POLITECNICO
MILANO 1863

SCUOLA DI INGEGNERIA INDUSTRIALE
E DELL'INFORMAZIONE

A game-theoretic approach to time inconsistent optimal control problems

Tesi di Laurea Magistrale in
Mathematical Engineering - Ingegneria Matematica

Author: **Alessandro Ballestrero**

Student ID: 946457
Advisor: Prof. Fulvia Confortola
Academic Year: 2020-21

Abstract

This thesis has the purpose of analyzing the main problems that arise in dealing with stochastic optimal control problems characterized by time-inconsistency, that is problems for which Bellman's principle of optimality does not hold. Three possible approaches to solving this class of control problems are presented: pre-commitment, naïve and sophisticated. In particular, the thesis explores the last of the three, the sophisticated one, which is based on a concept of optimality typical of game theory: the Nash equilibrium. For the time-inconsistent problem the equilibrium control strategy is constructed through a backward induction procedure, obtaining at the same time an equation characterizing the equilibrium value function of the problem. Finally, we use the presented results to study the generalized Merton's portfolio problem.

Keywords: time-inconsistency, HJB equation, BSDE equilibrium HJB equation

Abstract in lingua italiana

Questa tesi ha lo scopo di analizzare le principali problematiche che emergono nell'affrontare problemi di controllo ottimo stocastico caratterizzati da inconsistenza temporale, ovvero problemi per cui non vale il principio di ottimalità di Bellman. Vengono presentati tre possibili approcci alla risoluzione di questa classe di problemi: pre-commitment, naïve e sofisticato. In particolare, la tesi approfondisce l'ultimo dei tre, quello sofisticato, il quale si fonda su un concetto di ottimalità proprio della teoria dei giochi: l'equilibrio di Nash. Per il problema di controllo inconsistente nel tempo, viene costruita, attraverso una procedura di induzione a ritroso, la strategia di controllo di equilibrio, ottenendo allo stesso tempo un'equazione caratterizzante la funzione valore del problema all'equilibrio. Infine, utilizziamo i risultati ottenuti per studiare il problema di portafoglio di Merton generalizzato.

Parole chiave: inconsistenza temporale, equazione di HJB, BSDE, equazione di HJB all' equilibrio

Contents

Introduction

The aim of this thesis is to present and study a particular class of stochastic optimal control problems, known in the literature as time-inconsistent problems. The analysis of time-inconsistency goes back in time to the mid-nineties when Robert Strotz in [10] pointed out that, in the context of time-discounting methods in economics, any choice of discounting function, apart from the exponential case, leads to a dynamically inconsistent problem. More generally, in order to achieve dynamic consistency one needs to make specific and often restrictive assumptions about the objective functional the agent is minimizing (or maximizing). If the agent's objective does not satisfy these assumptions, time-consistency fails to hold and the usual concept of optimality does not apply. In simple terms, this means that the agent's tastes change over time, so that a plan for some future period considered optimal today is not necessarily optimal when that future period actually arrives. The natural question that arises is therefore how to handle time-inconsistency. In this work we present three different ways to deal with it that are commonly referred to as pre-commitment, naïve and sophisticated. The first two approaches are explained but not treated in depth since they do not properly keep into account the time-changing preferences of the agent. The work focuses instead on the last approach, the sophisticated, which provides an actual solution to the time-inconsistency of the tastes of the decision-maker, internalizing the incentives to deviate from a plan of action that is deemed optimal today and treating them as a constraint.

We briefly describe how the dissertation will be organised.
In Chapter 1 we will describe first what is a standard time-consistent stochastic optimal control problem, introducing the concepts of controlled state, control process and objective functional and highlighting the role of all these elements in the definition of the problem. For this class of optimisation problems we will derive, following [1], the Hamilton-Jacobi-Bellman (HJB) equation, which is a non-linear partial differential equation (PDE) whose solution gives the value function of the problem; in doing so we will remark the central role of Bellman's optimality principle in the construction of an optimal control law.

Then, we will point out what changes in the problem formulation lead to time-inconsistency, explaining in an intuitive manner the meaning of time-inconsistency and showing that the basic tool for the solution of time-consistent control problems, namely the HJB equation, can not be applied to this new class of problems. Here we will introduce also the pre-commitment and naïve approaches to time-inconsistency. For this introduction to time-inconsistency we refer mainly to [3].

We will conclude the chapter with an example of time-consistent control problem, i.e. the Exponential Utility Maximisation, which will allow us to show how the HJB equation is used in practise. After that we will make, for the same problem, a change in the model assumptions in order to show how the formulation becomes time-inconsistent and and how to solve it adopting the pre-commitment and naïve approaches.

In Chapter 2 we formalize the dissertation, we state the hypothesis on the coefficients of the state stochastic differential equation (SDE) and on the functions appearing in the objective functional that make the time-inconsistent problem well-posed. We proceed by illustrating the sophisticated approach to the control problem, which consists basically in viewing the control problem as a non-cooperative game where the interaction among the players is regulated by specific rules. Contextually, we will recall the notion of Nash equilibrium of a non-cooperative game which will be used to define a new concept of optimal control strategy. Moreover, following [4] and [8], it will be provided a very brief introduction to backward stochastic differential equation (BSDE) since some results regarding the forward-backward system will be used in the solution of the differential game. Finally, following [12], the backward induction solution of the game will be presented into details and we will obtain the equilibrium HJB equation giving the equilibrium value function of the time-inconsistent problem and the time-consistent equilibrium control strategy.

In Chapter 3 we examine the Generalized Merton's portfolio problem, that is the optimal investment-consumption allocation problem, over a finite time horizon, when the discounting functions have a very general form. This is generally a time-inconsistent problem, unless the discounting functions are exponential, hence a sophisticated solution is desirable. We will look at the form of the equilibrium HJB equation derived in Chapter 2 and, for completeness, we will also show the pre-commitment and naïve approaches.

1 | Stochastic optimal control problems

1.1. Problem description

Optimal control theory is the branch of mathematical optimization that addresses the problem of seeking a control for a dynamical system over a period of time, which might be finite or infinite, so that a certain objective function is optimized. It arises in decision-making problems and finds numerous and various applications in economics, management and finance.

In this thesis we are concerned with the study of dynamical systems whose evolution can be well described by means of a stochastic differential equation. Such representation allows to incorporate in the state dynamics the effects of the random forces which perturb the system. In particular, we will consider models in which the effects of the different and independent sources of uncertainty are synthesized by a white noise, i.e. Itô diffusion models. This description is very accurate for many phenomena such as unstable stock prices or physical systems subject to thermal fluctuations. Moreover we consider systems for which there is a controller who acts on the state taking his/her decisions (controls) on the basis of the information generated through time and with the goal of achieving the best expected result related to a final objective. Due to the fact that these systems change over time, we expect the relevant decisions, which are made based on the most updated information available, to be time dependent. Such optimization problems are called stochastic optimal control problems and solving them means to find what are the optimal value of the objective function and the optimal control which achieves it.

We introduce all the elements which take part to the definition of a stochastic optimal control problem in continuous time with a finite horizon T. First we list and briefly explain the essential features for the formulation:

- **State equation**: It is the equation which expresses the dynamics of the state variable of the system under investigation i.e. how the quantities which describe the system evolve in time.

- **Control**: The evolution of the system is influenced by a controller who at every time instant tries to steer the state on the basis of the available information.
- **Cost criterion**: The goal of the problem is to minimize (or maximize) a certain cost functional

More formally, let us consider a complete filtered probability space $(\Omega, F, \mathbb{F}, \mathbb{P})$, where the filtration $\mathbb{F}$ satisfies the usual conditions (i.e. right-continuous and complete). Let $W(\cdot)$ be a d-dimensional standard Brownian motion defined on the aforementioned filtered probability space.

The controlled dynamics of the n-dimensional state of the system is expressed through the stochastic differential equation:

$$\begin{cases} dX(s) = b(s, X(s), u(s))ds + \sigma(s, X(s), u(s))dW(s), \quad s \in [t, T] \\ X(t) = x, \end{cases} \tag{1.1}$$

where $b : [0, T] \times \mathbb{R}^n \times \mathbb{R}^m \mapsto \mathbb{R}^n$ and $\sigma : [0, T] \times \mathbb{R}^n \times \mathbb{R}^m \mapsto \mathbb{R}^{n \times d}$, are deterministic maps. Notice that the initial pair $(t, x) \in [0, T] \times \mathbb{R}^n$ of the state equation is deterministic.

The process $u : [0, T] \times \Omega \mapsto \mathbb{R}^m$ which appears in the state equation is the control process. Notice that, in general, the control acts on both the drift and diffusion of the state SDE. Moreover, the control process can take values in a generic metric space; for the exposition it has been chosen $\mathbb{R}^m$.

In order to evaluate the performance of the control process $u : [0, T] \times \Omega \mapsto \mathbb{R}^m$ we use a criterion based on a cost functional. The choice of the form of the cost functional turns out to be crucial to determine the nature of the optimization problem, namely to determine if it is time-consistent or time-inconsistent.

We start the dissertation introducing a functional which gives rise to a standard time-consistent problem. This step will allow us to shed some light on the concept of time-consistency and to present the PDE classical approach to stochastic control problems which is based on the HJB equation.

Then we will change the form of the cost functional showing how time-consistency can be lost and what are the consequences on the solution to the problem. Indeed this last class of problems, i.e. time-inconsistent problems, is what we eventually aim to solve in this thesis.

1.2. Time-consistent problems

The objective is to minimize with respect to the control process a functional of the following form

$$\mathbb{E}\left[\int_0^T C(s, X(s), u(s))ds + G(X(T))\right] \tag{1.2}$$

where the deterministic maps $C : [0, T] \times \mathbb{R}^n \times \mathbb{R}^m \mapsto \mathbb{R}$ and $G : \mathbb{R}^n \mapsto \mathbb{R}$ are referred to as the running cost and the terminal cost respectively.

In many concrete problems it is appropriate to require that the control process is adapted to the state process, namely at time t it is allowed to depend only on past observed values of $X(\cdot)$. In particular we restrict ourselves to consider adapted controls of the form $u(t) = u(t, X(t))$ where $u : [0, T] \times \mathbb{R}^n \mapsto \mathbb{R}^m$ is a deterministic function called the feedback control law. From now on we will denote both the control stochastic process and the deterministic control law by the letter u, the distinction will be clear from the context.

Moreover in many optimisation problems the control process has to satisfy some constraints imposed by the problem and this is taken into account by requiring that
$u(t, X(t)) \in U \subseteq \mathbb{R}^m$ for all t.

Definition 1.1. *A control law is admissible if the following conditions hold*

- $u(t, x) \in U$ *for all times* $t \geq 0$ *and all* $x \in \mathbb{R}^n$.
- *For every initial pair* (t, x) *the SDE* (1.1) *with* $u(s) = u(s, X(s))$ *has a unique solution.*
- *The functional* (1.2) *is well defined and finite.*

We denote by $\mathcal{U}$ *the class of admissible control laws*

Let us define the cost functional $\mathcal{J}_0 : \mathcal{U} \mapsto \mathbb{R}$ as

$$\mathcal{J}_0(u) = \mathbb{E}\left[\int_0^T C(s, X(s), u(s))ds + G(X(T))\right]$$

where $X(\cdot)$ is the solution to (1.1) with the initial pair $(0, x_0)$. The optimisation problem is therefore to minimize the cost functional $\mathcal{J}_0$ over all admissible control laws in $\mathcal{U}$. The solution to the optimization problem is a control law u which tells what the strategy/decision should be, given any possible value of the state process $X(\cdot)$.

We define the value function $\hat{\mathcal{J}}_0$ as

$$\hat{\mathcal{J}}_0 = \inf_{u \in \mathcal{U}} \mathcal{J}_0(u)$$

If we can find an admissible control law $\hat{u}$ such that $\mathcal{J}_0(\hat{u}) = \hat{\mathcal{J}}_0$ then $\hat{u}$ is an optimal control law for the problem.

In order to solve the above optimisation problem we use the methodology of dynamic programming which consists of investigating a wider family of optimisation problems denoted by $\mathcal{P}(t,x)$ and to connect them together with a partial differential equation known as Hamilton-Jacobi-Bellman equation. Solving the control problem is then equivalent to solving the HJB equation.
To be precise, the family of problems $\mathcal{P}(t,x)$ consists of the minimization of the cost functional given that the initial pair relative to the state equation (1.1) is (t,x). An intuitive interpretation of this class of problems is the following: the controller falls asleep at time zero, when he/she weaks up at time t the system is in state x and he/she tries to control it optimally for the remaining window of time [t,T]. Notice that our original problem corresponds to $\mathcal{P}(0,x_0)$.

We introduce the notations for the cost functional and the value function relative to problems $\mathcal{P}(t,x)$.

The cost functional $\mathcal{J} : [0,T] \times \mathbb{R}^n \times \mathcal{U} \mapsto \mathbb{R}$ is defined as

$$\mathcal{J}(t,x,u) = \mathbb{E}_{t,x}\left[\int_t^T C(s,X(s),u(s))ds + G(X(T))\right]$$

where the subscript relative to the expected value indicates that the state dynamics starts from the initial pair (t,x).

The value function $V : [0,T] \times \mathbb{R}^n \mapsto \mathbb{R}$ is defined as

$$V(t,x) = \inf_{u \in \mathcal{U}} \mathcal{J}(t,x,u)$$

Again, the control law $\hat{u}$ such that $\mathcal{J}(t,x,\hat{u}) = V(t,x)$ is the optimal control law for $\mathcal{P}(t,x)$.

Our objective is to obtain the HJB partial differential equation for the value function.

We make the following assumptions

Assumption 1.1. *Assume that*

1. *There exists an optimal control law* $\hat{u}$.
2. *The value function is regular, namely* $V \in C^{1,2}([0,T] \times \mathbb{R}^n)$
3. $\nabla_x V(t, X(t))\sigma(t, X(t), u(t)) \in \mathbb{H}^2(0,T)$

We fix a pair $(t,x) \in (0,T) \times \mathbb{R}^n$ and we choose a real quantity h to be interpreted as a small time increment such that $t+h < T$. Then we arbitrarily choose a control law $u : [0,T] \times \mathbb{R}^n \mapsto \mathbb{R}^m$ and we define a second one as

$$u^*(s,y) = \begin{cases} u(s,y), & (s,y) \in [t, t+h] \times \mathbb{R}^n \\ \hat{u}(s,y), & (s,y) \in (t+h, T] \times \mathbb{R}^n \end{cases}$$

The controller who adopts u^* uses the arbitrary control law u over the time interval $[t, t+h]$ and then switches to the optimal control law for the subsequent times.

The main property which allows us to build a control law like u^* for problems with an associated cost functional of the form (1.2) is **Bellman's principle of optimality**. According to this principle an optimal control law has the property that whatever the initial state and the initial decision are, the remaining decisions must constitute the optimal control law with regard to the state resulting from the first decision. Stated differently, the optimal control law $\hat{u}$ which solves the optimization problem $\mathcal{P}(t,x)$ is independent of the initial pair (t,x). The control law $\hat{u}$ which is the optimal solution for $\mathcal{P}(t,x)$ is still optimal for the problem $\mathcal{P}(t+h, X(t+h))$ whatever control law has driven the system up to time $t+h \geq t$. This very strong property is called time-consistency of the optimal solution. If it were not for the time consistency of the solution we would not have any guarantee that the optimal control law starting from $t+h$ was the same as the optimal control law starting from time t.

Before continuing the derivation of the HJB equation we give a proof of the Bellman's optimality principle

Proof (Bellman Optimality principle). Let $\hat{u}$ be the optimal control law found by solving the optimization problem $\mathcal{P}(0, x_0)$. Suppose now that for some $t > 0$ there exists a control law $\bar{u}$ on $[t,T]$ such that

$$\mathbb{E}\left[\int_t^T C(s, X(s), \bar{u}(s))ds + G(X(T)) \mid \mathcal{F}_t\right] \leq \mathbb{E}\left[\int_t^T C(s, X(s), \hat{u}(s))ds + G(X(T)) \mid \mathcal{F}_t\right]$$

$\mathbb{P}$-a.s. and with strict inequality for a set of positive measure.
We can thus construct a control law on $[0, T]$ of the form

$$\tilde{u}(s, y) = \begin{cases} \hat{u}(s, y), & (s, y) \in [0, t] \times \mathbb{R}^n \\ \bar{u}(s, y), & (s, y) \in (t, T] \times \mathbb{R}^n \end{cases}$$

Then it follows that

$$\begin{aligned}
&\mathcal{J}(0, x_0, \tilde{u}) = \mathbb{E}_{0,x_0}\left[\int_0^T C(s, X(s), \tilde{u}(s))ds + G(X(T))\right] \\
&= \mathbb{E}_{0,x_0}\left[\int_0^t C(s, X(s), \hat{u}(s))ds\right] + \mathbb{E}_{0,x_0}\left[\int_t^T C(s, X(s), \bar{u}(s))ds + G(X(T))\right] \\
&= \mathbb{E}_{0,x_0}\left[\int_0^t C(s, X(s), \hat{u}(s))ds\right] + \mathbb{E}_{0,x_0}\left[\mathbb{E}\left[\int_t^T C(s, X(s), \bar{u}(s))ds + G(X(T)) \mid \mathcal{F}_t\right]\right] \\
&< \mathbb{E}_{0,x_0}\left[\int_0^t C(s, X(s), \hat{u}(s))ds\right] + \mathbb{E}_{0,x_0}\left[\mathbb{E}\left[\int_t^T C(s, X(s), \hat{u}(s))ds + G(X(T)) \mid \mathcal{F}_t\right]\right] \\
&= \mathbb{E}_{0,x_0}\left[\int_0^t C(s, X(s), \hat{u}(s))ds\right] + \mathbb{E}_{0,x_0}\left[\int_t^T C(s, X(s), \hat{u}(s))ds + G(X(T))\right] \\
&= \mathbb{E}_{0,x_0}\left[\int_0^T C(s, X(s), \hat{u}(s))ds + G(X(T))\right] = \mathcal{J}(0, x_0, \hat{u})
\end{aligned}$$

where the properties of the conditional expectation have been exploited along with the initial assumption on the control law $\bar{u}$. Therefore we have found the inequality

$$\mathcal{J}(0, x_0, \tilde{u}) < \mathcal{J}(0, x_0, \hat{u})$$

which contradicts the optimality of the control law $\hat{u}$.

Let us go back to the formulation of the HJB equation. Starting from (t, x) we consider two different strategies over the interval $[t, T]$:

Strategy 1. Use the optimal control law $\hat{u}$.
Strategy 2. Use the control law u^*.

For each one of these strategies we compute the associated cost functional and using the fact that strategy 1. can not be worse than Strategy 2. and taking the limit as h approaches zero we obtain the desired PDE.

The cost functional for Strategy 1. is easily obtained since by definition $\hat{u}$ is the optimal control law so $\mathcal{J}(t,x,\hat{u}) = V(t,x)$.

