* $\| D_s u_t \|^p_{L^2([0,T]^2)} = \left(\int_0^T \int_0^T |D_s u_t|^2 \, ds dt \right)^{\frac{p}{2}}$ *has finite* $\mathbb{P}$*-expectation.*
In $\mathbb{L}^{1,p}$ *we define the norm*

$$\|u\|_{1,p} = \left(\|u\|^p_{L^p(\Omega \times [0,T])} + \mathbb{E} \left[\|Du\|^p_{L^2([0,T]^2)} \right] \right)^{\frac{1}{p}}. \tag{4.24}$$

If u and v are two processes in the space $\mathbb{L}^{1,2}$, then we have

$$\mathbb{E} \left(\delta(u) \delta(v) \right) = \int_0^T \mathbb{E} \left(u_t v_t \right) dt + \int_0^T \int_0^T \mathbb{E} \left(D_s u_t D_t v_s \right) \mu(ds) \mu(dt). \tag{4.25}$$

Suppose that $T = [0, \infty)$. Then, if both processes are adapted to the filtration generated by the Brownian motion, we have that $D_s u_t = 0$ for almost (s,t) such that $s > t$, since $\mathcal{F}_t = \mathcal{F}_{[0,t]}$. Consequently, the second summand in 4.25 is equal to zero, and we recover the usual isometry property of the Itô integral.

Proposition 4.5. *Suppose that* $u \in \mathbb{L}^{1,2}$*. Assume that for almost all* t *the process* $\{ D_t u_s, \, s \in [0,T] \}$ *is Skorohod integrable, and there is a version of the process* $\{ \int_0^T D_t u_s \, dW_s, \, t \in [0,T] \}$ *which is in* $L^2([0,T] \times \Omega)$*. Then* $\delta(u) \in \mathbb{D}^{1,2}$*, and we have*

$$D_t(\delta(u)) = u_t + \int_0^T D_t u_s \, dW_s. \tag{4.26}$$

Proof. For the proof see [58, pag 43]. □

Lemma 4.3. *Let* $W = \{ W(t), \, t \in [0,1] \}$ *be a one-dimensional Brownian motion. Consider a square integrable adapted process* $u = \{ u_t, \, t \in [0,1] \}$*, and set* $X_t = \int_0^t u_s \, dW_s$*. Then the process* u *belongs to the space* $\mathbb{L}^{1,2}$ *if and only if* X_1 *belongs to* $\mathbb{D}^{1,2}$*. In this case the process* X *belongs to* $\mathbb{L}^{1,2}$*, and we have*

$$\int_0^t \mathbb{E} \left(|D_s X_t|^2 \right) ds = \int_0^t \mathbb{E} \left(u_s^2 \right) ds + \int_0^t \int_0^s \mathbb{E} \left(|D_r u_s|^2 \right) dr ds, \tag{4.27}$$

for all $t \in [0,1]$.

Proof. For the proof see [58, pag 51]. □

4.4 Regularity of probability laws

We need to study the regularity of the probability law of a random vector defined on a probability space.
In particular, the random vector will be the solution of a stochastic differential equation.
To begin with, we define a new space.
Let V be a real Hilbert space. We consider the family $\mathcal{S}_V$ of V-valued smooth random variables of the form

$$F = \sum_{j=1}^{n} F_j v_j, \qquad v_j \in V, \qquad F_j \in \mathcal{S}.$$

Define $D^k F = \sum_{j=1}^{n} D^k F_j \otimes v_j$, $k \geq 1$. Then D^k is a closable operator from $\mathcal{S}_V \subset L^p(\Omega; V)$ into $L^p(\Omega; H^{\otimes k} \otimes V)$ for any $p \geq 1$. For any integer $k \geq 1$ and any real number $p \geq 1$ we can define the seminorm on $\mathcal{S}_V$

$$\| F \|_{k,p,V} = \left[\mathbb{E}\left(\| F \|_V^p\right) + \sum_{j=1}^{k} \mathbb{E}\left(\| D^j F \|_{H^{\otimes j} \otimes V}^p\right) \right]^{\frac{1}{p}}. \tag{4.28}$$

We define the space $\mathbb{D}^{k,p}(V)$ as the completion of $\mathcal{S}_V$ with respect to the norm $\| \cdot \|_{k,p,V}$.
Consider the intersection

$$\mathbb{D}^{\infty}(V) = \cap_{p \geq 1} \cap_{k \geq 1} \mathbb{D}^{k,p}(V).$$

Then $\mathbb{D}^{\infty}(V)$ is a complete, countably normed, metric space. We will write $\mathbb{D}^{\infty}(\mathbb{R}) = \mathbb{D}^{\infty}$. If F and G are random variables in $\mathbb{D}^{\infty}$, then the scalar product $\langle DF, DG \rangle_H$ is also in $\mathbb{D}^{\infty}$.

Proposition 4.6. *Suppose that $F = (F^1, F^2, ..., F^m)$ is a random vector whose components belong to $\mathbb{D}^{\infty}$. Let $\varphi \in C_p^{\infty}(\mathbb{R}^m)$. Then $\varphi(F) \in \mathbb{D}^{\infty}$, and we have*

$$D(\varphi(F)) = \sum_{i=1}^{m} \partial_i \varphi(F)\, DF^i.$$

Proof. For the proof see [58, pag]. □

Proposition 4.7. *Let F be a random variable in $\mathbb{D}^{k,\alpha}$ with $\alpha > 1$. Suppose that $D^i F$ belongs to $L^p(\Omega; H^{\otimes i})$ for $i = 0, 1, ..., k$ and for some $p > \alpha$. Then $F \in \mathbb{D}^{k,p}$, and there exists a sequence $G_n \in \mathcal{P}$ that converges to F in the norm $\| \cdot \|_{k,p}$.*

Proof. For the proof see [58, pag 76]. $\square$

Remark 4.5. *$\mathcal{P}$ is the class of the random variables of the form $p(W(h_1), W(h_2), ..., W(h_n))$ where $h_i \in H$ and p is a polynomial. Moreover $\mathcal{P}$ is dense in $L^r(\Omega)$ for all $r \geq 1$.*
The class $\mathcal{P}$ is also dense in $\mathbb{D}^{k,p}$ $\forall\, p \geq 1$ e $k \geq 1$. $\square$

Proposition 4.8. *Let u be an element of $\mathbb{D}^{1,p}(H)$, $p > 1$. Then we have*

$$\| \delta(u) \|_p \leq c_p \left(\| \mathbb{E}\,(u) \|_H + \| Du \|_{L^p(\Omega; H \otimes H)} \right).$$

Proof. For the proof [58, pag 80]. $\square$

We need some general criteria to study the probability law, the regularity of the density of a vector. We will apply these criteria to the solutions of stochastic differential equation.
We need some criteria to the solutions of stochastic differential equations.
Let $W = \{ W(h),\, h \in H \}$ be a process associated to a Hilbert space H and defined on complete probability space $(\Omega, \mathcal{F}, \mathbb{P})$. Assume also that $\mathcal{F}$ is generated by W.

Proposition 4.9. *Let F a random variable in the space $\mathbb{D}^{1,2}$. Suppose that $\frac{DF}{\| DF \|_H^2}$ belongs to the domain of the operator δ in $L^2(\Omega)$. Then the law of F has a continuous and bounded density given by*

$$p(x) = \mathbb{E}\left[\mathbf{1}_{\{F > x\}}\, \delta \left(\frac{DF}{\| DF \|_H^2} \right) \right]. \tag{4.29}$$

Proof. For the proof [58, pag 86]. $\square$

Suppose that $F = (F^1, F^2, ..., F^m)$ is a random vector whose components belong to the space $\mathbb{D}^{1,1}_{\text{loc}}$. We associate to F the following random symmetric nonnegative definite matrix:

$$\gamma_F = (\langle DF^i, DF^j \rangle_H)_{1 \leq i, j \leq m}.$$

This is the so called *Malliavin matrix* of a random vector F. Thus, we state the following principal theorem which has the proof on [58, pag 92]:

Theorem 4.2. *Let $F = (F^1, F^2, ..., F^m)$ be a random vector verifying the following conditions:*

(i) $F^i \in \mathbb{D}^{2,p}_{loc}$ *for all* $i = 1, 2, ..., m$, *for some* $p > 1$.

(ii) *The matrix* γ_F *is invertible a.s.*

Then the law of F *is absolutely continuous with respect to the Lebesgue measure on* $\mathbb{R}^m$.

4.4.1 Stochastic differential equation

We have stated the criteria to insure that a random vector has a law absolutely continuous.
Now we want to find a similar criteria for a stochastic process that is the solution of a stochastic differential equation, SDE.
In particular our SDE has the coefficient that also depend on a control process.
We see in details.

Suppose that $(\Omega, \mathcal{F}, \mathbb{P})$ is the canonical probability space associated with a d-dimensional Brownian motion $W = \{ W_t^i,\, t \in [0, T],\, 1 \leq i \leq d\}$ on a finite interval $[0, T]$.
This means that $\Omega = C_0([0, T]; \mathbb{R}^d)$, $\mathbb{P}$ is the Wiener measure, $\mathcal{F}$ is the completion of the Borel σ-field of Ω with respect to $\mathbb{P}$ and $\mathcal{F}_t$ is the filtration on $(\Omega, \mathcal{F})$ generated by W satisfying the usual conditions.
The underlying Hilbert space here is $H = L^2([0, T]; \mathbb{R}^d)$.
Let $U \subset \mathbb{R}^m$ be the set of all possible processes, and $\mathcal{U}_0$ the set all progressively measurable processes $u = \{ u_t,\, t \geq 0 \}$ U-valued.
We give the following Hypotheses on the coefficient and on the control processes.

Hypothesis 4.3.

(A1) *Let* $b : \mathbb{R}_+ \times \mathbb{R}^m \times U \longrightarrow \mathbb{R}^m$, $\sigma : \mathbb{R}_+ \times \mathbb{R}^m \times U \longrightarrow \mathbb{R}^{m \times d}$ *be measurable functions satisfying the following globally Lipschitz and boundness conditions:*
$| \sigma(t, x, u) - \sigma(t, y, v) | + | b(t, x, u) - b(t, y, v) |$
$\leq K | x - y | + H | u - v|$ *for any* $x, y \in \mathbb{R}^m$, $u, v \in \mathcal{U}$, $t \in [0, T]$;

(A2) $\exists\, \bar{u}$ *such that* $\bar{u} \in U$ *and*
$t \longrightarrow \sigma(t, 0, \bar{u})$ *and* $t \longrightarrow b(t, 0, \bar{u})$ *have linear growth (i.e.,* $|\sigma(t, 0, \bar{u})| + |b(t, 0, \bar{u})| \leq K(1 + |t|)$*)*

(Au) *the process* $u = \{ u_t,\, t \in [0,T] \} \in \mathcal{U} \subset \mathcal{U}_0$ *is such that* u_t *is a control process progressively measurable with respect to* $(\mathcal{F}_t)_{t\geq 0}$ *and admissible, i.e.,*

$$\mathbb{E}\left(\int_0^t |\, u_s \,|^p \, ds \right) < +\infty \qquad for\, p = 1,\, 2,\, \ldots\, . \tag{4.30}$$

where $\mathcal{U} \subset \mathcal{U}_0$ *is the set of the control processes* u_t *such that satisfies*

$$\mathbb{E}\left(\int_0^T |\, b(t,x,u_t)\,| + |\,\sigma(t,x,u_t)\,|^2 \, dt \right) < +\infty \quad for\, x \in \mathbb{R}^m. \tag{4.31}$$

□

Remark 4.6. *The condition* (4.31) *guarantees the existence of a controlled process for each given initial condition and control.* □

We denote by $X = \{ X_t,\, t \in [0,T]\}$ the solution of the following stochastic differential equation:

$$X_t = x_0 + \int_0^t b(s, X_s, u_s)\, ds + \int_0^t \sigma(s, X_s, u_s)\, dW_s, \tag{4.32}$$

written componentwise as

$$X_t^i = x_0^i + \int_0^t b_i(s, X_s, u_s)\, ds + \sum_{j=1}^{d} \int_0^t \sigma_{ij}(s, X_s, u_s)\, dW_s^j, \tag{4.33}$$

for $1 \leq i \leq m$, where $x_0 \in \mathbb{R}^m$ is a random variable $\mathcal{F}_0$-adapted.
So we state the following theorem:

Theorem 4.4. *Let* σ *and* b *be Lipshitz in the* x *variable uniformly in* $(t,u) \in [0,T] \times U$*, and* $\nu \in \mathcal{U}$ *and have linear growth. Then, for all* $\mathcal{F}_0$ *random variable* $\xi \in L^2(\Omega)$*, there exists a unique* $\mathcal{F}_t$*-adapted process* $X = \{X_t,\, t \in [0,T]\}$ *satisfying* (4.32) *together with the initial condition* $x_0 = \xi$*. Moreover,*

$$\mathbb{E}\left(\sup_{0\leq t\leq T} |X_t|^p \right) < +\infty.$$

Proof. For the proof to see [58, chapter 2, section 2]. □

Definition 4.6. *A strong solution of the stochastic differential equation* (4.33)*, on the given probability space* $(\Omega, \mathcal{F}, \mathbb{P})$ *and with respect to the fixed Brownian motion* W *and condition initial* ξ *where* ξ *is a random vector taking values in* $\mathbb{R}^m$*, is a process* $X = \{ X_t, \, t \in [0,T] \}$ *with continuous sample path and with the following properties:*

(i) *X is adapted to the filtration $\{\mathcal{F}_t\}$,*

(ii) $\mathbb{P}[X_0 = \xi] = 1$,

(iii) $\mathbb{P}\left[\int_0^t \left\{ |b_i(s, X_s, u_s)| + |\sigma_{ij}^2(s, X_s, u_s)| \right\} ds < +\infty \right] = 1$ *holds for every* $1 \le i \le m$, $1 \le j \le d$, *and* $t \in [0,T]$.

Definition 4.7. *Let* $b(t, X_t, u_t)$ *and* $\sigma(t, X_t, u_t)$ *be given. Suppose that, whenever* W *is an d-dimensional Brownian motion on some* $(\Omega, \mathcal{F}, \mathbb{P})$*,* ξ *is an independent, m-dimensional random vector,* $(\mathcal{F}_t)_{t \ge 0}$ *is the filtration on* $(\Omega, \mathcal{F})$ *generated by* W *satisfying the usual conditions, and* X*,* $\tilde{X}$ *are two strong solutions of* (4.32) *relative to* W *with initial condition* ξ*, then* $\mathbb{P}\left[X_t = \tilde{X}_t; \, t \in [0,T] \right] = 1$*. Under these conditions, we say that strong uniqueness holds for the pair* (b, σ)*.*

We have seen (Theorem 4.4) the conditions which assure the strong uniqueness of the solution to (4.32), such that X_t belongs to the space $\mathbb{D}^{1,p}$ for all $p \ge 2$.
Now, we want to find the criteria to assure that the process X_t has a probability law.
The idea is to use Theorem 4.2, and to prove that the conditions of the Theorem are satisfied so that we can guarantee the absolute continuity of the law of X_t, where $X_t = (X_t^1, X_t^2, ..., X_t^m)$.
First of all we have to calculate the stochastic derivative of the process X_t so we need to recall the following useful well-known theorem.
We relegate the proof to [58, pag 119].

Theorem 4.5. *Let* $X = \{ X_t, \, t \in [0,T] \}$ *be the solution of the*

$$X_t = x_0 + \sum_{j=1}^{d} \int_0^t \sigma_j(s, X_s) \, dW_s^j + \int_0^t b(s, X_s) \, ds \tag{4.34}$$

where the coefficients are supposed to be globally Lipschitz functions with linear growth and $x_0 \in \mathbb{R}^m$ *is the initial value of the process* X*. Then* X_t^i

belongs to $\mathbb{D}^{1,\infty}$ *for any* $t \in [0,T]$ *and* $i = 1, 2, ..., m$. *Moreover,*

$$\sup_{0 \le r \le t} \mathbb{E}\Big(\sup_{r \le s \le T} |\, D_r^j X_s^i \,|^p \Big) < +\infty, \tag{4.35}$$

and the derivative $D_r^j X_t$ *satisfies the following linear equation:*

$$\begin{aligned} D^j X_t = \sigma_j(r, X_r) + \int_r^t \bar{\sigma}_{k,\alpha}(s) D_r^j X_s^k \, dW_s^\alpha \\ + \int_r^t \bar{b}_k(s) D_r^j X_s^k \, ds \end{aligned} \tag{4.36}$$

for $r \le t$ *a.e., and*

$$D_r^j X_t = 0$$

for $r > t$ *a.e., where* $\bar{\sigma}_{k,\alpha}(s)$ *and* $\bar{b}_k(s)$ *are uniformly bounded and adapted* m*-dimensional processes.*

Chapter 5

Malliavin calculus with control processes

The coefficients of (4.34) do not depend on the control processes, thus we have to modify Theorem 4.5 to suit our needs.

