Applied Mathematics

MATH5551 STOCHASTICS AND FINANCE

Marek Rutkowski

School of Mathematics and Statistics

The University of Sydney

Semester 2, 2023

Contents

A linear *Backward Stochastic Differential Equation* (BSDE) was introduced by Jean–Michel Bismut in his paper [9] published in 1973 as an equation for the adjoint process in the stochastic version of Pontryagin maximum principle. In 1990, Etienne Pardoux and Shige Peng considered in [72] a general nonlinear BSDE of the following form, for a fixed $T > 0$ and every $t \in [0, T]$,

$$Y_t = \xi_T + \int_t^T g(s, Y_s, Z_s)\, ds - \int_t^T Z_s\, dW_s$$

where W is a d-dimensional Brownian motion, the terminal condition ξ_T is a predetermined random variable and the so-called *generator* $g : [0, T] \times \mathbb{R}^m \times \mathbb{R}^{m\times d} \to \mathbb{R}^m$ is a Lipschitz continuous (or merely continuous) map with respect to (y, z). By a *solution* to the BSDE, we mean a pair (Y, Z) of processes adapted to the filtration generated by the Brownian motion W and such that the equation is satisfied for all $t \in [0, T]$.

In 1990, Etienne Pardoux and Shige Peng established in their path-breaking paper [72] the existence and uniqueness theorem for a solution to BSDE with a uniformly Lipschitz continuous generator g. Their result was subsequently extended to the case of BSDEs with a continuous generator, as well as to BSDEs with either a single reflecting boundary (RBSDEs) or two reflecting boundaries (DRBSDEs). These developments were motivated by several applications of BSDEs in solving mathematical and practical problems in which the noise process appears in a natural way. During the past twenty years BSDEs have been widely used as a convenient mathematical tool to address problems in a wide range of areas including:

- an extension of the classic Feynman–Kac formula to semilinear parabolic PDEs via Markovian forward-backward stochastic differential equations,
- replication and super-replication problems for financial derivatives in market models either without frictions (using linear BSDEs) or with trading constraints (using nonlinear BSDEs),
- stochastic optimal control, where BSDEs can be seen an alternative to the more classic approach based on the Hamilton–Jacobi–Bellman (HJB) equation,
- Nash equilibria in zero-sum stochastic differential games, i.e., stochastic optimisation problems where two agents seek to attain antagonistic goals,
- optimal stopping problems and zero-sum Dynkin games via either reflected or doubly reflected BSDEs.

The goals of the course are twofold. First, we give a detailed introduction (with proofs of most results) to the general theory of BSDEs. This is followed by an overview of applications of BSDEs to various optimisation problems. We acknowledge that our course notes are partially based on the paper:

- Nicole El Karoui, Said Hamadène and Anis Matoussi: *Backward Stochastic Differential Equations and Applications*. In: *Indifference Pricing: Theory and Applications*, ed. René Carmona, Princeton University Press, 2009, pp. 267–320

as well as numerous papers by other authors.

The prerequisite for this course is some familiarity with the Itô stochastic calculus and basic results for stochastic differential equations. These topics are covered by the honours units of study MATH4511 *Arbitrage Pricing in Continuous Time* and MATH4512 *Stochastic Analysis*.

Preliminaries

Backward Stochastic Differential Equation

Let $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P})$ be a filtered probability space satisfying the usual hypotheses of right-continuity and completeness. Let $(W_t)_{t\in[0,T]}$ be an $\mathbb{R}^d$-valued standard Brownian motion. We assume that $\mathbb{F} = (\mathcal{F}_t)_{t\in[0,T]}$ is the filtration generated by the Brownian motion W and augmented by the class of $\mathbb{P}$-null sets of $\mathcal{F}_T$.

Our goal is to study the following **backward stochastic differential equation** (BSDE)

$$\begin{cases} -dY_t = g(t, Y_t, Z_t)\, dt - Z_t\, dW_t, \quad t \in [0, T], \\ Y_T = \xi_T, \end{cases}$$

where the processes Y, Z take values in $\mathbb{R}^m$ and $\mathbb{R}^{m\times d}$, respectively, and ξ_T is an $\mathbb{R}^m$-valued, $\mathcal{F}_T$-measurable random variable. The function $g : \Omega \times [0, T] \times \mathbb{R}^m \times \mathbb{R}^{m\times d} \to \mathbb{R}^m$ is called the *generator* (or *driver*). The following definition gives minimal conditions under which a pair (Y, Z) of stochastic processes can be interpreted as a solution to the backward stochastic differential equation. Let $\|\cdot\|$ stand for the usual Euclidean norm in either $\mathbb{R}^m$ or $\mathbb{R}^{m\times d}$.

Definition. A pair (Y, Z) is a *solution* to backward stochastic differential equation if the following conditions are satisfied:
(i) $(Y_t)_{t\in[0,T]}$ is an $\mathbb{R}^m$-valued, continuous and $\mathbb{F}$-adapted process;
(ii) $(Z_t)_{t\in[0,T]}$ is an $\mathbb{R}^{m\times d}$-valued, $\mathbb{F}$-predictable process and $\mathbb{P}\big(\int_0^T \|Z_s\|^2\, ds < \infty\big) = 1$;
(iii) the following equality is satisfied for all $t \in [0, T]$

$$Y_t = \xi_T + \int_t^T g(s, Y_s, Z_s)\, ds - \int_t^T Z_s\, dW_s. \tag{0.1}$$

Notice that the process $M_t := \int_0^t Z_s\, dW_s$ is a local martingale. If we assume that M is a martingale, then by taking the conditional expectation with respect to $\mathcal{F}_t$ we obtain the following consequence of (0.1)

$$Y_t = \mathbb{E}\Big(\xi_T + \int_t^T g(s, Y_s, Z_s)\, ds \,\Big|\, \mathcal{F}_t\Big),$$

which suggests that the concept of a BSDE is a far-reaching extension of the classical notion of conditional expectation combined with the predictable representation property of a Brownian motion.

Predictable Representation Property

For the proof of the next result, see Proposition 3.2 in Chapter 5 in Revuz and Yor [84] (for a more general version, see Theorem 4.2.15 in Bichteler [9]).