For what concerns the cost functional associated to Strategy 2. $\mathcal{J}(t,x,u^*)$ we evaluate one at a time the contributions over the disjoint intervals $[t, t+h)$ and $[t+h, T]$. On the first interval the cost functional is given by

$$\mathbb{E}_{t,x}\left[\int_t^{t+h} C(s, X(s), u(s))ds\right]$$

For the second interval we note that at time instant $t+h$ the system will be in the (stochastic) state $X(t+h)$ resulting from the control law u driving the dynamics over $[t, t+h]$ and that from that point on the optimal control law will be adopted. Hence at time $t+h$ the cost functional will be given by $V(t+h, X(t+h))$. Given that we stand at time t when the state of the system is x the cost functional for the future interval $[t+h, T]$ is given by

$$\mathbb{E}_{t,x}\Big[V(t+h, X(t+h))\Big]$$

The total cost functional associated with Strategy 2. is therefore

$$\mathbb{E}_{t,x}\left[\int_t^{t+h} C(s, X(s), u(s))ds + V(t+h, X(t+h))\right]$$

As we said before a comparison of the two strategies is straightforward giving us

$$V(t,x) \leq \mathbb{E}_{t,x}\left[\int_t^{t+h} C(s, X(s), u(s))ds + V(t+h, X(t+h))\right] \tag{1.3}$$

The inequality is given by the fact that the arbitrary control law u applied on the interval $[t, t+h]$ may not be the optimal one.

Remark 1.1. *Inequality* (1.3) *becomes an equality when the control law is the optimal one. This suggests us that the optimisation problem can be solved recursively, having described the relation between the value functions over successive time periods.*

Using the assumption on the regularity of the value function we apply Itô formula to rewrite $V(t+h, X(t+h))$ as

$$\begin{aligned} V(t+h, X(t+h)) = V(t,x) + &\int_t^{t+h} \left(\frac{\partial V}{\partial t}(s, X(s)) + \mathcal{A}^u V(s, X(s)) \right) ds \\ + &\int_t^{t+h} \nabla_x V(s, X(s)) \sigma(s, X(s), u(s)) dW(s) \end{aligned} \tag{1.4}$$

In the above equation we have used the infinitesimal generator associated to the SDE (1.1) defined as

$$\begin{aligned} \mathcal{A}^u &= \sum_{i=1}^n b_i(t,x,u) \frac{\partial}{\partial x_i} + \sum_{i,j=1}^n a_{ij}(t,x,u) \frac{\partial^2}{\partial x_i \partial x_j} \\ &= \langle b(t,x,u), \nabla_x \rangle + tr[a(t,x,u) H_{xx}] \end{aligned} \tag{1.5}$$

where $a(t,x,u) = \frac{1}{2}\sigma(t,x,u)\sigma(t,x,u)^T$, ∇_x and H_{xx} represent the gradient and the Hessian respectively, $\langle \cdot, \cdot \rangle$ denotes the scalar product and $tr[\cdot]$ denotes the trace of a square matrix. The superscript in $\mathcal{A}^u$ highlights the dependence on the control law.

Now if we insert (1.4) into (1.3) the expectation of the stochastic integral disappears thanks to the hypothesis of regularity of its argument and the terms $V(t,x)$ cancel out, leaving the inequality

$$\mathbb{E}_{t,x}\left[\int_t^{t+h} \left(C(s, X(s), u(s)) ds + \frac{\partial V}{\partial t}(s, X(s)) + \mathcal{A}^u V(s, X(s)) \right) ds \right] \geq 0$$

We divide both sides by h, move h inside the expectation and let it tend to zero. Assuming enough integrability it is possible to switch limit and expectation and by the fundamental theorem of integral calculus we get

$$C(t,x,u) + \frac{\partial V}{\partial t}(t,x) + \mathcal{A}^u V(t,x) \geq 0$$

In this inequality u represents the value of the control law u evaluated at (t,x). Moreover since the control law u was arbitrary the inequality holds for all $u \in U$ and it holds with equality if and only if $u = \hat{u}(t,x)$, i.e. if the control law is the optimal one. We have thus

$$\frac{\partial V}{\partial t}(t,x) + \inf_{u \in U} \{ C(t,x,u) + \mathcal{A}^u V(t,x) \} = 0 \tag{1.6}$$

Notice that the point (t, x) fixed at the beginning of the construction was arbitrary so the equation must hold for all $(t, x) \in [0, T] \times \mathbb{R}^n$. Equation (1.6) is a non linear PDE for which we have a boundary condition found by noticing that $V(T, x) = G(x)$ for all $x \in \mathbb{R}^n$.

We can now state formally the derived result

Theorem 1.1. *Let Assumption 1.1 hold then*

1. *The value function V satisfies the Hamilton-Jacobi-Bellman equation*

$$\begin{cases} \dfrac{\partial V}{\partial t}(t, x) + \inf\limits_{u \in U}\{C(t, x, u) + \mathcal{A}^u V(t, x)\} = 0, \quad (t, x) \in [0, T] \times \mathbb{R}^n \\ V(T, x) = G(x), \quad x \in \mathbb{R}^n \end{cases}$$

2. *For every $(t, x) \in [0, T] \times \mathbb{R}^n$ the infimum in the equation is attained by $u = \hat{u}(t, x)$.*

It is important to remark that Theorem 1.1 constitutes only a necessary condition for the optimisation problem. It only ensures that if V is the enough regular value function and $\hat{u}$ the optimal control law, then the HJB equation is satisfied by V and the infimum is realized by $\hat{u}$.

The next theorem we prove, called the Verification theorem shows that the Hamilton-Jacobi-Bellman equation also acts as a sufficient condition for the optimisation problem.

Theorem 1.2. *Let $H(t, x)$ and $g(t, x)$ be two functions such that*

- *H is enough regular (see Assumption 1.1) and solves the Hamilton-Jacobi-Bellman equation*

$$\begin{cases} \dfrac{\partial H}{\partial t}(t, x) + \inf\limits_{u \in U}\{C(t, x, u) + \mathcal{A}^u H(t, x)\} = 0, \quad (t, x) \in [0, T] \times \mathbb{R}^n \\ H(T, x) = G(x), \quad x \in \mathbb{R}^n \end{cases}$$

- *g is an admissible control law.*
- *For every $(t, x) \in [0, T] \times \mathbb{R}^n$ the infimum is attained for $u = g(t, x)$.*

Then the value function of the problem is given by $V = H(t, x)$ and there exists an optimal control law $\hat{u}$ given by $\hat{u} = g(t, x)$.

Proof. Take H and g as above, choose an arbitrary admissible control law u and fix a pair $(t, x) \in [0, T] \times \mathbb{R}^n$. Let us define the state process $X(\cdot)$ controlled by the process

$u(s) = u(s, X(s))$ over the time interval $[t, T]$ as the solution to the SDE:

$$\begin{cases} dX(s) = b(s, X(s), u(s))ds + \sigma(s, X(s), u(s))dW(s), \quad s \in [t, T] \\ X(t) = x, \end{cases}$$

We insert the process $X(\cdot)$ into the function H and use Itô formula

$$\begin{aligned} H(T, X(T)) = H(t, x) &+ \int_t^T \left(\frac{\partial H}{\partial t}(s, X(s)) + \mathcal{A}^u H(s, X(s)) \right) ds \\ &+ \int_t^T \nabla_x H(s, X(s)) \sigma(s, X(s), u(s)) dW(s) \end{aligned}$$

By hypothesis H solves the Hamilton-Jacobi-Bellman equation hence

$$C(t, x, u) + \frac{\partial H}{\partial t}(t, x) + \mathcal{A}^u H(t, x) \geq 0$$

for every $u \in U$ and thus it holds for every s and $\mathbb{P}$-a.s the inequality

$$\frac{\partial H}{\partial t}(s, X(s)) + \mathcal{A}^u H(s, X(s)) \geq -C(s, X(s), u(s))$$

From the boundary condition on H it follows that $H(T, X(T)) = G(X(T))$ $\mathbb{P}$-a.s so we obtain the inequality

$$H(t, x) \leq G(X(T)) + \int_t^T C(s, X(s), u(s)) - \int_t^T \nabla_x H(s, X(s)) \sigma(s, X(s), u(s)) dW(s)$$

Taking the expectation and by the assumption on the integrability of H we obtain

$$H(t, x) \leq \mathbb{E}_{t,x} \left[\int_t^T C(s, X(s), u(s)) ds + G(X(T)) \right] = \mathcal{J}(t, x, u)$$

By the fact that the control law u was arbitrary we get

$$H(t, x) \leq \inf_{u \in \mathcal{U}} \mathcal{J}(t, x, u) = V(t, x) \tag{1.7}$$

The reverse inequality is obtained by choosing the specif control law $u(t, x) = g(t, x)$ and repeating the same computations as above but using the fact that this time we have

$$C(t,x,g) + \frac{\partial H}{\partial t}(t,x) + \mathcal{A}^g H(t,x) = 0$$

which gives us the equality

$$H(t,x) = \mathbb{E}_{t,x}\left[\int_t^T C(s,X(s),g(s))ds + G(X(T))\right] = \mathcal{J}(t,x,g) \tag{1.8}$$

Exploiting the obvious inequality $V(t,x) \leq \mathcal{J}(t,x,g)$ and combining (1.7) and (1.8) we obtain

$$H(t,x) \leq V(t,x) \leq \mathcal{J}(t,x,g) = H(t,x)$$

which shows that actually $H(t,x) = V(t,x)$ and that the control law g is optimal.

We describe now how the Hamilton-Jacobi-Bellman equation is handled in practise. We consider therefore our stochastic optimal control problem $\mathcal{P}(t,x)$ and the corresponding PDE previously derived.

$$\begin{cases} \dfrac{\partial V}{\partial t}(t,x) + \inf_{u\in U}\{C(t,x,u) + \mathcal{A}^u V(t,x)\} = 0, \quad (t,x) \in [0,T]\times\mathbb{R}^n \\ V(T,x) = G(x), \quad x \in \mathbb{R}^n \end{cases}$$

We proceed as follow:

1. Consider the HJB equation as a partial differential equation where the function V is the unknown.

2. Fix arbitrarily a point $(t,x) \in [0,T]\times\mathbb{R}^n$ and consider, for that fixed pair, the optimization problem
$$\inf_{u\in U}[C(t,x,u) + \mathcal{A}^u V(t,x)]$$
In this problem u is the only variable while t and x are viewed as fixed parameters. Moreover, the functions C, b, σ and V are considered as given.

3. The optimal value of u, denoted by $\hat{u}$ depends obviously on the considered pair (t,x) but also on function V, which enters the optimization problem through its partial derivatives hidden in the expression $\mathcal{A}^u V(t,x)$. Such dependence is highlighted by

the expression

$$\hat{u} = \hat{u}(t, x, V)$$

4. The function $\hat{u}(t, x, V)$ is the candidate for the optimal control law. Of course we do not know the precise form of function V so this description is incomplete. Thus, we proceed by substituting the expression for $\hat{u}$ into the HJB partial differential equation, obtaining the equation

$$\begin{cases} \dfrac{\partial V}{\partial t}(t,x) + C(t,x,\hat{u}) + \mathcal{A}^{\hat{u}}V(t,x) = 0, \quad (t,x) \in [0,T] \times \mathbb{R}^n \\ V(T,x) = G(x), \quad x \in \mathbb{R}^n \end{cases}$$

5. We solve the above PDE. Assuming this can be carried out, once obtained a solution V we substitute it into the expression $\hat{u}(t, x, V)$ obtaining explicitly our candidate optimal control. Finally, exploiting the Verification theorem we can identify V as the value function and $\hat{u}$ as the optimal control law.

   It is important to remark that the hardest part of this procedure is point 5. since it requires to solve a highly nonlinear PDE. There are no standard analytic methods which can be used so very few optimal control problems can be solved analitically. What is usually done is to make an educated guess on the form of the solution function V. Such choice should be dictated by the observation that if an analytic solution exists then it should inherit some properties from the boundary condition function G and from the running cost C.

1.3. Time-inconsistent problems

We have seen in the previous section that the big advantage of time-consistency is that whatever the initial pair is, if an optimal control law can be constructed (for that initial pair) then it will remain optimal from that point on. However, this property of the optimal solution is indeed very ideal and we can expect that in real-life situations time-consistency could be lost. As we briefly mentioned in the first section, the nature of the optimisation problem is determined by the form of the cost functional. Problems with an associated cost functional of the form (1.2) are time-consistent in the sense that they admit time-consistent optimal control laws. Now we consider a new family of optimisation problems of the form

$$\inf_{u\in\mathcal{U}} \mathbb{E}_{t,x}\left[\int_t^T C(t,x,s,X(s),u(s))ds + G(t,x,X(T))\right] \tag{1.9}$$

with state dynamics given by

$$\begin{cases} dX(s) = b(s,X(s),u(s))ds + \sigma(s,X(s),u(s))dW(s), & s\in[t,T] \\ X(t) = x, \end{cases}$$

We notice that in the cost functional appearing in (1.9) the running cost and the terminal cost present a dependence on the initial pair (t,x) that the cost functional (1.2) did not present. This family of optimisation problems is much more complex and interesting than the one considered in the previous section because it allows to model those decision-making problems in which the perspective over the objective function varies as time flows and as the state of the system changes.

It can be shown that Bellman's principle of optimality does not hold for such collection of optimisation problems. Intuitively this is due to the fact that as time goes on and the state of the system evolves also the cost functional to optimize changes. If we interpret the dependence on the initial condition as a particular preference of the controller regarding the costs to be minimized, it is clear that at different times and states the decision maker will have different perspectives on the optimization problem. The controller at time $s>t$ has to face a different optimization problem with respect to the one faced at time t, a problem which is not simply the restriction of the latter on the interval $[s,T]$ but in which his/her new perception of the costs is taken into account.

If we try to solve optimisation problem (1.9) for a given pair (t,x) with the procedure developed in the previous section, treating (t,x) as fixed parameters, we come up with a control law which explicitly depends on the initial pair (t,x). In this case the optimal solution is said to be time-inconsistent. Time-inconsistency of the optimal solution means that the control law $\hat{u}$ which is the optimal solution to the optimization problem for the initial pair (t,x) is no longer optimal at some later point $(s,X(s))$. In other words, the restriction of the control law $\hat{u}$, optimal for the pair (t,x), on a following time interval $[s,T]$ does not coincide with the control function u' optimal for the pair $(s,X(s))$.

The optimization process is therefore less immediate and intuitive. The decision maker believes that he/she should follow an optimal decision rule which he/she has derived at time t when the state process is equal to x, but after time t his/her decision rule is no longer optimal and he/she should switch to a different decision rule.

The fact that Bellman's principle of optimality does not hold when facing optimisation problems of the kind (1.9) implies that the Hamilton-Jacobi-Bellman equation cannot be derived. We recall that the starting point for the derivation of the Hamilton-Jacobi-Bellman PDE was the recursive relation between the value functions over successive time intervals expressed by (1.3). We show that, due to the changes in the cost functional, that relation does not hold anymore. Let us denote by $\hat{u}_{t,x}$ the optimal control process for problem (1.9) relative to the initial pair (t, x). The corresponding value function can therefore be expressed as

$$V(t,x) = \mathbb{E}_{t,x}\left[\int_t^T C(t,x,s,X(s),\hat{u}_{t,x}(s))ds + G(t,x,X(T))\right]$$

If we split the time integral inside the expectation over the disjoint intervals $[t, t+h]$ and $(t+h, T]$ we can rewrite the same equivalence as

$$\begin{aligned} V(t,x) =& \mathbb{E}_{t,x}\left[\int_t^{t+h} C(t,x,s,X(s),\hat{u}_{t,x}(s))ds \right. \\ &\left. + \int_{t+h}^T C(t,x,s,X(s),\hat{u}_{t,x}(s))ds + G(t,x,X(T))\right] \end{aligned}$$

If the last two terms inside the expectation (conditioned on $\mathcal{F}_{t+h}$) represented the value function over the interval $[t+h, T]$ we would obtain the desired recursive relation and could proceed with the derivation of the HJB equation. Unfortunately this is not the case since the value function over $[t+h, T]$ depends now explicitly on the initial state $(t+h, X(t+h)$ and so does the relative optimal control process, which will be denoted by $\hat{u}_{t+h,X(t+h)}$. Hence we have

$$\begin{aligned} V(t+h, X(t+h)) =& \mathbb{E}\left[\int_{t+h}^T C(t+h, X(t+h), s, X(s), \hat{u}_{t+h,X(t+h)}(s))ds \right. \\ &\left. + G(t+h, X(t+h), X(T)) \middle| \mathcal{F}_{t+h}\right] \\ &\neq \mathbb{E}\left[\int_{t+h}^T C(t,x,s,X(s),\hat{u}_{t,x}(s))ds + G(t,x,X(T)) \middle| \mathcal{F}_{t+h}\right] \end{aligned}$$

Remark 1.2. *Notice that if the cost functional is independent of the initial state, that is if we are facing a time-consistent optimisation problem, the above equality is indeed verified.*

The fact that for the family of optimization problems (1.9) the principle of optimality does not hold poses the question of how to define the optimal control law. We analyze now two distinct approaches to deal with time-inconsistency united by the fact that both ignore the time-varying nature of the cost functional.

The first approach leads to the so called **pre-commitment** solution. We fix a point, for instance $(0, x_0)$ and try to find the control law $\hat{u}_{0,x_0}$ which minimizes

$$\mathcal{J}(0, x_0, u) = \mathbb{E}_{0,x_0}\left[\int_0^T C(0, x_0, s, X(s), u(s))ds + G(0, x_0, X(T))\right]$$

We accomplish this by solving the relative HJB equation with the running cost and the terminal cost evaluated and fixed at $(0, x_0)$, i.e. solving

$$\begin{cases} \dfrac{\partial V}{\partial t}(s, y) + \inf_{u \in U}\{C(0, x_0, s, y, u) + \mathcal{A}^u V(s, y)\} = 0, \quad (s, y) \in [0, T] \times \mathbb{R}^n \\ V(T, y) = G(0, x_0, y), \quad y \in \mathbb{R}^n \end{cases}$$

and defining $\hat{u}_{0,x_0}$ as

$$\hat{u}_{0,x_0}(t, x) = arg \min_{u \in U}\{C(0, x_0, t, x, u) + \mathcal{A}^u V(t, x)\} \quad (t, x) \in [0, T] \times \mathbb{R}^n$$

Apparently this could seem in contradiction with what we have previously asserted, namely that the Hamilton-Jacobi-Bellman equation cannot be derived for time-inconsistent problems. However, as we anticipated before this approach completely disregards the fact that the cost functional changes as time goes by. The optimisation problem, which is actually time-varying, is frozen at the preferences of the controller at time $t = 0$. In light of this, the dependence of the costs C and G on the initial conditions is taken out of the equation and we simply interpret $(0, x_0)$ in the functional as fixed parameters. This simplification allows us to obtain the classical HJB equation just as if we were facing a standard time-consistent stochastic optimal control problem.

The optimal control law $\hat{u}_{0,x_0}$ obtained with this method is hence optimal assuming that the time-preferences of the decision-maker do not change. The controller decides on a plan of action that is optimal today and commits to it, ignoring the incentives to revise it in the future. Obviously this approach does not provide an actual solution to the problem posed by time-inconsistency but rather tries to circumvent it.

The other approach we take into consideration provides a solution which is referred to as **naïve solution**. In this case at every time instant t we view our problem as the pre-committed problem and compute the optimal pre-committed control for today. At

$t + dt$ we look at a new pre-commited problem and so on. We are thus rolling over a continuously updated sequence of pre-committed optimal controls. The naïve agent fails to recognize the time inconsistency issue and his/her strategies are myopic and constantly changing.