5.1 Malliavin derivative of the solution of SDE

In the following we assert that the set U is the closure of an open set of $\mathbb{R}^m$. Moreover, we recall that we consider the control processes satisfying Hypothesis 4.3-*(Au)*.
We finally can state the following

Theorem 5.1. *Let* $b : [0,T] \times \mathbb{R}^m \times U \longrightarrow \mathbb{R}^m$,
$\sigma : [0,T] \times \mathbb{R}^m \times U \longrightarrow \mathbb{R}^{m\times d}$ *be measurable functions, satisfying Hypotheses 4.3 of the subsection* (4.4.1).
Let $u \in \mathcal{U}$ *such that* $u \in \mathbb{L}^{1,\infty}$ *and* u_t *satisfies Hypotheses 4.3-(Au).*
Let $X = \{ X_t,\, t \in [0,T] \}$ *be the solution of the*

$$X_t = x_0 + \int_0^t b(s, X_s, u_s)\, ds + \sum_{j=1}^{d} \int_0^t \sigma_j(s, X_s, u_s)\, dW_s^j \tag{5.1}$$

and $x_0 \in \mathbb{R}^m$ *is the initial value of the process* X.
Then X_t^i *belongs to* $\mathbb{D}^{1,\infty}$ *for any* $t \in [0,T]$ *and* $i = 1, ..., m$.
Moreover,

$$\sup_{0\le s\le t} \mathbb{E}\left[\left(\int_0^s |D_r^j X_s^i|^2 dr \right)^{\frac{p}{2}} \right] < +\infty \tag{5.2}$$

for $j = 1, ..., d$, *and the derivative* $D_r^j X_t^i$ *satisfies the following linear equation:*

$$D_r^j X_t^i = \sigma_j^i(r, X_r, u_r) + \int_r^t \bar{\sigma}_{k,\alpha,X}(s) D_r^j X_s^k \, dW_s^\alpha + \int_r^t \bar{\sigma}_{k,\alpha,u}(s) D_r^j u_s^k \, dW_s^\alpha + \int_r^t \bar{b}_{k,X}(s) D_r^j X_s^k \, ds + \int_r^t \bar{b}_{k,u}(s) D_r^j u_s^k \, ds \tag{5.3}$$

for $r \leq t$ *a.e., and*

$$D_r^j X_t = 0$$

for $r > t$ *a.e., where* $\bar{\sigma}_{k,\alpha,X}(s)$, $\bar{\sigma}_{k,\alpha,u}(s)$, *and* $\bar{b}_{k,X}(s)$, $\bar{b}_{k,u}(s)$ *are uniformly bounded and* $(\mathcal{F}_t)_{t\geq 0}$*-adapted m-dimensional processes.*

Remark 5.1. *If the coefficients of the equation* (5.1) *are continuously differentiable, then we can write*

$$\bar{\sigma}_{k,X}^i(s) = (\partial_k \sigma_{l,X}^i)(s, X_s, u_s),$$
$$\bar{\sigma}_{k,u}^i(s) = (\partial_k \sigma_{l,u}^i)(s, X_s, u_s),$$
$$\bar{b}_{k,X}(s) = (\partial_k b_X^i)(s, X_s, u_s)$$

and

$$\bar{b}_{k,u}(s) = (\partial_k b_u^i)(s, X_s, u_s)$$

□

Remark 5.2. *The* (5.2) *implies that* X_t^i *belongs to* $\mathbb{L}^{1,p}$ *for any* $p \geq 1$, $t \in [0,T]$ *and* $i = 1, ..., m$. *This follows directly from Jensen's inequality, that is*

$$\mathbb{E}\left[\left(\int_0^T \int_0^s (D_r X_s)^2 ds dr\right)^{\frac{p}{2}}\right] \leq \frac{c}{T} \int_0^T \mathbb{E}\left[\left(\int_0^s (D_r X_s)^2 dr\right)^{\frac{p}{2}}\right] ds < +\infty$$

by (5.2). □

For the sake to simplify the proof of the theorem, we first will prove it in the 1-dimensional case, so, in order, we rewrite the environment of the problem.
Let $H = L^2([0,T];\mathbb{R})$ and $(W_t)_{t\geq 0}$ be a $(\mathcal{F}_t)_{t\geq 0}$ 1-dimensional standard Brownian motion.
We rewrite new hypotheses on the coefficients.

Hypothesis 5.2.

(A1') *The diffusion coefficient is defined as*

$$\sigma \,:\, [0,T] \times \mathbb{R} \times U \longrightarrow \mathbb{R},$$

and is such that $|\,\sigma(t,\,y,\,u) - \sigma(t,\,y',\,u')\,| \,\leq\, K\,(|\,y' - y\,| + |\,u' - u\,|)$ *for any* $t \in [0,T]$, $y,\, y' \in \mathbb{R}$, $u,\, u' \in \mathcal{U}$*;*

(A2') *the drift coefficient is defined as*

$$b \,:\, [0,T] \times \mathbb{R} \times U \longrightarrow \mathbb{R}$$

and is such that $|\,b(t,\,y,\,u) - b(t,\,y',\,u')\,| \,\leq\, K\,(|\,y' - y\,| + |\,u' - u\,|)$ *for any* $t \in [0,T]$, $y,\, y' \in \mathbb{R}$, $u,\, u' \in \mathcal{U}$*;*

(A3') $\exists\, \bar{u}$ *such that* $\bar{u} \in U$ *and* $t \longrightarrow \sigma(t,\,0,\,\bar{u})$ *and* $t \longrightarrow b(t,\,0,\,\bar{u})$ *have linear growth (i.e.,* $|\sigma(t,\,0,\,\bar{u})| + |b(t,\,0,\,\bar{u})| \,\leq\, K(1 + |t|)$*).* □

for a positive constant K.

Proof of the Theorem 5.1. Consider the Picard approximation given by

$$\begin{aligned} X_t^0 &= x_0, \\ X_t^{n+1} &= x_0 + \int_0^t \sigma(s, X_s^n, u_s)\, dW_s + \int_0^t b(s, X_s^n, u_s)\, ds \end{aligned} \tag{5.4}$$

if $n \,\geq\, 0$, n integer and $t \,\in\, [0,T]$. For details see [21, chapter 1] We will prove the following property by induction on n:
$(*)$ $X_t^n \,\in\, \mathbb{D}^{1,\infty}$ for $n \geq 0$ and $t \,\in\, [0,T]$; furthermore, for all $p > 1$ we have

$$f_n(t) := \sup_{0 \leq s \leq t} \mathbb{E}\left[\left(\int_0^s |D_r X_s^n|^2 dr \right)^{\frac{p}{2}} \right] < +\infty \tag{5.5}$$

and

$$f_{n+1}(t) \leq c_1 + c_2 \int_0^t f_n(s)\, ds, \tag{5.6}$$

for some positive constants c_1 and c_2.

Clearly, $(*)$ holds for $n \,=\, 0$. Suppose it is true for n.
Applying the Remark 4.4 to the random vector $(X_s^n,\, u_s)$ whose components

belong to the space $\mathbb{D}^{1,p}$ and to the functions σ and b, we can deduce that the random variables $\sigma(s, X_s^n, u_s)$ and $b(s, X_s^n, u_s)$ belong to $\mathbb{D}^{1,p}$ $\forall\ s \in [0,T]$ and that there exist adapted processes $\bar{\sigma}_X(s)$, $\bar{b}_X(s)$, and $\bar{\sigma}_u(s)$, $\bar{b}_u(s)$, uniformly bounded respectively by K_X and K_u, such that

$$D_r\sigma(s, X_s^n, u_s) = \bar{\sigma}_X(s) D_r X_s^n \mathbf{1}_{r\le s} + \bar{\sigma}_u(s) D_r u_s \mathbf{1}_{r\le s}, \tag{5.7}$$

and

$$D_r b(s, X_s^n, u_s) = \bar{b}_X(s) D_r X_s^n \mathbf{1}_{r\le s} + \bar{b}_u(s) D_r u_s \mathbf{1}_{r\le s}. \tag{5.8}$$

From Proposition 4.7 we deduce that the random variables $\sigma(s, X_s^n, u_s)$ and $b(s, X_s^n, u_s)$ belong to $\mathbb{D}^{1,\infty}$.
Thus, the processes $\{\, D_r\sigma(s, X_s^n, u_s)\,\}_s$ and $\{\, D_r b(s, X_s^n, u_s)\,\}_s$ are square integrable and adapted $\forall\ r$.
From (5.7) and (5.8) we get the following inequalities:

$$|D_r\sigma(s, X_s^n, u_s)| \le K_X\,|D_r X_s^n| + K_u\,|D_r u_s|, \tag{5.9}$$

$$|D_r b(s, X_s^n, u_s)| \le K_X\,|D_r X_s^n| + K_u\,|D_r u_s|. \tag{5.10}$$

Now, to differentiate the integral $\int_0^{(\cdot)} \sigma(s, X_s^n, u_s)\, dW_s$, it suffices that the process $\{\sigma(s, X_s^n, u_s)\}_s$ belongs to $\mathbb{L}^{1,2}$, so we have to verify this property.
By Definition 4.4, we have to prove that for almost all t

$$\mathbb{E}\left[\int_0^T |\,\sigma(t, X_t^n, u_t)\,|^2\, dt\right] < +\infty \tag{5.11}$$

and

$$\mathbb{E}\left[\int_0^T\int_0^T |\,D_r\sigma(s, X_s^n, u_s)\,|^2\, dr\, ds\right] < +\infty. \tag{5.12}$$

By Hypotheses 5.2 we have

$$\begin{aligned}
\mathbb{E}\Big[\int_0^T |\,\sigma(t,X_t^n, u_t)\,|^2\, dt\Big] &\le C_2\Big(\mathbb{E}\Big[\int_0^T |\,\sigma(t, X_t^n, u_t) - \sigma(t,0,\bar{u})|^2\, dt \\
&\quad + \mathbb{E}\int_0^T |\,\sigma(t,0,\bar{u})\,|^2\, dt\Big]\Big) \\
&\le C_2'\left(\mathbb{E}\int_0^T K^2(|X_t^n|^2 + |u_t - \bar{u}|^2)\, dt + \mathbb{E}\int_0^T K^2|t|^2\, dt\right) \\
&\le C_2''\left(\mathbb{E}\int_0^T K^2(|X_t^n|^2 + |u_t|^2 + |\bar{u}|^2)\, dt + \mathbb{E}\int_0^T K^2|t|^2\, dt\right) \\
&\le \tilde{C}\left(\mathbb{E}\int_0^T |X_t^n|^2\, dt + \mathbb{E}\int_0^T |u_t|^2\, dt\right) < +\infty
\end{aligned} \tag{5.13}$$

because $X_t^n \in \mathbb{D}^{1,\infty}$ by induction on n and $u \in \mathbb{L}^{1,2}$. $\tilde{C}$ depend on K, T and on $\bar{u}$.
Now we have to prove (5.12).
By (5.9) we have

$$\begin{aligned} &\mathbb{E}\left[\int_0^T \int_0^T \mid D_r\sigma(s, X_s^n, u_s)\mid^2 dr\, ds\right] \\ &\leq K_X^2\, \mathbb{E}\left[\int_0^T \int_0^T \mid D_r X_s^n \mid^2 dr\, ds\right] + K_u^2\, \mathbb{E}\left[\int_0^T \int_0^T \mid D_r u_s \mid^2 dr\, ds\right], \end{aligned} \tag{5.14}$$

so (5.12) holds by induction hypothesis on n and by the fact that $u \in \mathbb{L}^{1,2}$. Lemma 4.3 implies that $\int_0^{(\cdot)} \sigma(s, X_s^n, u_s) dW_s$ belongs to $\mathbb{D}^{1,2}$ and, by Hypotheses 5.2, and by Proposition 4.5, for any $r \leq t$, we have

$$\begin{aligned} D_r\Big(\int_0^t \sigma(s, X_s^n, u_s)\, dW_s\Big) &= \sigma(r, X_r^n, u_r) + \int_r^t D_r\sigma(s, X_s^n, u_s)\, dW_s \\ &= \sigma(r, X_r^n, u_r) + \int_r^t \bar{\sigma}_X(s) D_r X_s^n\, dW_s \\ &\quad + \int_r^t \bar{\sigma}_u(s) D_r u_s\, dW_s. \end{aligned} \tag{5.15}$$

By a similar argument we can prove that $\int_0^t b(s, X_s^n, u_s) ds \in \mathbb{D}^{1,2}$, and for any $r \leq t$, it holds

$$\begin{aligned} D_r\left(\int_0^t b(s, X_s^n, u_s)\, ds\right) &= \int_r^t D_r b(s, X_s^n, u_s)\, ds \\ &= \int_r^t \bar{b}_X(s) D_r X_s^n\, ds + \int_r^t \bar{b}_u(s) D_r u_s\, ds. \end{aligned} \tag{5.16}$$

From these equalities and equation (5.4), we deduce that $X_t^{n+1} \in \mathbb{D}^{1,2}\ \forall t \in [0, T]$.
By invoking Proposition 4.7, in order to prove that $X_t^{n+1} \in \mathbb{D}^{1,p}$ with $p > 2$, we have to prove that $DX_t^{n+1} \in L^p(\Omega; H)$, that is,

$$\mathbb{E}\left[\left[\int_0^T \mid D_r X_t^{n+1} \mid^2 dr\right]^{\frac{p}{2}}\right] < +\infty. \tag{5.17}$$

By (5.9), (5.10), (5.15) and (5.16) we have

$$
\begin{aligned}
&\mathbb{E}\Big[\Big[\int_0^T |\, D_r X_t^{n+1} |^2\, dr\Big]^{\frac{p}{2}}\Big] \\
&\le C_{p,2}\Big\{\mathbb{E}\Big[\Big(\int_0^T |\sigma(r, X_r^n, u_r)|^2\, dr\Big)^{\frac{p}{2}}\Big] + \mathbb{E}\Big[\Big(\int_0^T \int_r^t K_X^2 |D_r X_s^n|^2\, ds\, dr\Big)^{\frac{p}{2}}\Big] \\
&+ \mathbb{E}\Big[\Big(\int_0^T \int_r^t K_u |D_r u_s|^2\, ds\, dr\Big)^{\frac{p}{2}}\Big] + \mathbb{E}\Big[\Big(\int_0^T \int_r^t K_X^2 |D_r X_s^n|^2\, ds\, dr\Big)^{\frac{p}{2}}\Big] \\
&+ \mathbb{E}\Big[\Big(\int_0^T \int_r^t K_u^2 |D_r u_s|^2\, ds\, dr\Big)^{\frac{p}{2}}\Big]\Big\} \\
&\le \tilde{C}\Big\{\mathbb{E}\Big[\Big(\int_0^T |\sigma(r, X_r^n, u_r)|^2\, dr\Big)^{\frac{p}{2}}\Big] + \mathbb{E}\Big[\Big(\int_0^T \int_r^t |D_r X_s^n|^2\, ds\, dr\Big)\Big]^{\frac{p}{2}} \\
&+ \mathbb{E}\Big[\Big(\int_0^T \int_r^t |D_r u_s|^2\, ds\, dr\Big)\Big]^{\frac{p}{2}}\Big\}.
\end{aligned}
\tag{5.18}
$$

Now, we have to show that the first expectation on the right hand-side of the last inequality is finite, so applying conditions Hypotheses 5.2-(A1')-(A3'), we have

$$
\begin{aligned}
&\mathbb{E}\Big[\Big(\int_0^T |\sigma(r, X_r^n, u_r)|^2\, dr\Big)^{\frac{p}{2}}\Big] \\
&\qquad \le \mathbb{E}\left[\left(\int_0^T K^p\, (1 + |X_r^n| + |u_r - \bar{u}| + |r|)^2\, dr\right)^{\frac{p}{2}}\right] \\
&\qquad \le \tilde{K}^p\left[C_{t,p} + \mathbb{E}\left(\int_0^T |X_r^n|^2\, dr\right)^{\frac{p}{2}} + \mathbb{E}\left(\int_0^T |\, u_r\,|^2\, dr\right)^{\frac{p}{2}}\right] \\
&\qquad < +\infty
\end{aligned}
\tag{5.19}
$$

because $X_t^n \in \mathbb{D}^{1,\infty}$ and by Hölder

$$
\mathbb{E}\left(\int_0^T |\, u_r\,|^2\, dr\right)^{\frac{p}{2}} \le \mathbb{E}\left(\int_0^T |\, u_r\,|^p\, dr\right) \le ||\, u\,||^p_{\mathbb{L}^{1,p}} \tag{5.20}
$$

and $u_t \in \mathbb{L}^{1,\infty}$.
$\tilde{K}$ is a constant depending on K, p, and t and $C_{t,p}$ is a constant depending on t and p.