Theorem. *Let W be a d-dimensional Brownian motion and let $\mathbb{F} = \mathbb{F}^W$ be its natural filtration. Assume that ξ is $\mathcal{F}_T$-measurable and square-integrable, that is, $\xi \in L^2(\mathcal{F}_T)$. Then there exists a unique $\mathbb{R}^d$-valued process $\gamma \in \mathcal{H}^2_m(0, T)$ such that*

$$\xi = \mathbb{E}(\xi) + \int_0^T \gamma_t\, dW_t. \tag{0.2}$$

Consequently, the square-integrable martingale $M_t := \mathbb{E}(\xi \mid \mathcal{F}_t)$ has the representation

$$M_t = \mathbb{E}(\xi) + \int_0^t \gamma_s \, dW_s. \tag{0.3}$$

More generally, any local martingale M with respect to $\mathbb{F}^W$ has the representation

$$M_t = c + \int_0^t \gamma_s \, dW_s$$

where c is a constant and the process γ is locally in $\mathcal{H}^2_m(0,T)$.

Note that the uniqueness of γ means that if γ and $\widetilde{\gamma}$ are any two processes from $\mathcal{H}^2_m(0,T)$ yielding the integral representation of ξ, then $\|\gamma - \widetilde{\gamma}\|_2 = 0$.

Notice that (0.3) gives for $t = T$

$$M_T = \mathbb{E}(\xi) + \int_0^T \gamma_s \, dW_s \tag{0.4}$$

and thus, by subtracting (0.3) from (0.4), we obtain

$$M_T - M_t = \int_t^T \gamma_s \, dW_s$$

which also means that

$$M_t = \xi - \int_t^T \gamma_s \, dW_s \tag{0.5}$$

since we also know that $M_T = \xi$.

Notice that equality (0.5) can also be interpreted as a *backward equation* where the terminal condition ξ is given and a pair (M, γ) of stochastic processes is not known. Then the theorem shows that if $\xi \in L^2(\mathcal{F}_T)$, then a unique solution to the backward equation (0.5) is given by the square-integrable martingale $M_t = \mathbb{E}(\xi \mid \mathcal{F}_t)$ and the process $\gamma \in \mathcal{H}^2_m(0,T)$, which is given by (0.2).

Optional and Predictable Processes

Let $T > 0$ be either a finite or infinite horizon date and let $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P})$ be a probability space where the filtration $\mathbb{F} = (\mathcal{F}_t)_{t\in[0,T]}$ is assumed to satisfy the usual conditions of right-continuity and $\mathbb{P}$-completeness.

Definition. A stochastic process $X = (X_t)_{t\in[0,T]}$ on $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P})$ is said to be $\mathbb{F}$-*progressively measurable* if for every $t \in [0,T]$ the restriction of the mapping $(\omega, s) \to X_s(\omega)$ to $\Omega \times [0,t]$ is $\mathcal{B}[0,t] \otimes \mathcal{F}_t$-measurable. A set $A \subset \Omega \times [0,T]$ is $\mathbb{F}$-*progressively measurable* if the process $\mathbb{1}_A$ is $\mathbb{F}$-progressively measurable. We denote by $\mathcal{P}r$ the σ-algebra of $\mathbb{F}$-progressively measurable sets.

We denote by $\mathcal{O}$ (resp., $\mathcal{P}$) the σ-algebra of $\mathbb{F}$-*optional* (resp., $\mathbb{F}$-*predictable*) sets in $\Omega \times [0,T]$. We say that a process is càdlàg if it is RCLL.

Definition. The σ-algebra $\mathcal{O}$ (resp., $\mathcal{P}$) in $\Omega \times [0,T]$ is generated by all càdlàg (resp., left-continuous) and $\mathbb{F}$-adapted stochastic processes. A stochastic process $X = (X_t)_{t\in[0,T]}$ on $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P})$

is said to be $\mathbb{F}$-*optional* (resp., $\mathbb{F}$-*predictable*) if the mapping $X : \Omega \times [0,T] \to \mathbb{R}$ is $\mathcal{O}$-measurable (resp., $\mathcal{P}$-measurable).

Any $\mathbb{F}$-adapted, left-continuous process is $\mathbb{F}$-predictable and every such process can be written as a limit of $\mathbb{F}$-adapted, càdlàg processes so it is also $\mathbb{F}$-optional, which shows that $\mathcal{P} \subset \mathcal{O}$. Furthermore, it is easy to show that every $\mathbb{F}$-adapted, càdlàg process is $\mathbb{F}$-progressively measurable and thus any $\mathbb{F}$-optional process is $\mathbb{F}$-progressively measurable. We conclude that $\mathcal{P} \subset \mathcal{O} \subset \mathcal{P}r$.

If the filtration $\mathbb{F}$ is generated by a Brownian motion W and satisfies the usual conditions of right-continuity and completeness, then $\mathcal{P} = \mathcal{O}$ and thus any $\mathbb{F}$-optional process is also $\mathbb{F}$-predictable. Consequently, any $\mathbb{F}$-stopping time is also an $\mathbb{F}$-predictable stopping time. However, the inclusion $\mathcal{O} \subset \mathcal{P}r$ is strict for the Brownian filtration. For instance, consider the set $\{(\omega,t) \in \Omega \times [0,T] : W_t \neq 0\}$, which is a union of disjoint open *excursion intervals*. Then the set of left endpoints of excursion intervals is $\mathbb{F}$-progressively measurable but not $\mathbb{F}$-optional. The standard Poisson process is optional in its own filtration, but not predictable, so the inclusion $\mathcal{O} \subset \mathcal{P}r$ is also strict, in general.

Let us recall a general result on a relationship between the σ-algebras $\mathcal{O}$ and $\mathcal{P}$.