In particular at time t, given that the state process is $X(t) = x$ the decision-maker uses $\hat{u}_{t,x}(t, x)$ where $\hat{u}_{t,x}$ is the control law obtained by solving

$$\begin{cases} \dfrac{\partial V}{\partial t}(s, y) + \inf\limits_{u \in U}\{C(t, x, s, y, u) + \mathcal{A}^u V(s, y)\} = 0, \quad (s, y) \in [0, T] \times \mathbb{R}^n \\ \qquad\qquad\qquad\qquad\qquad\qquad V(T, y) = G(t, x, y), \quad y \in \mathbb{R}^n \end{cases}$$

and defining $\hat{u}_{t,x}$ as

$$\hat{u}_{t,x}(s, y) = arg \min_{u \in U}\{C(t, x, s, y, u) + \mathcal{A}^u V(s, y)\} \quad (s, y) \in [0, T] \times \mathbb{R}^n$$

This procedure is iterated at every time instant in the interval $[0, T]$. While the naïve solution may appear to take into account the time-varying nature of the cost functional as the solution is constantly updated, this is not the case. Actually the naive solution to the optimization problem for the initial pair (t, x) is derived under the assumption that all future controllers will use the same costs $C(t, x, \cdot, \cdot, \cdot)$ and $G(t, x, \cdot)$ as the controller who search for the optimal decision rule at time t. Hence, future changes in the costs are still not modelled.

1.4. Example: Exponential utility maximization

We present now an example of time-consistent stochastic optimal control problem, taken from financial mathematics, to illustrate how Hamilton-Jacobi-Bellman equation is used in practice to determine the value function and the optimal control law. In particular we follow the steps presented at the end of section 1.2. After that we will change the form of the utility functional, making it dependent on the initial state, i.e. time-inconsitent and we will solve the problem exploiting the two methodologies explained above, namely the pre-commitment and the naïve.

We consider a financial market consisting of two assets: a risk free asset $B = (B(t), 0 \leq t \leq T)$ which can be interpreted as the price of a bond, assuming the short rate on interest r being a constant and a risky asset $S = (S(t), 0 \leq t \leq T)$, interpreted as the price of a stock. Furthermore, we assume that the value of the bond price is constant, $B(t) = 1$ for all $t \in [0, T]$, i.e. we assume that the short rate of interest is zero.

We model the price of the risky stock with the geometric Brownian motion

$$dS(t) = S(t)\mu dt + S(t)\sigma dW(t) \quad t \in [0, T]$$

The problem we face is that one of an investor who, starting from an initial wealth of x_0, needs to decide how to allocate it optimally among the two assets in order to maximize his/her expected utility associated to the final wealth reached. Therefore, the investor sets up a self-financing portfolio, without consumption, whose generated wealth is denoted by X^π, where π denotes the amount of money invested in the risky asset. The dynamics of the wealth process is given by the following SDE

$$\begin{cases} dX^\pi(t) = \pi(t)(\mu dt + \sigma dW(t)), \quad s \in [0, T] \\ X(0) = x_0, \end{cases}$$

The amount of money invested in the stock π is the control process through which the decision-maker can steer the state process, i.e. the wealth process.

The goal of the investor is to maximize his/her expected exponential utility from the terminal wealth so the optimization problem is

$$\sup_{\pi \in \mathcal{U}} \mathbb{E}\left[-e^{-\gamma X^\pi(T)}\right]$$

where γ denotes the risk aversion coefficient of the investor.

The Hamilton-Jacobi-Bellman equation for this optimization problem becomes

$$\begin{cases} \frac{\partial V}{\partial t}(t, x) + \sup_{\pi \in U}\{\pi\mu \frac{\partial V}{\partial x}(t, x) + \frac{1}{2}\pi^2\sigma^2 \frac{\partial V^2}{\partial^2 x}\} = 0, \quad (t, x) \in [0, T] \times \mathbb{R} \\ \qquad V(T, y) = -e^{-\gamma x}, \quad x \in \mathbb{R} \end{cases}$$

To begin with, we tackle the static optimization problem. Assuming V as given we maximize with respect to π the quantity

$$\pi\mu \frac{\partial V}{\partial x}(t, x) + \frac{1}{2}\pi^2\sigma^2 \frac{\partial V^2}{\partial^2 x}$$

and assuming an interior solution, by the first order condition we find

$$\pi^*(t,x,V) = \frac{-\mu V_x(t,x)}{\sigma^2 V_{xx}(t,x)}$$

We see that in order to implement the optimal investment plan we need to know the value function V. We therefore make an ansatz for V assuming $V(t,x) = -e^{h(t)x+k(t)}$. For such value function we have

$$V_t = -(h'x+k')e^{hx+k}, \quad V_x = -he^{hx+k}, \quad V_{xx} = -h^2 e^{hx+k}$$

where we have used the shorthand notation $h = h(t), k = k(t)$ and the apex to denote the time derivative. Now we can substitute the derivatives into the optimal control law obtaining

$$\pi^*(t,x) = -\frac{\mu}{\sigma^2 h}$$

Inserting the above control law into the HJB equation we end up with

$$-(h'x+k')e^{hx+k} - \frac{\mu}{\sigma^2 h}\mu(-he^{hx+k}) + \frac{1}{2}\frac{\mu^2}{\sigma^4 h^2}\sigma^2(-h^2 e^{hx+k}) = 0$$

which can be written as

$$e^{hx+k}(-h'x - k' + \frac{\mu^2}{2\sigma^2}) = 0$$

Taking into account the final condition for the value function we know that $h(T) = -\gamma$ and $k(T) = 0$, thus the functions h and k are the solutions to

$$\begin{cases} -h'x = 0 \\ h(T) = -\gamma \end{cases} \qquad \begin{cases} k' - \dfrac{\mu^2}{2\sigma^2} = 0 \\ k(T) = 0 \end{cases}$$

which result in $h(t) = -\gamma$ and $k(t) = \frac{\mu^2}{2\sigma^2}(t-T)$. The optimal control law is thus given by $\pi^* = \frac{\mu}{\sigma^2\gamma}$.

The hypothesis of a constant risk aversion coefficient is however very strong and does not fully represent the reality. Many observations suggest instead that the risk aversion coefficient of an investor is greatly influenced by the conditions of the economy, see [5]. In a market that is on the rise and where the conditions of the economy are generally favorable, i.e. a bull market, investors are willing to take more risk. For example, in the case of equity markets, a bull market denotes a context characterized by the rise in the prices of companies' shares. In such times, investors often have faith that the uptrend will continue over the long term. This optimistic perspective should be modeled with a lower risk aversion coefficient. On the other hand, in a market in decline, where share prices are continuously dropping, investors are willing to take less risk, which should be modeled with a higher risk aversion coefficient.

Also, empirical evidence suggests that the risk aversion attitude depends on prior gains and losses and more generally on current wealth. Following a win on a previous bet investors are more risk seeking than usual, while after a loss they become more risk averse; see [11] for a deeper understanding of the effect. Consequently, for a description more relevant to reality, we should investigate a portfolio selection problem for an investor with risk aversion coefficient depending on the investor's current wealth, that is

$$\sup_{\pi \in \mathcal{U}} \mathbb{E}\left[-e^{-\gamma(x)X^{\pi}(T)}\right]$$

where $\gamma : \mathbb{R} \mapsto (0, \infty)$ is a map on wealth. Clearly, the optimal control law for an investor maximizing the expected exponential utility from terminal wealth is now time-inconsistent since it depends on the risk aversion coefficient which varies with wealth. The optimal control law for an investor who at time t possesses a wealth x and has risk aversion coefficient $\gamma(x)$ will not be optimal anymore when at time $s > t$ the wealth process changes and the risk aversion coefficient becomes $\gamma(X^{\pi}(s))$.

The pre-commitment solution for an investor starting at time $t = 0$ with wealth x_0 is given by $\pi^*(t) = \frac{\mu}{\gamma(x_0)\sigma^2}$. As explained earlier, this solution derives from the assumption that the investor freezes his/her risk aversion coefficient at the value $\gamma(x_0)$ neglecting its evolution as the wealth changes.

On the other side, the naïve control law is given by $\pi^*(t) = \frac{\mu}{\gamma(X^{\pi^*}(t))\sigma^2}$. The naïve solution tries to bond together the strategies which are optimal for all controllers.

The advantage of the pre-commitment and naïve solution is that they are derived by solving classical HJB equations and are based on the notion of optimality described in Section 1.2. However, both the pre-commitment and the naive solution completely ignore the key feature of time-inconsistent optimization problems which is the time-varying nature of the cost/utility functional.

2 | Time-inconsistent optimal controls

2.1. Problem formulation and hypotheses

As we have shown in the previous chapter and, in particular, in the example developed, both the pre-commitment and naïve approaches do not take properly into account the time-changing feature of the cost functional. On the other hand we would like to find the optimal strategy for the controller at time t, characterized by costs of the form $C(t, x, \cdot, \cdot, \cdot)$ and $G(t, x, \cdot)$ assuming that he/she knows that the future controllers will have different preferences, depending on the future state of the system and time and most likely will apply different decision rules, according to their new costs. Such a solution is called sophisticated and it requires a different concept of optimality which will be made clear in the sequel. In this chapter we aim to find a sophisticated solution adopting a game-theoretic approach to the stochastic optimization problem.

Before starting the actual description of the game-theoretic approach to the stochastic optimal control problem let us describe briefly some changes in the modeling assumptions required in order to produce a sophisticated solution. First, we consider from now on a complete filtered probability space $(\Omega, F, \mathbb{F}, \mathbb{P})$ where $\mathbb{F}$ is the natural filtration generated by the d-dimensional Brownian motion $W(\cdot)$. This choice is necessary because in the following we will exploit some results on the existence of the solution to some BSDEs which are based on the martingale representation theorem.
Moreover, we restrict the class of time-inconsistent problems that we are studying: in section 1.3. of chapter 1 we have introduced time-inconsistent optimal control problems in a very general form, considering a cost functional where the cost functions depends on both initial time and initial state of the system, i.e. of the type:

$$\mathcal{J}(t, x, u) = \mathbb{E}_{t,x}\left[\int_t^T C(t, x, s, X(s), u(s))ds + G(t, x, X(T))\right]$$

where the state dynamics, which we report again for convenience, is the same introduced in the first chapter

$$\begin{cases} dX(s) = b(s, X(s), u(s))ds + \sigma(s, X(s), u(s))dW(s), \quad s \in [t, T] \\ X(t) = x, \end{cases} \tag{2.1}$$

From now on we will restrict ourselves to those optimisation problems in which the cost functional is allowed to depend only on the initial time. Moreover we will adopt henceforth the following formulation for the cost functional

$$\mathcal{J}(t, x, u) = \mathbb{E}_t\left[\int_t^T C(t, s, X(s), u(s))ds + G(t, X(T))\right] \tag{2.2}$$

where $\mathbb{E}_t[\,\cdot\,] = \mathbb{E}[\,\cdot \mid \mathcal{F}_t]$.

Remark 2.1. *It is important to remark that this new formulation of the cost functional does not prevent us from using the results and theorems derived in the previous chapter, all of which still apply to optimisation problems formulated in terms of functionals of the form* (2.2)*. Actually, the two formulations coincide. To be precise, given that the coefficients b and σ of the SDE* (2.1) *are deterministic functions of t,x and u we know that the strong solution is adapted to the natural filtration and the Markov property holds. Thanks to this property we know that for any Borelian bounded function g on $\mathbb{R}^n$ and for all $t \leq \theta$ in $[0, T]$ we have*

$$\mathbb{E}\left[g(X(\theta)) \mid \mathcal{F}_t\right] = \nu_\theta(t, X(t))$$

where ν_θ is the function over $[0, T] \times \mathbb{R}^n$ defined by

$$\nu_\theta(t, x) = \mathbb{E}_{t,x}\left[g(X(\theta))\right]$$

where the subscripts in $\mathbb{E}_{t,x}(\cdot)$ denotes as usual that the state dynamics is the one given by (2.1)*. Basically, when the state process is Markovian, conditioning with respect to the information generated up to time t does not add information to the only knowledge of the state of the system at time t.*

Functionals of the form (2.2) frequently appear in optimisation problems in the field of finance and insurance where the discounting functions for future cash flows enter the objective functional. Traditionally, cash flows have always been discounted with exponential functions with constant rate, giving rise to problems of the type

$$\sup_{u\in\mathcal{U}} \mathbb{E}_t\left[\int_t^T e^{-\rho(s-t)}C(X(s),u(s))ds + e^{-\rho(T-t)}G(X(T))\right]$$

which are equivalent to

$$\sup_{u\in\mathcal{U}} \mathbb{E}_t\left[\int_t^T e^{-\rho s}C(X(s),u(s))ds + e^{-\rho T}G(X(T))\right]$$

having factored out the t-term in the exponential function. Such discounted utilities do not depend explicitly on the initial time, thus the optimal solution is time-consistent. However, when we adopt an exponential function as discounting term we assume that the investor assigns the same weighting factor at time t_1 and at time $t_2 > t_1$ to value the cash flow at time $t_3 > t_2$. As a consequence the optimal decision rule at time t remains optimal at a later time $s > t$. Such an assumption is quite strong and indeed experimental studies have shown that people's preferences over future cash flows are not constant but rather time-varying. It is a well known fact that, for instance, people prefer two oranges in 21 days to one orange in 20 days, but they also prefer one orange now to two oranges tomorrow; the discount applied to the same delay of 1 day is different for different starting times. This shall lead investors to be more patient in choices regarding distant future consumption but impulsive when considering near-term consumption. Similarly, this attitude shall prevent them to take advantage of very promising opportunities only because they require a small delay in the near term, potentially regretting it in the future. Such a behaviour is called the common difference effect and it can not be modelled with exponential discounting functions. In the economic literature, see [6] and [7], we can find evidence that people discount the future income with non-constant rates of time preferences and the real-life rates of time preference tend to decline in time. In other words, people's valuation tends to decrease rapidly for short period delays and less rapidly for longer period delays. Such feature can be described for example by hyperbolic discounting or other non-exponential discounting functions. Generically, the discounting terms $e^{-\rho(s-t)}$ and $e^{-\rho T}$ might be replaced by some functions $\lambda(t,s)$ and $\nu(t,T)$ thus leading to the problem

$$\sup_{u\in\mathcal{U}} \mathbb{E}_t\left[\int_t^T \lambda(t,s)C(X(s),u(s))ds + \nu(t,T)G(X(T))\right]$$

for which the functional is of type (2.2) and the dependence of the discounted utilities on time t cannot be removed leading to a time-inconsistent optimization problem.

In order to formulate as precisely as possible the time-inconsistent stochastic control problem we want to make more explicit the class of admissible control processes that we introduced in Chapter 1. This class of processes was described by Definition 1.1 in section 1.2 as the family of processes for which there exist a strong solution to the state SDE (2.1). However, the existence of a strong solution to the controlled state equation can be guaranteed by posing some measurability hypothesis on the class of control processes and some hypothesis on the form of the drift and diffusion of the state SDE. We define the class of admissible control processes as:

$$\begin{aligned}\mathcal{U}[t,T] = \Big\{ u : [t,T]\times\Omega \mapsto U \subseteq \mathbb{R}^m \;\Big|\; & u(\cdot) \text{ is } \mathbb{F}\text{-progressively measurable,} \\ & \mathbb{E}\int_t^T |u(s)|^2 < \infty \Big\}\end{aligned} \tag{2.3}$$

where $U \subseteq \mathbb{R}^m$ is a closed subset, which could be either bounded or unbounded. Then we introduce the following assumption

Assumption 2.1. *The maps $b : [0,T]\times\mathbb{R}^n\times U \mapsto \mathbb{R}^n$, $\sigma : [0,T]\times\mathbb{R}^n\times U \mapsto \mathbb{R}^{n\times d}$ are continuous and there exist a constant L such that*

$$\begin{cases} |b(t,x,u) - b(t,y,u)| \leq L|x-y|, \\ |\sigma(t,x,u) - \sigma(t,y,u)| \leq L|x-y|, \end{cases} \qquad (t,u)\in[0,T]\times U, \;\; x,y\in\mathbb{R}^n \tag{2.4}$$

and

$$|b(t,0,u)| + |\sigma(t,0,u)| \leq L(1+|u|) \qquad (t,u)\in[0,T]\times U \tag{2.5}$$

Condition (2.4) requires the coefficients b and σ to be Lipschitz continuous in x uniformly in t and u, whereas condition (2.5) combined with (2.4) requires that they have a sublinear growth at infinity. Indeed,

$$\begin{aligned}&|b(t,x,u)| + |\sigma(t,x,u)| = |b(t,x,u) \pm b(t,0,u)| + |\sigma(t,x,u) \pm \sigma(t,0,u)| \\ &\leq |b(t,x,u) - b(t,0,u)| + |b(t,0,u)| + |\sigma(t,x,u) - \sigma(t,0,u)| + |\sigma(t,0,u)| \\ &\leq L|x-0| + L|x-0| + L(1+|u|) \\ &= L(1+|x|+|u|)\end{aligned}$$

The following standard result concerns the well-posedness of the system state equation (2.1).

Proposition 2.1. *Let Assumption 2.1 hold, then for any* $(t,x) \in [0,T] \times \mathbb{R}^n$ *and* $u(\cdot) \in \mathcal{U}[t,T]$*, state equation* (2.1) *admits a unique strong solution* $X(\cdot) = X(\cdot,t,x,u(\cdot))$ *and the following estimate holds:*

$$\mathbb{E}|X(s,t,x,u(\cdot))|^2 \leq K(1+|x|^2+\mathbb{E}\int_t^s |u(r)|^2 dr), \qquad s \in [t,T] \tag{2.6}$$

We proceed with an hypothesis on the functions which appear in the cost functional (2.2) in order to have a functional which is finite for any admissible control process.

Remark 2.2. *We will denote by* $D[0,T]$ *the set of time pairs* $\{(t,s) \in [0,T]^2 \mid 0 \leq t \leq s \leq T\}$*. This notation is useful for the description of the domain of definition of the running cost function* C*.*

Assumption 2.2. *The maps* $C : D[0,T] \times \mathbb{R}^n \times U \mapsto \mathbb{R}$ *and* $G : [0,T] \times \mathbb{R}^n \mapsto \mathbb{R}$ *are continuous and there exist a constant* L *such that*

$$\begin{cases} 0 \leq C(\tau,t,x,u) \leq L(1+|x|^2+|u|^2), \\ 0 \leq G(\tau,x) \leq L(1+|x|^2), \end{cases} \qquad (\tau,t,x,u) \in D[0,T] \times \mathbb{R}^n \times U \tag{2.7}$$

Thanks to the a priori estimate (2.6) we see that for any $(t,x) \in [0,T] \times \mathbb{R}^n$ and $u(\cdot) \in \mathcal{U}[t,T]$ it holds:

$$\begin{aligned} \mathbb{E}|C(\tau,s,X(s),u(s))| &\leq L(1+\mathbb{E}|X(s)|^2+\mathbb{E}|u(s)|^2) \\ &\leq L(1+|x|^2+\mathbb{E}|u(s)|^2+\mathbb{E}\int_t^s |u(r)|^2 dr) \end{aligned}$$

and

$$\mathbb{E}|G(\tau,X(T))| \leq L(1+\mathbb{E}|X(T)|^2) \leq L(1+|x|^2+\mathbb{E}\int_t^T |u(r)|^2 dr)$$

so that the cost functional $\mathcal{J}(t,x,u)$ is finite for any $u(\cdot) \in \mathcal{U}[t,T]$.

Before stating rigorously the time-inconsistent optimal control problem which we are going to study in the subsequent sections let us introduce some notations which will be used extensively from now on. Let $\mathbb{S}^n$ be the set of all $(n \times n)$ symmetric real matrices. We recall that

$$a(t,x,u) = \frac{1}{2}\sigma(t,x,u)\sigma(t,x,u)^T, \qquad (t,x,u) \in [0,T] \times \mathbb{R}^n \times U$$

and define the Hamiltonian

$$\mathbb{H}(\tau,t,x,u,p,P) = \langle\, b(t,x,u), p\,\rangle + tr[a(t,x,u)P] + C(\tau,t,x,u), \qquad (\tau,t,x,u) \in D[0,T] \times \mathbb{R}^n \times U \times \mathbb{R}^n \times \mathbb{S}^n \tag{2.8}$$

where again we use the notation $\langle \cdot, \cdot \rangle$ to denote the scalar product and $tr[\cdot]$ to denote the trace of a square matrix.