The second expectation is finite, in fact we have

$$
\begin{aligned}
\mathbb{E}\Big[\Big(\int_0^T \int_r^t |D_r X_s^n|^2\, dsdr\Big)^{\frac{p}{2}}\Big] \\
&= \mathbb{E}\left[\left(\int_0^T \int_0^s |D_r X_s^n|^2\, drds\right)^{\frac{p}{2}}\right] \\
&\leq c\,\mathbb{E}\left[\left(\int_0^T \left(\int_0^s |D_r X_s^n|^2\, dr\right)^{p'} ds\right)^{\frac{1}{p'}\frac{p}{2}}\right] \\
&= c\,\mathbb{E}\left[\int_0^t \left[\int_0^s |D_r X_s^n|^2\, dr\right]^{\frac{p}{2}} ds\right]
\end{aligned}
\tag{5.21}
$$

where we have applied Hölder for $p' = \dfrac{p}{2}$, $p > 2$ so $p' > 1$. So, by induction on n the expectation is finite.
The third expectation is finite by Definition 4.5.
So, it follows that (5.17) is true, consequently $X_t^{n+1} \in \mathbb{D}^{1,p}$ with $p \geq 2$. Moreover by Hölder inequality it follows that $X_t^{n+1} \in \mathbb{D}^{1,p}$, $p < 2$, this finally implies that $X_t^{n+1} \in \mathbb{D}^{1,\infty}$.
Finally we have to show that X_t^n converges to the process X.
Now, let

$$
\beta_p = \sup_n \mathbb{E}\left[\left(\int_0^t |\sigma(r, X_r^n, u_r)|^2\, dr\right)^{\frac{p}{2}}\right]. \tag{5.22}
$$

Because the equation (5.19) yields for each n, it implies that $\beta_p < \infty$.
Now, applying Corollary E.1 we obtain

$$
\begin{aligned}
f_{n+1}(t) &= \sup_{0\leq s\leq t} \mathbb{E}\left[\left(\int_0^s |D_r X_s^{n+1}|^2\, dr\right)^{\frac{p}{2}}\right] \\
&\leq k_{p,t}\Big\{\beta_p + c_p\,\mathbb{E}\left[\left(\int_0^t \int_r^t |D_r X_s^n|^2\, dsdr\right)^{\frac{p}{2}}\right] \\
&+ c_p'\,\mathbb{E}\left[\left(\int_0^t \int_r^t |D_r u_s|^2\, dsdr\right)^{\frac{p}{2}}\right]\Big\} \\
&\leq c_1 + \tilde{K}'\,\mathbb{E}\left[\left(\int_0^t \int_r^t |D_r X_s^n|^2\, dsdr\right)^{\frac{p}{2}}\right]
\end{aligned}
\tag{5.23}
$$

because $u_t \in \mathbb{L}^{1,\infty}$ and by Definition 4.5 we have

$$\eta_p = \mathbb{E}\left[\left(\int_0^t \int_r^t |D_r u_s|^2 \, dr ds\right)^{\frac{p}{2}}\right] \leq \| u \|^p_{\mathbb{L}^{1,p}} < +\infty. \tag{5.24}$$

c_1 and $\tilde{K}'$ depend on β_p, c_p, c'_p, η_p and on $k_{p,t}$.
Now, applying Hölder to the right side of the equation (5.23) and recalling the equations (5.21), we have

$$\begin{aligned} \mathbb{E}\Big[\Big(\int_0^t \int_r^t |D_r X^n_s|^2 \, ds dr\Big)^{\frac{p}{2}}\Big] & \\ &\leq c\,\mathbb{E}\left[\left(\int_0^t \int_0^s |D_r X^n_s|^2 \, dr ds\right)^{\frac{p}{2}}\right] \\ &\leq c \int_0^t \mathbb{E}\left(\int_0^s |D_r X^n_s|^2 \, dr\right)^{\frac{p}{2}} ds. \end{aligned} \tag{5.25}$$

Finally, we obtain

$$\begin{aligned} f_{n+1}(t) &\leq c_1 + \tilde{K}' c \int_0^t \mathbb{E}\left(\int_0^s |D_r X^n_s|^2 \, dr\right)^{\frac{p}{2}} ds \\ &\leq c_1 + c_2 \int_0^t f_n(s) \, ds. \end{aligned} \tag{5.26}$$

So (5.5) and (5.34) hold for $n + 1$.
Now we consider

$$\begin{aligned} \mathbb{E}\left(\sup_{0\leq t\leq T} |X^{n+1}_t - X^n_t|^p\right) = \mathbb{E}\Big[\sup_{0\leq t\leq T}\Big| &\int_0^t \big(\sigma(s, X^{n+1}_s, u_s) - \sigma(s, X^n_s, u_s) dW_s\big) \\ &+ \int_0^t \big(b(s, X^{n+1}_s, u_s) - b(s, X^n_s, u_s)\, ds\big)\Big|^p\Big]. \end{aligned} \tag{5.27}$$

Then, by the Corollary E.1 and by Hypotheses 5.2-(A1')-(A1'), we obtain

$$\mathbb{E}\left(\sup_{0\leq t\leq T} |X^{n+1}_t - X^n_t|^p\right) \leq c_p K^p T^{p-1} \int_0^T \mathbb{E}\left(| X^n_s - X^{n-1}_s|^p\right) ds. \tag{5.28}$$

Then, it follows inductively that the preceding expression is bounded by

$$\frac{1}{n!}(c_p K^p T^{p-1})^{n+1} \sup_{0\leq s\leq T} |X^1_s|^p, \tag{5.29}$$

so, consequently, we have

$$\sum_{n=0}^{\infty} \mathbb{E}\left(\sup_{0\le t\le T} |X_t^{n+1} - X_t^n|^p\right) < +\infty, \tag{5.30}$$

which implies the existence of a continuous process X satisfying (5.1) and such that $\mathbb{E}\left(\sup_{0\le t\le T} |X_t|^p\right) \le C$ for all $p > 1$ and C is a positive constant depending on p, T and K [21, section 1.2].
So we know that

$$\mathbb{E}\left(\sup_{s\le T} |\, X_s^n - X_s\,|^p\right) \longrightarrow 0 \tag{5.31}$$

as n tends to infinity.
By Gromwall's lemma applied to (5.23) we deduce that the derivatives of the X_t^n are bounded in $L^p(\Omega \times [0,T])$ uniformly in n for all $p \ge 2$.
Due to Lemma 4.2 we can assert that the random variable X_t belongs to $\mathbb{D}^{1,\infty}$.
Finally, applying the operator D to equation (5.1) and using Proposition 4.4, we deduce the linear stochastic differential equation (5.3) for the derivative of X_t. □

Remark 5.3. *If $u \in \mathbb{L}^{1,p}$ with p fixed, $p \ge 2$, then $X_t \in \mathbb{D}^{1,q}$ $\forall\, q \le p$.* □

5.1.1 The n-dimensional case

For the sake of continuity, in this subsection we will prove the Theorem (5.1) in the n-dimensional case.
We will proceed as well as for the 1-dimensional case, using the same techniques, taking into account that now we will deal matrices in the place of the vectors.

Remark 5.4. *We want to remark that the reason which we repeat the proof of the Theorem although using the same instruments, is that to prove the main result of this work in the next section, we will use the matrix notation, so we think that it is better to be used...* □

Proof of the Theorem 5.1. We will consider the Picard approximation (5.4), taking into account that X_t is stochastic process m dimensional, that is, we have

$$\begin{aligned} X_t^{i,n} &= x_0^i, \\ X_t^{i,n+1} &= x_0^i + \sum_{j=1}^d \int_0^t \sigma_j^i(s, X_s^n, u_s)\, dW_s^j + \int_0^t b^i(s, X_s^n, u_s)\, ds \end{aligned} \tag{5.32}$$

for $i = 1, \ldots, m$, $j = 1, \ldots, d$. We will prove the following property by induction on n:

$(**)$ $X_t^{i,n} \in \mathbb{D}^{1,\infty}$ for all $i = 1, \ldots, m$, $n \geq 0$ and $t \in [0, T]$; furthermore, for all $p > 1$ and $j = 1, \ldots, d$ we have

$$f_{j,n}(t) := \sup_{0 \leq s \leq t} \mathbb{E}\left[\left(\int_0^s |D_r^j X_s^n|^2 dr\right)^{\frac{p}{2}}\right] < +\infty \tag{5.33}$$

and

$$f_{j,n+1}(t) \leq c_1 + c_2 \int_0^t f_{j,n}(s) ds, \tag{5.34}$$

for some positive constants c_1 and c_2.

Clearly, $(**)$ holds for $n = 0$. Suppose it is true for n.
Applying the Remark 4.4 to the random vector $(X_s^n,\, u_s)$ whose components belong to the space $\mathbb{D}^{1,p}$ and to the functions σ_j^i and b^i, we can deduce that the random variables $\sigma_j^i(s, X_s^n, u_s)$ and $b^i(s, X_s^n, u_s)$ belong to $\mathbb{D}^{1,p}$ $\forall\ s \in [0, T]$, $i = 1, \ldots, m$, $j = 1, \ldots, d$, and that there exist m-dimensional adapted processes

$$\begin{aligned} \bar{\sigma}_{j,X}^{n,i}(s) &= \{\bar{\sigma}_{j,X,1}^{n,i}(s), \ldots, \bar{\sigma}_{j,X,m}^{n,i}(s)\}, \\ \bar{\sigma}_{j,u}^{n,i}(s) &= \{\bar{\sigma}_{j,u,1}^{n,i}(s), \ldots, \bar{\sigma}_{j,u,m}^{n,i}(s)\}, \\ \bar{b}_X^{n,i}(s) &= \{\bar{b}_{X,1}^{n,i}(s), \ldots, \bar{b}_{X,m}^{n,i}(s)\}, \\ \bar{b}_u^{n,i}(s) &= \{\bar{b}_{u,1}^{n,i}(s), \ldots, \bar{b}_{X,m}^{n,i}(s)\} \end{aligned} \tag{5.35}$$

uniformly bounded respectively by K_X and K_u, such that

$$D_r \sigma_j^i(s, X_s^n, u_s) = \bar{\sigma}_{j,X,k}^{n,i}(s) D_r X_s^{n,k} \mathbf{1}_{r \leq s} + \bar{\sigma}_{j,u,k}^{n,i}(s) D_r u_s^k \mathbf{1}_{r \leq s}, \tag{5.36}$$

and

$$D_r b^i(s, X_s^n, u_s) = \bar{b}_{X,k}^{n,i}(s) D_r X_s^{n,k} \mathbf{1}_{r \leq s} + \bar{b}_{u,k}^{n,i}(s) D_r u_s^K \mathbf{1}_{r \leq s}. \tag{5.37}$$

From Proposition 4.7 we deduce that the random variables $\sigma_j^i(s, X_s^n, u_s)$ and $b^i(s, X_s^n, u_s)$ belong to $\mathbb{D}^{1,\infty}$.

Thus, the processes $\{ D_r\sigma^i_j(s, X^n_s, u_s) \}_s$ and $\{ D_r b^i(s, X^n_s, u_s) \}_s$ are square integrable and adapted $\forall\, r$ and $\forall\, i = 1, \dots, m,\ j = 1, \dots, d$.
Therefore, from (5.36) and (5.37) we get the following inequalities:

$$|D_r\sigma^i_j(s, X^n_s, u_s)| \le K_X\, |D_r X^{n,k}_s| + K_u\, |D_r u_s|, \tag{5.38}$$

$$|D_r b^i(s, X^n_s, u_s)| \le K_X\, |D_r X^n_s| + K_u\, |D_r u_s|. \tag{5.39}$$

We have to differentiate the integral $\int_0^{(\cdot)} \sigma^i_j(s, X^n_s, u_s)\, dW^j_s$. To this aim, it suffices to prove that the processes $\{\sigma^i_j(s, X^n_s, u_s)\}_s$, belong to $\mathbb{L}^{1,2}$ $\forall\, i = 1, \dots, m,\ j = 1, \dots, d$. Due to Definition 4.4, we have to prove that for almost all t and for $i = 1, \dots, m,\ j = 1, \dots, d$ it holds

$$\mathbb{E}\left[\int_0^T |\, \sigma^i_j(t, X^n_t, u_t)\,|^2\, dt \right] < +\infty \tag{5.40}$$

and

$$\mathbb{E}\left[\int_0^T \int_0^T |\, D_r\sigma^i_j(s, X^n_s, u_s)\,|^2\, dr\, ds \right] < +\infty. \tag{5.41}$$

By Hypotheses 4.3 we have

$$\begin{aligned}
\mathbb{E}\Big[\int_0^T |\, \sigma^i_j(t, X^n_t, u_t)\,|^2\, dt \Big] &\le C_2\Big(\mathbb{E}\Big[\int_0^T |\, \sigma^i_j(t, X^n_t, u_t) - \sigma^i_j(t, 0, \bar u)|^2\, dt \\
&\quad + \mathbb{E} \int_0^T |\, \sigma^i_j(t, 0, \bar u)\,|^2\, dt \Big]\Big) \\
&\le C_2' \left(\mathbb{E} \int_0^T K^2 (|X^n_t|^2 + |u_t - \bar u|^2)\, dt + \mathbb{E} \int_0^T K^2 |t|^2\, dt \right) \\
&\le C_2'' \left(\mathbb{E} \int_0^T K^2 (|X^n_t|^2 + |u_t|^2 + |\bar u|^2)\, dt + \mathbb{E} \int_0^T K^2 |t|^2\, dt \right) \\
&\le \tilde C \left(\mathbb{E} \int_0^T |X^n_t|^2\, dt + \mathbb{E} \int_0^T |u_t|^2\, dt \right) < +\infty
\end{aligned} \tag{5.42}$$

because $X^n_t \in \mathbb{D}^{1,\infty}$ by induction on n and $u \in \mathbb{L}^{1,2}$. $\tilde C$ depend on K, T and on $\bar u$.
Now we have to prove (5.41).
By (5.38) we have

$$\begin{aligned}
&\mathbb{E}\left[\int_0^T \int_0^T |\, D^l_r \sigma^i_j(s, X^n_s, u_s)\,|^2\, dr\, ds \right] \\
&\le K_X^2\, \mathbb{E}\left[\int_0^T \int_0^T |\, D^l_r X^n_s\,|^2\, dr\, ds \right] + K_u^2\, \mathbb{E}\left[\int_0^T \int_0^T |\, D^l_r u_s\,|^2\, dr\, ds \right],
\end{aligned} \tag{5.43}$$

so (5.41) holds by induction hypothesis on n and by the fact that $u \in \mathbb{L}^{1,2}$. Lemma 4.3 implies that $\int_0^{(\cdot)} \sigma_j^i(s, X_s^n, u_s) dW_s^j$ belongs to $\mathbb{D}^{1,2}$ and, by Hypotheses 4.3 and by Proposition 4.5, for any $r \leq t$, we have

$$\begin{aligned} D_r^l\Big(\int_0^t \sigma_j^i(s, X_s^n, u_s)\, dW_s^j\Big) &= \sigma_l^i(r, X_r^n, u_r) + \int_r^t D_r^l \sigma_j^i(s, X_s^n, u_s)\, dW_s^j \\ &= \sigma_l^i(r, X_r, u_r) + \sum_{k=1}^m \sum_{\alpha=1}^d \int_r^t \bar{\sigma}_{\alpha,X,k}^{n,i}(s) D_r^l X_s^{n,k} dW_s^\alpha \\ &+ \sum_{k=1}^m \sum_{\alpha=1}^d \int_r^t \bar{\sigma}_{\alpha,X,k}^{n,i}(s) D_r^l u_s^{n,k} dW_s^\alpha \end{aligned} \tag{5.44}$$

By a similar argument we can prove that $\int_0^t b^i(s, X_s^n, u_s) ds \in \mathbb{D}^{1,2}$ with $i = 1, \ldots, m$, and for any $r \leq t$, it holds

$$\begin{aligned} D_r^l \left(\int_0^t b^i(s, X_s^n, u_s)\, ds\right) &= \int_r^t D_r^l b^i(s, X_s^n, u_s)\, ds \\ &= \sum_{k=1}^m \int_r^t \bar{b}_{X,k}^{n,i}(s) D_r^l X_s^{n,k}\, ds + \sum_{k=1}^m \int_r^t \bar{b}_{u,k}^{n,i}(s) D_r^l u_s^k\, ds. \end{aligned} \tag{5.45}$$

From these equalities and equation (5.4), we deduce that $X_t^{n+1} \in \mathbb{D}^{1,2}\ \forall t \in [0, T]$.

By invoking Proposition 4.7, in order to prove that $X_t^{n+1} \in \mathbb{D}^{1,p}$ with $p > 2$, we want that $DX_t^{n+1} \in L^p(\Omega; H)$, taking into account that since X_t^n is a matrix, then DX_t^{n+1} is a matrix too.

So it has to be true

$$\mathbb{E}\left[\left[\int_0^T |\, D_r X_t^{n+1} |^2\, dr\right]^{\frac{p}{2}}\right] < +\infty. \tag{5.46}$$

By (5.38), (5.39), (5.44) and (5.45) we have

$$
\begin{aligned}
&\mathbb{E}\Big[\Big[\int_0^T |\,D_r^l X_t^{n+1,i}\,|^2\,dr\Big]^{\frac{p}{2}}\Big] \\
&\le\ C_{p,2}\Big\{\mathbb{E}\Big[\Big(\int_0^T |\sigma_l^i(r,X_r^n,u_r)|^2\,dr\Big)^{\frac{p}{2}}\Big]+\sum_{k=1}^m\sum_{\alpha=1}^d\mathbb{E}\Big[\Big(\int_0^T\int_r^t K_{k,X,\alpha}^2|D_r^l X_s^{n,k}|^2\,ds\,dr\Big)^{\frac{p}{2}}\Big] \\
&+\sum_{k=1}^m\sum_{\alpha=1}^d\mathbb{E}\Big[\Big(\int_0^T\int_r^t K_{k,u,\alpha}|D_r^l u_s^k|^2\,ds\,dr\Big)^{\frac{p}{2}}\Big]+\sum_{k=1}^m\mathbb{E}\Big[\Big(\int_0^T\int_r^t K_{k,X}^2|D_r^l X_s^{n,k}|^2\,ds\,dr\Big)^{\frac{p}{2}}\Big] \\
&+\sum_{k=1}^m\mathbb{E}\Big[\Big(\int_0^T\int_r^t K_{k,u}^2|D_r^l u_s^k|^2\,ds\,dr\Big)^{\frac{p}{2}}\Big]\Big\} \\
&\le C\Big\{\,\mathbb{E}\Big[\Big(\int_0^T |\sigma_l^i(r,X_r^n,u_r)|^2\,dr\Big)^{\frac{p}{2}}\Big]+\mathbb{E}\Big[\Big(\int_0^T\int_r^t |D_r^l X_s^n|^2\,ds\,dr\Big)\Big]^{\frac{p}{2}} \\
&+\mathbb{E}\Big[\Big(\int_0^T\int_r^t |D_r^l u_s|^2\,ds\,dr\Big)\Big]^{\frac{p}{2}}\Big\}
\end{aligned}
\tag{5.47}
$$

for $i=1,\dots,m$, $l=1,\dots,d$, where the constant C also depends on dimensions of the Brownian motion and of the stochastic process X_t, respectively d and m.
Now, we have to prove that the right hand-side of the last inequality is finite. Consider the first expectation of the right hand-side: applying Hypotheses 4.3, we have