Proposition A. *We have $\mathcal{O} = \mathcal{P} \vee \mathcal{M}$ where the σ-algebra $\mathcal{M}$ is generated by jumps of bounded $\mathbb{F}$-martingales, that is, by the processes $M - M_-$ where M is a bounded $\mathbb{F}$-martingale.*

Definition. A random time $\tau : \Omega \to [0,T]$ is called an $\mathbb{F}$-*stopping time* if the process $A_t := \mathbb{1}_{\{\tau \le t\}}$ is $\mathbb{F}$-optional (or, equivalently, $\mathbb{F}$-adapted) and it is an $\mathbb{F}$-*predictable stopping time* if the càdlàg process A is $\mathbb{F}$-predictable. We denote by $\mathcal{T}_{[0,T]}$ (resp., $\mathcal{T}^p_{[0,T]}$) the class of all $\mathbb{F}$-*stopping times* (resp., $\mathbb{F}$-*predictable stopping times*) with values in $[0,T]$ so that $\mathcal{T}^p_{[0,T]} \subset \mathcal{T}_{[0,T]}$.

We define the σ-fields $\mathcal{F}_\tau$ and $\mathcal{F}_{\tau-}$ of all events occurring up to a stopping time τ and strictly before a stopping time τ, respectively.

Definition. If τ is an $\mathbb{F}$-stopping time, then the σ-field $\mathcal{F}_\tau$ is defined by

$$\mathcal{F}_\tau = \{F \in \mathcal{F}_\infty \,|\, F \cap \{\tau \le t\} \in \mathcal{F}_t \text{ for all } t \in \mathbb{R}_+\}$$

and the σ-field $\mathcal{F}_{\tau-}$ is generated by all events from $\mathcal{F}_0$ and all events of the form $B \cap \{\tau > t\}$ with $B \in \mathcal{F}_t$ and $t \ge 0$.

We are ready to state some basic results on σ-fields $\mathcal{F}_\tau$ and $\mathcal{F}_{\tau-}$.

Proposition B.1 *Let τ be an $\mathbb{F}$-stopping time. A random variable Y is $\mathcal{F}_\tau$-measurable (resp., $\mathcal{F}_{\tau-}$-measurable) if and only if there exists an $\mathbb{F}$-optional (resp., $\mathbb{F}$-predictable) process X such that $Y = X_\tau$ on the set $\{\tau < \infty\}$.*

Proposition B.2 *If X is an $\mathbb{F}$-optional process, then for every $\tau \in \mathcal{T}_{[0,T]}$:*
(i) the random variable $X_\tau \mathbb{1}_{\{\tau<\infty\}}$ is $\mathcal{F}_\tau$-measurable;
(ii) the stopped process X^τ is optional.

Proposition B.3 *If X is an $\mathbb{F}$-predictable process, then for every $\tau \in \mathcal{T}_{[0,T]}$:*
(i) the random variable $X_\tau \mathbb{1}_{\{\tau<\infty\}}$ is $\mathcal{F}_{\tau-}$-measurable;
(ii) the stopped process X^τ is $\mathbb{F}$-predictable.

Section Theorem

The *section theorem* is an important result belonging to the realm of the general theory of stochastic processes, which was developed in 1960s and 1970s by French mathematicians Paul-André Meyer, Claude Dellacherie, Catherine Doléans-Dade, Jean Jacod and their co-workers from the *Strasbourg School*.

Their main goal was to extend the Itô integration theory to general martingales and hence also semimartingales. Some of their final results are the celebrated Bichteler-Dellacherie theorem characterising stochastic integrators as semimartingales and the Itô-Tanaka-Meyer formula for semimartingales. Both results can be extended to the case of a filtration that is not right-continuous in the so-called *optional stochastic calculus*.

For the proof of the section theorem, we refer to the monograph by Dellacherie [16] or Theorem 6.2 in Karatzas [45].

Section Theorem. *Let B be an $\mathbb{F}$-optional (resp., $\mathbb{F}$-predictable) set and let π be the canonical projection of $\Omega \times [0,T]$ onto Ω, that is, $\pi(\omega,t) = \omega$. For every $\varepsilon > 0$, there exists an $\mathbb{F}$-stopping time (resp., an $\mathbb{F}$-predictable stopping time) $\tau : \Omega \to [0,\infty]$ such that:*
(i) the graph $[\tau]$ of τ is a subset of B;
(ii) $\mathbb{P}(\tau < \infty) \geq \mathbb{P}(\pi(B)) - \varepsilon$.

The following consequences of the optional and predictable section theorem are frequently used.

Proposition C.1 *Let X and Y be two $\mathbb{F}$-optional (resp., $\mathbb{F}$-predictable) processes. If for every finite $\mathbb{F}$-stopping time (resp., every finite $\mathbb{F}$-predictable stopping time) τ we have that $X_\tau = Y_\tau, \mathbb{P}-$a.s., then the processes X and Y are indistinguishable, that is, $\mathbb{P}(X_t = Y_t,\ \forall\, t \in [0,T]) = 1$.*

Proof. We will apply the section theorem to the $\mathbb{F}$-optional (resp., $\mathbb{F}$-predictable) set $B = \{(t,\omega) \in \mathbb{R}_+ \times \Omega : X_t(\omega) \neq Y_t(\omega)\}$. Assume that X and Y are not indistinguishable, so that the set B is not evanescent. By the section theorem, there exists an $\mathbb{F}$-stopping time with a non-evanescent graph contained in B. This in turn implies the existence of some $t \in \mathbb{R}_+$ such that $\mathbb{P}(X_{\tau\wedge t} = Y_{\tau\wedge t}) < 1$ where $\widehat{\tau} := \tau \wedge t$ is a finite $\mathbb{F}$-stopping time. □

Proposition C.2 *Let X and Y be two $\mathbb{F}$-optional (resp., $\mathbb{F}$-predictable) processes. If for every stopping time (resp., every $\mathbb{F}$-predictable stopping time)*

$$\mathbb{E}\big(X_\tau \mathbb{1}_{\{\tau<\infty\}}\big) = \mathbb{E}\big(Y_\tau \mathbb{1}_{\{\tau<\infty\}}\big),$$

then the processes X and Y are indistinguishable.