Remark 2.3. *Notice that the HJB equation (Theorem 1.1) can be equivalently formulated exploiting the Hamiltonian function. Indeed, since*

$$\mathcal{A}^u = \sum_{i=1}^{n} b_i(t,x,u)\frac{\partial}{\partial x_i} + \sum_{i,j=1}^{n} a_{ij}(t,x,u)\frac{\partial^2}{\partial x_i \partial x_j} = \langle\, b(t,x,u), p\,\rangle \; + \; tr[a(t,x,u)P]$$

the optimisation part in the equation can be written as

$$\inf_{u\in U}\{C(t,x,u) + \mathcal{A}^u V(t,x)\} = \inf_{u\in U} \mathbb{H}(t,x,u,V_x(t,x),V_{xx}(t,x))$$

where we have used the same $\mathbb{H}$ defined in (2.8) ignoring τ and where V_x and V_{xx} denote respectively the Gradient and the Hessian of the value function. From now on we will formulate the relevant HJB equations in terms of the Hamiltonian function.

As we showed at the end of Chapter 1. a fundamental step to tackle the HJB equation consists of finding a candidate optimal control law by solving the optimisation part of the equation. To make sure that this step can be carried out in the time-inconsistent context in which we will operate we formulate the following assumption

Assumption 2.3. *There exists a map $\psi : D[0,T] \times \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{S}^n \mapsto U$ with needed regularity such that*

$$\psi(\tau,t,x,p,P) \in argmin\, \mathbb{H}(\tau,t,x,\cdot,p,P)$$
$$\equiv \Big\{ \bar{u} \in U \;\Big|\; \mathbb{H}(\tau,t,x,\bar{u},p,P) = \min_{u\in U} \mathbb{H}(\tau,t,x,u,p,P) \Big\},$$
$$(\tau,t,x,p,P) \in D[0,T] \times \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{S}^n$$

This hypothesis is needed since $\inf_{u\in U} \mathbb{H}(\tau,t,x,u,p,P)$ is not necessarily finite on the whole space $D[0,T] \times \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{S}^n$ when U is unbounded. Note that when U is bounded, since it was assumed to be closed, it is also compact and the above infimum is finite over the entire space.

We can now formally state the time-inconsistent optimal control problem that we will analyze in details.

For any initial pair $(t,x) \in [0,T) \times \mathbb{R}^n$, find a control process $\bar{u}(\cdot) \in \mathcal{U}[t,T]$ such that

$$\mathcal{J}(t,x,\bar{u}) = \inf_{u\in\mathcal{U}[t,T]} \mathcal{J}(t,x,u)$$

with $\mathcal{J}$ defined in (2.2).
In the next section we will present the game theoretic-approach to time inconsistency which will allow us to build a sophisticated solution to the above stochastic control problem.

2.2. Game-theoretic approach to time-inconsistency

Let us interpret the control problem as a game played by a continuum of agents over the time interval $[0,T]$. The t-agent only chooses the strategy at time t. When the costs are constant, i.e. independent of the specific state and time at which the optimization process is performed, the future agents will simply solve the remaining part of the same optimization problem faced by the agent at time t. This is what happens in a time-consistent optimization problem. However, in a time-inconsistent problem where the players' costs are time-varying, the residual players will not solve the restriction of the same optimization problem faced by the agent at time t, since the cost functional changes constantly. In particular, the decision maker, given that at time t the state is x, aims at minimizing the cost functional:

$$\mathcal{J}(t,x,u) = E_{t,x}\left[\int_t^T C(t,s,X(s),u(s))ds + G(t,X(T))\right]$$

In this game-theoretic framework, the value function of the problem at time t depends not only on the strategy chosen by the t-agent, but also on the strategies chosen by all future agents. If the t-agent follows the naïve approach and looks for the best strategy according to his/her current preferences, i.e. solving the optimisation problem with fixed costs $C(t,\cdot,\cdot,\cdot)$, $G(t,\cdot)$ exploiting dynamic programming, then his/her optimal strategy will not be adopted by the future players. Consequently, the true value function of the agent at time t will be higher than the one resulting from his/her naïve optimisation. What the t-agent should do is to sacrifice his/her short-term utility in favor of a longer-term benefit. In light of this it seems reasonable to look for a control strategy which constitutes a sub-game perfect Nash equilibrium.
In game theory a Nash equilibrium is a solution concept of a non cooperative game involving multiple players in which each player knows the equilibrium strategies of the other players, and every player does not benefit by changing only his/her own strategy. In other words, if each player has chosen a strategy and no player can improve his/her reward by changing the strategy while the other players keep their strategies unchanged, then the current set of strategies constitutes the Nash equilibrium. A sub-game perfect equilibrium is a refinement of a Nash equilibrium used in sequential games. A strategy profile is a sub-game perfect equilibrium if it represents a Nash equilibrium of every sub-game of the original game. Intuitively, this means that at any point in the game, the players' behavior from that point onward should represent a Nash equilibrium of the continuation game (i.e. of the sub-game), no matter what happened before.

Now, starting from the description of the game theoretic approach to a discrete optimization problem, which is a simpler situation, we arrive to a continuous time problem which is the one of our interest. In a discrete time optimisation problem the controls are chosen at discrete times $0 = t_0, t_1, ..., t_{N-1}$ with $t_N = T$ and kept fixed in between. The cost functional in this case has the form of a summation instead of an integral

$$\mathcal{J}(t_n, x, \pi) = \sum_{k=n}^{N-1} C(t_n, t_k, X(t_k), \pi_{t_k}) + G(t_n, X(T))$$

The Nash equilibrium is the best response to all other control strategies in that equilibrium and it can be determined in a backward induction fashion.

- Let the last agent to take decision, i.e. the agent at time t_{N-1}, optimize the cost functional $\mathcal{J}(t_{N-1}, x, \pi_{t_{N-1}})$ over the admissible controls $\pi_{t_{N-1}}$ at time t_{N-1} for every $x \in \mathbb{R}^n$.
- Let the agent at time t_{N-2} optimize the cost functional $\mathcal{J}(t_{N-2}, x, \pi_{t_{N-2}})$ over the

admissible controls $\pi_{t_{N-2}}$ at time t_{N-2} for every $x \in \mathbb{R}^n$, assuming that the agent at time t_{N-1} will adopt his/her optimal control determined at the previous step.

- Proceed recursively by induction

We can formalize the definition of a Nash equilibrium strategy in a discrete time optimisation problem

Definition 2.1. *Let us consider a control strategy $\hat{\pi} = (\hat{\pi}_{t_0}, \hat{\pi}_{t_1}, ..., \hat{\pi}_{t_{N-1}})$. Consider an arbitrary point $(t_k, x) \in \{t_0, t_1, ..., t_{N-1}\} \times \mathbb{R}^n$ and an arbitrary control law π admissible at time t_k. Define a new strategy*

$$\pi^* = \begin{cases} \pi_t, & t = t_k \\ \hat{\pi}_t, & t \in \{t_{k+1}, ..., t_{N-1}\} \end{cases}$$

If

$$\inf_{\pi \in \mathcal{U}} \mathcal{J}(t, x, \pi^*) = \mathcal{J}(t, x, \hat{\pi})$$

then $\hat{\pi}$ is called a Nash equilibrium control strategy and $\mathcal{J}(t, x, \hat{\pi})$ gives the Nash equilibrium value function of the optimization problem.

Clearly, we can not apply the above definition in a continuous time problem, since a change in the control law at a single time t does not affect neither the controlled state dynamics neither the cost functional to be minimized. To deal with continuous time optimisation problems we adopt the following approach: we introduce a sequence of differential games, each one associated to a different partition of the interval $[0, T]$ in N sub-intervals, with N arbitrary natural number. Formally, let Π the partition defined by:

$$\Pi : 0 = t_0 < t_1 < ... < t_{N-1} < t_N = T$$

with mesh size $\|\Pi\|$ given by

$$\|\Pi\| = \max_{1 \leq k \leq N} (t_k - t_{k-1})$$

The differential game associated to partition Π consists of N players, each one controlling the system on a distinct time interval and with the goal of minimizing a sophisticated cost functional built by using some properties of forward-backward stochastic differential equations. In such differential game the interaction among the agents is the following:

- The pair $(t_k, X(t_k))$, which is the terminal pair of player k who controls the system on $[t_{k-1}, t_k)$, is the initial pair for player $k+1$ who controls the state on the following time interval $[t_k, t_{k+1})$.

- All the players know that each player looks for an optimal control relative to his/her own problem.
- Each player will maintain his/her own cost preferences even over the time intervals in which the other players control the system.

Let us assume that each player has an optimal control process over his/her time interval of influence, denoted by $\bar{u}^k$ for all $k \in \{1, 2, ..., N\}$ as well as his/her own value function $V^k(\cdot, \cdot)$ defined on $[t_{k-1}, t_k] \times \mathbb{R}^n$. We can therefore define the Nash equilibrium control $\bar{u}^\Pi(\cdot)$ and the Nash equilibrium value function $V^\Pi(\cdot, \cdot)$ relative to the Π-differential game as follows:

$$\bar{u}^\Pi(t) = \sum_{k=1}^{N} \bar{u}^k(t) \mathbb{1}_{[t_{k-1}, t_k)}(t), \quad t \in [0, T)$$

and

$$V^\Pi(t, x) = \sum_{k=1}^{N} V^k(t, x) \mathbb{1}_{[t_{k-1}, t_k)}(t), \quad (t, x) \in [0, T) \times \mathbb{R}^n$$

If the following limits exist

$$\begin{cases} \lim_{\|\Pi\| \to 0} \bar{u}^\Pi(t) = \bar{u}(t), & t \in [0, T] \\ \lim_{\|\Pi\| \to 0} V^\Pi(t, x) = V(t, x), & (t, x) \in [0, T] \times \mathbb{R}^n \end{cases}$$

for some process $\bar{u} \in \mathcal{U}[0, T]$ and for some function $V : [0, T] \times \mathbb{R}^n \mapsto \mathbb{R}$ we call them time-consistent equilibrium control and time-consistent equilibrium value function of the time-inconsistent optimisation problem.

In particular, we give the following definition:

Definition 2.2. *A continuous map $\Psi : [0, T] \times \mathbb{R}^n \mapsto U$ is called a time-consistent equilibrium strategy of the time-inconsistent optimisation problem if for any $(t, x) \in [0, T] \times \mathbb{R}^n$,*

$$\begin{cases} d\bar{X}(s) = b(s, \bar{X}(s), \Psi(s, \bar{X}(s)))ds + \sigma(s, \bar{X}(s), \Psi(s, \bar{X}(s)))dW(s), & s \in [t, T] \\ \bar{X}(t) = x, \end{cases}$$

admits a unique solution $\bar{X}(\cdot) \equiv \bar{X}(\cdot, t, x, \Psi)$ *such that for any partition* $\Pi = \{t_k, 0 \le k \le N\}$,

$$\begin{aligned}
&\mathcal{J}\big(t_{k-1}, \bar{X}(t_{k-1}, 0, x, \Psi), \Psi\big|_{[t_{k-1},T]}\big) \\
&\le \mathcal{J}\big(t_{k-1}, \bar{X}(t_{k-1}, 0, x, \Psi), u^k \oplus \Psi\big|_{[t_k,T]}\big) + R(\|\Pi\|), \\
&\qquad\qquad\qquad\qquad\qquad\qquad a.s., \quad \forall u^k \in \mathcal{U}[t_{k-1}, t_k]
\end{aligned}$$

where $R(\|\Pi\|)$ *is a generic remainder term satisfying* $R(\|\Pi\|) \to 0$ *as* $\|\Pi\| \to 0$, *and*

$$\big(u^k \oplus \Psi\big|_{[t_k,T]}\big)(t) = \begin{cases} u^k(t), & t \in [t_{k-1}, t_k) \\ \Psi(t, X^k(t)), & t \in [t_k, T] \end{cases}$$

with

$$\begin{cases} dX^k(s) = b(s, X^k(s), u^k(s))ds + \sigma(s, X^k(s), u^k(s))dW(s), & s \in [t_{k-1}, t_k) \\ dX^k(s) = b(s, X^k(s), \Psi(s, X^k(s)))ds + \sigma(s, X^k(s), \Psi(s, X^k(s)))dW(s), & s \in [t_k, T] \\ X^k(t_{k-1}) = \bar{X}(t_{k-1}, 0, x, \Psi) \end{cases}$$

In this case, the process $\bar{X}(\cdot, 0, x, \Psi)$ is called **time-consistent equilibrium state process**, the control $\bar{u}(\cdot) \equiv \Psi(\cdot, \bar{X}(\cdot))$ is called **time-consistent equilibrium control** for the initial state x and as a pair $(\bar{X}(\cdot), \bar{u}(\cdot))$ is called a **time-consistent equilibrium pair**. Moreover, the function $V : [0,T] \times \mathbb{R}^n \mapsto \mathbb{R}$ is called the **time-consistent equilibrium value function** of the time-inconsistent problem if:

$$V(t, \bar{X}(t, 0, x, \Psi)) = \mathcal{J}(t, \bar{X}(t, 0, x, \Psi), \Psi\big|_{[t,T]}), \qquad a.s., \quad t \in [0,T]$$

Note that the definition of time-consistent equilibrium strategy we just gave is of closed loop type, since the control function is a feedback law, i.e. at time t it depends only on the value of the state at that time instant.

In words, what definition 2.2 tells us is that once a time-consistent equilibrium strategy is found it is never convenient to deviate from it on any time interval of whatever partition of $[0, T]$. Consequently, the value function at any time is obtained using the time-consistent equilibrium control $\bar{u}$, restricted from that instant onward.

2.3. The forward-backward system

As we anticipated in the previous section, a key step in the game theoretical approach to the optimal control problem consists of minimizing, for each player involved in the N-game, a sophisticated cost functional built using some techniques of forward-backward stochastic differential equations. In this section we provide a brief introduction to backward stochastic differential equations and we present a result which will be exploited extensively in the iterative procedure that we will adopt to solve the differential game later on.

Let $W = (W(s))_{0\leq s\leq T}$ be the standard d-dimensional Brownian motion we considered so far, defined on the filtered probability space $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P})$ where $\mathbb{F}$ is the natural filtration of $W(\cdot)$ and T is a finite time horizon.
We denote by $\mathbb{S}^2(0,T)$ the set of all real-valued progressively measurable processes $Y(\cdot)$ such that

$$\mathbb{E}\big[\sup_{0\leq s\leq T} |Y(s)|^2\big] < \infty$$

and by $\mathbb{H}^2(0,T)^d$ the set of $\mathbb{R}^d$-valued progressively measurable processes $Z(\cdot)$ such that

$$\mathbb{E}\bigg[\int_0^T |Z(s)|^2 ds\bigg] < \infty$$

Given a pair (ξ, f) called the terminal condition and generator and the following hypothesis

Assumption 2.4. *Let (ξ, f) satisfy*

- *ξ is a real-valued random variable belonging to $L^2(\Omega, \mathcal{F}_T, \mathbb{P})$.*
- *$f : \Omega \times [0,T] \times \mathbb{R} \times \mathbb{R}^d \mapsto \mathbb{R}$ such that:*
  - *$f(\cdot, t, y, z)$ is progressively measurable for all y, z.*
  - *$f(\cdot, t, 0, 0) \in \mathbb{H}^2(0,T)$.*
  - *f satisfies a uniform Lipschitz condition in (y,z), i.e. there exists a constant C s.t*

$$|f(t,y_1,z_1) - f(t,y_2,z_2)| \leq C(|y_1 - y_2| + |z_1 - z_2|), \quad a.s.\ \forall\, y_1, y_2,\ \ \forall\, z_1, z_2,\ \ \forall\, t$$

we consider the following backward stochastic differential equation

$$\begin{cases} dY(s) = -f(s, Y(s), Z(s))ds + Z(s)dW(s), & s \in [0, T] \\ Y(T) = \xi \end{cases} \tag{2.9}$$

Finding a solution to the above BSDE means to find a pair of processes $(Y(\cdot), Z(\cdot)) \in \mathbb{S}^2(0, T) \times \mathbb{H}^2(0, T)^d$ which satisfy

$$Y(t) = \xi + \int_t^T f(s, Y(s), Z(s))ds - \int_t^T Z(s)dW(s) \qquad a.s., \ \ 0 \leq t \leq T$$

The following standard theorem provides an existence and uniqueness result for the above BSDE.

Theorem 2.1. *Given a pair (ξ, f) satisfying Assumption 2.4, there exists a unique solution $(Y(\cdot), Z(\cdot))$ to the BSDE (2.9).*

The proof, which makes use of a fixed point method, can be found in [9].

In the sequel we will deal with BSDE where the terminal condition and the generator have a particular form, depending on another stochastic process $X(\cdot)$ given in advance. In particular we suppose $X(\cdot)$ is a process with values on $\mathbb{R}^n$ and belonging to $\mathbb{S}^p(0, T)$ for every $p \in [1, \infty)$. We consider therefore the following BSDE

$$\begin{cases} dY(s) = \upsilon(s, X(s), Y(s), Z(s))ds + Z(s)dW(s), & s \in [0, T] \\ Y(T) = \phi(X(T)) \end{cases} \tag{2.10}$$

where $\upsilon : [0, T] \times \mathbb{R}^n \times \mathbb{R} \times \mathbb{R}^d \mapsto \mathbb{R}$ and $\phi : \mathbb{R}^n \mapsto \mathbb{R}$ are given Borel-measurable functions. Again we look for a solution pair $(Y(\cdot), Z(\cdot)) \in \mathbb{S}^2(0, T) \times \mathbb{H}^2(0, T)^d$. Let us assume the following

Assumption 2.5. *For every $s \in [0, T]$, $x \in \mathbb{R}^n$, $y, y' \in \mathbb{R}$, $z, z' \in \mathbb{R}^d$, we have, for some constants $K \geq 0$ and $m \geq 0$,*

- $|\upsilon(s, x, y, z) - \upsilon(s, x, y', z')| \leq K(|y - y'| + |z - z'|)$
- $|\phi(x)| + |\upsilon(s, x, 0, 0)| \leq K(1 + |x|^m)$

Hence setting $f(t, y, z) = \upsilon(t, X(t), y, z)$ and $\xi = \phi(X(T))$ we can check that ξ and f satisfy the conditions of Assumption 2.4, thus there exists a unique solution $(Y(\cdot), Z(\cdot))$ to (2.10).

We will face the case in which $X(\cdot)$ is given as the solution to a stochastic differential equation that will be called "forward". To be precise, let us consider the SDE

$$\begin{cases} dX(s) = b(s, X(s))ds + \sigma(s, X(s))dW(s), & s \in [t, T] \\ X(t) = x \end{cases} \tag{2.11}$$

where the functions $b : [t, T] \times \mathbb{R}^n \mapsto \mathbb{R}^n$ and $\sigma : [t, T] \times \mathbb{R}^n \mapsto \mathbb{R}^{n \times d}$ are continuous functions satisfying the usual hypothesis of Lipschitz continuity with respect to x and sublinear growth at infinity expressed by Assumption 2.1.