$$
\begin{aligned}
&\mathbb{E}\Big[\Big(\int_0^T |\sigma_l^i(r,X_r^n,u_r)|^2\,dr\Big)^{\frac{p}{2}}\Big] \\
&\qquad\le\ \mathbb{E}\left[\left(\int_0^T K^p\,(1+|X_r^n|+|u_r-\bar u|+|r|)^2\,dr\right)^{\frac{p}{2}}\right] \\
&\qquad\le\ \tilde K^p\left[C_{t,p}+\mathbb{E}\left(\int_0^T |X_r^n|^2\,dr\right)^{\frac{p}{2}}+\mathbb{E}\left(\int_0^T |\,u_r\,|^2\,dr\right)^{\frac{p}{2}}\right] \\
&\qquad<\ +\infty
\end{aligned}
\tag{5.48}
$$

because $X_t^n\in\mathbb{D}^{1,\infty}$ and by Hölder

$$
\mathbb{E}\left(\int_0^T |\,u_r\,|^2\,dr\right)^{\frac{p}{2}}\le\mathbb{E}\left(\int_0^T |\,u_r\,|^p\,dr\right)\le\|\,u\,\|_{\mathbb{L}^{1,p}}^p \tag{5.49}
$$

and $u_t\in\mathbb{L}^{1,\infty}$.
$\tilde K$ is a constant depending on K, p, and t and $C_{t,p}$ is a constant depending

on t and p.
The second expectation is finite, in fact we have

$$\begin{aligned}
\mathbb{E}\Big[\Big(\int_0^T\int_r^t |D_r^l X_s^n|^2\,dsdr\Big)^{\frac{p}{2}}\Big] \\
&= \mathbb{E}\left[\left(\int_0^T\int_0^s |D_r^l X_s^n|^2\,drds\right)^{\frac{p}{2}}\right] \\
&\leq c\,\mathbb{E}\left[\left(\int_0^T\left(\int_0^s |D_r^l X_s^n|^2\,dr\right)^{p'} ds\right)^{\frac{1}{p'}\frac{p}{2}}\right] \\
&= c\,\mathbb{E}\left[\int_0^t\left[\int_0^s |D_r^l X_s^n|^2\,dr\right]^{\frac{p}{2}} ds\right]
\end{aligned} \tag{5.50}$$

for $l = 1,\dots,d$, where we have applied Hölder for $p' = \dfrac{p}{2}$, $p > 2$ so $p' > 1$.
So, by induction on n the expectation is finite.
By Definition 4.5 we can assert that also third expectation is finite.
So, it follows that (5.46) is true, consequently $X_t^{n+1} \in \mathbb{D}^{1,p}$ with $p \geq 2$.
Moreover by Hölder inequality it follows that $X_t^{n+1} \in \mathbb{D}^{1,p}$, $p < 2$, this finally implies that $X_t^{n+1} \in \mathbb{D}^{1,\infty}$.
Finally we have to show that X_t^n converges to the process X.
Let

$$\beta_p = \sup_n \mathbb{E}\left[\left(\int_0^t |\sigma_l^i(r, X_r^n, u_r)|^2\,dr\right)^{\frac{p}{2}}\right]. \tag{5.51}$$

Since the equation (5.48) yields for each n, it implies that $\beta_p < \infty$ for $i = 1,\dots,m$, and $l = 1,\dots,d$.
Due to (5.47) and applying Corollary E.1 we obtain

$$\begin{aligned}
f_{l,n+1}(t) &= \sup_{0\leq s\leq t} \mathbb{E}\left[\left(\int_0^s |D_r^l X_s^{n+1}|^2\,dr\right)^{\frac{p}{2}}\right] \\
&\leq k_{p,t}C\Big\{\beta_p + c_p\,\mathbb{E}\left[\left(\int_0^t\int_r^t |D_r^l X_s^n|^2\,dsdr\right)^{\frac{p}{2}}\right] \\
&+ c_p'\,\mathbb{E}\left[\left(\int_0^t\int_r^t |D_r^l u_s|^2\,dsdr\right)^{\frac{p}{2}}\right]\Big\} \\
&\leq c_1 + \tilde{K}'\,\mathbb{E}\left[\left(\int_0^t\int_r^t |D_r^l X_s^n|^2\,dsdr\right)^{\frac{p}{2}}\right]
\end{aligned} \tag{5.52}$$

because $u_t \in \mathbb{L}^{1,\infty}$ and by Definition 4.5 we have

$$\eta_p = \mathbb{E}\left[\left(\int_0^t \int_r^t |D_r^l u_s|^2 \, drds\right)^{\frac{p}{2}}\right] \leq \| u \|_{\mathbb{L}^{1,p}}^p < +\infty. \tag{5.53}$$

c_1 and $\tilde{K}'$ depend on C, β_p, c_p, c_p', η_p and on $k_{p,t}$.
We conclude the proof of the Theorem applying Hölder to the right side of the equation (5.52) and recalling the equations (5.50). Then, we have

$$\begin{aligned}\mathbb{E}\left[\left(\int_0^t \int_r^t |D_r^l X_s^n|^2 \, dsdr\right)^{\frac{p}{2}}\right] &\leq c\,\mathbb{E}\left[\left(\int_0^t \int_0^s |D_r^l X_s^n|^2 \, drds\right)^{\frac{p}{2}}\right] \\ &\leq c \int_0^t \mathbb{E}\left(\int_0^s |D_r^l X_s^n|^2 \, dr\right)^{\frac{p}{2}} ds.\end{aligned} \tag{5.54}$$

Finally, we obtain

$$\begin{aligned} f_{l,n+1}(t) &\leq c_1 + \tilde{K}'c \int_0^t \mathbb{E}\left(\int_0^s |D_r^l X_s^n|^2 \, dr\right)^{\frac{p}{2}} ds \\ &\leq c_1 + c_2 \int_0^t f_{l,n}(s) \, ds. \end{aligned} \tag{5.55}$$

So (5.33) and (5.21) hold for $n+1$.
Although the argumentations to conclude the proof of the theorem are the same used to prove the 1-dimensional case, we do not want to omit any calculations, to well-understand the problem.
Calculate

$$\begin{aligned}\mathbb{E}\left(\sup_{0\leq t\leq T} |X_t^{n+1} - X_t^n|^p\right) = \mathbb{E}\Big[\sup_{0\leq t\leq T}\Big|\sum_{j=1}^d \int_0^t \big(\sigma_j(s, X_s^{n+1}, u_s) - \sigma_j(s, X_s^n, u_s) dW_s^j\big) \\ + \int_0^t \big(b(s, X_s^{n+1}, u_s) - b(s, X_s^n, u_s)\, ds\big)\Big|^p\Big].\end{aligned} \tag{5.56}$$

Then, recalling that we indicate with X_t the $m \times 1$ vector, by the Corollary E.1 and by Hypothesis-(A1), we obtain

$$\mathbb{E}\left(\sup_{0\leq t\leq T} |X_t^{n+1} - X_t^n|^p\right) \leq c_p d K^p T^{p-1} \int_0^T \mathbb{E}\left(| X_s^n - X_s^{n-1}|^p\right) ds. \tag{5.57}$$

Then, it follows inductively that the preceding expression is bounded by

$$\frac{1}{n!}(c_p d K^p T^{p-1})^{n+1} \sup_{0\leq s\leq T} |X_s^1|^p, \tag{5.58}$$

so, consequently, we have

$$\sum_{n=0}^{+\infty} \mathbb{E}\left(\sup_{0\le t\le T} |X_t^{n+1} - X_t^n|^p\right) < \infty, \tag{5.59}$$

which implies the existence of a continuous process X satisfying (5.1) and such that $\mathbb{E}\left(\sup_{0\le t\le T} |X_t|^p\right) \le C$ for all $p > 1$ and C is a positive constant depending on p, T and K [21, section 1.2].
So we know that

$$\mathbb{E}\left(\sup_{s\le T} |\, X_s^n - X_s\,|^p\right) \longrightarrow 0 \tag{5.60}$$

as n tends to infinity.
By Gromwall's lemma applied to (5.23) we deduce that the derivatives of the X_t^n are bounded in $L^p(\Omega \times [0, T])$ uniformly in n for all $p \ge 2$.
Due to Lemma 4.2 we can assert that the random variable X_t belongs to $\mathbb{D}^{1,\infty}$.
Finally, applying the operator D to equation (5.1) and using Proposition 4.4, we deduce the linear stochastic differential equation (5.3) for the derivative of X_t. $\square$

Chapter 6

Absolute continuity of the probability law

The main aim of this chapter is to find conditions on the coefficients of equation (5.1) which guarantee that the solution X_t at any time $t \in (0, T]$ has an absolute continuous law with respect to the Lebesgue measure.
In particular, we want to find conditions weaker than the well-known hypoellipticity property of the operator $\mathcal{L}$: consider the second-order differential operator

$$\mathcal{L} = A_0 + \frac{1}{2}\sum_{j=1}^{d}(A_j)^2.$$

Hörmander's theorem [47] states that if the Lie algebra generated by the vector fields $A_0, A_1, \ldots, A_d$ has full rank at each point of $\mathbb{R}^m$, then the operator $\mathcal{L}$ is hypoelliptic. We want to find a condition weaker of this assumption.
Following the idea of Nualart, we will calculate the Malliavin Matrix, and we will find the conditions assuring the invertibility of the matrix: in this way, if the Malliavin matrix of X_t is strictly positive, due to Theorem 4.2, we can state that the law of X_t is absolute continuous with respect to the Lebesgue measure.

6.1 Hörmander condition in standard cases

In order to proceed, we recall the well-known Hörmander conditions usually applied to find a probability law for a stochastic process X_t solution of a SDE without depending on a control process.

Let $X = \{ X_t, \, t \in [0,T] \}$ be the solution of the

$$X_t = x_0 + \sum_{j=1}^{d} \int_0^t \sigma_j(X_s) dW_s^j + \int_0^t b(X_s)\, ds \tag{6.1}$$

where the coefficients are assumed to be infinitely differentiable functions with bounded derivatives of all orders that do not depend on the time. $x_0 \in \mathbb{R}^m$ is the initial value of the process X.
Consider the following vector fields in $\mathbb{R}^m$ associated with the coefficients of (6.1):

$$\begin{aligned} \sigma_j &= \sigma_j^i(x) \frac{\partial}{\partial x_i}, \qquad j = 1, ..., d, \\ b &= b^i(x) \frac{\partial}{\partial x_i} \end{aligned}$$

that is the vectors σ_j $(=\sigma_j^i)$ for $j = 1, ..., d$ and b $(= b^i)$ are identified with the first order differential operator.
The covariant derivative of σ_k in the direction of σ_j is defined as the vector field $\sigma_j^{\nabla} \sigma_k = \sigma_j^l \partial_l \sigma_k^i \frac{\partial}{\partial x_i}$, and the Lie bracket between the vector fields σ_j and σ_k is defined by

$$[\sigma_j, \, \sigma_k] = \sigma_j^{\nabla} \sigma_k - \sigma_k^{\nabla} \sigma_j.$$

∇ is called *connection.*
Set

$$\begin{aligned} \sigma_0 &= \left[b^i(x) - \frac{1}{2} \sigma_l^j(x)\, \partial_j \sigma_l^i(x) \right] \frac{\partial}{\partial x_i} \\ &= b - \frac{1}{2} \sum_{l=1}^{d} \sigma_l^{\nabla} \sigma_l. \end{aligned} \tag{6.2}$$

The vector σ_0 appears when we write the stochastic differential equation (6.1) in terms of the Stratonovich integral of the Ito integral:

$$X_t = \sum_{j=1}^{d} \int_0^t \sigma_j(X_s) \circ dW_s^j + \int_0^t \sigma_0(X_s) ds.$$

Hörmander's condition can be stated as follows:

(H) The vector space spanned by the vector fields

$$\sigma_1, \sigma_2, ..., \sigma_d, \quad [\sigma_i, \sigma_j], \, 0 \leq i, j \leq d, \quad [\sigma_i, [\sigma_j, \sigma_k]], \, 0 \leq i, j, k \leq d, ...$$

at point x_0 is $\mathbb{R}^m$.

Theorem 6.1. *Assume that Hörmander's condition* (H) *holds. Then for any* $t > 0$ *the random vector* X_t *has a probability distribution that is absolutely continuous with respect to the Lebesgue measure.*

The proof of this result can be found in [58, pag 131].
We also recall that in the 1-dimensional case, condition **(H)** asserts that $\sigma(X_0) \neq 0$ or $b(X_0)\partial_x^n \sigma(X_0) \neq 0$ for some $n \geq 1$.

Our aim is to establish what are the conditions that insure that the process X_t has a probability distribution that is absolutely continuous with respect to the Lebesgue measure, where the stochastic process X_t has coefficients depending also on control processes.

6.1.1 Malliavin matrix for feedback control processes

Before to proceed with our objective, we want to remark a basilar aspect in mathematical financial environment.
In most applications in finance the control processes are feedback controls.
This means that the controller is allowed to know the past history of states X_r for $r \leq s$ when the control u_s is chosen. So, from now on, we will concentrate on this type of processes.
Moreover, as for the diffusion and drift coefficients, we assume that $\tilde{u}_t$ does not depend on time.
So we will therefore consider Markovian control processes, that is, control processes u_t that can be written as $u_t = \tilde{u}(X_t)$ where $\tilde{u} : \mathbb{R}^m \longrightarrow \mathcal{U} \subseteq \mathbb{R}^m$.
We suppose that $\tilde{u}$ is a lipschitz function.
Then, by chain rule 4.3, it holds

$$D_r^l u_s^k = \sum_{h=1}^{m} \partial_h \tilde{u}^k(X_s) D_r^l X_s^h.$$

By Theorem 4.2, we know that if the Malliavin matrix γ_{X_t} is invertible, then the law of X_t is absolutely continuous. Then, the idea to proceed is to calculate the Malliavin matrix of stochastic process X_t and then to establish if and when it is invertible.
First of all we have to calculate the stochastic derivative of X_t.
We recall that X_t is define in (6.1), and the coefficients are assumed to be infinitely differentiable functions with bounded derivatives of all orders that do not depend on the time.

We have

$$D_r^l X_t^i = \sigma_l^i(X_r, u_r) + \sum_{k,h=1}^{m} \int_r^t \Big(\partial_k b_X^i(X_s, u_s) D_r^l X_s^k + \partial_k b_u^i(X_s, u_s) \partial_h \tilde{u}^k(X_s) D_r^l X_s^h\Big) ds$$
$$+ \sum_{k,h=1}^{m} \sum_{\theta=1}^{d} \int_r^t \Big(\partial_k \sigma_{\theta,X}^i(X_s, u_s) D_r^l X_s^k + \partial_k \sigma_{\theta,u}^i(X_s, u_s) \partial_h \tilde{u}^k(X_s) D_r^l X_s^h\Big) dW_s^\theta \tag{6.3}$$

for $r \leq t$ a.e., and $D_r^j X_t = 0$ for $r > t$ a.e..
The Malliavin matrix γ_{X_t} is defined as

$$\gamma_t^{i,j} = \langle D_r X_t^i,\, D_r X_t^j \rangle_H = \sum_{l=1}^{d} \int_0^t D_r^l X_t^i D_r^l X_t^j \, dr. \tag{6.4}$$

We know that to study the strictly positive of γ_{X_t} we would have to calculate the scalar product but the computations will not give us wanted results.
To overcome the problem we have to find another way to write γ_{X_t}.
We calculate the stochastic differential of $\gamma_{X_t}^{i,j}$. We have

$$\begin{aligned} d(\gamma_t^{i,j}) &= \sum_{l=1}^{d} d\Big(\int_0^t D_r^l X_t^i D_r^l X_t^j \, dr \Big) \\ &= \sum_{l=1}^{d} \Big[\sigma_l^i(t) \sigma_l^j(t) dt + \int_0^t d(D_r^l X_t^i D_r^l X_t^j) dr \Big]. \end{aligned} \tag{6.5}$$

Now, for $i, j = 1, ..., m$

$$d(D_r^l X_t^i D_r^l X_t^j) = d(D_r^l X_t^i) D_r^l X_t^j + D_r^l X_t^i d(D_r^l X_t^j) + d[D_r^l X^i, D_r^l X^j]_t. \tag{6.6}$$

We will write separately the dynamics of the terms of the above equation to avoid to mix up indices of columns and of the rows. So, for $i = 1, ...m$, and $j = 1, \ldots, d$, we have

$$d(D_r^l X_t^i) = \sum_{k,h=1}^{m} \sum_{\theta=1}^{d} \Big(\partial_k \sigma_{\theta,X}^i(t) D_r^l X_t^k + \partial_k \sigma_{\theta,u}^i(t) \partial_h \tilde{u}^k(X_t) D_r^l X_t^h\Big) dW_t^\theta$$
$$+ \sum_{k,h=1}^{m} \Big(\partial_k b_X^i(s) D_r^l X_s^k + \partial_k b_u^i(s) \partial_h \tilde{u}^k(X_t) D_r^l X_s^h\Big) dt, \tag{6.7}$$

$$\begin{aligned} d(D_r^l X_t^j) = & \sum_{k,h=1}^{m} \sum_{\theta=1}^{d} \Big(\partial_k \sigma_{\theta,X}^j(t) D_r^l X_t^k + \partial_k \sigma_{\theta,u}^j(t) \partial_h \tilde{u}^k(X_t) D_r^l X_t^h \Big) dW_t^\theta \\ & + \sum_{k,h=1}^{m} \Big(\partial_k b_X^j(s) D_r^l X_s^k + \partial_k b_u^j(s) \partial_h \tilde{u}^k(X_t) D_r^l X_s^h \Big) dt, \end{aligned} \tag{6.8}$$

and

$$\begin{aligned} & d[D_r^l X^i, D_r^l X^j]_t \\ & = \sum_{\theta=1}^{d} \sum_{k,h,\alpha,\nu=1}^{m} \Big[\Big(\partial_k \sigma_{\theta,X}^i(t) D_r^l X_t^k + \partial_k \sigma_{\theta,u}^i(t) \partial_h \tilde{u}^k(X_t) D_r^l X_t^h \Big) \Big(\partial_\nu \sigma_{\theta,X}^j(t) D_r^l X_t^\nu \\ & + \partial_\nu \sigma_{\theta,u}^j(t) \partial_\alpha \tilde{u}^\nu(X_t) D_r^l X_t^\alpha \Big) \Big] dt. \end{aligned} \tag{6.9}$$