Proof. It suffices to apply the section theorem to the $\mathbb{F}$-optional (resp., $\mathbb{F}$-predictable) sets

$$B = \{(\omega,t) \in \Omega \times [0,T] : X_t(\omega) < Y_t(\omega)\}$$

and

$$D = \{(\omega,t) \in \Omega \times [0,T] : X_t(\omega) > Y_t(\omega)\}.$$

Then we argue along similar lines as in the proof of Proposition C.1. □

1 BSDEs with Lipschitz Continuous Generators

Our first goal is to establish the existence and uniqueness of a solution to BSDE

$$Y_t = \xi_T + \int_t^T g(s, Y_s, Z_s)\, ds - \int_t^T Z_s\, dW_s \tag{1.1}$$

in a suitable space of stochastic processes when the generator g is assumed to be uniformly Lipschitz continuous.

Banach Spaces of Stochastic Processes

The following definitions are standard in the theory of backward stochastic differential equations.

- $L^p_m(\mathcal{F}_T)$: all $\mathbb{R}^m$-valued, square-integrable, $\mathcal{F}_T$-measurable random variables with the norm for $p > 0$
$$\|\xi\|_p := [\mathbb{E}(\|\xi\|^p)]^{1/p}.$$
- $\mathcal{S}^2_m(0,T)$: the space of all càdlàg, $\mathbb{F}$-adapted (or merely $\mathbb{F}$-optional) processes $X : \Omega \times [0,T] \to \mathbb{R}^m$ such that
$$\|X^*_T\|_2 := \left[\mathbb{E}\left(\sup_{0\le t\le T} \|X_t\|^2\right)\right]^{1/2} = \left[\mathbb{E}_{\mathbb{P}}\left(\operatorname*{ess\,sup}_{\tau\in\mathcal{T}_{[0,T]}} \|X_\tau\|^2\right)\right]^{1/2} < \infty$$
where we denote $X^*_T := \sup_{t\in[0,T]} \|X_t\|$. The set of all continuous processes from $\mathcal{S}^2_m(0,T)$ is denoted by $\mathcal{S}^{2,c}_m(0,T)$.
- $\mathcal{H}^2_m(0,T)$: the space of all $\mathbb{F}$-progressively measurable processes $X : \Omega \times [0,T] \to \mathbb{R}^m$ such that
$$\|X\|_2 := \left[\mathbb{E}\left(\int_0^T \|X_t\|^2\, dt\right)\right]^{1/2} < \infty.$$
It is worth noting that $\mathcal{S}^2_m(0,T) \subset \mathcal{H}^2_m(0,T)$. The set of nondecreasing (resp., nondecreasing and continuous) processes from $\mathcal{H}^2_m(0,T)$ is denoted by $\mathcal{A}^2_m(0,T)$ (resp., $\mathcal{A}^{2,c}_m(0,T)$).
- $\mathcal{H}^1_m(0,T)$: the space of all $\mathbb{F}$-progressively measurable processes $X : \Omega \times [0,T] \to \mathbb{R}^m$ such that
$$\|X\|_1 := \mathbb{E}\left[\left(\int_0^T \|X_t\|^2\, dt\right)^{1/2}\right] < \infty.$$

It is clear that $\mathcal{H}^2_m(0,T) \subset \mathcal{H}^1_m(0,T)$. For a fixed, but arbitrary, real number $\alpha \ge 0$, we define the space $\mathcal{H}^{2,\alpha}_m(0,T)$ as the space $\mathcal{H}^2_m(0,T)$ endowed with the norm $\|\cdot\|_{2,\alpha}$ where for any $X \in \mathcal{H}^2_m(0,T)$

$$\|X\|_{2,\alpha} := \left[\mathbb{E}\left(\int_0^T e^{\alpha t}\|X_t\|^2\, dt\right)\right]^{1/2}.$$

It is easy to check that the norms $\|\cdot\|_2 = \|\cdot\|_{2,0}$ and $\|\cdot\|_{2,\alpha}$ for any fixed α are equivalent and thus, in particular, the class $\mathcal{H}^{2,\alpha}_m(0,T)$ does not depend on the choice of α. Recall that the equivalence of $\|\cdot\|_2$ and $\|\cdot\|_{2,\alpha}$ means that there exist two strictly positive constants c_α and

C_α such that for any $X \in \mathcal{H}^2_m(0,T)$ we have $c_\alpha \|X\|_2 \le \|X\|_{2,\alpha} \le C_\alpha \|X\|_2$. We take for granted the following result.

Lemma. *The space $(\mathcal{H}^{2,\alpha}_m(0,T), \|\cdot\|_{2,\alpha})$ is a Banach space, that is, a complete normed linear space.*

Let us also recall the fixed point theorem due to Banach [1].

Theorem. *Let $(E, \|\cdot\|)$ be the Banach space and let $\Phi : E \to E$ is a contraction, that is, there exists a constant $0 \le c < 1$ such that $\|\Phi(X)\| \le c\|X\|$ for all $X \in E$. Then there exists a unique fixed point $\widehat{X}$, that is, a unique $\widehat{X}$ in E such that $\Phi(\widehat{X}) = \widehat{X}$. Moreover, $\lim_{n\to\infty} \|\widehat{X} - X_n\| = 0$ where the sequence $(X_n)_{n=1}^\infty$ is defined by recurrence: $X_0 \in E$ and $X_n = \Phi(X_{n-1})$ for all $n \in \mathbb{N}$.*

1.1 Standard Parameters

In this section, we work under the following assumption (H.1) in which $\mathcal{P}$ stands for the σ-field of predictable sets in $\Omega \times [0,T]$.

Assumption 1.1. We say that g satisfies *assumption (H.1)* if:
(i) the mapping $g : \Omega \times [0,T] \times \mathbb{R}^m \times \mathbb{R}^{m\times d} \to \mathbb{R}^m$ is $\mathcal{P} \otimes \mathcal{B}(\mathbb{R}^m) \otimes \mathcal{B}(\mathbb{R}^{m\times d})$-measurable;
(ii) the process $g(\cdot, 0, 0)$ belongs to $\mathcal{H}^2_m(0,T)$;
(iii) the mapping g is uniformly Lipschitz continuous: there exists a constant $L > 0$ such that for all $y_1, y_2 \in \mathbb{R}^m$ and $z_1, z_2 \in \mathbb{R}^{m\times d}$

$$\|g(\omega,t,y_1,z_1) - g(\omega,t,y_2,z_2)\| \le L\big(\|y_1 - y_2\| + \|z_1 - z_2\|\big), \quad d\mathbb{P} \otimes dt - \text{a.e.} \tag{1.2}$$

Assumption (H.1) needs to be complemented by some conditions regarding the terminal condition ξ_T. It is common to postulate that ξ_T is a square-integrable random variable. This leads to the following definition of *standard parameters*.