Choosing for (2.10) the process $X(\cdot)$ given by (2.11) we obtain the so called "forward-backward" system (FBSDE):

$$\begin{cases} dX(s) = b(s, X(s))ds + \sigma(s, X(s))dW(s), & s \in [t, T] \\ dY(s) = \upsilon(s, X(s), Y(s), Z(s))ds + Z(s)dW(s), & s \in [t, T] \\ X(t) = x \\ Y(T) = \phi(X(T)) \end{cases} \tag{2.12}$$

From the assumption on the coefficients of the SDE (2.11) and Assumption 2.5 it follows the existence of a unique strong solution to the FBSDE, which will be denoted $\{X(s, t, x), Y(s, t, x), Z(s, t, x) \ s \in [t, T]\}$ to stress the dependence on the initial condition.

Notice that $Y(t, t, x)$ is deterministic. Given that a solution exists, since we require it progressively measurable, in particular adapted, then $Y(t, t, x)$ is $\mathcal{F}_t$-measurable. At the same time the solution $Y(\cdot, t, x)$ is $\mathcal{F}_{[t,T]} = \sigma\{W_s - W_t, t \leq s \leq T\}$-measurable due to its dependence on the process $X(\cdot, t, x)$. Since $\mathcal{F}_t$ is independent of $\mathcal{F}_{[t,T]}$ and since $Y(t, t, x)$ is measurable with respect to both, the conclusion follows.

In solving the differential game it will be useful to represent the process $Y(\cdot)$ of the solution of a FBSDE in terms of an associated PDE. Therefore, let us consider the problem

$$\begin{cases} \dfrac{\partial u}{\partial s}(s, x) + \mathcal{A}u(s, x) = \upsilon(s, x, u(s, x), u_x(s, x)\sigma(s, x)) \\ u(T, x) = \phi(x), \qquad s \in [t, T], \ x \in \mathbb{R}^n \end{cases} \tag{2.13}$$

where $\mathcal{A}$ is the infinitesimal operator (1.5) introduced in Chapter 1 and where the functions $\upsilon(\cdot)$ and $\phi(\cdot)$ are the the coefficients of the BSDE in the forward-backward system (2.12). We look for solutions $u : [t, T] \times \mathbb{R}^n \mapsto \mathbb{R}$ in $C^{1,2}([t, T] \times \mathbb{R}^n)$.

Theorem 2.2. *Let Assumption 2.1 and Assumption 2.5 hold. If* $u \in C^{1,2}([t,T] \times \mathbb{R}^n)$ *is a solution to the PDE* (2.13) *then*

$$u(t,x) = Y(t,t,x)$$

where (X,Y,Z) *denote a solution to the forward-backward system eq.* (2.12).

Proof. We apply Itô formula to the process $\{u(s,X(s)),\ s \in [t,T]\}$ obtaining

$$du(s,X(s)) = \left(\frac{\partial u}{\partial s}(s,X(s)) + \mathcal{A}u(s,X(s))\right)ds + u_x(s,X(s))\sigma(s,X(s))dW(s)$$

since u solves the PDE, this can be written as

$$\begin{aligned} du(s,X(s)) = \upsilon(s,X(s),u(s,X(s)),u_x(s,X(s))\sigma(s,X(s)))ds \\ + u_x(s,X(s))\sigma(s,X(s))dW(s) \end{aligned}$$

Setting $Y'(s) = u(s,X(s)),\ Z'(s) = u_x(s,X(s))\sigma(s,X(s)),\ s \in [t,T]$ we obtain

$$\begin{cases} dY'(s) = \upsilon(s,X(s),Y'(s),Z'(s))ds + Z'(s)dW(s), \quad s \in [t,T] \\ Y'(T) = u(T,X(T)) = \phi(X(T)) \end{cases}$$

Hence, both (Y,Z) and (Y',Z') solve the same BSDE and by uniqueness they must coincide a.s. for every $s \in [t,T]$. In particular, $u(t,x) = u(t,X(t)) = Y'(t) = Y(t)$. By the above proof we also have that the following hold

$$u(s,X(s)) = Y(s,t,x), \qquad a.s.,\ s \in [t,T]$$

2.4. Multi-person differential game

In this section we solve the differential game associated to the time-inconsistent control problem for an arbitrary partition $\Pi = \{t_k \mid 0 \le k \le N\}$ of $[0,T]$, obtaining the equilibrium control and the equilibrium value function for that particular partition. Then we will take the limit as the mesh of the partition goes to zero to obtain the time-consistent equilibrium control and value function.

Let us consider first Player N who controls the system on $[t_{N-1},t_N)$. In details, for each $(t,x) \in [t_{N-1},t_N] \times \mathbb{R}^n$, consider the controlled SDE

$$\begin{cases} dX^N(s) = b(s, X^N(s), u^N(s))ds + \sigma(s, X^N(s), u^N(s))dW(s), & s \in [t, t_N] \\ X^N(t) = x \end{cases}$$

with associated cost functional

$$\mathcal{J}^N(t, x, u^N) = \mathbb{E}_t\Big[\int_t^{t_N} C(t_{N-1}, s, X^N(s), u^N(s))ds + G(t_{N-1}, X^N(T))\Big]$$

Notice that

$$\mathcal{J}^N(t_{N-1}, x, u^N) = \mathcal{J}(t_{N-1}, x, u^N), \qquad x \in \mathbb{R}^n,\ u^N \in \mathcal{U}[t_{N-1}, t_N] \tag{2.14}$$

In this way we have constructed a cost functional customized for player N, where his/her cost preferences are synthetized by the time term t_{N-1} in the costs $C(t_{N-1}, \cdot, \cdot, \cdot)$ and $G(t_{N-1}, \cdot)$. Moreover, we notice that such preferences are fixed for every initial pair (t, x) belonging to the influence region of Player N, in other words, the functional $\mathcal{J}^N$ is not time varying on $[t_{N-1}, t_N]$.
We pose the following problem for Player N:

Problem P_N. For any $(t, x) \in [t_{N-1}, t_N) \times \mathbb{R}^n$, find a control $\bar{u}^N = \bar{u}^N(\cdot, t, x) \in \mathcal{U}[t, t_N]$ such that

$$\mathcal{J}^N(t, x, \bar{u}^N) = \inf_{u^N \in \mathcal{U}[t, t_N]} \mathcal{J}^N(t, x, u^N) = V^\Pi(t, x)$$

The above defines the value function $V^\Pi(\cdot, \cdot)$ on $[t_{N-1}, t_N] \times \mathbb{R}^n$ and if we could find an optimal control $\bar{u}^N$ then by (2.14) we would get

$$\mathcal{J}(t_{N-1}, x, \bar{u}^N) = V^\Pi(t_{N-1}, x), \qquad x \in \mathbb{R}^n$$

As we highlighted before the cost functional $\mathcal{J}^N$ is independent of the initial time of the optimisation procedure and this makes Problem P_N time-consistent. Therefore, the value function $V^\Pi(\cdot, \cdot)$ is given by the classical solution of the following HJB equation

$$\begin{cases} V_t^\Pi(t, x) + \inf_{u \in U} \mathbb{H}(t_{N-1}, t, x, u, V_x^\Pi(t, x), V_{xx}^\Pi(t, x)) = 0, & (t, x) \in [t_{N-1}, t_N] \times \mathbb{R}^n \\ V^\Pi(t_N, x) = G(t_{N-1}, x), \quad x \in \mathbb{R}^n \end{cases} \tag{2.15}$$

where we have used the notation with the Hamiltonian function (2.8). By the definition of $\psi : D[0, T] \times \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{S}^n \mapsto U$ in Assumption 2.3 we can write the above equation as follows

$$\begin{cases} V_t^{\Pi}(t,x) + \mathbb{H}(t_{N-1}, t, x, \psi(t_{N-1}, t, x, V_x^{\Pi}(t,x), V_{xx}^{\Pi}(t,x)), V_x^{\Pi}(t,x), V_{xx}^{\Pi}(t,x)) = 0, \\ \qquad\qquad (t,x) \in [t_{N-1}, t_N] \times \mathbb{R}^n \\ V^{\Pi}(t_N, x) = G(t_{N-1}, x), \quad x \in \mathbb{R}^n \end{cases} \tag{2.16}$$

With a solution $V^{\Pi}(\cdot,\cdot)$ to (2.15) (or (2.16)) let us assume that the following closed-loop system admits a unique solution $\bar{X}^N = \bar{X}^N(\cdot, t_{N-1}, x)$, i.e. that ψ is an admissible control law

$$\begin{cases} d\bar{X}^N(s) = b(s, \bar{X}^N(s), \psi(t_{N-1}, s, \bar{X}^N(s), V_x^{\Pi}(s, \bar{X}^N(s)), V_{xx}^{\Pi}(s, \bar{X}^N(s))))ds \\ \qquad + \sigma(s, \bar{X}^N(s), \psi(t_{N-1}, s, \bar{X}^N(s), V_x^{\Pi}(s, \bar{X}^N(s)), V_{xx}^{\Pi}(s, \bar{X}^N(s))))dW(s), \\ \qquad\qquad s \in [t_{N-1}, t_N] \\ \bar{X}^N(t_{N-1}) = x \end{cases} \tag{2.17}$$

then by Assumption 2.3 and Theorem 1.2 (Verification thm.), we have that an optimal control process $\bar{u}^N(\cdot)$ of Problem P_N for the initial pair (t_{N-1}, x) is given by

$$\begin{aligned} \bar{u}^N(s) = \bar{u}^N(s, t_{N-1}, x) &= \psi(t_{N-1}, s, \bar{X}^N(s), V_x^{\Pi}(s, \bar{X}^N(s)), V_{xx}^{\Pi}(s, \bar{X}^N(s))) \\ &= \psi(t_{N-1}, s, \bar{X}^N(s, t_{N-1}, x), V_x^{\Pi}(s, \bar{X}^N(s, t_{N-1}, x)), V_{xx}^{\Pi}(s, \bar{X}^N(s, t_{N-1}, x))) \end{aligned}$$

and $\bar{X}^N(\cdot)$ is the corresponding optimal controlled state process.

We move on now to consider the optimal control problem for Player $N-1$ on $[t_{N-2}, t_{N-1})$. For any initial pair $(t,x) \in [t_{N-2}, t_{N-1}] \times \mathbb{R}^n$ the state equation, controlled by such agent is given by

$$\begin{cases} dX^{N-1}(s) = b(s, X^{N-1}(s), u^{N-1}(s))ds + \sigma(s, X^{N-1}(s), u^{N-1}(s))dW(s), \\ \qquad\qquad s \in [t, t_{N-1}) \\ X^{N-1}(t) = x \end{cases} \tag{2.18}$$

To determine an appropriate cost functional for Player $N-1$ we rely on the knowledge that this player can control the system only on the interval $[t_{N-2}, t_{N-1})$ and that Player N takes over at time t_{N-1} to control the system from that point on. In addition, Player $N-1$ knows that the following player will play optimally based on the initial pair $(t_{N-1}, X^{N-1}(t_{N-1}))$ which is Player $N-1$'s terminal pair. Therefore, the sophisticated cost functional of Player $N-1$ should be

$$\begin{aligned}\mathcal{J}^{N-1}(t,x,u^{N-1}) = \mathbb{E}_t\Big[\int_t^{t_{N-1}} C(t_{N-2},s,X^{N-1}(s),u^{N-1}(s))ds \\ + \int_{t_{N-1}}^{t_N} C(t_{N-2},s,\bar{X}^N(s,t_{N-1},X^{N-1}(t_{N-1})),\bar{u}^N(s,t_{N-1},X^{N-1}(t_{N-1})))ds \\ + G(t_{N-2},\bar{X}^N(t_N,t_{N-1},X^{N-1}(t_{N-1})))\Big]\end{aligned} \tag{2.19}$$

Notice that, as we anticipated in section 2.2., even if Player $N-1$ knows that Player N will control the system on $[t_{N-1},t_N]$ he/she still maintain his/her own preferences on the future costs (see the term t_{N-2} appearing in the running cost $C(\cdot)$ on $[t_{N-1},t_N]$ and in the terminal cost $G(\cdot)$ at t_N). Let us denote

$$\begin{aligned}h^{N-1}(x) = \mathbb{E}_{t_{N-1}}\Big[\int_{t_{N-1}}^{t_N} C(t_{N-2},s,\bar{X}^N(s,t_{N-1},x),\bar{u}^N(s,t_{N-1},x))ds \\ + G(t_{N-2},\bar{X}^N(t_N,t_{N-1},x))\Big]\end{aligned}$$

then the cost functional (2.19) can be written as

$$\begin{aligned}\mathcal{J}^{N-1}(t,x,u^{N-1}) = \mathbb{E}_t\Big[\int_t^{t_{N-1}} C(t_{N-2},s,X^{N-1}(s),u^{N-1}(s))ds \\ + h^{N-1}(X^{N-1}(t_{N-1}))\Big]\end{aligned} \tag{2.20}$$

The optimal control problem associated to the state equation (2.18) and to the cost functional (2.20) has the form of a standard time-consistent problem. However, the map $x \mapsto h^{N-1}(x)$ seems too implicit, which is difficult for us to pass to the limits later on. We now would like to make it more explicit in some sense. Exploiting the results for FBSDE with deterministic coefficients presented in section 2.3. we proceed as follows. For the optimal state process $\bar{X}^N = \bar{X}^N(\cdot,t_{N-1},x)$ determined by (2.17) on $[t_{N-1},t_N]$ we introduce the following BSDE

$$\begin{cases} dY^N(s) = -C(t_{N-2},s,\bar{X}^N(s),\psi(t_{N-1},s,\bar{X}^N(s),V_x^{\Pi}(s,\bar{X}^N(s)),V_{xx}^{\Pi}(s,\bar{X}^N(s))))ds \\ \qquad\qquad + Z^N(s)dW(s), \qquad s\in[t_{N-1},t_N] \\ Y^N(t_N) = G(t_{N-2},\bar{X}^N(t_N)) \end{cases}$$

which is equivalent to the following

$$\begin{cases} dY^N(s) = -C(t_{N-2}, s, \bar{X}^N(s), \bar{u}^N(s))ds + Z^N(s)dW(s), \quad s \in [t_{N-1}, t_N] \\ Y^N(t_N) = G(t_{N-2}, \bar{X}^N(t_N)) \end{cases} \tag{2.21}$$

Note that t_{N-2} appears both in the generator and in the terminal condition of the BSDE. Moreover, the generator does not depend on $Y(\cdot)$ neither on $Z(\cdot)$ so it is a particular case of the more general scenario presented in section 2.3. Nonetheless, both driver and terminal condition satisfy Assumption 2.5 so that this BSDE admits a unique adapted solution $(Y^N, Z^N) = (Y^N(\cdot, t_{N-1}, x), Z^N(\cdot, t_{N-1}, x))$ depending on the initial condition of (2.17). Further, it holds

$$Y^N(t_{N-1}) = \mathbb{E}_{t_{N-1}}\left[\int_{t_{N-1}}^{t_N} C(t_{N-2}, s, \bar{X}^N(s), \bar{u}^N(s))ds + G(t_{N-2}, \bar{X}^N(t_N))\right] = h^{N-1}(x)$$

It is seen that (2.17) and (2.21) form a forward-backward system for which we have the following representation result derived in section 2.3.

$$Y^N(s) = \Theta^N(s, \bar{X}^N(s)), \qquad a.s., \; s \in [t_{N-1}, t_N]$$

as long as $\Theta^N(\cdot, \cdot)$ is a classical solution to the following PDE

$$\begin{cases} \Theta^N_s(s,x) + \langle \Theta^N_x(s,x), b(s,x,\psi(t_{N-1}, s, x, V^\Pi_x(s,x), V^\Pi_{xx}(s,x)))\rangle \\ \; + tr(a(s,x,\psi(t_{N-1}, s, x, V^\Pi_x(s,x), V^\Pi_{xx}(s,x)))\Theta^N_{xx}(s,x)) \\ \; + C(t_{N-2}, s, x, \psi(t_{N-1}, s, x, V^\Pi_x(s,x), V^\Pi_{xx}(s,x))) = 0, \quad (s,x) \in [t_{N-1}, t_N] \times \mathbb{R}^n \\ \Theta^N(t_N, x) = G(t_{N-2}, x), \qquad x \in \mathbb{R}^n \end{cases}$$

or more compactly, exploiting the Hamiltonian defined in (2.8)

$$\begin{cases} \Theta^N_s(s,x) + \mathbb{H}(t_{N-2}, s, x, \psi(t_{N-1}, s, x, V^\Pi_x(s,x), V^\Pi_{xx}(s,x)), \Theta^N_x(s,x), \Theta^N_{xx}(s,x)) = 0, \\ \qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad (s,x) \in [t_{N-1}, t_N] \times \mathbb{R}^n \\ \Theta^N(t_N, x) = G(t_{N-2}, x), \qquad x \in \mathbb{R}^n \end{cases}$$

It is important to point out that in general

$$\begin{aligned}
&\Theta^N(t_{N-1},x) = Y^N(t_{N-1}) = h^{N-1}(x) \\
&= \mathbb{E}_{t_{N-1}}\left[\int_{t_{N-1}}^{t_N} C(t_{N-2},s,\bar{X}^N(s),\bar{u}^N(s))ds + G(t_{N-2},\bar{X}^N(T))\right] \\
&\neq \mathbb{E}_{t_{N-1}}\left[\int_{t_{N-1}}^{t_N} C(t_{N-1},s,\bar{X}^N(s),\bar{u}^N(s))ds + G(t_{N-1},\bar{X}^N(T))\right] \\
&= V^{\Pi}(t_{N-1},x).
\end{aligned} \tag{2.22}$$

Thanks to the above representation $\Theta^N(\cdot,\cdot)$ of $Y^N(\cdot)$, we can formulate the cost functional (2.20) as follows

$$\begin{aligned}
\mathcal{J}^{N-1}(t,x,u^{N-1}) =& \mathbb{E}_t\Bigg[\int_t^{t_{N-1}} C(t_{N-2},s,X^{N-1}(s),u^{N-1}(s))ds \\
&+ \Theta^N(t_{N-1},X^{N-1}(t_{N-1}))\Bigg]
\end{aligned} \tag{2.23}$$

In other terms, the cost functional (2.23) separates the distinct contributions of the two players. In particular, the first time integral represents the portion of costs which is under direct control of Player $N-1$ while the contribution of Player N is enclosed in the function $\Theta^N(\cdot,\cdot)$. Having expressed $\mathcal{J}^{N-1}$ in a more easy way, we are now ready to formulate the control problem relative to Player $N-1$

Problem P_{N-1}**.** For any $(t,x) \in [t_{N-2},t_{N-1}) \times \mathbb{R}^n$, find a $\bar{u}^{N-1}(\cdot) = \bar{u}^{N-1}(\cdot,t,x) \in \mathcal{U}[t,t_{N-1}]$ such that

$$\mathcal{J}^{N-1}(t,x,\bar{u}^{N-1}) = \inf_{u^{N-1}\in[t,t_{N-1}]} \mathcal{J}^{N-1}(t,x,u^{N-1}) = V^{\Pi}(t,x)$$

The above problem defines the value function $V^{\Pi}(\cdot,\cdot)$ on $[t_{N-2},t_{N-1}) \times \mathbb{R}^n$. Again, we see that cost functional (2.23) is independent on the initial time of the optimisation so it gives rise to a time-consistent control problem for which, under proper conditions, the value function V^{Π} is the classical solution to the following HJB equation

$$\begin{cases}
V_t^{\Pi}(t,x) + \inf\limits_{u\in U} \mathbb{H}(t_{N-2},t,x,u,V_x^{\Pi}(t,x),V_{xx}^{\Pi}(t,x)) = 0, & (t,x) \in [t_{N-2},t_{N-1}) \times \mathbb{R}^n \\
V^{\Pi}(t_{N-1}-0,x) = \Theta^N(t_{N-1},x), & x \in \mathbb{R}^n
\end{cases}$$

Again, exploiting the map $\psi(\cdot)$ defined in Assumption 2.3 we can write the above as

$$\begin{cases} V_t^{\Pi}(t,x) + \mathbb{H}(t_{N-2},t,x,\psi(t_{N-2},t,x,V_x^{\Pi}(t,x),V_{xx}^{\Pi}(t,x)),V_x^{\Pi}(t,x),V_{xx}^{\Pi}(t,x)) = 0, \\ \qquad\qquad\qquad\qquad\qquad\qquad (t,x) \in [t_{N-2},t_{N-1}) \times \mathbb{R}^n \\ V^{\Pi}(t_{N-1}-0,x) = \Theta^N(t_{N-1},x), \quad x \in \mathbb{R}^n \end{cases} \tag{2.24}$$

From (2.22) we see that in general

$$V^{\Pi}(t_{N-1}-0,x) = \Theta^N(t_{N-1},x) \neq V^{\Pi}(t_{N-1},x)$$

Thus, the value function $V^{\Pi}(\cdot,\cdot)$ which is now defined on $[t_{N-2},t_N] \times \mathbb{R}^n$ may have a jump at t_{N-1} (for whatever $x \in \mathbb{R}^n$).