Now we substitute (6.7), (6.8) and (6.9) into equation (6.6), with respect to the requested calculus. We have

$$\begin{aligned} & d(D_r^l X_t^i D_r^l X_t^j) \\ & = \Big[\sum_{k,h=1}^{m} \sum_{\theta=1}^{d} \Big(\partial_k \sigma_{\theta,X}^i(t) D_r^l X_t^k + \partial_k \sigma_{\theta,u}^i(t) \partial_h \tilde{u}^k(X_t) D_r^l X_t^h \Big) dW_t^\theta \\ & + \sum_{k,h=1}^{m} \Big(\partial_k b_X^i(t) D_r^l X_t^k + \partial_k b_u^i(t) \partial_h \tilde{u}^k(X_t) D_r^l X_t^h \Big) dt \Big] D_r^l X_t^j \\ & + D_r^l X_t^i \Big[\sum_{k,h=1}^{m} \sum_{\theta=1}^{d} \Big(\partial_k \sigma_{\theta,X}^j(t) D_r^l X_t^k + \partial_k \sigma_{\theta,u}^j(t) \partial_h \tilde{u}^k(X_t) D_r^l X_t^h \Big) dW_t^\theta \\ & + \sum_{k,h=1}^{m} \Big(\partial_k b_X^j(t) D_r^l X_t^k + \partial_k b_u^j(t) \partial_h \tilde{u}^k(X_t) D_r^l X_t^h \Big) dt \Big] \\ & + \sum_{\theta=1}^{d} \sum_{k,h,\alpha,\nu=1}^{m} \Big[\Big(\partial_k \sigma_{\theta,X}^i(t) D_r^l X_t^k + \partial_k \sigma_{\theta,u}^i(t) \partial_h \tilde{u}^k(X_t) D_r^l X_t^h \Big) \Big(\partial_\nu \sigma_{\theta,X}^j(t) D_r^l X_t^\nu \\ & + \partial_\nu \sigma_{\theta,u}^j(t) \partial_\alpha \tilde{u}^\nu(X_t) D_r^l X_t^\alpha \Big) \Big] dt. \end{aligned} \tag{6.10}$$

Finally, substitute the above equation into (6.5), so that we will have the exactly and explicit form of the differential stochastic of the Malliavin matrix for any $i, j = i, ..., m$. We have

$$
\begin{aligned}
d(\gamma^{i,j}_{X_t}) &= \sum_{l=1}^{d} \Big\{ \sigma^i_l(t)\sigma^j_l(t)dt + \int_0^t d(D^l_r X^i_t D^l_r X^j_t)dr \Big\} \\
&= \sum_{l=1}^{d} \Bigg\{ \sigma^i_l(t)\sigma^j_l(t)dt \\
&+ \int_0^t \Bigg[\Bigg[\sum_{k,h=1}^{m} \sum_{\theta=1}^{d} \Big(\partial_k \sigma^i_{\theta,X}(t) D^l_r X^k_t + \partial_k \sigma^i_{\theta,u}(t) \partial_h \tilde{u}^k(X_t) D^l_r X^h_t \Big) dW^\theta_t \\
&+ \sum_{k,h=1}^{m} \Big(\partial_k b^i_X(t) D^l_r X^k_t + \partial_k b^i_u(t) \partial_h \tilde{u}^k(X_t) D^l_r X^h_t \Big) dt \Bigg] D^l_r X^j_t \\
&+ D^l_r X^i_t \Bigg[\sum_{k,h=1}^{m} \sum_{\theta=1}^{d} \Big(\partial_k \sigma^j_{\theta,X}(t) D^l_r X^k_t + \partial_k \sigma^j_{\theta,u}(t) \partial_h \tilde{u}^k(X_t) D^l_r X^h_t \Big) dW^\theta_t \\
&+ \sum_{k,h=1}^{m} \Big(\partial_k b^j_X(t) D^l_r X^k_t + \partial_k b^j_u(t) \partial_h \tilde{u}^k(X_t) D^l_r X^h_t \Big) dt \Bigg] \\
&+ \sum_{\theta=1}^{d} \sum_{k,h,\alpha,\nu=1}^{m} \Bigg[\Big(\partial_k \sigma^i_{\theta,X}(t) D^l_r X^k_t + \partial_k \sigma^i_{\theta,u}(t) \partial_h \tilde{u}^k(X_t) D^l_r X^h_t \Big) \Big(\partial_\nu \sigma^j_{\theta,X}(t) D^l_r X^\nu_t \\
&+ \partial_\nu \sigma^j_{\theta,u}(t) \partial_\alpha \tilde{u}^\nu(X_t) D^l_r X^\alpha_t \Big) \Bigg] dt \Bigg] dr \Bigg\} \\
&= \sum_{l=1}^{d} \sigma^i_l(t)\sigma^j_l(t)dt + \sum_{\theta=1}^{d} \Big(\sum_{k=1}^{m} \partial_k \sigma^i_{\theta,X}(t) \gamma^{k,j}_{X_t} + \sum_{k,h=1}^{m} \partial_k \sigma^i_{\theta,u}(t) \partial_h \tilde{u}^k(t) \gamma^{h,j}_{X_t} \Big) dW^\theta_t \\
&+ \Big(\sum_{k=1}^{m} \partial_k b^i_X(t) \gamma^{k,j}_{X_t} + \sum_{k,h=1}^{m} \partial_k b^i_u(t) \partial_h \tilde{u}^k(t) \gamma^{h,j}_{X_t} \Big) dt + \sum_{\theta=1}^{d} \sum_{k=1}^{m} \partial_k \sigma^j_{\theta,X}(t) \gamma^{i,k}_{X_t} dW^\theta_t \\
&+ \sum_{\theta=1}^{d} \sum_{k,h=1}^{m} \partial_k \sigma^j_{\theta,u}(t) \partial_h \tilde{u}^k(t) \gamma^{i,h}_{X_t} dW^\theta_t + \Big(\sum_{k=1}^{m} \partial_k b^j_X(t) \gamma^{i,k}_{X_t} + \sum_{k,h=1}^{m} \partial_k b^j_u(t) \partial_h \tilde{u}^k(t) \gamma^{i,h}_{X_t} \Big) dt \\
&+ \sum_{\theta=1}^{d} \Big(\sum_{k,\nu=1}^{m} \partial_k \sigma^i_{\theta,X}(t) \partial_\nu \sigma^j_{\theta,X}(t) \gamma^{k,\nu}_{X_t} + \sum_{k,\nu,\alpha=1}^{m} \partial_k \sigma^i_{\theta,X}(t) \partial_\nu \sigma^j_{\theta,u}(t) \partial_\alpha \tilde{u}^k(t) \gamma^{k,\alpha}_{X_t} \Big) dt \\
&+ \sum_{\theta=1}^{d} \sum_{k,h,\nu=1}^{m} \partial_k \sigma^i_{\theta,u}(t) \partial_h \tilde{u}^k(t) \partial_\nu \sigma^j_{\theta,X}(t) \gamma^{h,\nu}_{X_t} dt \\
&+ \sum_{\theta=1}^{d} \sum_{k,h,\nu,\alpha=1}^{m} \partial_k \sigma^i_{\theta,u}(t) \partial_h \tilde{u}^k(t) \partial_\nu \sigma^j_{\theta,u}(t) \partial_\alpha \tilde{u}^k(t) \gamma^{h,\alpha}_{X_t} dt.
\end{aligned}
\tag{6.11}
$$

For the sake to simplify the notations, we set $\ell^i_{k,\theta,X}(t) = \partial_k \sigma^i_{\theta,X}(t)$, $\ell^j_{k,\theta,X}(t) = \partial_k \sigma^j_{\theta,X}(t)$ and $g^i_{k,X}(t) = \partial_k b^i_X(t)$, $g^j_{k,X}(t) = \partial_k b^j_X(t)$ for all i, j, $k = 1, ..., m$. Obviously when we consider the derivatives with respect to the control process u, we will use the same functions ℓ and g, that is, $\ell^i_{\theta,u}(t) = \partial_k \sigma^i_{\theta,u}(t)$, $\ell^j_{\theta,u}(t) = \partial_k \sigma^j_{\theta,u}(t)$ and $g^i_{k,u}(t) = \partial_k b^i_u(t)$, $g^j_{k,u}(t) = \partial_k b^j_u(t)$ for all i, j, $k = 1, ..., m$. Moreover we set $h(t) = \partial_h \tilde{u}^k(t)$, for k, $h = 1, \dots, m$.
Then, we write the Malliavin matrix defined in equation (6.11) as

$$
\begin{aligned}
&d(\gamma^{i,j}_{X_t}) \\
&= \sum_{\theta=1}^{d} \sigma^i_\theta(t)\sigma^j_\theta(t)dt + \sum_{\theta=1}^{d}\sum_{k=1}^{m}\Big(\ell^i_{k,\theta,X}(t)\gamma^{k,j}_{X_t} + \ell^j_{k,\theta,X}(t)\gamma^{i,k}_{X_t}\Big)dW^\theta_t \\
&+ \sum_{\theta=1}^{d}\sum_{k,h=1}^{m}\Big(\ell^i_{k,\theta,u}(t)h^k_h(t)\gamma^{h,j}_{X_t} + \ell^j_{k,\theta,u}(t)h^k_h(t)\gamma^{i,h}_{X_t}\Big)dW^\theta_t \\
&+ \sum_{k=1}^{m}\Big(g^i_{k,X}(t)\gamma^{k,j}_{X_t} + g^j_{k,X}(t)\gamma^{i,k}_{X_t}\Big)dt + \sum_{k,h=1}^{m}\Big(g^i_{k,u}(t)h^k_h(t)\gamma^{h,j}_{X_t} + g^j_{k,u}(t)h^k_h(t)\gamma^{i,h}_{X_t}\Big)dt \\
&+ \sum_{\theta=1}^{d}\sum_{k,\nu=1}^{m}\ell^i_{k,\theta,X}(t)\ell^j_{\nu,\theta,X}(t)\gamma^{k,\nu}_{X_t}dt + \sum_{\theta=1}^{d}\sum_{k,\nu,\alpha=1}^{m}\ell^i_{k,\theta,X}(t)\ell^j_{\nu,\theta,u}(t)h^k_\alpha(t)\gamma^{k,\alpha}_{X_t}dt \\
&+ \sum_{\theta=1}^{d}\sum_{k,h,\nu=1}^{m}\ell^i_{k,\theta,u}(t)h^k_h(t)\ell^j_{\nu,\theta,X}(t)\gamma^{h,\nu}_{X_t}dt + \sum_{\theta=1}^{d}\sum_{k,h,\nu,\alpha=1}^{m}\ell^i_{k,\theta,u}(t)h^k_h(t)\ell^j_{\nu,\theta,u}(t)\gamma^{h,\alpha}_{X_t}dt
\end{aligned}
\tag{6.12}
$$

so

$$
\begin{aligned}
&d(\gamma^{i,j}_{X_t}) \\
&= \sum_{\theta=1}^{d} \sigma^i_\theta(t)\sigma^j_\theta(t)dt + \sum_{\theta=1}^{d}\Big[\Big(\ell_{\theta,X}(t)\gamma_{X_t}\Big)^{i,j} + \Big(\ell_{\theta,X}(t)\gamma_{X_t}\Big)^{j,i}\Big]dW^\theta_t \\
&+ \sum_{\theta=1}^{d}\Big[\Big(\ell_{\theta,u}(t)h(t)\gamma_{X_t}\Big)^{i,j} + \Big(\ell_{\theta,u}(t)h(t)\gamma_{X_t}\Big)^{j,i}\Big]dW^\theta_t \\
&+ \Big[\Big(g_X(t)\gamma_{X_t}\Big)^{i,j} + \Big(g_X(t)\gamma_{X_t}\Big)^{j,i}\Big]dt + \Big[\Big(g_u(t)h(t)\gamma_{X_t}\Big)^{i,j} + \Big(g_u(t)h(t)\gamma_{X_t}\Big)^{j,i}\Big]dt \\
&+ \sum_{\theta=1}^{d}\Big(\ell_{\theta,X}(t)\gamma_{X_t}\ell^T_{\theta,X}(t)\Big)^{i,j}dt + \sum_{\theta=1}^{d}\Big[\ell_{\theta,X}(t)\gamma_{X_t}\Big(\ell_{\theta,u}(t)h(t)\Big)^T\Big]^{i,j}dt \\
&+ \sum_{\theta=1}^{d}\Big(\ell_{\theta,u}(t)h(t)\gamma_{X_t}\ell^T_{\theta,X}(t)\Big)^{i,j}dt + \sum_{\theta=1}^{d}\Big[\ell_{\theta,u}(t)h(t)\gamma_{X_t}\Big(\ell_{\theta,u}(t)h(t)\Big)^T\Big]^{i,j}dt
\end{aligned}
\tag{6.13}
$$

and, consequently, using matrix notation, we have

$$\begin{aligned} d\gamma_{X_t} = & \sum_{\theta=1}^{d} \left[\left(\ell_{\theta,X}(t) + \ell_{\theta,u}(t)h(t)\right)\gamma_{X_t} + \gamma_{X_t}\left(\ell_{\theta,X}(t) + \ell_{\theta,u}(t)h(t)\right)^T \right] dW_t^\theta \\ & + \left[\sigma_\theta(t)\sigma_\theta^T(t) + \left(g_X(t) + g_u(t)h(t)\right)\gamma_{X_t} + \gamma_{X_t}\left(g_X(t) + g_u(t)h(t)\right)^T \right] dt \\ & + \sum_{\theta=1}^{d} \left[\left(\ell_{\theta,X}(t) + \ell_{u,X}(t)h(t)\right)\gamma_{X_t}\left(\ell_{\theta,X}(t) + \ell_{\theta,u}(t)h(t)\right)^T \right] dt \end{aligned} \tag{6.14}$$

The system defined by equation (6.14) represents a system of m stochastic differential equations.

6.1.2 Malliavin matrix written as product of matrices

We recall that our objective is to prove that the Malliavin matrix is strictly positive to apply Theorem 4.2.
The presence of the control process in the stochastic differential equation (5.1), and consequently in the Malliavin matrix, renders not possible to apply the standard techniques to the problem in its current form ([58]).
A possible way to tackle this problem consists in representing the Malliavin matrix in such way that it is more sample to study.
Our idea to solve the problem is the following:
we will define a matrix γ_t having the same dimension of the Malliavin matrix as a product of vectors of the stochastic process, taking into account the dimensionality. Then we will calculate the stochastic differential of γ_t, showing that the system obtained represents a system of m stochastic differential equations, having the same form of (6.14). Then we will find the solution of the obtained system: obviously it will be the same of the (6.14), that is, it will represent the Malliavin matrix defined in (6.4), but it will be define in such way that it is possible to prove that it is strictly positiveness.

We will begin proving the following

Proposition 6.1. *Let σ and b the coefficients of the stochastic differential equation* (5.1)*, satisfying Hypotheses 4.3 of the subsection 4.4.1.*
Let

$$\gamma_t = e^{A_t}\left(\int_0^t \left(e^{A_s}\right)^{-1}\sigma(X_s,u_s)\left(\left(e^{A_s}\right)^{-1}\sigma(X_s,u_s)\right)^* ds\right)\left(e^{A_t}\right)^* \tag{6.15}$$

be a system of equations where A_t *is a* $m \times m$ *matrix whose elements are stochastic differential equations.*
Then (6.15) *is the solution of* (6.14).
In particular, the processes A_t *are stochastic differential equations whose coefficients are linear combinations of the partial derivatives of the coefficients of* (6.14).

Proof. We will proceed by first calculating the stochastic differential of (6.15), and then proving that it has the same form of the (6.14).