Definition 1.1. We say that (g, ξ_T) are *standard parameters* if:
(i) the mapping $g : \Omega \times [0,T] \times \mathbb{R}^m \times \mathbb{R}^{m\times d} \to \mathbb{R}^m$ satisfies assumption (H.1);
(ii) the terminal condition ξ_T belongs to $L^2_m(\mathcal{F}_T)$.

1.2 BSDE with the Null Generator

As a preliminary step, we examine the case of a BSDE with a null generator, that is, when the mapping g vanishes.

Lemma 1.1. *Assume that $\xi_T \in L^2(\mathcal{F}_T)$. Then the BSDE with the generator $g \equiv 0$*

$$Y_t = \xi_T - \int_t^T Z_s \, dW_s \tag{1.3}$$

has a unique solution (Y,Z) in the space $\mathcal{H}^2_m(0,T) \times \mathcal{H}^2_{m\times d}(0,T)$. Moreover, Y belongs to $\mathcal{S}^{2,c}_m(0,T)$.

Proof. We search for a solution (Y,Z) in the space $\mathcal{H}^2_m(0,T) \times \mathcal{H}^2_{m\times d}(0,T)$. The process Z (if it exists) belongs to $\mathcal{H}^2_{m\times d}(0,T)$ and thus the process $\int_0^t Z_s \, dW_s$ is a square-integrable martingale so that for every $t \in [0,T]$

$$\mathbb{E}\Big(\int_0^T Z_s \, dW_s \,\Big|\, \mathcal{F}_t\Big) = \int_0^t Z_s \, dW_s.$$

Therefore, for all $t \in [0,T]$
$$\mathbb{E}\Big(\int_t^T Z_s \, dW_s \,\Big|\, \mathcal{F}_t \Big) = 0$$
and thus, by taking the conditional expectation of both sides of (1.3), we obtain the equality $Y_t = \mathbb{E}(\xi_T \,|\, \mathcal{F}_t)$ for all $t \in [0,T]$. Since $\xi_T \in L^2(\mathcal{F}_T)$, from the Doob inequality, we deduce that Y belongs to the space $\mathcal{S}^{2,c}_m(0,T)$ (and thus also to $\mathcal{H}^2_m(0,T)$). This proves the existence and uniqueness of the component Y of a solution (Y,Z). It remains to establish the existence and uniqueness of Z in the space $\mathcal{H}^2_{m\times d}(0,T)$.

We already know that if Z belongs to $\mathcal{H}^2_{m\times d}(0,T)$ then Y equals $Y_t = \mathbb{E}(\xi_T \,|\, \mathcal{F}_t)$. By taking (1.3) with $t=0$ and subtracting from (1.3), we obtain
$$Y_t = \mathbb{E}(\xi_T \,|\, \mathcal{F}_t) = Y_0 + \int_0^t Z_s \, dW_s. \tag{1.4}$$

By applying the predictable representation property to the martingale M given by $M_t := \mathbb{E}(\xi_T \,|\, \mathcal{F}_t)$ we obtain the existence of a unique process $Z \in \mathcal{H}^2_{m\times d}(0,T)$ such that the second equality in (1.4) holds for all $t \in [0,T]$. Moreover,
$$\xi_T = Y_T = Y_0 + \int_0^T Z_s \, dW_s = Y_0 + \int_0^t Z_s \, dW_s + \int_t^T Z_s \, dW_s = Y_t + \int_t^T Z_s \, dW_s$$
so that (1.3) is satisfied for all $t \in [0,T]$. We conclude that the pair (Y,Z) is a solution to BSDE (1.3) and it is the unique solution to this equation in $\mathcal{H}^2_m(0,T) \times \mathcal{H}^2_{m\times d}(0,T)$. We have also shown that if $Z \in \mathcal{H}^2_{m\times d}(0,T)$, then necessarily $Y \in \mathcal{S}^{2,c}_m(0,T)$. □

1.3 BSDE with a Fixed Generator

We extend Lemma 1.1 by considering the case of a constant generator, that is, a generator that does not depend on (y,z).

Lemma 1.2. *Assume that $\xi_T \in L^2(\mathcal{F}_T)$ and let $(u,v) \in \mathcal{H}^2_m(0,T) \times \mathcal{H}^2_{m\times d}(0,T)$ be fixed. Then the BSDE with the generator $\widehat{g}_t := g(t,u_t,v_t)$*
$$Y_t = \xi_T + \int_t^T \widehat{g}_s \, ds - \int_t^T Z_s \, dW_s \tag{1.5}$$
has a unique solution (Y,Z) in the space $\mathcal{H}^2_m(0,T) \times \mathcal{H}^2_{m\times d}(0,T)$ and Y belongs to $\mathcal{S}^{2,c}_m(0,T)$.

Proof. **Step 1.** Since $(u,v) \in \mathcal{H}^2_m(0,T) \times \mathcal{H}^2_{m\times d}(0,T)$, the process $g(\cdot,0,0)$ belongs to $\mathcal{H}^2_m(0,T)$, and the mapping g is assumed to be uniformly Lipschitz continuous, it is easy to check that the process $\widehat{g}$ belongs to $\mathcal{H}^2_m(0,T)$ and $(\widehat{g},\xi_T)$ are standard parameters.