Now, suppose that the state dynamics driven by the control law ψ admits a unique solution $\bar{X}^{N-1}(\cdot)$

$$\begin{cases} d\bar{X}^{N-1}(s) = b(s,\bar{X}^{N-1}(s),\psi(t_{N-2},s,\bar{X}^{N-1}(s),V_x^{\Pi}(s,\bar{X}^{N-1}(s)),.. \\ \qquad ..V_{xx}^{\Pi}(s,\bar{X}^{N-1}(s))))ds + \sigma(s,\bar{X}^{N-1}(s),\psi(t_{N-2},s,.. \\ \qquad ..\bar{X}^{N-1}(s),V_x^{\Pi}(s,\bar{X}^{N-1}(s)),V_{xx}^{\Pi}(s,\bar{X}^{N-1}(s))))dW(s), \\ \qquad\qquad\qquad\qquad s \in [t_{N-2},t_{N-1}) \\ \bar{X}^{N-1}(t_{N-2}) = x \end{cases} \tag{2.25}$$

Then, again by Theorem 1.2 (the Verification thm.) we have an optimal control process for Problem P_{N-1} with initial pair (t_{N-2},x), given by

$$\begin{aligned} \bar{u}^{N-1}(s) &= \bar{u}^{N-1}(s,t_{N-2},x) \\ &= \psi(t_{N-2},s,\bar{X}^{N-1}(s),V_x^{\Pi}(s,\bar{X}^{N-1}(s)),V_{xx}^{\Pi}(s,\bar{X}^{N-1}(s))) \\ &= \psi(t_{N-2},s,\bar{X}^{N-1}(s,t_{N-2},x),V_x^{\Pi}(s,\bar{X}^{N-1}(s,t_{N-2},x)),V_{xx}^{\Pi}(s,\bar{X}^{N-1}(s,t_{N-2},x))) \end{aligned}$$

Now, for the optimal pair $(\bar{X}^{N-1}(\cdot),\bar{u}^{N-1}(\cdot)) = (\bar{X}^{N-1}(\cdot,t_{N-2},x),\bar{u}^{N-1}(\cdot,t_{N-2},x))$ of Problem P_{N-1} we make a natural extension on $[t_{N-1},t_N]$ as follows

$$\begin{cases} \bar{X}^{N-1}(s) = \bar{X}^N(s,t_{N-1},\bar{X}^{N-1}(t_{N-1})), \\ \bar{u}^{N-1}(s) = \bar{u}^N(s,t_{N-1},\bar{X}^{N-1}(t_{N-1})), \end{cases} \qquad s \in [t_{N-1},t_N]$$

So that the extended optimal pair $(\bar{X}^{N-1}(\cdot),\bar{u}^{N-1}(\cdot))$ satisfies

$$
\begin{cases}
d\bar{X}^{N-1}(s) = b(s, \bar{X}^{N-1}(s), \psi(t_{N-2}, s, \bar{X}^{N-1}(s), V_x^{\Pi}(s, \bar{X}^{N-1}(s)), .. \\
\qquad ..V_{xx}^{\Pi}(s, \bar{X}^{N-1}(s))))ds + \sigma(s, \bar{X}^{N-1}(s), \psi(t_{N-2}, s, .. \\
\qquad ..\bar{X}^{N-1}(s), V_x^{\Pi}(s, \bar{X}^{N-1}(s)), V_{xx}^{\Pi}(s, \bar{X}^{N-1}(s))))dW(s), \\
\qquad\qquad s \in [t_{N-2}, t_{N-1}) \\
d\bar{X}^{N-1}(s) = b(s, \bar{X}^{N-1}(s), \psi(t_{N-1}, s, \bar{X}^{N-1}(s), V_x^{\Pi}(s, \bar{X}^{N-1}(s)), .. \\
\qquad ..V_{xx}^{\Pi}(s, \bar{X}^{N-1}(s))))ds + \sigma(s, \bar{X}^{N-1}(s), \psi(t_{N-1}, s, .. \\
\qquad ..\bar{X}^{N-1}(s), V_x^{\Pi}(s, \bar{X}^{N-1}(s)), V_{xx}^{\Pi}(s, \bar{X}^{N-1}(s))))dW(s), \\
\qquad\qquad s \in [t_{N-1}, t_N] \\
\bar{X}^{N-1}(t_{N-2}) = x
\end{cases}
\tag{2.26}
$$

where we have highlighted with different colors the changes in the dynamics of the state over the two distinct time intervals. Basically, over $[t_{N-2}, t_{N-1})$ the control process driving the system is the optimal one for Player $N-1$, namely $\bar{u}^{N-1}(\cdot)$, while over $[t_{N-1}, t_N]$ the system is driven by the optimal control for Player N, that is $\bar{u}^N(\cdot)$. From now on we refer to the extended pair $(\bar{X}^{N-1}(\cdot), \bar{u}^N(\cdot))$ as a sophisticated equilibrium pair on $[t_{N-2}, t_N]$.

Let us introduce the following notation

$$
l^{\Pi}(s) = \sum_{k=1}^{N} t_{k-1} \mathbb{1}_{[t_{k-1}, t_k)}(s), \qquad s \in [0, T] \tag{2.27}
$$

for which it is easy to check that

$$
0 \le s - l^{\Pi}(s) \le \|\Pi\|, \qquad s \in [0, T]
$$

Then, thanks to this notation it is possible to rewrite (2.26) in a more compact way as

$$
\begin{cases}
d\bar{X}^{N-1}(s) = b(s, \bar{X}^{N-1}(s), \psi(l^{\Pi}(s), s, \bar{X}^{N-1}(s), V_x^{\Pi}(s, \bar{X}^{N-1}(s)), .. \\
\qquad ..V_{xx}^{\Pi}(s, \bar{X}^{N-1}(s))))ds + \sigma(s, \bar{X}^{N-1}(s), \psi(l^{\Pi}(s), s, .. \\
\qquad ..\bar{X}^{N-1}(s), V_x^{\Pi}(s, \bar{X}^{N-1}(s)), V_{xx}^{\Pi}(s, \bar{X}^{N-1}(s))))dW(s), \\
\qquad\qquad s \in [t_{N-2}, t_N] \\
\bar{X}^{N-1}(t_{N-2}) = x
\end{cases}
\tag{2.28}
$$

where we have highlighted the usage of $l^{\Pi}(\cdot)$.

Also, it holds

$$
\begin{aligned}
\mathcal{J}(t_{N-2}, x, \bar{u}^{N-1}) = \mathbb{E}_{t_{N-2}}\Bigg[\int_{t_{N-2}}^{t_{N-1}} & C(t_{N-2}, s, \bar{X}^{N-1}(s), \bar{u}^{N-1}(s))ds \\
&+ \int_{t_{N-1}}^{t_N} C(t_{N-2}, s, \bar{X}^{N-1}(s), \bar{u}^{N-1}(s))ds + G(t_{N-2}, \bar{X}^{N-1}(T))\Bigg] \\
&= \mathcal{J}^{N-1}(t_{N-2}, x, \bar{u}^{N-1}) = V^{\Pi}(t_{N-2}, x)
\end{aligned}
$$

In general it can happen that

$$
\mathcal{J}(t_{N-2}, x, \bar{u}^{N-1}) > \inf_{u \in \mathcal{U}[t_{N-2}, t_N]} \mathcal{J}(t_{N-2}, x, u)
$$

i.e. the sophisticated equilibrium pair might not be the optimal pair (for the given initial conditions). This is not surprising, as a matter of fact $\bar{u}^{N-1}$ is just the equilibrium control relative to the sub-game (starting at t_{N-2}) associated to the specific Π-partition. Intuition suggests that if we could solve the sub-game for a finer partition, that is adding more players and restricting the time intervals on which they control the system, we would obtain an equilibrium control which approaches closer the true optimal control. This is why we will pass to the limit later on to obtain the time consistent equilibrium control.

In order to state an optimal control problem for Player $N-2$ we consider, for any $(t, x) \in [t_{N-3}, t_{N-2}) \times \mathbb{R}^n$, the state equation

$$
\begin{cases}
dX^{N-2}(s) = b(s, X^{N-2}(s), u^{N-2}(s))ds + \sigma(s, X^{N-2}(s), u^{N-2}(s))dW(s), \\
\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad s \in [t, t_{N-2}) \\
X^{N-2}(t) = x
\end{cases}
$$

which expresses the dynamics of the system governed by Player $N-2$.

Now, to move forward we have to construct a suitable cost functional for Player $N-2$ and following what we already did for Player $N-1$ we construct $\mathcal{J}^{N-2}$ on $[t_{N-3}, t_{N-2}) \times \mathbb{R}^n \times \mathcal{U}[t_{N-3}, t_{N-2}]$ as follows

$$\mathcal{J}^{N-2}(t,x,u^{N-2}) = \mathbb{E}_t\Big[\int_t^{t_{N-2}} C(t_{N-3},s,X^{N-2}(s),u^{N-2}(s))ds$$
$$+\int_{t_{N-2}}^{t_N} C(t_{N-3},s,\bar{X}^{N-1}(s,t_{N-2},X^{N-2}(t_{N-2})),\bar{u}^{N-1}(s,t_{N-2},X^{N-2}(t_{N-2})))ds$$
$$+G(t_{N-3},\bar{X}^{N-1}(t_N,t_{N-2},X^{N-2}(t_{N-2})))\Big]$$

where, as before, the contributions to the cost functional coming from different players have been separated. In particular, the part that is under the control of Player $N-2$ is the first time integral while the second one and the terminal cost depend on the sophisticated equilibrium pair $(\bar{X}^{N-1}(\cdot),\bar{u}^{N-1}(\cdot))$ for Players $N-1$ and N on $[t_{N-2},t_N]$ which is not influenced by Player $N-2$. Again, notice that the cost preferences of Player $N-2$ are maintained also on the time intervals where he/she does not control the system.

Following the reasoning presented above we introduce the following BSDE on $[t_{N-2},t_N]$:

$$\begin{cases} dY^{N-1}(s) = -C(t_{N-3},s,\bar{X}^{N-1}(s),\bar{u}^{N-1}(s))ds + Z^{N-1}(s)dW(s), \\ \qquad\qquad\qquad\qquad\qquad\qquad\qquad s\in[t_{N-2},t_N] \\ Y(t_N) = G(t_{N-3},\bar{X}^{N-1}(t_N)) \end{cases} \tag{2.29}$$

Let $(Y^{N-1}(\cdot),Z^{N-1}(\cdot)) = (Y^{N-1}(\cdot,t_{N-2},x),Z^{N-1}(\cdot,t_{N-2},x))$ be the adapted solution, uniquely depending on the initial condition (t_{N-2},x). Then (2.28) and (2.29) form a FBSDE. Again, we represent $Y^{N-1}(\cdot)$ as

$$Y^{N-1}(s) = \Theta^{N-1}(s,\bar{X}^{N-1}(s)), \qquad a.s.,\ s\in[t_{N-2},t_N]$$

where $\Theta^{N-1}(\cdot,\cdot)$ is a classical solution to the following PDE

$$\begin{cases} \Theta_s^{N-1}(s,x) + \mathbb{H}(t_{N-3},s,x,\psi(l^{\Pi}(s),s,x,V_x^{\Pi}(s,x),V_{xx}^{\Pi}(s,x)),.. \\ \qquad ..,\Theta_x^{N-1}(s,x),\Theta_{xx}^{N-1}(s,x)) = 0, \qquad (s,x)\in[t_{N-2},t_N]\times\mathbb{R}^n \\ \Theta^{N-1}(t_N,x) = G(t_{N-3},x), \qquad x\in\mathbb{R}^n \end{cases}$$

This representation allows us to rewrite the cost functional $\mathcal{J}^{N-2}$ as follows

$$\mathcal{J}^{N-2}(t,x,u^{N-2}) = \mathbb{E}_t\Big[\int_t^{t_{N-2}} C(t_{N-3},s,X^{N-2}(s),u^{N-2}(s))ds \\ + \Theta^{N-1}(t_{N-2},X^{N-2}(t_{N-2}))\Big]$$

We can now pose the control problem for Player $N-2$

Problem P_{N-2}**.** For any $(t,x) \in [t_{N-3},t_{N-2})\times\mathbb{R}^n$, find a $\bar{u}^{N-2} = \bar{u}^{N-2}(\cdot,t,x) \in \mathcal{U}[t,t_{N-2}]$ such that

$$\mathcal{J}^{N-2}(t,x,\bar{u}^{N-2}) = \inf_{u^{N-2}\in\mathcal{U}[t,t_{N-2}]} \mathcal{J}^{N-2}(t,x,u^{N-2}) = V^{\Pi}(t,x)$$

This problem defines the value function $V^{\Pi}(\cdot,\cdot)$ on the interval $[t_{N-3},t_{N-2})$ and together with the previous definitions we see that the equilibrium value function is now defined on $[t_{N-3},t_N]\times\mathbb{R}^n$. As for the other players above, Player $N-2$ faces a time-consistent control problem which, taking advantage of the map $\psi(\cdot)$ defined in Assumption 2.3 gives the value function as the classical solution to the following HJB equation

$$\begin{cases} V_t^{\Pi}(t,x) + \mathbb{H}(t_{N-3},t,x,\psi(t_{N-3},t,x,V_x^{\Pi}(t,x),V_{xx}^{\Pi}(t,x)),V_x^{\Pi}(t,x),V_{xx}^{\Pi}(t,x)) = 0, \\ \qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad (t,x)\in[t_{N-3},t_{N-2})\times\mathbb{R}^n \\ V^{\Pi}(t_{N-2}-0,x) = \Theta^{N-1}(t_{N-2},x), \quad x\in\mathbb{R}^n \end{cases}$$

We point out that in general, similar to (2.22)

$$V^{\Pi}(t_{N-2}-0,x) \neq V^{\Pi}(t_{N-2},x)$$

The procedure described so far for players N,$N-1$ and $N-2$ can be continued recursively. By backward induction we can construct sophisticated cost functional $\mathcal{J}^k(t,x,u^k)$ for Player $k,\ k\in\{1,2,...,N\}$ and

$$V^{\Pi}(t,x) = \inf_{u^k\in\mathcal{U}[t,t_k]} \mathcal{J}^k(t,x,u^k), \qquad (t,x)\in[t_{k-1},t_k)\times\mathbb{R}^n$$

with the equilibrium value function $V^{\Pi}(\cdot,\cdot)$ satisfying the following HJB equations on the time intervals associated to the partition Π

$$\begin{cases} V_t^{\Pi}(t,x) + \mathbb{H}(l^{\Pi}(t), t, x, \psi(l^{\Pi}(t), t, x, V_x^{\Pi}(t,x), V_{xx}^{\Pi}(t,x)), V_x^{\Pi}(t,x), V_{xx}^{\Pi}(t,x)) = 0, \\ \qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad (t,x) \in [t_{N-1}, t_N) \times \mathbb{R}^n \\ V^{\Pi}(t_N, x) = G(t_{N-1}, x), \quad x \in \mathbb{R}^n \end{cases}$$

and for $k = 1, 2, ..., N-1$

$$\begin{cases} V_t^{\Pi}(t,x) + \mathbb{H}(l^{\Pi}(t), t, x, \psi(l^{\Pi}(t), t, x, V_x^{\Pi}(t,x), V_{xx}^{\Pi}(t,x)), V_x^{\Pi}(t,x), V_{xx}^{\Pi}(t,x)) = 0, \\ \qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad (t,x) \in [t_{k-1}, t_k) \times \mathbb{R}^n \\ V^{\Pi}(t_k - 0, x) = \Theta^{k+1}(t_k, x), \quad x \in \mathbb{R}^n \end{cases}$$

where we have used the notation with $l^{\Pi}(\cdot)$. Notice that the difference between the HJB equations relative to each player lies in the different control laws $\psi(l^{\Pi}(\cdot), \cdot, ..)$ and in the final conditions. We recall that for $k = 1, 2, ..., N-1$, $\Theta^{k+1}(\cdot,\cdot)$ is a classical solution to the following PDE

$$\begin{cases} \Theta_s^{k+1}(s,x) + \mathbb{H}(t_{k-1}, s, x, \psi(l^{\Pi}(s), s, x, V_x^{\Pi}(s,x), V_{xx}^{\Pi}(s,x)), .. \\ \qquad\qquad .., \Theta_x^{k+1}(s,x), \Theta_{xx}^{k+1}(s,x)) = 0, \qquad (s,x) \in [t_k, t_N] \times \mathbb{R}^n \\ \Theta^{k+1}(t_N, x) = G(t_{k-1}, x), \qquad x \in \mathbb{R}^n \end{cases}$$

Let us define now the function

$$\Psi^{\Pi}(t,x) = \psi(l^{\Pi}(t), t, x, V_x^{\Pi}(t,x), V_{xx}^{\Pi}(t,x)), \quad (t,x) \in [0,T] \times \mathbb{R}^n \tag{2.30}$$

which is just the link of the optimal control laws for the different players over the entire interval $[0,T]$.

Then for any $x \in \mathbb{R}^n$, let $\bar{X}^{\Pi}(\cdot)$ be the unique solution to the following stochastic differential equation driven by the feedback control law $\Psi^{\Pi}(\cdot,\cdot)$

$$\begin{cases} d\bar{X}^{\Pi}(s) = b(s, \bar{X}^{\Pi}(s), \Psi^{\Pi}(s, \bar{X}^{\Pi}(s)))ds + \sigma(s, \bar{X}^{\Pi}(s), \Psi^{\Pi}(s, \bar{X}^{\Pi}(s)))dW(s), \quad s \in [0,T] \\ \bar{X}^{\Pi}(0) = x \end{cases}$$

and denote by

$$\bar{u}^{\Pi}(s) = \Psi^{\Pi}(s, \bar{X}^{\Pi}(s)), \qquad s \in [0,T]$$

the control process associated to $\Psi^{\Pi}(\cdot,\cdot)$ appearing in the above SDE.