Let M_t a $m \times m$ matrix of stochastic differential equation defined as

$$\begin{cases} M_0 &= I, \\ dM_t &= M_t(U_t dt + \sum_{\theta=1}^d R_{\theta,t} dW_t^\theta) \end{cases} \tag{6.16}$$

where U_t represents the matrix of drift coefficients, R_t represents the matrix of the diffusion coefficients and $(W_t)_{t\geq 0}$ represents the $(\mathcal{F}_t)_{t\geq 0}$ d-dimensional standard Brownian motion, defined on $(\Omega, F, \mathbb{P})$.
Now we set

$$M_t = e^{A_t}$$

and

$$N_t = \int_0^t \left(e^{A_s}\right)^{-1} \sigma(X_s, u_s) \left(\left(e^{A_s}\right)^{-1} \sigma(X_s, u_s)\right)^* ds,$$

so we can write equation (6.15) as the product of stochastic matrix defined as

$$\gamma_t = M_t N_t M_t^*. \tag{6.17}$$

Now we want to calculate the stochastic differential of (6.17).
We recall the formula to calculate the stochastic differential of a matrices product $D_t N_t C_t$, where D_t, N_t, C_t are stochastic differential equations $\mathbb{R}^d$-valued such that

$$dD_t = U_t dt + \sum_{\theta=1}^d R_{\theta,t} dW_t^\theta$$

$$dN_t = H_t dt + \sum_{\theta=1}^d K_{\theta,t} dW_t^\theta$$

$$dC_t = V_t dt + \sum_{\theta=1}^d Z_{\theta,t} dW_t^\theta,$$

where the coefficients are assumed to be infinitely differentiable functions with bounded derivatives of all orders that do not depend on time, and $(W_t)_{t\geq 0}$ is a d-dimensional standard Brownian motion. It holds true

$$\begin{aligned}
&d(D_tN_tC_t)\\
&= \left[\left(U_tN_t + D_tH_t + \sum_{\theta=1}^{d} R_{\theta,t}K_{\theta,t}\right)C_t + D_tN_tV_t + \sum_{\theta=1}^{d}\left(R_{\theta,t}N_t + D_tK_{\theta,t}\right)Z_{\theta,t}\right]dt\\
&+ \sum_{\theta=1}^{d}\left[\left(R_{\theta,t}N_t + D_tK_{\theta,t}\right)C_t + D_tN_tZ_{\theta,t}\right]dW_t^{\theta}\\
&= \left\{U_tN_tC_t + D_tH_tC_t + D_tN_tV_t + \sum_{\theta=1}^{d}\left[R_{\theta,t}K_{\theta,t}C_t + R_{\theta,t}N_tZ_{\theta,t} + D_tK_{\theta,t}Z_{\theta,t}\right]\right\}dt\\
&+ \sum_{\theta=1}^{d}\left[R_{\theta,t}N_tC_t + D_tK_{\theta,t}C_t + D_tN_tZ_{\theta,t}\right]dW_t^{\theta}.
\end{aligned} \tag{6.18}$$

We use above result to calculate the stochastic differential of (6.17), then we have

$$\begin{aligned}
d\gamma_t =\ & d(M_tN_tM_t^*) = \Bigg[M_tU_tN_tM_t^* + M_tM_t^{-1}\sigma(X_t)\left(M_t^{-1}\sigma(X_s)\right)^* M_t^*\\
&+ M_tN_t\left(M_tU_t\right)^* + M_t\sum_{\theta=1}^{d} R_{\theta,t}N_t\left(M_t\sum_{\theta=1}^{d} R_{\theta,t}\right)^*\Bigg]dt\\
&+ \sum_{\theta=1}^{d}\left[M_tR_{\theta,t}N_tM_t^* + M_tN_t\left(M_t\sum_{\theta=1}^{d} R_{\theta,t}\right)^*\right]dW_t^{\theta}\\
&= \left(M_tU_tN_tM_t^* + \sigma(X_t)\sigma^*(X_t) + M_tN_tU_t^*M_t^* + \sum_{\theta=1}^{d} M_tR_{\theta,t}N_tR_{\theta,t}^*M_t^*\right)dt\\
&+ \sum_{\theta=1}^{d}\left(M_tR_{\theta,t}N_tM_t^* + M_tN_tR_{\theta,t}^*M_t^*\right)dW_t^{\theta}\\
&= \left(U_t\gamma_t + \sigma(X_t)\sigma^*(X_t) + \gamma_tU_t^* + \sum_{\theta=1}^{d} R_{\theta,t}\gamma_tR_{\theta,t}^*\right)dt\\
&+ \sum_{\theta=1}^{d}\left(R_{\theta,t}\gamma_t + \gamma_tR_{\theta,t}^*\right)dW_t^{\theta}.
\end{aligned} \tag{6.19}$$

Equation (6.19) has exactly the same form as equation (6.14). In facts, setting

$$\begin{aligned} \sigma(X_t)\sigma^*(X_t) &= \sum_{\theta=1}^{d} \sigma_\theta(X_t)\sigma_\theta^*(X_t), \\ U_t &= g_X(t) + g_u(t)h(t), \\ R_{\theta,t} &= \ell_{\theta,X}(t) + \ell_{\theta,u}(t)h(t), \end{aligned} \tag{6.20}$$

we obtain the same equation (6.14). This means that the solution of equation (6.14) is exactly (6.15) and the process $M_t = e^{A_t}$ of (6.16) is exactly

$$\begin{cases} M_0 &= e^{A_0} = I, \\ dM_t &= e^{A_t} \left[\left(g_X(t) + g_u(t)h(t)\right) dt \; + \; \sum_{\theta=1}^{d} \left(\ell_{\theta,X}(t) + \ell_{\theta,u}(t)h(t)\right) dW_t^\theta \right]. \end{cases} \tag{6.21}$$

Thus the matrix γ_t is exactly the Malliavin matrix γ_{X_t} defined in (6.4), and we have proved the claim. □

It remains to prove main result, that is the positivity of γ_t.
We can again simplify the proof of the matrix positivity recalling the exponential matrix positivity, so that, to our aim, it suffice to prove that

$$Y_t = \int_0^t e^{-A_s}\sigma_\theta(X_s, u_s) \left(e^{-A_s}\sigma_\theta(X_s, u_s)\right)^* ds \; > \; 0 \tag{6.22}$$

for $t \geq \; 0$ a.e..

6.2 Strictly Positiveness of Malliavin Matrix

In this section we will prove the main result of this work, that is, we will assure the existence of the density function of the process X_t defined in (4.32).
The idea is to consider the standard case which X_t does not depend on the control process u_t and to find and verify an alternative condition to that well-known of Hörmander (see condition **(H)** in Subsection 6.1). Then we will prove that the found condition will also holds when the drift and diffusion coefficients of X_t are depending on the control process.
For the sake to simplify the computations we will begin studying the 1-dimensional case.

6.2.1 The 1-dimensional case

We derive an expression of $\sigma(X_s, u_s)$ using the Taylor's expansion of $\sigma(\cdot, u_s)$ at X_0:

$$\begin{aligned}\sigma(X_s, u_s) = \; & \sigma(X_0, u_s) + \frac{\partial \sigma}{\partial x}(X_0, u_s)(X_s - X_0) + \ldots + \frac{\partial^{(n)} \sigma}{\partial x^n}(X_0, u_s)\frac{(X_s - X_0)^n}{n!} \\ & + \frac{\partial^{(n+1)} \sigma}{\partial x^{n+1}}(\xi_s, u_s)\frac{(X_s - X_0)^{n+1}}{(n+1)!}\end{aligned} \tag{6.23}$$

where ξ_s belongs to the interval in the points X_0 and X_s.
We define for $n \geq 1$

$$\begin{aligned}H_n(X, u) := \; & \sigma(X, u) + \frac{\partial \sigma}{\partial x}(X, u)(X - X_0) + \ldots + \frac{\partial^{(n)} \sigma}{\partial x^n}(X, u)\frac{(X - X_0)^n}{n!} \\ & + \frac{\partial^{(n)} \sigma}{\partial x^n}(X, u)\frac{(X - X_0)^n}{n!}\end{aligned} \tag{6.24}$$

We state the following

Proposition 6.2. *We suppose that there exists an integer $n \geq 0$ such that $H_n(X, u) \neq 0 \; \forall X \sim X_0, \; X \neq X_0, \; u \sim u_0$.*
Let Y_t the Malliavin matrix defined as

$$Y_t = \int_0^t \left(e^{-A_s} \sigma(X_s, u_s) \right)^2 ds. \tag{6.25}$$

Then, Y_t is strictly positive for $t \geq 0$ a.e.

Proof. We suppose that there exists a X_0 such that $\sigma(X_0, u(X_0)) \neq 0$. Define the stopping time

$$\tau^1 = \begin{cases} \inf\{t \in [0,T] : |X_t - X_0| \geq \epsilon\} \wedge T, & \text{if } \{t \in [0,T] : |X_t - X_0| \geq \epsilon\} \neq \emptyset; \\ T, & \text{if } \{t \in [0,T] : |X_t - X_0| \geq \epsilon\} = \emptyset. \end{cases}$$

If $s < \tau^1$, since $\sigma(X_0, u_0) \neq 0$, we can state it holds true

$$|\sigma(X_s, u_s) - \sigma(X_0, u_0)| \leq C$$

where C is a Lipschitz constant.
Then, we have to prove that $P(\tau^1 > 0) = 1$, otherwise, there exists a set A with positive probability such that $\tau^1(\omega) = 0$ for any $\omega \in A$.
For all $\omega \in A$ there exists a sequence $t_n \in [0, T], \; t_n = t_n(\omega)$ such that

$$t_n \longrightarrow 0 : \quad |X_{t_n} - X_0| \geq \epsilon.$$

Now we can assume, without loss of generality that

$$[0,T] \ni t \mapsto X_t(\omega) \in \mathbb{R}$$

is continuous $\forall\, \omega \in A$, ω a.s.
This means that $\forall\, \omega \in A$

$$X_t(\omega) \longrightarrow X_0,\ t \longrightarrow 0$$

so that

$$\lim_{t_n \longrightarrow 0} X_{t_n}(\omega) = X_0, \qquad \forall\, \omega \in A$$

but this implies a contradiction with $|X_{t_n} - X_0| \geq \epsilon$.
We conclude that $P(\tau^1 > 0) = 1$.
Now, fix ϵ_0 such that $0 < \epsilon_0$ and let C_{ϵ_0} be a constant such that

$$C_{\epsilon_0} \geq \left\| \frac{\partial \sigma^2}{\partial x} \right\|_{L^\infty(X_0-\epsilon_0, X_0+\epsilon_0)} + \left\| \frac{\partial \sigma^2}{\partial u} \right\|_{L^\infty(X_0-\epsilon_0, X_0+\epsilon_0)} \tilde{L},$$

where $\tilde{L}$ is the Lipschitz constant of the function $\tilde{u}(\cdot)$. So we can state

$$|\sigma^2(X_s, \tilde{u}(X_s)) - \sigma^2(X_s, \tilde{u}(X_0))| \leq C_{\epsilon_0}|X_s - X_0|,$$

$\forall\, 0 \leq s < \tau^1(\omega)$, $\forall\, \omega \in \Omega$ P a.s. Then, ω a.s and $0 < \epsilon < \epsilon_0$ we have

$$\int_0^T e^{-A_s}\sigma^2(X_s, \tilde{u}(X_s))ds \geq \int_0^{\tau^1(\omega)} e^{-A_s}\sigma^2(X_0, \tilde{u}(X_0))ds - \int_0^{\tau^1(\omega)} e^{-A_s}C_{\epsilon_0}|X_s - X_0|ds.$$

Finally, since for all $s < \tau^1(\omega)$ it holds $|X_s - X_0| < \epsilon$, we deduce that

$$\int_0^T e^{-A_s}\sigma^2(X_s, \tilde{u}(X_s))ds \geq \sigma^2(X_0, \tilde{u}(X_0)) \int_0^{\tau^1_\epsilon(\omega)} e^{-A_s}ds - C_{\epsilon_0}\epsilon \int_0^{\tau^1_\epsilon(\omega)} e^{-A_s}ds.$$

Therefore choosing $\epsilon < \min\{\sigma^2(X_0, \tilde{u}(X_0))/2C_{\epsilon_0}, \epsilon_0\}$, then we have

$$\int_0^T e^{-A_s}\sigma^2(X_s, \tilde{u}(X_s))ds \geq \frac{\sigma^2(X_0, \tilde{u}(X_0))}{2} \int_0^{\tau^1_\epsilon(\omega)} e^{-A_s}ds > 0$$

for all ω a.s..
Then we can state that

$$P(\tau^1 > 0) = 1. \tag{6.26}$$

Define the stopping time

$$\tau^2 = \begin{cases} \inf\{t \in [0,T] : X_t \neq X_0\} \wedge T, & \text{if } \{t \in [0,T] : X_t \neq X_0\} \neq \emptyset; \\ T, & \text{if } \{t \in [0,T] : X_t \neq X_0\} = \emptyset. \end{cases}$$

We will prove that $P(\tau^2(\omega) = 0) = 1$.
Suppose that $P(\tau^2(\omega) > 0) > 0$, that is, there exists a set A with $P(A) > 0$ such that, it holds $\tau^2(\omega) > 0$ for all $\omega \in A$.
We have

$$X_0 = X_t = X_0 + \int_0^t b(X_s, u_s)ds + \int_0^t \sigma(X_s, u_s)dW_s,$$

for all t such that $0 < t < \tau^2(\omega)$, with $\omega \in A$. Hence

$$0 = \int_0^t b(X_0, u_0)ds = b(X_0, u_0)t > 0.$$

This is an obvious contradiction, implying that

$$P(\tau^2 = 0) = 1. \tag{6.27}$$

Thus, we easily argue that, defining

$$h_T^{(k)} = \int_0^T |X_s - X_0|^k ds$$

it satisfies

$$P(h_T^{(k)} > 0) = 1.$$

By the last considerations, there exists a $(\mathcal{F})$-measurable set C such that

(i) $P(C) = 1$;

and $\forall\, \omega \in C$ hold

(ii) $[0, T] \ni t \longrightarrow X_t(\omega)) \in \mathbb{R}$ is continuous;

(iii) $\tau^1(\omega) > 0$, $\tau^2(\omega) = 0$.

As a consequence it holds that integral $h_T^{(k)}$ is positive in C.
By hypotheses on $\tilde{u}(x)$, we know that $\tilde{u}(x) \in K$, where K is a compact set.
We fix $\epsilon_0 > 0$. Let $N > 0$ and $\forall\, k$ such that $0 < k < N$ we have

$$\Big|\frac{\partial^k \sigma(x, y)}{\partial x^k} - \frac{\partial^k \sigma(x', y')}{\partial x^k}\Big| \le M\big(|x - x'| + |y - y'|\big)$$

where x, $x' \in [X_0 - \epsilon_0, X_0 + \epsilon_0]$, y, $y' \in K$, and M is Lipschitz constant.
For every s, $0 < s < \tau^1(\omega)$, $\omega \in C$, we have $X_s(\omega) \in [X_0 - \epsilon_0, X_0 + \epsilon_0]$.
Our aim is to prove that

$$Y_t = \int_0^t \Big(e^{-A_s}\sigma(X_s, u_s)\Big)^2 ds$$

ω a.s.. We know that $\omega \in C\ \tau_\epsilon^1(\omega) > 0\ \tau^2(\omega) = 0$ and the trajectories are continuous. So, we can write by hypothesis that $H_n(X, u) \neq 0\ \forall\, |X - X_0| < \epsilon$, $|u - u_0| < \epsilon$.
Hence we can fix

$$\sigma^2(X_s, u_s) = H_n^2(X_s, u_s) + o\left((X_s - X_0)^n\right);$$

in facts by hypothesis on $u_s = \tilde{u}(X_s)$ we have $|u_s - u_0| \leq \tilde{L}\epsilon$ so that $o\left((X_s - X_0)^n\right) \leq M\epsilon(1 + \tilde{L})$. So, fixed $\bar{\epsilon} = \min\{\tilde{L}, 1\}\epsilon$, if $t < \tau_{\bar{\epsilon}}^1$ we obtain

$$\forall\ s < t \quad \sigma^2(X_s, u_s) > 0 \implies Y_t > 0.$$

that proves the claim.

$\square$

Part III

Appendices

Appendix A

Semigroups Theory

A.1

Proposition A.1. *Let* $S(t) := e^{At}$ *for some* $A \in M_n(\mathbb{C})$*. Then the function* $S(\cdot) : \mathbb{R}_+ \longrightarrow M_n(\mathbb{C})$ *is differentiable and satisfies the differential equation*

(DE)

$$\begin{cases} \dfrac{d}{dt}S(t) = AS(t) \quad for\, t \geq 0, \\ \\ S(0) = I. \end{cases}$$

Conversely, every differentiable function $S(\cdot) : \mathbb{R}_+ \longrightarrow M_n(\mathbb{C})$ *satisfying (DE) is already of the form* $S(t) = e^{At}$ *for some* $A \in M_n(\mathbb{C})$*. Finally, we observe that* $A = \dot{S}(0)$*.*

Definition A.1. *Find all maps* $S(\cdot) : \mathbb{R}_+ \longrightarrow \mathcal{L}(X)$ *satisfying the functional equation*

(FE)

$$\begin{cases} S(t+s) = S(t)S(s) \quad \text{for all } s, t \geq 0, \\ S(0) = I. \end{cases}$$

The problem was investigated by E. Hille [46] and K. Yosida [73] independently by each other in 1948.
Although the object of this study is not dynamical systems per se (or semigroups), it is however necessary to precisely define our problem. For the proof of Proposition (A.1) and an in-depth discussion the reader is referred to [28].

Definition A.2. *A family $(S(t))_{t\geq 0}$ of bounded linear operators on a Banach space X is called a (one-parameter) semigroup (or linear dynamical system) on X if it satisfies the functional equation (FE). If (FE) holds even for all $t, s \in \mathbb{R}$, we call $(S(t))_{t\in\mathbb{R}}$ a (one-parameter) group on X.*

The "typical" example of one-parameter semigroups of operators on a Banach space X is the following: we take any operator $A \in \mathcal{L}(X)$. We can define an operator-valued exponential function by

$$e^{tA} := \sum_{k=0}^{\infty} \frac{t^k A^k}{k!},$$

where the convergence of this series takes place in the Banach algebra $\mathcal{L}(X)$. One can show that $(e^{At})_{t\geq 0}$ satisfies the functional equation (FE) and the the differential equation (DE).

Theorem A.1. *Every uniformly continuous semigroup $(S(t))_{t\geq 0}$ on a Banach space X is of the form*

$$S(t) = e^{tA}, \quad t \geq 0$$

for some bounded operator $A \in \mathcal{L}(X)$.