Step 2. We are now going to "reduce" BSDE (1.5) to BSDE with the null generator. To this end, we define
$$\widehat{\xi_T} = \xi_T + \int_0^T \widehat{g}_s \, ds. \tag{1.6}$$

Using the Cauchy-Schwarz inequality for the Hilbert space $L^2((0,T),\mathbb{R})$ of square-integrable functions on $[0,T]$, we obtain

$$\mathbb{E}\bigg(\Big\|\int_0^T \widehat{g}_t\,dt\Big\|^2\bigg) \leq \mathbb{E}\bigg(\int_0^T \|\widehat{g}_t\|\,dt\bigg)^2 \leq T\,\mathbb{E}\bigg(\int_0^T \|\widehat{g}_t\|^2\,dt\bigg) < \infty.$$

Hence the integral in (1.6) belongs to $L^2(\mathcal{F}_T)$ and thus $\widehat{\xi_T} \in L^2(\mathcal{F}_T)$. Let us write

$$\widehat{Y}_t := Y_t + \int_0^t \widehat{g}_s\,ds = Y_t + \widehat{G}_t \tag{1.7}$$

where $\widehat{G}_t := \int_0^t \widehat{g}_s\,ds$. Then (1.5) can be represented as follows

$$\widehat{Y}_t = \widehat{\xi_T} - \int_t^T Z_s\,dW_s. \tag{1.8}$$

Equality (1.8) is a BSDE with null generator where the pair $(\widehat{Y}, Z)$ is unknown.

Step 3. We claim that (Y,Z) is a solution to (1.5) in $\mathcal{H}^2_m(0,T)\times\mathcal{H}^2_{m\times d}(0,T)$ if and only if $(\widehat{Y},Z)$ is a solution to (1.8) in $\mathcal{H}^2_m(0,T)\times\mathcal{H}^2_{m\times d}(0,T)$. To verify this statement, it suffices to show that a process Y belongs to $\mathcal{H}^2_m(0,T)$ if and only if $\widehat{Y}$ given by (1.7) belongs to $\mathcal{H}^2_m(0,T)$.

To establish this equivalence, we observe that $\widehat{G}$ belongs to $\mathcal{H}^2_m(0,T)$ since $\widehat{g}\in\mathcal{H}^2_m(0,T)$ and thus

$$\begin{aligned}\mathbb{E}\bigg(\int_0^T \|\widehat{G}_t\|^2\,dt\bigg) &= \mathbb{E}\bigg(\int_0^T \Big\|\int_0^t \widehat{g}_s\,ds\Big\|^2 dt\bigg) \leq \mathbb{E}\bigg(\int_0^T \Big(\int_0^t \|\widehat{g}_s\|\,ds\Big)^2 dt\bigg)\\ &\leq \mathbb{E}\bigg(\int_0^T \Big(t\int_0^t \|\widehat{g}_s\|^2\,ds\Big)dt\bigg) \leq \mathbb{E}\bigg(T\int_0^T \|\widehat{g}_t\|^2\,dt\bigg) < \infty.\end{aligned}$$

Step 4. In the last step of the proof, we will show that BSDE (1.8) has a unique solution $(\widehat{Y},Z)$ in the space $\mathcal{H}^2_m(0,T)\times\mathcal{H}^2_{m\times d}(0,T)$. To this end, we argue along the same lines as in the proof of Lemma 1.1.

Since, by assumption, a yet unknown process Z belongs to $\mathcal{H}^2_{m\times d}(0,T)$, the process $\widehat{Y}$ necessarily satisfies $\widehat{Y}_t = \mathbb{E}(\widehat{\xi_T}\,|\,\mathcal{F}_t)$ and thus Y equals

$$Y_t = \mathbb{E}(\widehat{\xi_T}\,|\,\mathcal{F}_t) - \int_0^t \widehat{g}_s\,ds = \mathbb{E}\Big(\xi_T + \int_t^T \widehat{g}_s\,ds\,\Big|\,\mathcal{F}_t\Big). \tag{1.9}$$

Note that the process $\widehat{Y}$ (and thus also the process Y) belongs to $\mathcal{H}^2_m(0,T)$. Moreover, $\widehat{Y}$ (and thus also Y) is unique. To obtain the existence and uniqueness of Z, we define the square-integrable martingale $M_t := \mathbb{E}(\widehat{\xi_T}\,|\,\mathcal{F}_t)$.

From the predictable representation property of the Brownian filtration, there exists a unique $\mathbb{F}$-predictable process $Z\in\mathcal{H}^2_{m\times d}(0,T)$ such that

$$\widehat{M_t} = \widehat{M_0} + \int_0^t Z_s\,dW_s, \quad t\in[0,T].$$

By combining (1.6) with (1.8), we confirm that the pair (Y,Z) with Y given by (1.9) is indeed a solution to BSDE (1.5). □

1.4 BSDEs with Lipschitz Continuous Generators

We will search for a solution (Y, Z) to BSDE (1.1) in the space $\mathcal{H}^2_m(0,T) \times \mathcal{H}^2_{m\times d}(0,T)$. In the case of a uniformly Lipschitz continuous generator, it was shown by Pardoux and Peng in 1990 that a solution in the above-mentioned space exists and is unique (see also Theorem 2.1 in El Karoui et al. [30]).

Theorem 1.1 (Pardoux and Peng [72]). *If (g, ξ_T) are standard parameters, then there exists a unique solution (Y, Z) to BSDE (1.1) in the space $\mathcal{H}^2_m(0,T) \times \mathcal{H}^2_{m\times d}(0,T)$ and Y belongs to $\mathcal{S}^{2,c}_m(0,T)$.*

Proof. **Step 1.** We define the mapping Φ

$$\mathcal{H}^{2,\alpha}_m(0,T) \times \mathcal{H}^{2,\alpha}_{m\times d}(0,T) \ni (u,v) \overset{\Phi}{\mapsto} (Y^{u,v}, Z^{u,v}) \in \mathcal{H}^{2,\alpha}_m(0,T) \times \mathcal{H}^{2,\alpha}_{m\times d}(0,T)$$

where $\Phi(u,v) = (Y^{u,v}, Z^{u,v})$ is defined as a solution (Y, Z) of the following BSDE

$$Y_t = \xi_T + \int_t^T g(s, u_s, v_s)\, ds - \int_t^T Z_s\, dW_s. \tag{1.10}$$

We deduce from Lemma 1.2 that this mapping is well defined, that is, for any $(u,v) \in \mathcal{H}^{2,\alpha}_{m\times d}(0,T)$ there exists a unique solution (Y, Z) to BSDE (1.10) in the space $\mathcal{H}^{2,\alpha}_{m\times d}(0,T)$. Our goal is to show that for sufficiently large $\alpha > 0$ the mapping Φ is a contraction. We will then conclude the proof by an application of the Banach fixed point theorem.