Then $\bar{u}^{\Pi} = (\bar{u}^1, \bar{u}^2, ..., \bar{u}^N)$ is a Nash equilibrium of the N-person non-cooperative differential game associated to partition Π. In fact, according to our construction, we have

$$\begin{aligned}
&\mathcal{J}(t_{k-1}, \bar{X}^{\Pi}(t_{k-1}), \Psi^{\Pi}\big|_{[t_{k-1},T]}) = \mathcal{J}(t_{k-1}, \bar{X}^{\Pi}(t_{k-1}), \bar{u}^{\Pi}\big|_{[t_{k-1},T]}) \\
&= V^{\Pi}(t_{k-1}, \bar{X}^{\Pi}(t_{k-1})) = \mathcal{J}^k(t_{k-1}, \bar{X}^{\Pi}(t_{k-1}), \bar{u}^{\Pi}\big|_{[t_{k-1},t_k]}) \\
&= \inf_{u^k \in \mathcal{U}[t_{k-1},t_k]} \mathcal{J}^k(t_{k-1}, \bar{X}^{\Pi}(t_{k-1}), u^k) \leq \mathcal{J}^k(t_{k-1}, \bar{X}^{\Pi}(t_{k-1}), u^k) \\
&= \mathcal{J}(t_{k-1}, \bar{X}^{\Pi}(t_{k-1}), u^k \oplus \Psi^{\Pi}\big|_{[t_k,T]}), \qquad a.s., \ \forall\, u^k \in \mathcal{U}[t_{k-1}, t_k], \ 1 \leq k \leq N
\end{aligned} \tag{2.31}$$

where

$$\big(u^k \oplus \Psi^{\Pi}\big|_{[t_k,T]}\big)(t) = \begin{cases} u^k(t), & t \in [t_{k-1}, t_k) \\ \Psi^{\Pi}(t, X^k(t)), & t \in [t_k, T] \end{cases}$$

2.5. Limiting procedure

We now would like to examine the situation when $\|\Pi\| \to 0$. Let us recall that in the previous section we presented a procedure to find the Nash equilibrium control strategy and the equilibrium value function to the differential game associated to an arbitrary, but fixed partition Π of $[0,T]$. Moreover, having adopted a backward induction construction, we have guaranteed that the equilibrium control determined is optimal for any sub-game of the complete sequential game. This means that the restriction of the equilibrium strategy to a smaller section of the original partition is still optimal for the remaining players involved in the game. However, as we have pointed out in the previous section, in general for $k = 1, 2, .., N-1$, it holds $\mathbb{P}$-a.s.

$$\mathcal{J}(t_{k-1}, \bar{X}^{\Pi}(t_{k-1}), \bar{u}^{\Pi}\big|_{[t_{k-1},T]}) > \inf_{u \in \mathcal{U}[t_{k-1},T]} \mathcal{J}(t_{k-1}, \bar{X}^{\Pi}(t_{k-1}), u)$$

which means that the control process that realizes the minimum of the cost functional evaluated at the time instants that constitute the partition might differ from the equilibrium control process.

This is due to the fact that the the equilibrium strategy determined so far is optimal only assuming that there are N players (or controller's future selves) and they undertake to solve the optimal control problem following the interaction rules that we explained in section 2.2.

Nevertheless, the N-players differential game is just an approximation of the actual control problem which could be more realistically seen as a dynamic game played by a continuum of agents, each one facing an optimisation problem with a particular cost preference. This should give us the intuition that the finer is the partition Π for which we solve the differential game, the closer the equilibrium strategy will be to the optimal control strategy.

Remark 2.4. *It is interesting to notice that we can look at the differential game associated to a time-inconsistent optimal control problem, as a succession of agents with a naïve approach, each one assuming that the future players will adopt his/her personal cost preferences but with the additional knowledge that the future agents will adopt the control strategy optimal for their specific problem.*

Suppose now that we have the following

$$\lim_{\|\Pi\|\to 0}\Big(\big|V^{\Pi}(t,x)-V(t,x)\big|+\big|V_x^{\Pi}(t,x)-V_x(t,x)\big|+\big|V_{xx}^{\Pi}-V_{xx}(t,x)\big|\Big)=0$$

uniformly for (t,x) in any compact set, for some $V(\cdot,\cdot)$. Under Assumption 2.3 we also have

$$\lim_{\|\Pi\|\to 0}\big|\Psi^{\Pi}(t,x)-\Psi(t,x)\big|=0$$

uniformly for (t,x) in any compact set, where $\Psi^{\Pi}(\cdot,\cdot)$ has been defined in (2.30) and with

$$\Psi(t,x)=\psi(t,t,x,V_x(t,x),V_{xx}(t,x)),\quad (t,x)\in[0,T]\times\mathbb{R}^n$$

Then the following limit exists

$$\lim_{\|\Pi\|\to 0}\|\bar{X}^{\Pi}(\cdot)-\bar{X}(\cdot)\|_{\mathbb{S}^2(0,T)^n}$$

for $\bar{X}(\cdot)$ solving the following SDE

$$\begin{cases} d\bar{X}(s)=b(s,\bar{X}(s),\bar{u}(s))ds+\sigma(s,\bar{X}(s),\bar{u}(s))dW(s), & s\in[0,T]\\ \bar{X}(0)=x \end{cases}\tag{2.32}$$

with

$$\bar{u}(s)=\Psi(s,\bar{X}(s)),\quad s\in[0,T]\tag{2.33}$$

where we recall that

$$\mathbb{S}^2(0,T)^n = \Big\{ X : \Omega \times [0,T] \mapsto \mathbb{R}^n \;\Big|\; X(\cdot) \text{ is progressively measurable and } \mathbb{E}\Big[\sup_{s\in[0,T]} \big|X(s)\big|^2\Big] < \infty \Big\}$$

also we have,

$$\lim_{\|\Pi\|\to 0} \|\bar{u}^\Pi(\cdot) - \bar{u}(\cdot)\|_{\mathcal{U}^2[0,T]} = 0$$

By (2.31) we have

$$\mathcal{J}(l^\Pi(t), \bar{X}^\Pi(l^\Pi(t)), \bar{u}^\Pi) = V^\Pi(l^\Pi(t), \bar{X}^\Pi(l^\Pi(t))), \qquad a.s.,\ t \in [0,T]$$

which passing to the limit shows that $V(\cdot,\cdot)$ is the time-consistent equilibrium value function of the time-inconsistent optimal control problem. Moreover we have the following:

$$\begin{aligned}
&\mathcal{J}(t_{k-1}, \bar{X}(t_{k-1}), \Psi\big|_{[t_{k-1},T]}) = \mathcal{J}(t_{k-1}, \bar{X}(t_{k-1}), \bar{u}\big|_{[t_{k-1},T]}) \\
&= \mathbb{E}_{t_{k-1}}\Big[\int_{t_{k-1}}^{T} C(t_{k-1}, s, \bar{X}(s), \bar{u}(s))ds + G(t_{k-1}, \bar{X}(T))\Big] \\
&\le \mathbb{E}_{t_{k-1}}\Big[\int_{t_{k-1}}^{T} C(t_{k-1}, s, \bar{X}^\Pi(s), \bar{u}^\Pi(s))ds + G(t_{k-1}, \bar{X}^\Pi(T))\Big] + R(\|\Pi\|) \\
&= V^\Pi(t_{k-1}, \bar{X}^\Pi(t_{k-1})) + R(\|\Pi\|) = \mathcal{J}^k(t_{k-1}, \bar{X}^\Pi(t_{k-1}), \bar{u}^\Pi\big|_{[t_{k-1},t_k]}) + R(\|\Pi\|) \\
&\le \mathcal{J}^k(t_{k-1}, \bar{X}^\Pi(t_{k-1}), u^k) + R(\|\Pi\|) \\
&= \mathcal{J}(t_{k-1}, \bar{X}(t_{k-1}), u^k \oplus \Psi\big|_{[t_k,T]}) + R(\|\Pi\|), \qquad u^k \in \mathcal{U}[t_{k-1}, t_k]
\end{aligned},$$

Thus, by Definition 2.2, $\Psi(\cdot,\cdot) = \bar{u}(\cdot,\cdot)$ is a time-consistent equilibrium strategy to our original time-inconsistent optimal control problem.

In the remaining part of the section we will pass to the limit to find the equation that can be used to characterize the equilibrium value function $V(\cdot,\cdot)$.
In order to achieve this, let us explore the relationship between the equilibrium value function of the N-person differential game $V^\Pi(\cdot,\cdot)$ and the auxiliary functions $\Theta^{k+1}(\cdot,\cdot)$.

First, let us write the equations for $\Theta^{k+1}(\cdot,\cdot)$ in integral form

$$\begin{aligned}\Theta^{k+1}(t,x) = G(t_{k-1},x) + \int_t^T \mathbb{H}(t_{k-1},s,x,\psi(l^{\Pi}(s),s,x,V_x^{\Pi}(s,x),V_{xx}^{\Pi}(s,x)),..\\ ..,\Theta_x^{k+1}(t,x),\Theta_{xx}^{k+1}(t,x))ds, \quad (t,x)\in[t_k,T]\times\mathbb{R}^n\end{aligned} \tag{2.34}$$

Let us make a first observation; as we have explained in details in the previous section, the function $\Theta^{k+1}(\cdot,\cdot)$ is introduced to formulate Problem P_k relative to the k-th player and it represents the part of the cost functional that cannot be optimized by player k, since it depends on the sophisticated equilibrium strategy determined by the following players. Notice indeed that $\Theta^{k+1}(\cdot,\cdot)$ is defined on $[t_k,T]$ which is the time interval successive to the one over which player k controls the system. However, once the k-th player has solved the corresponding HJB equation and hence determined his/her optimal control law $\psi(t_{k-1},\cdot)$ defined on $[t_{k-1},t_k]$, we can extend (2.34) to the wider interval $[t_{k-1},T]$ since the equilibrium value function $V^{\Pi}(\cdot,\cdot)$ and the sophisticated control law $\Psi^{\Pi}(\cdot,\cdot)$ defined in (2.30) are now well defined also on $[t_{k-1},t_k]$.

We can now define

$$\Theta^{\Pi}(\tau,t,x) = \sum_{k=1}^{N-1}\Theta^{k+1}(t,x)\mathbb{1}_{[t_{k-1},t_k)}(\tau), \qquad (\tau,t,x)\in D[0,T]\times\mathbb{R}^n$$

and

$$\begin{cases} G^{\Pi}(\tau,x) = \sum_{k=1}^{N-1} G(t_{k-1},x)\mathbb{1}_{[t_{k-1},t_k)}(\tau), & (\tau,x)\in[0,T]\times\mathbb{R}^n \\ C^{\Pi}(\tau,s,x,u) = \sum_{k=1}^{N-1} C(t_{k-1},s,x,u)\mathbb{1}_{[t_{k-1},t_k)}(\tau), & (\tau,s,x,u)\in D[0,T]\times\mathbb{R}^n\times U \end{cases}$$

then

$$\begin{aligned}\Theta^{\Pi}(\tau,t,x) = G^{\Pi}(\tau,x) + \int_t^T \mathbb{H}(l^{\Pi}(\tau),s,x,\psi(l^{\Pi}(s),s,x,V_x^{\Pi}(s,x),V_{xx}^{\Pi}(s,x)),..\\ ..\Theta_x^{\Pi}(\tau,s,x),\Theta_{xx}^{\Pi}(\tau,s,x))ds\end{aligned} \tag{2.35}$$

where the presence of $C^{\Pi}(\cdot)$ is hidden in the Hamiltonian function.

Let us look at $V^{\Pi}(\cdot,\cdot)$. For $(t,x)\in[t_{N-1},t_N]\times\mathbb{R}^n$, we have

$$V^{\Pi}(t,x)=G(t_{N-1},x)+\int_t^T \mathbb{H}(t_{N-1},s,x,\psi(l^{\Pi}(s),s,x,V^{\Pi}_x(s,x),V^{\Pi}_{xx}(s,x)),..\\ ..V^{\Pi}_x(s,x),V^{\Pi}_{xx}(s,x))ds$$

and for $(t,x)\in[t_{k-1},t_k)\times\mathbb{R}^n,\ k=1,2,...N-1$.

$$V^{\Pi}(t,x)=\Theta^{k+1}(t_k,x)+\int_t^{t_k} \mathbb{H}(t_{k-1},s,x,\psi(l^{\Pi}(s),s,x,V^{\Pi}_x(s,x),V^{\Pi}_{xx}(s,x)),..\\ ..V^{\Pi}_x(s,x),V^{\Pi}_{xx}(s,x))ds$$

On the other hand, for $(t,x)\in[t_{k-1},t_k)\times\mathbb{R}^n$ we have

$$\Theta^{\Pi}(t_{k-1},t,x)=\Theta^{\Pi}(t_{k-1},t_k,x)+\int_t^{t_k} \mathbb{H}(t_{k-1},s,x,\psi(l^{\Pi}(s),s,x,V^{\Pi}_x(s,x),V^{\Pi}_{xx}(s,x)),..\\ ..\Theta^{\Pi}_x(t_{k-1},s,x),\Theta^{\Pi}_{xx}(t_{k-1},s,x))ds$$

so that we have

$$V^{\Pi}(t,x)=\Theta^{\Pi}(l^{\Pi}(t),t,x),\qquad (t,x)\in[0,t_{N-1})\times\mathbb{R}^n \tag{2.36}$$

Let us assume now that

$$\lim_{\|\Pi\|\to 0}\Big(\big|\Theta^{\Pi}(\tau,t,x)-\Theta(\tau,t,x)\big|+\big|\Theta^{\Pi}_x(\tau,t,x)-\Theta_x(\tau,t,x)\big|\\ +\big|\Theta^{\Pi}_{xx}(\tau,t,x)-\Theta_{xx}(\tau,t,x)\big|\Big)=0,\qquad (\tau,s,x)\in D[0,T]\times\mathbb{R}^n$$

for some $\Theta(\cdot,\cdot,\cdot)$.
Furthermore, let us assume that the derivative of the cost functions $C(\tau,\cdot,\cdot,\cdot)$ and $G(\tau,\cdot)$ with respect to the first argument τ are bounded for any $(\tau,t,x,u)\in D[0,T]\times\mathbb{R}^n\times U$ so that

$$\lim_{\|\Pi\|\to 0}\big|G(\tau,x)-G^{\Pi}(\tau,x)\big|=\lim_{\|\Pi\|\to 0}\sum_{k=1}^{N}\big|G(\tau,x)-G(t_{k-1},x)\big|\mathbb{1}_{[t_{k-1},t_k)}(\tau)\le\lim_{\|\Pi\|\to 0}K\|\Pi\|=0$$

and similarly for the running cost we have

$$\lim_{\|\Pi\|\to 0} C^{\Pi}(\tau, s, x, \psi(l^{\Pi}(s), s, x, V^{\Pi}_x(s,x), V^{\Pi}_{xx}(s,x))) = \\ C(\tau, s, x, \psi(s, s, x, V_x(s,x), V_{xx}(s,x)))$$

As a consequence, we obtain the following integro-partial differential equation for $\Theta(\cdot,\cdot,\cdot)$

$$\Theta(\tau,t,x) = G(\tau,x) + \int_t^T \mathbb{H}(\tau, s, x, \psi(s,s,x,\Theta_x(s,s,x),\Theta_{xx}(s,s,x)), .. \\ ..\Theta_x(\tau,s,x), \Theta_{xx}(\tau,s,x))ds \tag{2.37}$$

which is the limit of (2.35) where we have taken into account (2.36). Moreover, it holds the relation

$$V(t,x) = \Theta(t,t,x), \qquad (t,x) \in [0,T]\times\mathbb{R}^n \tag{2.38}$$

We can write equation (2.37) in differential form as

$$\begin{cases} \Theta_t(\tau,t,x) + \mathbb{H}(\tau,t,x,\psi(t,t,x,\Theta_x(t,t,x),\Theta_{xx}(t,t,x)),\Theta_x(\tau,t,x),\Theta_{xx}(\tau,t,x)) = 0, \\ \qquad\qquad\qquad\qquad (\tau,t,x)\in D[0,T]\times\mathbb{R}^n \\ \Theta(\tau,T,x) = G(\tau,x) \end{cases}$$

or more explicitly as

$$\begin{cases} \Theta_t(\tau,t,x) + \langle b(t,x,\psi(t,t,x,\Theta_x(t,t,x),\Theta_x(t,t,x))),\Theta_x(\tau,t,x)\rangle \\ \quad + tr[a(t,x,\psi(t,t,x,\Theta_x(t,t,x),\Theta_x(t,t,x)))\Theta_{xx}(\tau,t,x)] \\ \quad + C(\tau,t,x,\psi(t,t,x,\Theta_x(t,t,x),\Theta_x(t,t,x))) = 0, \quad (\tau,t,x)\in D[0,T]\times\mathbb{R}^n \\ \Theta(\tau,T,x) = G(\tau,x), \qquad (\tau,x)\in[0,T]\times\mathbb{R}^n \end{cases} \tag{2.39}$$

Such an equation is called the **equilibrium Hamilton-Jacobi-Bellman equation** of the time-inconsistent optimal control problem. If a solution $\Theta(\cdot,\cdot,\cdot)$ to the above can be found then the time-consistent equilibrium value function $V(\cdot,\cdot)$ can be determined by (2.38) and the time-consistent equilibrium pair can be obtained by (2.32) and (2.33).

It is interesting to note that both $\Theta(\tau,t,x)$ and $\Theta(t,t,x)$ enter the equation where the latter is the restriction of the former to $\tau = t$. Let us remark that (2.39) is fully non linear and due to the fact that $\Theta(\tau,t,x)$ is different from $\Theta(t,t,x)$ the existing theory for non linear parabolic equation cannot be applied. Notice however that if we could determine $\Theta(t,t,x)$ in an independent way, then (2.39) would actually become linear in $\Theta(\tau,t,x)$, considering τ as a mere parameter.

In this work we will not analyze the problem of Well-Posedness of the equilibrium HJB equation which, being technically hard, would require additional tools and assumptions and, above all, would go beyond the scope of this thesis, that is to present a solution to the time-inconsistency problem in stochastic optimal control theory emphasizing the constructive part of the method and not all the downstream details.

What will be done in the next chapter is to illustrate a specific problem to which we can apply the equilibrium HJB derived from the methodology presented in the previous sections.

3 | Generalized Merton Problem

In this chapter we show how to apply the equilibrium HJB equation to a financial problem known in the literature as the generalized Merton problem. We consider an investor who possesses at time t the initial capital x and whose problem is to determine the optimal investment and consumption strategies in order to maximize his/her utility function over a fixed time horizon. For simplicity, especially in the notation of the resulting equations, we will consider a market with only two assets:

- A bank account with the deterministic short rate of interest r, i.e. a risk free asset B whose dynamics is given by

$$dB(t) = rB(t)dt.$$

- A risky asset with price process given by a standard Black-Scholes model

$$dS(t) = \mu S(t)dt + \sigma S(t)dW(t).$$

In the following we adopt the same formulation as in [12] to describe the wealth process of the investor who sets up a self-financing portfolio, investing in the two assets and consuming part of his wealth

$$\begin{cases} dX(t) = [rX(t) + (\mu - r)u(s) - c(s)]ds + \sigma u(s)dW(s), & s \in [t, T] \\ X(t) = x \end{cases}$$

where we have denoted by $u(\cdot)$ the process representing the absolute wealth invested in the stock and by $c(\cdot)$ the consumption rate process.