Definition A.3. *A family $(S(t))_{t\geq 0}$ of bounded linear operators on a Banach space X is called a strongly continuous (one-parameter) semigroup (or C_0-semigroup, where the symbol C_0 abbreviates "Cesáro summable" of order 0), if it satisfies the functional equation (FE) and is strongly continuous.*

Hence, $(S(t))_{t\geq 0}$ is a strongly continuous semigroup if the functional equation (FE) holds and the orbit maps

$$\text{(SC)} \qquad \xi_x := t \mapsto \xi_x(t) := S(t)x$$

are continuous from $\mathbb{R}_+$ into X for every $x \in X$.
If these properties hold for $\mathbb{R}$ instead of $\mathbb{R}_+$, we call $(S(t))_{t\in R}$ a strongly continuous (one-parameter) group (or C_0 group) on X.
The operator A is the infinitesimal generator of $S(\cdot)$. It will be a linear, but generally unbounded, operator defined only on a dense subspace $D(A)$ of the Banach space X. So we write

$$\begin{cases} \mathcal{D}(A) = \left\{ x \in X \,\middle|\, \exists \lim\limits_{h \longrightarrow 0^+} \dfrac{S(t)x - x}{h} \right\} \\ \\ Ax = \lim\limits_{h \longrightarrow 0^+} \dfrac{S(t)x - x}{h}, \; \forall\, x \in \mathcal{D}(A) \end{cases}$$

Appendix B

Canonical $\mathbb{R}$-valued integration and $W_r^{1,2}$-valued integration

B.1 Bochner Integral

This integral is the generalization of Lebesgue integral to Banach-valued functions, i.e, is used for integrations of functions f from some finite measure space $[a, b]$ to a Banach space X equipped with a norm $\|\cdot\|_X$.
Let X be a $\mathbb{R}$-valued Banach space and a, b $\in, \mathbb{R}$, $a < b$. We denote with $\|\cdot\|$ the norm on X and we consider the functions $f : [a, b] \longrightarrow X$. We recall some definitions and properties.

Definition B.1. *We denote by $\mathcal{S}([a, b]; X)$ the vectorial space of the simple functions, that is, of the functions $\varphi : [a, b] \longrightarrow X$ such that*

$$\varphi(t) = \sum_{i=1}^{k} x_i \chi_{A_i}(t), \quad t \in [a, b]$$

where $k \in N^+$, $x_i \in X$, and A_i are Lebesgue measurable subsets of $[a, b]$: Furthermore A_i are disjointed.

Definition B.2. *If $\varphi \in \mathcal{S}([a, b]; X)$, $\varphi = \sum_{i=1}^{k} x_i \chi_{A_i}$, then the integral of φ on $[a, b]$ belongs to X and it holds*

$$\int_a^b \varphi(t)dt = \sum_{i=1}^{k} x_i m(A_i).$$

We note that $\|\varphi(\cdot)\|$ is a $\mathbb{R}$-valued simple function on $[a, b]$.

Definition B.3. *A function* $f : [a, b] \longrightarrow X$ *is said to be* strongly measurable *if there exists a sequence* $(\varphi_n)_{n\in\mathbb{N}} \subseteq \mathcal{S}(X)$ *such that*

$$\varphi_n(t) \longrightarrow f(t) \quad on\, X\ \forall\, t \in [a, b].$$

The function f *is said to be* weakly measurable *if for any* $F \in X^*$ *the function* $t \mapsto Ff(t)$ *is Lebesgue measurable.*

Remark B.1. *If a function* f *is* strongly measurable *then the real function* $t \to \|f(t)\|$ *is also Lebesgue measurable.*

Definition B.4. *Let* $f : [a, b] \longrightarrow X$ *be a* strongly measurable *function. Then f is said summable on* $[a, b]$ *if*

$$\int_a^b \|f(t)\|dt < \infty.$$

Definition B.5. *For* $1 \leq p < \infty$ *we denote by* $L^p([a, b]; X)$ *the space of the functions* $f : [a, b] \longrightarrow X$ strongly measurable *such that* $\int_a^b \|f(t)\|^p dt < \infty$. *If* $p = \infty$ *we denote by* $L^\infty([a, b]; X)$ *the space of the functions* $f : [a, b] \longrightarrow X$ strongly measurable *such that* $\sup\operatorname{ess}_{t\in[a,b]} \|f(t)\| < \infty$.

Definition B.6. *Let X be a Banach space and* $f : [a, b] \longrightarrow X$ *be a summable function. The* ***Bochner*** *integral of f on* $[a, b]$ *is the element in X defined as*

$$\int_a^b f(t)dt = \lim_{n\to\infty} \int_a^b \varphi_n(t)dt,$$

where we take the limit with respect to the norm of X, and $(\varphi_n)_{n\in\mathbb{N}}$ *is a sequence of simple functions such that* $\lim_{n\to\infty} \int_a^b \|f(t) - \varphi_n(t)\|dt = 0$.

Let f be a Bochner integrable function on $[a, b]$. We define a seminorm, called **Bochner norm**, defined as $\|f\|_1 = \|f\|_{L^1} := \int_a^b \|f\|_X dt$.

We will use above definitions to link the $W_r^{1,2}$-valued integration with the canonical $\mathbb{R}$-valued integration.
The Banach space considered is the Sobolev space $X = W_r^{1,2}$, with the norm $\|\cdot\|_{W_r^{1,2}}$. The space of simple functions is defined as $\mathcal{S}([a, b]; W_r^{1,2})$, and it is dense in $L^2([a, b]; W_r^{1,2})$, endowed with the norm

$$\|f\|_{L^2\left([a,b];W_r^{1,2}\right)} = \left(\int_a^b \|f(t)\|^2_{W_r^{1,2}}\right)^{\frac{1}{2}}.$$

Given $f \in W_r^{1,2}$, we can define for all $\xi \in [-r, 0]$, the $\mathbb{R}$-valued function $f(\xi)$, .

Lemma B.1. *Let $f \in W_r^{1,2}$; then $f(\xi) \in \mathbb{R}$ for all $\xi \in [-r, 0]$.*

We have to show that if we have a function f defined in $W_r^{1,2}$, the same function calculated in a point t belonging to an interval is a function defined in $\mathbb{R}$, that is $f(t) \in \mathbb{R}$.

□

Lemma B.2. *Let $f, g \in W_r^{1,2}$; suppose that, for all $\xi \in [-r, 0]$ we have $f(\xi) = g(\xi)$ in $\mathbb{R}$. Then $f = g$ in $W_r^{1,2}$.*

Proof. We know that, for all $\xi \in [-r, 0]$, it holds

$$(f - g)(\xi) = (f(\xi) - g(\xi)) \quad \text{in}\, \mathbb{R}.$$

So, as we have argued in the preceding lemma, we can state

$$\begin{aligned}
\|f - g\|^2_{W_r^{1,2}} &= \int_{-r}^{0} |(f - g)(\xi)|^2 d\xi + \int_{-r}^{0} |(f - g)'(\xi)|^2 d\xi \\
&= \int_{-r}^{0} |f(\xi) - g(\xi)|^2 d\xi + \int_{-r}^{0} |f'(\xi) - g'(\xi)|^2 d\xi \\
&= 0
\end{aligned}$$

since we know that for all $f, g \in W_r^{1,2}$ if

$$f - g \longrightarrow 0, \quad \text{then} \quad f' - g' \longrightarrow 0$$

in L_r^2.

□

Finally we can state the link between the $W_r^{1,2}$-valued integration and the $\mathbb{R}$-valued integration.

Lemma B.3. *Let $f \in L^2\left([a, b]; W_r^{1,2}\right)$. Then we have, for all $\xi \in [-r, 0]$ the following equality in $\mathbb{R}$*

$$\left(\int_a^b f(t)dt\right)(\xi) = \int_a^b f(t)(\xi)dt.$$

Proof. Let $f \in L^2([a, b]; W_r^{1,2})$ and let $(\varphi_n)_{n\geq 0} \subseteq \mathcal{S}([a, b]; W_r^{1,2})$ such that $\varphi_n \longrightarrow f$ in $L^2([a, b]; W_r^{1,2})$, that is

$$\int_a^b \|f(t) - \varphi_n(t)\|_{W_r^{1,2}} dt \longrightarrow 0 \quad \text{for } n \longrightarrow \infty.$$

We have the equality in $\mathbb{R}$ for $n \in \mathbb{N}$, in facts, by Definition B.2 we know that $\int_a^b \varphi_n(t)dt$ belongs to $W_r^{1,2}$, so that, for what we have above asserted, $\left(\int_a^b \varphi_n(t)dt\right)(\xi)$ is $\mathbb{R}$-valued for $\xi \in [-r, 0]$. So we have

$$\begin{gathered} \left(\int_a^b \varphi_n(t)dt\right)(\xi) = \int_a^b \varphi_n(t)(\xi)dt \\ \text{and} \\ \left(\int_a^b \varphi_n(t)'dt\right)(\xi) = \int_a^b \varphi_n(t)'(\xi)dt. \end{gathered} \tag{B.1}$$

We want to pass to the limit this equality to get the claim. We know that $\varphi_n \longrightarrow f$ in $L^2([a, b]; W_r^{1,2})$, so we have

$$\begin{aligned} &\int_a^b dt \int_{-r}^0 |\varphi_n(t)(\xi) - f(t)(\xi)|^2 d\xi + \int_a^b dt \int_{-r}^0 |\varphi_n(t)'(\xi) - f(t)'(\xi)|^2 d\xi \\ &= \int_a^b \|\varphi_n(t) - \varphi(t)\|^2_{W_r^{1,2}} dt \longrightarrow 0, \end{aligned}$$

so that

$$\int_{-r}^0 d\xi \int_a^b |\varphi_n(t)(\xi) - f(t)(\xi)|^2 dt \longrightarrow 0$$

and

$$\int_{-r}^0 d\xi \int_a^b |\varphi_n(t)'(\xi) - f(t)'(\xi)|^2 dt \longrightarrow 0.$$

Thus, this means that we can suppose that, for all $\xi \in [-r, 0]$, we have the convergences $\varphi_n(\cdot)(\xi) \longrightarrow f(\cdot)(\xi)$, and $\varphi_n(\cdot)'(\xi) \longrightarrow f(\cdot)'(\xi)$, in the space $L^2([a, b]; \mathbb{R})$, i.e., for all $\xi \in [-r, 0]$, the convergences in $\mathbb{R}$,

$$\begin{gathered} \int_a^b \varphi_n(t)(\xi)dt \longrightarrow \int_a^b f(t)(\xi)dt, \\ \int_a^b \varphi_n(t)'(\xi)dt \longrightarrow \int_a^b f(t)'(\xi)dt. \end{gathered} \tag{B.2}$$

Furthermore, since $\varphi_n \to f$ in $L^2([a, b]; W_r^{1,2})$, we have the convergences in $W_r^{1,2}$

$$\begin{gathered} \int_a^b \varphi_n(t)dt \longrightarrow \int_a^b f(t)dt, \\ \int_a^b \varphi_n(t)'dt \longrightarrow \int_a^b f(t)'dt. \end{gathered}$$

So, we can write

$$\int_{-r}^{0}\left[\left(\left(\int_a^b f(t)dt\right)(\xi)-\left(\int_a^b \varphi_n(t)dt\right)(\xi)\right)^2\right]d\xi \longrightarrow 0$$

and, obviously,

$$\int_{-r}^{0}\left[\left(\left(\int_a^b f(t)'dt\right)(\xi)-\left(\int_a^b \varphi_n(t)'dt\right)(\xi)\right)^2\right]d\xi \longrightarrow 0.$$

Finally, without loss of generality, we can conclude that for all $\xi\,,\in\,[-r,0]$,

$$\begin{aligned}\left(\int_a^b \varphi_n(t)dt\right)(\xi) &\longrightarrow \left(\int_a^b f(t)dt\right)(\xi),\\ \left(\int_a^b \varphi_n(t)'dt\right)(\xi) &\longrightarrow \left(\int_a^b f(t)'dt\right)(\xi)\end{aligned} \tag{B.3}$$

in $\mathbb{R}$. So, combining (B.1), (B.2) and (B.3), we obtain the desired equality.

□

Appendix C

Linear Deterministic Equations

C.1

We are here concerned with the initial value problem in a Banach space E:

$$\begin{cases} u'(t) = Au(t) + f(t), \ f \in [0,t], \\ \\ u(0) = x \ , x \in E, \end{cases} \tag{C.1}$$

where A is the infinitesimal generator of a C_0- semigroup $S(\cdot)$ in E and $f \in L^1([0,T];E)$ is measurable.
A *strict* solution of problem (C.1) in $L^p([0,T];E)$, $p \in [1,\infty]$, is a function u that belongs to $W_p^1([0,T];E) \cap \mathrm{L}^p([0,T];\mathcal{D}(A))$ and fulfill (C.1).
A *strict* solution of problem (C.1) in $C([0,T];E)$, is a function u that belongs to $C^1([0,T];E) \cap C([0,T];\mathcal{D}(A))$ and fulfills (C.1).
A *weak* solution of problem (C.1) is a function $u \in C([0,T];E)$ such that

$$\varphi(u(t)) = \varphi(x) + \int_0^t (A^*\varphi)(u(s))ds + \int_0^t f(s)ds, \ \forall \varphi \in \mathcal{D}(A^*).$$

Obviously, a strict solution is also a weak solution; but not conversely.
The following result is proved in [16].

Proposition C.1. *Let A be the infinitesimal generator of a C_0-semigroup $S(\cdot)$ in E, and $f \in L^1([0,T];E)$. Then there exists a unique weak solution u of equation* (C.1) *and it is given by the variation of constants formula*

$$u(t) = S(t)x + \int_0^t S(t-s)f(s)ds, \quad f \in [0,T]. \tag{C.2}$$

The function $u(\cdot)$ defined by (C.2) is called the *mild* solution of problem (C.1).

Before proving a sufficient condition for the existence of strict solutions, it is convenient to introduce the approximating problem

$$\begin{cases} u'(t) = A_n u_n(t) + f(t), \quad t \in [0,T], \\ \\ u(0) = x, \, x \in E. \end{cases} \tag{C.3}$$

where A_n are the Yosida approximations of A [11, chapter II, pag 101]. Clearly problem (C.3) has a unique solution $u_n \in W_1^1([0,T];E)$, given by the variation of constants formula

$$u_n(t) = S_n(t)x + \int_0^t S_n(t-s)f(s)ds, \quad f \in [0,T] \tag{C.4}$$

where $S_n(t) = e^{tA_n}$, $t > 0$, and moreover

$$\lim_{n \longrightarrow \infty} u_n = u \quad \text{in}\, C([0,T];E). \tag{C.5}$$

Proposition C.2. *Let A be the infinitesimal generator of a C_0-semigroup $S(\cdot)$ in E.*

(i) If $x \in \mathcal{D}(A)$ and $f \in W_p^1([0,T];E)$ with $p \geq 1$, then problem (C.1) *has a unique strict solution u in $L^p([0,T];E)$, given by formula* (C.4) *and moreover $u \in C^1([0,T];E) \cap C([0,T];\mathcal{D}(A))$.*

(ii) If $x \in \mathcal{D}(A)$ and $f \in L^p([0,T];\mathcal{D}(A))$, then problem (C.1) *has a unique strict solution $u \in C([0,T];E)$, given by formula* (C.2) *and moreover $u \in W_p^1([0,T];E) \cap C([0,T];\mathcal{D}(A))$.*

Appendix D

The Legendre Transform

The Legendre transform is a classical topic of convex analysis. The Legendre transform for differentiable functions defines a correspondence which, for convex functions, is intimately connected with the conjugacy correspondence. A comprehensive treatment of this transform can be found in many textbooks (e.g. [64]); here we give for the convenience of the reader a short and self-contained exposition of the properties which are used in this work.
Let us first recall some basic notions about convex functions. Let $h : \mathbb{R}^n \to \mathbb{R}$ abe a convex function and $u \in \mathbb{R}^n$. We recall that the subdifferential of h at u, denoted by $D^-h(u)$, is the set defined as

$$D^-h(u) := \left\{ p \in \mathbb{R}^n \mid \frac{h(v) - h(u)}{v - u} \geq p, \quad \forall\, v \in \mathbb{R}^n \right\}.$$

If the function is convex, then this definition is equivalent to the one of Fréchet subdifferential.
We deduce the following

Proposition D.1. *Let $h : \mathbb{R}^n \to \mathbb{R}$ be a convex function. Then $D^-h(u)$ is non-empty for every $u \in \mathbb{R}^n$. In addition h is differentiable at u if and only if $D^-h(u)$ is a singleton. If h is differentiable at all points then its differential $Dh(u)$ is continuous.* □

Let us now give the definition of the Legendre transform. We restrict ourselves to convex functions defined in the whole space and superlinear.

Definition D.1. *Let $h : \mathbb{R}^n \to \mathbb{R}$ be a convex function which satisfies*

$$\lim_{|u| \longrightarrow +\infty} \frac{h(u)}{|u|} = +\infty. \tag{D.1}$$

The Legendre transform of h is the function

$$h^*(p) := \sup_{u \in \mathbb{R}^n} \{up - h(u)\}, \qquad p \in \mathbb{R}^n. \tag{D.2}$$

Theorem D.1. *Let $h : \mathbb{R}^n \to \mathbb{R}$ be a convex function satisfying* (D.1).

(i) *For every p there exists at least one point u_p where the supremum in* (D.2) *is attained. In addition for every bounded set $C \subset \mathbb{R}^n$ there exists $R > 0$ such that $|u_p| < R$ for all $p \in C$.*

(ii) *The function h^* is convex and satisfies* $\lim_{|p| \to +\infty} \dfrac{h^*(p)}{|p|} = +\infty$.

(iii) $h^{**} = h$.

(iv) *Given $\bar{u}$, $\bar{p} \in \mathbb{R}^n$ we have*

$$\bar{p} \in D^- h(\bar{u}) \iff \bar{u} \in D^- h^*(\bar{p}) \iff h(\bar{u}) + h^*(\bar{p}) = \bar{u} \cdot \bar{p}.$$

Theorem D.2. *h is strictly convex if and only if h^* is of class C^1.*

Proof. For the proofs of Theorem D.1 and Theorem D.2 we refer the reader to [18, Theorem A.1.3, Theorem A.1.4]. □

Appendix E

Inequalities and Properties Martingales

E.1

Definition E.1. ***Feynman-Kack formula*** *Consider continuous functions* $f : \mathbb{R}^d \longrightarrow \mathbb{R}$, $k : \mathbb{R}^d \longrightarrow [0, \infty)$, *and* $g : [0,T] \times \mathbb{R}^d \longrightarrow \mathbb{R}$. *Suppose that* v *is a continuous, real-valued function on* $[0,T] \times \mathbb{R}^d$, *of class* $C^{1,2}$ *on* $[0,T] \times \mathbb{R}^d$ *and satisfies*

$$-\frac{\partial v}{\partial t} + kv = \frac{1}{2} \bigtriangleup v + g; \quad on\, [0,T) \times \mathbb{R}^d, \tag{E.1}$$

$$v(T,x) = f(x); \quad x \in \mathbb{R}^d. \tag{E.2}$$

Then the function v *is said to be solution of the Cauchy problem for the backward heat equation* (E.1) *with potential* k *and Lagrangian* g, *subject to the terminal condition* (E.2).