Step 2. Let us denote $(Y,Z) := (Y^{u,v}, Z^{u,v})$, $(Y', Z') := (Y^{u',v'}, Z^{u',v'})$, $g_t := g(t, u_t, v_t)$ and $g'_t := g(t, u'_t, v'_t)$ so that $dY_t = -g_t\, dt + Z_t\, dW_t$ and $dY'_t = -g'_t\, dt + Z'_t\, dW_t$. By applying the Itô formula (see Propositions 12.1 and 12.2) to the function $f(x,t) = e^{\alpha t}\, \|x\|^2$ and the process $X_t := Y_t - Y'_t$, we obtain

$$\begin{aligned} &e^{\alpha t}\|Y_t - Y'_t\|^2 + \int_t^T e^{\alpha s}\|Z_s - Z'_s\|^2\, ds \\ &= M_T - M_t - \alpha \int_t^T e^{\alpha s}\|Y_s - Y'_s\|^2\, ds + 2\int_t^T e^{\alpha s}(Y_s - Y'_s)(g_s - g'_s)\, ds \end{aligned} \tag{1.11}$$

where the real-valued continuous local martingale M is given by

$$M_t = 2\int_0^t e^{\alpha s}(Y_s - Y'_s)(Z_s - Z'_s)\, dW_s.$$

Our goal is to show that $\mathbb{E}(M_T - M_t) = 0$ for every $t \in [0,T]$. We note that the quadratic variation of M equals

$$\langle M, M\rangle_t = 4\int_0^t e^{2\alpha s}\|(Y_s - Y'_s)(Z_s - Z'_s)\|^2\, ds.$$

Consequently,

$$\begin{aligned} \mathbb{E}\big[\langle M, M\rangle_T^{1/2}\big] &= 4\,\mathbb{E}\bigg[\bigg(\int_0^T e^{2\alpha t}\|(Y_t - Y'_t)(Z_t - Z'_t)\|^2\, dt\bigg)^{1/2}\bigg]. \\ &\le 4\,\mathbb{E}\bigg[\bigg(\int_0^T e^{2\alpha t}\|Y_t - Y'_t\|^2\|Z_t - Z'_t\|^2\, dt\bigg)^{1/2}\bigg]. \end{aligned}$$

Using the inequality $2ab \le a^2 + b^2$ for any real numbers a, b, we obtain

$$\begin{aligned}
&\mathbb{E}\Big[\Big(\int_0^T e^{2\alpha t}\|Y_t - Y_t'\|^2\|Z_t - Z_t'\|^2\,dt\Big)^{1/2}\Big] \\
&\le \mathbb{E}\Big(\sup_{t\in[0,T]}\|Y_t - Y_t'\|\,\Big(\int_0^T e^{2\alpha t}\|Z_t - Z_t'\|^2\,dt\Big)^{1/2}\Big) \\
&\le \frac{1}{2}\,\mathbb{E}\Big(\sup_{t\in[0,T]}\|Y_t - Y_t'\|^2\Big) + \frac{c}{2}\,\mathbb{E}\Big(\int_0^T \|Z_t - Z_t'\|^2\,dt\Big) < \infty
\end{aligned}$$

where $c = e^{2\alpha T}$, since, by assumption, the random variable $\sup_{t\in[0,T]}\|Y_t - Y_t'\|$ is square-integrable and $Z - Z'$ belongs to the space $\mathcal{H}^2_{m\times d}(0,T)$. Hence the process γ given by $\gamma_t := e^{\alpha t}(Y_t - Y_t')(Z_t - Z_t')$ belongs to $\mathcal{H}^1_d(0,T)$, which means that $\mathbb{E}\big[\langle M, M\rangle_T^{1/2}\big] < \infty$.

Using the Burkholder-Davis-Gundy inequality with $p = 1$ (see Theorem 12.2)

$$\mathbb{E}(M_T^*) \le C_1\,\mathbb{E}\big[\langle M, M\rangle_T^{1/2}\big],$$

we conclude that the continuous local martingale M satisfies $\mathbb{E}(M_T^*) < \infty$ and thus, in view of Lemma 12.3, M is a uniformly integrable martingale. It is thus clear that $\mathbb{E}(M_T - M_t) = \mathbb{E}(M_T) - \mathbb{E}(M_t) = 0$ for every $t \in [0,T]$.

Step 3. Our goal in this step is to show there exists a constant K, which may depend on T and L, such that for every $\alpha > 0$

$$\|(Y,Z) - (Y',Z')\|^2_{2,\alpha} := \|Y - Y'\|^2_{2,\alpha} + \|Z - Z'\|^2_{2,\alpha} \le \frac{K}{\alpha}\|(u - u', v - v')\|^2_{2,\alpha}. \tag{1.12}$$

The following elementary inequality holds for $\alpha > 0$ and arbitrary real numbers L, a, b

$$-\alpha a^2 + 2Lab = -\alpha\Big(a - \frac{L}{\alpha}\,b\Big)^2 + \frac{L^2}{\alpha}\,b^2 \le \frac{L^2}{\alpha}\,b^2. \tag{1.13}$$