We consider a utility functional of power type, with a very general form for the discounting factors so that it can include various types of non-exponential discounting functions, such as hyperbolic discounting, quasi-exponential discounting and many more, all of which make the original Merton problem time-inconsistent.

$$\mathcal{J}(t,x,u,c) = \mathbb{E}_t\left[\int_t^T \nu(t,s)c(s)^\beta ds + \rho(t)X(T)^\beta\right]$$

where $\nu : D[0,T] \mapsto (0,\beta]$ with $\beta \in (0,1)$ and $\rho : [0,T] \mapsto (0,\infty)$ are the continuous discounting functions associated to the consumption strategy and to the terminal wealth respectively. Notice that both these functions present a dependence on the initial time of the optimisation thus making the problem in general time-inconsistent.
In order to formulate a well posed problem we assume in addition that as soon as the investor has no remaining wealth his/her activity terminates, with an associated value function equal to zero. As a convention, we define $x^\beta = -\infty$ if $x < 0$.
It is easy to check that both the coefficients of the wealth SDE and those of the utility functional, satisfy Assumption 2.1 and Assumption 2.2 respectively. Therefore we are working under the hypothesis which allowed us to obtain the equilibrium HJB equation.

Before turning to the actual application of the equilibrium HJB equation to our case study we analyze the **pre-commitment** and **naïve** solutions.

As we explained in section 1.3 a possibility for the investor is to commit to stick to the same discounting functions that he/she applies at time t even after that time instant. Under this assumption the problem becomes time-consistent, with the time t in the utility functional interpreted as a fixed parameter.

The HJB equation associated to the problem of the pre-committed agent becomes

$$\begin{cases} V_s + \sup\limits_{c\geq 0,\, u\in\mathbb{R}} \mathbb{H}(s,x,u,c,V_x,V_{xx}) = 0, & (s,x) \in [t,T]\times(0,\infty) \\ V(s,0) = 0, & s \in [t,T] \\ V(T,x) = \rho(t)x^\beta, & x \in (0,\infty) \end{cases}$$

We look for the maximum of the Hamiltonian

$$\mathbb{H}(s,x,u,c,V_x,V_{xx}) = rxV_x + (\mu - r)uV_x - cV_x + \frac{1}{2}\sigma^2u^2V_{xx} + \nu(t,s)$$

which, assuming an interior solution, is attained by the first order conditions, giving us

$$\bar{u}(s,x) = -\frac{(\mu-r)V_x}{\sigma^2 V_{xx}}, \qquad \bar{c}(s,x) = \left(\frac{\beta\nu(t,s)}{V_x}\right)^{\frac{1}{1-\beta}}$$

Then, by looking at its final condition, we make the following ansatz on the form of the value function

$$V(s,x) = \alpha(s)x^{\beta} \tag{3.1}$$

and substituting its explicit derivatives into $\bar{u}$ and $\bar{c}$ we obtain the control laws

$$\bar{u}(s,x) = \frac{(\mu - r)}{\sigma^2(1-\beta)}x, \qquad \bar{c}(s,x) = \left(\frac{\nu(t,s)}{\alpha(s)}\right)^{\frac{1}{1-\beta}} x \tag{3.2}$$

where $\alpha(\cdot)$ is given by the solution of the HJB equation where we have substituted the explicit derivatives of $V(\cdot,\cdot)$ and the optimal control laws $\bar{u}$ and $\bar{c}$.

$$\begin{cases} \alpha'(s) + \left(r\beta + \dfrac{\beta(\mu-r)^2}{(1-\beta)\sigma^2}\right)\alpha(s) + (1-\beta)\nu(t,s)^{\frac{1}{1-\beta}}\alpha(s)^{-\frac{\beta}{1-\beta}} = 0 \\ \alpha(T) = \rho(t) \end{cases}$$

The above ODE equation is known as a Bernoulli equation and it can be solved introducing a new function $y(s) = \alpha(s)^{1-(-\frac{\beta}{1-\beta})} = \alpha(s)^{\frac{1}{1-\beta}}$ which satisfies the linear ODE

$$\begin{cases} y'(s) + \left(\dfrac{r\beta}{1-\beta} + \dfrac{\beta(\mu-r)^2}{(1-\beta)^2\sigma^2}\right)y(s) + \nu(t,s)^{\frac{1}{1-\beta}} = 0 \\ y(T) = \rho(t)^{\frac{1}{1-\beta}} \end{cases}$$

From the solution $y(\cdot)$ we can thus obtain $\alpha(\cdot)$ and substituting in (3.1) and (3.2) we obtain the value function and the optimal control laws.

Alternatively the investor can have a naïve approach to the problem, in which case he/she keeps updating the optimal control laws by rolling over the time the pre-commitment strategies. Thus, to determine the optimal controls at time $w > t$ he/she solves the HJB equation

$$\begin{cases} V_s + \sup\limits_{c\geq 0,\, u\in\mathbb{R}} \mathbb{H}(s,x,u,c,V_x,V_{xx}) = 0, \quad (s,x)\in[w,T]\times(0,\infty) \\ V(s,0) = 0, \quad s\in[t,T] \\ V(T,x) = \rho(w)x^{\beta}, \quad x\in(0,\infty) \end{cases}$$

where the Hamiltonian function has the following form

$$\mathbb{H}(s, x, u, c, V_x, V_{xx}) = rxV_x + (\mu - r)uV_x - cV_x + \frac{1}{2}\sigma^2 u^2 V_{xx} + \nu(w, s)$$

Assuming the same assumption on the form of the cost functional, the optimal control laws are then given by

$$\bar{u}(w, x) = \frac{(\mu - r)}{\sigma^2(1-\beta)}x, \qquad \bar{c}(w, x) = \left(\frac{\nu(w, w)}{\alpha_w(w)}\right)^{\frac{1}{1-\beta}} x \tag{3.3}$$

where $\alpha_w(\cdot)$ denotes a family of ODE indexed by $w \in [0, T]$ and $\alpha_w(w)$ is given by the solution of

$$\begin{cases} \alpha_w'(s) + \left(r\beta + \dfrac{\beta(\mu - r)^2}{(1-\beta)\sigma^2}\right)\alpha_w(s) + (1-\beta)\nu(w, s)^{\frac{1}{1-\beta}}\alpha_w(s)^{-\frac{\beta}{1-\beta}} = 0, \quad s \in [w, T] \\ \alpha(T) = \rho(w) \end{cases}$$

evaluated at time $s = w$.

At last we consider the sophisticated approach which is the one that takes actually into account the time-changing feature of the discounting functions. The investor must solve the equilibrium HJB equation derived in the previous chapter

$$\begin{cases} \Theta_s(t, s, x) + \mathbb{H}(t, s, \psi(s, s, x, \Theta_x(s, s, x), \Theta_{xx}(s, s, x)), \Theta_x(t, s, x), \Theta_{xx}(t, s, x)) = 0, \\ \qquad\qquad\qquad\qquad\qquad\qquad (t, s, x) \in D[0, T] \times (0, \infty) \\ \Theta(t, s, 0) = 0, \quad (t, s) \in D[0, T] \\ \Theta(t, T, x) = \rho(t)x^\beta, \quad (t, x) \in [0, T] \times (0, \infty) \end{cases}$$

where we recall that the function $\Theta(\cdot, \cdot, \cdot)$ evaluated on the diagonal equals the value function of the problem $\Theta(s, s, x) = V(s, x)$ and the function ψ maximizes the Hamiltonian function with respect to the control variables

$$\mathbb{H}(t, s, x, u, c, p, P) = p[rx + (\mu - r)u - c] + \frac{1}{2}\sigma^2 u^2 P + \nu(t, s)c^\beta$$

where for simplicity we have used the letters p and P to denote $\Theta_x(t, s, x)$ and $\Theta_{xx}(t, s, x)$. respectively. Again, by the first order conditions we obtain

$$\psi(t, s, x, p, P) = (\bar{u}, \bar{c}) = \left(-\frac{(\mu - r)p}{\sigma^2 P}, \left[\frac{\beta\nu(t, s)}{p}\right]^{\frac{1}{1-\beta}}\right)$$

in particular

$$\psi(s,s,x,p,P) = (\bar{u},\bar{c}) = \left(-\frac{(\mu-r)\bar{p}}{\sigma^2\bar{P}}, \left[\frac{\beta\nu(s,s)}{\bar{p}}\right]^{\frac{1}{1-\beta}}\right)$$

with $\bar{p} = \Theta_x(s,s,x)$ and $\bar{P} = \Theta_{xx}(s,s,x)$. and

$$\mathbb{H}(\tau,s,x,\psi(s,s,x,\bar{p},\bar{P}),p,P) = rxp + \frac{(\mu-r)^2\bar{p}P}{2\sigma^2\bar{P}}\left(\frac{\bar{p}}{\bar{P}} - 2\frac{p}{P}\right) + \frac{\big(\beta\nu(s,s)\big)^{\frac{\beta}{1-\beta}}}{\bar{p}^{\frac{1}{1-\beta}}}(\nu(\tau,s)\bar{p} - \beta\nu(s,s)p)$$

so that the equilibrium HJB equation reads

$$\begin{cases} \Theta_s(\tau,s,x) + \dfrac{(\mu-r)^2\Theta_x(s,s,x)\Theta_{xx}(\tau,s,x)}{2\sigma^2\Theta_{xx}(s,s,x)}\left(\dfrac{\Theta_x(s,s,x)}{\Theta_{xx}(s,s,x)} - 2\dfrac{\Theta_x(\tau,s,x)}{\Theta_{xx}(\tau,s,x)}\right) \\ \quad + rx\Theta_x(\tau,s,x) + \dfrac{\big(\beta\nu(s,s)\big)^{\frac{\beta}{1-\beta}}}{\Theta_x(s,s,x)^{\frac{1}{1-\beta}}}(\nu(\tau,s)\Theta_x(s,s,x) - \beta\nu(s,s)\Theta_x(\tau,s,x)) = 0, \\ \qquad\qquad (\tau,s,x) \in D[0,T]\times(0,\infty) \\ \Theta(\tau,s,0) = 0, \qquad (\tau,s)\in D[0,T] \\ \Theta(\tau,T,x) = \rho(\tau)x^\beta, \qquad (\tau,x)\in[0,T]\times(0,\infty) \end{cases}$$

Now we make the following guess on the form of the solution

$$\Theta(\tau,s,x) = \varphi(\tau,s)x^\beta, \qquad (\tau,s,x)\in D[0,T]\times(0,\infty)$$

which after some substitutions and rearrangements allows us to write the equilibrium HJB equation in the form

$$\begin{cases} \varphi_s(\tau,s) + \left[\lambda - \beta\left(\dfrac{\nu(s,s)}{\varphi(s,s)}\right)^{\frac{1}{1-\beta}}\right]\varphi(\tau,s) + \left(\dfrac{\nu(s,s)}{\varphi(s,s)}\right)^{\frac{\beta}{1-\beta}}\nu(\tau,s) = 0 \\ \qquad\qquad (\tau,s)\in D[0,T] \\ \varphi(\tau,T) = \rho(\tau) \end{cases} \tag{3.4}$$

where $\lambda = \frac{[2r\sigma^2(1-\beta)+(\mu-r)^2]\beta}{2\sigma^2(1-\beta)}$ is a constant.

We notice that, considering $\varphi(s,s)$ as a known function, (3.4) is a linear first order differential equation in the unknown $\varphi(\tau,s)$, thus

$$\varphi(\tau,s) = e^{\lambda(T-s)-\beta\int_s^T\left(\frac{\nu(l,l)}{\varphi(l,l)}\right)^{\frac{1}{1-\beta}}dl}\rho(\tau) + \int_s^T e^{\lambda(l-s)-\beta\int_s^l\left(\frac{\nu(w,w)}{\varphi(w,w)}\right)^{\frac{1}{1-\beta}}dw}\left(\frac{\nu(l,l)}{\varphi(l,l)}\right)^{\frac{\beta}{1-\beta}}\nu(\tau,l)dl \quad (\tau,s)\in D[0,T]$$

In particular we obtain the following integral equation for the map $s \mapsto \varphi(s,s)$:

$$\varphi(s,s) = e^{\lambda(T-s)-\beta\int_s^T\left(\frac{\nu(l,l)}{\varphi(l,l)}\right)^{\frac{1}{1-\beta}}dl}\rho(s) + \int_s^T e^{\lambda(l-s)-\beta\int_s^l\left(\frac{\nu(w,w)}{\varphi(w,w)}\right)^{\frac{1}{1-\beta}}dw}\left(\frac{\nu(l,l)}{\varphi(l,l)}\right)^{\frac{\beta}{1-\beta}}\nu(s,l)dl \quad s\in[0,T] \tag{3.5}$$

If equation (3.5) admits a unique solution then we can obtain the time-consistent equilibrium value function

$$V(t,x) = \Theta(t,t,x) = \varphi(t,t)x^\beta, \qquad (t,x)\in[0,T]\times(0,\infty)$$

and the time-consistent equilibrium optimal control

$$\bar{u}(t) = \frac{(\mu-r)}{\sigma^2(1-\beta)}\bar{X}(t), \qquad \bar{c}(t) = \left(\frac{\nu(t,t)}{\varphi(t,t)}\right)^{\frac{1}{1-\beta}}\bar{X}(t), \qquad t\in[0,T]$$

To evaluate the well-posedness of (3.5) we proceed as follows: first, we make a change of variable introducing

$$z(s) = \frac{\varphi(s,s)}{\nu(s,s)}, \qquad s\in[0,T]$$

so that we can rewrite (3.5) as

$$z(s) = e^{\lambda(T-s)-\beta\int_s^T z(w)^{\frac{1}{\beta-1}}dw}\rho(s) + \int_s^T e^{\lambda(\tau-s)-\beta\int_s^\tau z(w)^{\frac{1}{\beta-1}}dw} z(\tau)^{\frac{\beta}{\beta-1}}\nu(s,\tau)d\tau, \quad s\in[0,T] \tag{3.6}$$

For $z(\cdot)$ we have the following result, which the reader can find in [12].

Proposition 3.1. *Suppose $z:[0,T]\mapsto(0,\infty)$ is a solution to (3.6). Then*

$$e^{(\lambda-\bar{\lambda})(T-s)}\rho_0 \le z(s) \le e^{\lambda(T-s)}\rho_1, \qquad s\in[0,T]$$

with

$$\rho_0 = \min_{t\in[0,T]} \rho(t), \qquad \rho_1 = \max_{t\in[0,T]} \rho(t), \qquad \bar{\lambda} = \sup_{0\leq t\leq s\leq T} \frac{-ln(\nu(t,s))}{s-t} < \infty$$

Thanks to the above Proposition 3.1 it is possible to show that the operator $T : (C[0,T], \|\cdot\|_\infty) \mapsto (C[0,T], \|\cdot\|_\infty)$ defined by the right-hand side of (3.6), i.e.

$$T : v(t) \mapsto e^{\lambda(T-t)-\beta\int_t^T v(w)^{\frac{1}{\beta-1}}dw}\rho(t) + \int_t^T e^{\lambda(\tau-t)-\beta\int_t^\tau v(w)^{\frac{1}{\beta-1}}dw} v(\tau)^{\frac{\beta}{\beta-1}}\nu(t,\tau)d\tau$$

is a contraction, at least locally. This guarantees the existence of a local solution thanks to contraction mapping theorem. Indeed, the operator T is a map on a complete metric space into itself hence, being a contraction, it admits a unique fixed point belonging to the space of continuous functions. Finally, exploiting a continuation argument the solution can be extended to the whole interval $[0,T]$.

4 | Conclusions and future developments

In this work we have confronted with the study of time inconsistency in stochastic optimal control problems. Starting from a standard time-consistent problem we have analyzed the complications that arise when we modify the cost functional of the control problem, introducing an explicit dependence on the initial pair time-state of the optimisation, thus making the problem time-inconsistent. We have shown how the HJB equation, which is one of the most central results in the theory of stochastic optimal control, can not be directly applied to problems that are time-inconsistent, since the principle of dynamic programming does not apply to this family. Using the HJB equation to solve this class of problems would in fact give rise to optimal control laws from which, as time goes by, the controller himself/herself would be encouraged to depart. We have thus examined into details a possible approach to deal with time-inconsistency, namely the game-theoretic approach, which consists in interpreting the control problem as a differential game played by a continuum of agents, each one trying to minimize his/her cost functional. This new perspective on the control problem allowed us to introduce a new paradigm for the optimal decision rules based on the concept of Nash equilibrium strategy of the game. In particular, we have first discretized the original sequential game, introducing an arbitrary N-partition of $[0, T]$ with N associated players, each one governing the system over the relative sub-interval. Proceeding backward in time we have then constructed a Nash equilibrium control strategy for the N-game, i.e. a strategy profile for which none of the players has any incentives to unilaterally deviate from it. Next, we have considered the limit as the mesh of the partition goes to zero in order to take us back to the situation of a game played by a continuum of agents. This limiting procedure has allowed us to obtain the equilibrium control strategy and an equation for the equilibrium value function of our original time-inconsistent optimal control problem which is called equilibrium HJB equation. To conclude the work we have shown an example of a time-inconsistent problem which arises in finance and that can be dealt with exploiting the results presented in this thesis.

There are several possible developments and insights for this study of time-inconsistency in stochastic optimal control problems which have not been treated in this work. First of all, it could be carried out an analysis on the well-posedness of the equilibrium HJB equation (2.39) in order to provide some strong analytical guarantees of the applicability of the equation, at least to certain classes of problems. As suggested in [12] this is a very difficult question to tackle in the general case while a viable path is to restrict to control problems in which the control does not enter the diffusion term of the state equation. This would indeed simplify the final form of the equilibrium HJB making it easier to assess its well-posedness. Also, it would be interesting to study the connections between the results presented in this work, in particular the equilibrium HJB equation, which we recall was derived following [12] and other analogous results in the literature, such as those presented in [2] where the authors find a system of non-linear equations which is the counterpart of the equilibrium HJB equation. Another line of development of this study of time-inconsistency might be to deal with other kinds of problems such as time-inconsistent stopping problems.

Bibliography

[1] T. Björk. *Arbitrage theory in continuous time*. Oxford university press, 2009.

[2] T. Bjork and A. Murgoci. A general theory of markovian time inconsistent stochastic control problems. *Available at SSRN 1694759*, 2010.

[3] Ł. Delong et al. Time-inconsistent stochastic optimal control problems in insurance and finance. *Warsaw School of Economics, Collegium of Economic Analysis Annals*, 51:229–254, 2018.

[4] M. A. Fuhrman and G. Tessitore. Backward stochastic differential equations in finite and infinite dimensions. applications to optimal control and hedging. 2004.

[5] S. Gordon and P. St-Amour. A preference regime model of bull and bear markets. *American Economic Review*, 90(4):1019–1033, 2000.

[6] G. Loewenstein and D. Prelec. Anomalies in intertemporal choice: Evidence and an interpretation. *The Quarterly Journal of Economics*, 107(2):573–597, 1992.

[7] E. G. Luttmer and T. Mariotti. Subjective discounting in an exchange economy. *Journal of Political Economy*, 111(5):959–989, 2003.

[8] É. Pardoux. Backward stochastic differential equations and viscosity solutions of systems of semilinear parabolic and elliptic pdes of second order. In *Stochastic Analysis and Related Topics VI*, pages 79–127. Springer, 1998.

[9] H. Pham. *Continuous-time stochastic control and optimization with financial applications*, volume 61. Springer Science & Business Media, 2009.

[10] R. H. Strotz. Myopia and inconsistency in dynamic utility maximization. *The review of economic studies*, 23(3):165–180, 1955.

[11] R. H. Thaler and E. J. Johnson. Gambling with the house money and trying to break even: The effects of prior outcomes on risky choice. *Management science*, 36 (6):643–660, 1990.

[12] J. Yong. Time-inconsistent optimal control problems and the equilibrium hjb equation. *arXiv preprint arXiv:1204.0568*, 2012.