Theorem E.1. ***(Feynman(1948), Kack (1949))*** *Let* v *be as Definition (E.1) and assume that*

$$\max_{0 \le t \le T} |v(t,x)| + \max_{0 \le t \le T} |g(t,x)| \le K e^{a\|x\|^2};\ \forall\, x \in \mathbb{R}^d, \tag{E.3}$$

for some constants $K > 0$ *and* $0 < 1/(2Td)$. *Then* v *admits the stochastic representation*

$$\begin{aligned} v(t,x) = \mathbb{E}^x \Bigg[& f(W_{T-t}) \exp\left\{ \int_0^{T-t} k(W_s) ds \right\} \\ & + \int_0^{T-t} g(t+\theta, W_\theta) \exp\left\{ - \int_0^{\theta} k(W_s) ds \right\} d\theta \Bigg]; \quad 0 \le t \le T,\, x \in \mathbb{R}^d. \end{aligned} \tag{E.4}$$

In particular, such a solution is unique.

For the proof see [48, pag 269].

Proposition E.1. ***(Doob's maximal inequality)***
Let $\{ X_t, (\mathcal{F}_t)_{t \geq 0};\ t \in [0,T] \}$ *be a submartingale whose every path is right-continuous, let* $[\sigma, \tau]$ *be a subinterval of* $[0, \infty)$*. We have the following result:*

$$\mathbb{E}\left(\sup_{\alpha \leq t \leq \tau} X_t \right)^p \leq \left(\frac{p}{p-1} \right)^p \mathbb{E}\,(X_\tau^p), \quad p > 1,$$

provided $X_t \leq 0$ *a.s.* P *for every* $t \leq 0$*, and* $\mathbb{E}(X_\tau^p) < \infty$*.*

Corollary E.1. *Let* $[\sigma, \tau]$ *be a subinterval of* $[0, \infty)$*.*
If $\{ X_t, (\mathcal{F}_t)_{t \geq 0};\ t \in [0,T] \}$ *is a* $(\mathcal{F}_t)_{t \geq 0}$ *continuous martingale, then*

$$\mathbb{E}\left(\sup_{\alpha \leq t \leq \tau} |X_t| \right)^p \leq \left(\frac{p}{p-1} \right)^p \mathbb{E}\,(|X_\tau|^p) \quad p > 1.$$

Theorem E.2. ***(Burkholder-Davis-Gundy inequalities)***
Let $M \in \mathcal{M}^{c,loc}$ *where* $\mathcal{M}^{c,loc}$ *is the space of continuous local martingale and use the convention*

$$M_t^* \triangleq \max_{0 \leq s \leq t} |M_s|.$$

For every $m > 0$ *there exists universal positive constants* k_m, K_m *(depending only on m) such that*

$$k_m\, \mathbb{E}\,([M]_T^m) \leq \mathbb{E}\,((M_T^*)^{2m}) \leq K_m\, \mathbb{E}\,([M]_T^m)$$

holds for every stopping times T*. In particular, if we have* $\mathbb{E}\sqrt{[M]_a} < \infty$ *for every* $0 < a < \infty$*, then* M *is a martingale.*

Definition E.2. *We say that a stochastic process* $W = \{W(h),\ h \in H\}$ *defined in a complete probability space* $(\Omega, \mathcal{F}, P)$ *is an isonormal Gaussian process (or a Gaussian process on H) if W is a centered Gaussian family of random variables such that* $\mathbb{E}(W(h)W(g)) = \langle h, g \rangle_H$ *for all* $h, g \in H$*.*

E.2 Lagrange Multipliers

Definition E.3. *Let* X *be a Banach space,* $F \in C^1(X, \mathbb{R})$ *a set of constraints:*

$$S := \{ v \in X;\ F(v) = 0 \}.$$

Suppose that for any $u \in S$, it holds $F'(u) \neq 0$. If $J \in C^1(X, \mathbb{R})$ (or C^1 on a neighborhood of S or C^1 on S) then it says that $c \in \mathbb{R}$ is a critical point of J on S if there exists $u \in S$, and $\lambda \in \mathbb{R}$ such that $J(u) = c$ et $J'(u) = \lambda F'(u)$. The point u is a critical point of J on S and the valued-real λ is named ***Lagrange multiplier*** *for the critical value c (or critical point u).*

If X is a functional space and the equation $J'(u) = \lambda F'(u)$ correspond to a partial differential equation, it says that $J'(u) = \lambda F'(u)$ is the Euler?Lagrange equation (or Euler-equation) satisfied by the critical point u on the constraint S.
The Definition E.3 is justified by the following result that establish the existence of Lagrange multipliers.

Theorem E.3. *Let X, Y, Z be Banach spaces, Ω an open set of $X \times Y$ and $f \in C^1(\Omega, Z)$. Suppose that $(x_0, y_0) \in \Omega$ in such a way that $f(x_0, y_0) = 0$ and that $\partial_y f(x_0, y_0)$ is an homomorphism (linear) from Y to Z. Then there exists $\omega \subset X$ connected neighborhood of x_0 and an unique application $\varphi \in C^1(\omega, Y)$ such that $\varphi(x_0) = y_0$ and such that for all $x \in \omega$ it holds $f(x, \varphi(x)) = 0$. Moreover, if $x \in \omega$ and $f(x, y_*) = 0$, then $y_* = \varphi(x)$. The derivative φ' is given by:*

$$\varphi'(x) = -(\partial_y f(x, \varphi(x)))^{-1} \circ (\partial_x f(x, \varphi(x))).$$

Bibliography

[1] 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.

[2] Bambi M., *Endogenous growth and time to build: the AK case*, Journal of Economic Dynamics and Control, Vol. 32, pp. 1015–1040, 2008.

[3] Bambi M., Fabbri G., Gozzi F., *Optimal policy and consumption smoothing effects in the time-to-build AK model*, Munich Personal RePEc Archive, Deposited on September 2009.

[4] Barbu V., Da Prato G., *Hamilton-Jacobi Equations in Hilbert Spaces*, Pitman, London, 1983.

[5] Barbu V., Da Prato G., *Hamilton-Jacobi equations in Hilbert spaces; variational and semigroup approach*, Annali di Matematica Pura e Applicata, Vol. 142, No. 1, pp. 303–349, 1985.

[6] Barbu V., Da Prato G., *A note on a Hamilton-Jacobi equation in Hilbert space*, Nonlinear Analysis, Vol. 9, No. 12, pp. 1337–1345, 1985.

[7] Barbu V., Da Prato G., Popa C., *Existence and uniqueness of the dynamic programming equation in Hilbert spaces*, Nonlinear Analalysis, Vol. 7, No. 3, pp. 283–299, 1983.

[8] Barbu V., Precupanu Th., *Convexity and Optimization in Banach Spaces*, Acta Applicandae Mathematicae, Bucharest, Vol. 24, No. 3, 1986.

[9] Bardi M., Capuzzo Dolcetta I., *Optimal Control and Viscosity Solutions of Hamilton-Jacobi-Bellman Equations*, Birkhäuser, 2nd Edition, Boston, 1997.

[10] Bekes R., *The range convolution operators*, Pacific Journal of Mathematics, Vol. 110, No. 2, pp. 257–271, 1984.

[11] Bensoussan A., Da Prato G., Delfour M. C., Mitter S. K., *Representation and Control of Infinite Dimensional System*, Birkhäuser Boston, 2nd Edition, 2007.

[12] Boucekkine R., Licandro O., Puch Luis A., Fernando Del Rio F., *Vintage capital and the dynamicsof the AK model*, Journal of Economic Theory, Vol. 120, pp. 39–72, 2005.

[13] Bouleau N. and Hirsch F., *Proprietes d'absolue continuité dans le espaces de Dirichlet et applications aux équations différentielles stochastiques*, Lecture Notes in Mathematics, Springer, Vol. 1204, pp. 131–161, 1986.

[14] Brezis H., *Analisi Funzionale*, Liguori Editori, Napoli, 1986.

[15] Bruckner A., *Differentiation of real functions*, Lecture Notes in Mathematics, Vol. 659, Springer-Verlag, Berlin-Heidelberg-New York, 1978.

[16] Cannarsa P., Di Blasio G., *A direct approach to infinite dimensional Hamilton-Jacobi equations and applications to convex control with state constraints*, Differential and Integral Equations, Vol. 8, No. 2, pp. 225–246, 1995.

[17] 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.

[18] Cannarsa P., Sinestrari C., *Semiconcave functions, Hamilton-Jacobi equations, and optimal control*, Birkhäuser, Boston, 2004.

[19] Cannarsa P., Soner H.M. *Generalized one-sided estimates for solutions of Hamilton-Jacobi equations and applications*, Nonlinear Analysis, Vol. 13, pp. 305–323, 1989.

[20] Crandall M. G. and Lions P.L., *Viscosity Solutions of Hamilton-Jacobi equations*, Transactions of the American Mathematical Society, Vol. 277, No.1, pp. 1–42, 1984.

[21] Collins P. J., *Differential and Integral Equations*, Oxford University Press, 2006.

[22] Da Prato G., *Introduction to Stochastic Analysis and Mallavin Calculus*, Edizioni della Normale, Pisa, ISBN 978-88-7642-313-0, 2007.

[23] Da Prato G., Zabczyk J., *Stochastic Equations in Infinite Dimensions*, Cambridge University Press, 1992.

[24] Davies E. B., *One Parameter Semigroups*, Academic Press, 1980.

[25] 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.

[26] 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.

[27] Ekeland I., Temam R., *Convex Analysis and Variational Problem*, North Holland Company, 1976.

[28] Engel K. J., Nagel R., *One-Parameter Semigroups for Linear Evolution Equations*, Spinger-Verlag, 2000.

[29] Fabbri G., Gozzi F., Swiech A., *Verification theorem and construction of ε-optimal controls for control of abstract evolution equations*, Journal of Convex Analysis, ISSN: 0944-6532, 2009.

[30] Faggian S., *Regular solutions of Hamilton-Jacobi equations arising in Economics*, Applied Mathematics and Optimization, Vol. 51, No. 2, pp. 123–162, 2005.

[31] Faggian S., *Regular Solutions of First-Order HamiltonñJacobi Equations for Boundary Control Problems and Applications to Economics*, Applied Mathematic Optimization 51, pp. 123ñ-162, 2005.

[32] Faggian S., *Infinite dimensional Hamilton–Jacobi-Bellman equations and applications to boundary control with state constraints*, SIAM Journal on Control and Optimization, Vol. 47, No. 4, pp. 2157–2178, 2008.

[33] Federico S., *Stochastic Control Problems for Pension Funds Management,* PhD thesis, Scuola Normale Superiore, Pisa, 2009.

[34] Federico S., Goldys B., Gozzi F., *HJB Equations for the Optimal Control of Differential Equations with Delays and State Constraints, I: Regularity of Viscosity*, March 16, 2010.

[35] Federico S., Goldys B., Gozzi F., *HJB Equations for the Optimal Control of Differential Equations with Delays and State Constraints, II: Verification and Optimal*, March 26, 2010.

[36] Feichtinger G., Hartl R., Sethi S., *Dynamic Optimal Control Models in Advertising: Recent Developments*, Management Science, Vol. 40, No. 2, pp. 195–226, 1994.

[37] Fleming W.H., Soner M., *Controlled Markov processes and Viscosity solutions*, Springer-Verlang, 2nd Edition, 2006.

[38] Gozzi F., *Some results for an optimal control problem with a semilinear state equation II*, SIAM Journal on Control and Optimization, Vol. 29, No.4, pp. 751–768, 1991.

[39] Gozzi F., *Some results for an infinite horizon control problem governed by a semilinear state equation*, Proceedings Vorau, July 10-16 1988; editors Kappel F., Kunisch K., Schappacher W., International Series of Numerical Mathematics, Vol. 91, Birkäuser - Verlag, Basel, pp. 145–163, 1989.

[40] Gozzi F., Fabbri G., *Solving optimal growth models with vintage capital: The dynamic programming approach*, Journal of Economic Theory 143, pp. 331–373, 2008.

[41] Gozzi F., Marinelli C., *Stochastic optimal control of delay equations arising in advertising models*, In: Stochastic PDEís and Applications VII, Levico, Terme, Italy, January 5ñ10, Vol. 245, pp. 133–148, December 2004.

[42] 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, ISSN: 0022-3239.

[43] Hale J.K., Verduyn Lunel S.M., *Introduction to Functional Differential Equations*, Applied Mathematical Sciences, Springer-Verlag, Vol. 99, 1993.

[44] Hartl R. F., *Optimal dynamic advertising policies for hereditary processes*, Journal of Optimization Theory and Applications, Vol. 43, pp. 51ñ-72, 1984.

[45] Hbid M.L., Ezzinbi K., *Delay Differential Equations and Applications*, Springer Netherlands, 2006.

[46] Hille E., *Functional Analysis and Semigroups*, American Mathematical Society, Vol. 21, 1948.

[47] Hörmander L., *Hypoelliptic second order differential equations*, Acta Mathematica, Vol. 119, No.1, pp. 147.-171, 1967.

[48] Karatzas I., Shreve S., *Brownian Motion and Stochastic Calculus*, Springer, 1991.

[49] Karatzas I., Shreve S., *Methods of Mathematical Finance, Applications of Mathematics - Stochastic Modelling and Applied Probability*, Vol. 39, Springer-Verlag, 1998.

[50] Kavian O., *Introduction à la théorie des points critiques*, Springer, 1994.

[51] Koivo H.N., Lee E.B., *Controller synthesis for linear system with retarded state and control variables and quadratic cost*, Automatica, Vol. 8, pp. 203–208, 1972.

[52] Kydland F. E., Prescott E. C., *Time-to-build and aggregate fluctuations*, Econometrica, Vol. 50, No. 6, pp. 1345–1370, 1982.

[53] Li X., Yong J., *Optimal Control Theory for Infinite-Dimensional Systems*, Birkhauser Boston, 1995.

[54] Maz'ja Vladimir G., *Sobolev Spaces*, Springer-Verlag Berlin Heidelberg, 1985.

[55] Merton R.C., *Optimum consumption and portfolio rules in a continuous-time model*, Journal of Economic Theory, Vol.3, pp. 373–413, 1971.

[56] Nerlove M., and Arrow J. K., *Optimal advertising policy under dynamic conditions*, Economica, Vol. 29, pp. 129–142, 1962.

[57] Neveu J., *Discrete-parameter Martingales*, North Holland, Amsterdam, 1975.

[58] Nualart D., *The malliavin calculus and related topics*, Springer, 2nd Edition, 2006.

[59] Papi M., *On the domain of the implicit function and applications*, Journal of Inequalities and Applications, Vol. 2005, Issue 3, pp. 221–234.

[60] Pauwels W., *Optimal Dynamic Advertising Policies in the Presence of Continuously Distributed Time Lags*, Journal of Optimization Theory and Applications, Vol. 22, No 1, pp. 79–89, May 1977.

[61] Phelps R., *Convex Functions, Monotone operators and Differentiability (2nd edition)*, Lecture Notes in Mathematics, Springer-Verlag, Vol. 1364, 1993.

[62] Preiss D., *Differentiability of Lipschitz functions on Banach spaces*, Journal of Functional Analysis, Vol. 91, pp. 312–345, 1990.

[63] Revuz D. and Yor M., *Continuous Martingales and Brownian Motion*, Springer Verlag, 1991.

[64] Rockafellar R.T., *Convex Analysis*, Princeton University Press, 1970.

[65] Saks S., *Theory of the Integral*, Dover Phoenix Editions, 1964.

[66] Sethi S. P., *Sufficient Conditions for the Optimal Control of a Class of Systems with Continuous Lags*, Journal of Optimization Theory and Applications, Vol. 13, No. 5, 1974.

[67] Takeuchi K., *Comparison theorems for solutions of stochastic differential equation*, (Received July 2, 1980); Memoirs of the Faculty of Science, Kyushu Universit, Ser. A., Vol. 35, No. 1, 1981.

[68] Touzi N., *Stochastic control problems, viscosity solutions, and application to finance, special research semester on financial markets: Mathematical, statistical and economic analysis*, Pisa, april 29-July 15, 2002.

[69] Tsoukalas John D., *Time to Build Capital: Revisiting Investment-Cash Flow Sensitivities*, Journal of Economic Dynamics & Control, November 2010.

[70] Vinter R. B. and Kwong R. H., *The infinite time quadratic control problem for linear systems with state and control delays: an evolution equation approach*, SIAM Journal on Control and Optimization, Vol. 19, No. 1, pp. 139–153, 1981.

[71] Yong J., Zhou X.Y., *Stochastic Controls - Hamiltonian Systems and HJB equations*, Springer-Verlag, Berlin-New York, 1999.

[72] Zhou X. Y., Yong J. and Li X., *Stochastic Verification theorems within the framework of viscosity solutions*, SIAM Journal on Control and Optimization, Vol. 37, pp. 243–253, 1997.

[73] Yosida K., *On the differentiability and the representation of one-parameter semigroups of linear Operators*, Journal of the Mathematical Society of Japan, Vol. 1, pp. 15–21, 1948.

[74] Young L. C., *Lectures on the Calculus of Variations*, Chelsea, New York, Second edition, 1980.