Using (1.11), the martingale property of M, the Cauchy-Schwarz inequality for the inner product in $\mathbb{R}^m$, the Lipschitz continuity of g, and (1.13), we get for any $t \in [0,T]$

$$\begin{aligned}
&\mathbb{E}\big(e^{\alpha t}\|Y_t - Y_t'\|^2\big) + \mathbb{E}\Big(\int_t^T e^{\alpha s}\|Z_s - Z_s'\|^2\,ds\Big) \\
&\overset{(1.11)}{\le} \mathbb{E}\Big(-\alpha\int_t^T e^{\alpha s}\|Y_s - Y_s'\|^2\,ds + 2\int_t^T e^{\alpha s}|(Y_s - Y_s')(g_s - g_s')|\,ds\Big) \\
&\overset{(1.2)}{\le} \mathbb{E}\Big(-\alpha\int_t^T e^{\alpha s}\|Y_s - Y_s'\|^2\,ds + 2L\int_t^T e^{\alpha s}\|Y_s - Y_s'\|\big(\|u_s - u_s'\| + \|v_s - v_s'\|\big)\,ds\Big) \\
&\overset{(1.13)}{\le} \frac{L^2}{\alpha}\,\mathbb{E}\Big(\int_t^T e^{\alpha s}\big(\|u_s - u_s'\| + \|v_s - v_s'\|\big)^2\,ds\Big) \\
&\le \frac{2L^2}{\alpha}\,\mathbb{E}\Big(\int_t^T e^{\alpha s}\big(\|u_s - u_s'\|^2 + \|v_s - v_s'\|^2\big)\,ds\Big) \le \frac{2L^2}{\alpha}\,\|(u,v) - (u',v')\|^2_{2,\alpha}
\end{aligned}$$

where we also used $(a + b)^2 \le 2a^2 + 2b^2$. By taking $t = 0$, we obtain

$$\|Z - Z'\|^2_{2,\alpha} \le \frac{2L^2}{\alpha}\|(u,v) - (u',v')\|^2_{2,\alpha}. \tag{1.14}$$

Moreover, for every $t \in [0,T]$

$$\mathbb{E}\big(e^{\alpha t}\|Y_t - Y'_t\|^2\big) \le \frac{2L^2}{\alpha}\|(u,v)-(u',v')\|^2_{2,\alpha}$$

and thus

$$\|Y - Y'\|^2_{2,\alpha} = \mathbb{E}\bigg(\int_0^T e^{\alpha t}\|Y_t - Y'_t\|^2\,dt\bigg) \le \frac{2L^2T}{\alpha}\|(u,v)-(u',v')\|^2_{2,\alpha}. \tag{1.15}$$

By combining (1.14) with (1.15), we obtain

$$\|(Y,Z)-(Y',Z')\|^2_{2,\alpha} = \|Y-Y'\|^2_{2,\alpha} + \|Z-Z'\|^2_{2,\alpha} \le \frac{2L^2(T+1)}{\alpha}\|(u,v)-(u',v')\|^2_{2,\alpha}$$

so that inequality (1.12) holds with $K = 2L^2(T+1)$.

Step 4. In Step 1, we introduced the mapping Φ on the Banach space $E = \mathcal{H}^{2,\alpha}_m(0,T) \times \mathcal{H}^{2,\alpha}_{m\times d}(0,T)$ endowed with the norm $\|\cdot\|_E = \|(\cdot,\cdot)\|_{2,\alpha}$ defined in (1.12). In Step 3, we have shown that

$$\|\Phi(u,v) - \Phi(u',v')\|^2_{2,\alpha} \le \frac{K}{\alpha}\|(u,v)-(u',v')\|^2_{2,\alpha}.$$

It is thus clear that the mapping Φ is a contraction on $(E, \|\cdot\|_E)$ as soon as α satisfies $\alpha > K$. From the Banach fixed point theorem, Φ has a unique fixed point $(\widehat{Y},\widehat{Z})$. The equality $\Phi(\widehat{Y},\widehat{Z}) = (\widehat{Y},\widehat{Z})$ means that $(\widehat{Y},\widehat{Z})$ is a unique solution to the BSDE (1.1) in the space $\mathcal{H}^{2,\alpha}_m(0,T) \times \mathcal{H}^{2,\alpha}_{m\times d}(0,T)$.

Step 5. Since the norms $\|(\cdot,\cdot)\|_{2,\alpha}$ for $\alpha \ge 0$ are equivalent, the processes $\widehat{Y}$ and $\widehat{Z}$ belong to the spaces $\mathcal{H}^2_m(0,T)$ and $\mathcal{H}^2_{m\times d}(0,T)$, respectively. Moreover, by construction, the process $\widehat{Y}$ is continuous. To prove that the process $\widehat{Y}$ belong to the space $\mathcal{S}^{2,c}_m(0,T)$, we need to show that

$$\mathbb{E}\left(\sup_{0\le t\le T}\|\widehat{Y}_t\|^2\right) < \infty.$$

Let us consider any $(Y,Z) \in \mathcal{H}^{2,\alpha}_m(0,T) \times \mathcal{H}^{2,\alpha}_{m\times d}(0,T)$. From (1.1), it follows that

$$\|Y_t\| \le \|\xi_T\| + \int_0^T \|g(s,Y_s,Z_s)\|\,ds + \sup_{0\le t\le T}\left\|\int_t^T Z_s\,dW_s\right\|.$$

Since (g,ξ_T) are standard parameters, the real-valued random variable

$$\|\xi_T\| + \int_0^T \|g(s,Y_s,Z_s)\|\,ds$$

is square-integrable. Furthermore, the Burkholder-Davis-Gundy inequality (see Theorem 12.2) yields

$$\begin{aligned}\mathbb{E}\bigg(\sup_{0\le t\le T}\bigg\|\int_t^T Z_s\,dW_s\bigg\|^2\bigg) &\le 2\,\mathbb{E}\bigg(\bigg\|\int_0^T Z_s\,dW_s\bigg\|^2\bigg) + \mathbb{E}\bigg(\sup_{0\le t\le T}\bigg\|\int_0^t Z_s\,dW_s\bigg\|^2\bigg)\\ &\le 4\,\mathbb{E}\bigg(\int_0^T \|Z_s\|^2\,ds\bigg).\end{aligned}$$

We thus conclude that Y belongs to $\mathcal{S}^{2,c}_m(0,T)$ and thus the proof of Theorem 1.1 is completed. $\square$

1.5 Picard's Iteration Scheme

The Banach fixed point theorem leads to the Picard approximation scheme where one constructs a sequence $(Y^n, Z^n)_{n=0}^{\infty}$ of processes converging to the unique solution (Y, Z) to BSDE (1.1).

Proposition 1.1. *Define the Picard sequence $(Y^n, Z^n)_{n=0}^{\infty}$ by $(Y^0, Z^0) := (0, 0)$ and for every $n = 0, 1, \ldots$*

$$\begin{cases} -dY_t^{n+1} = g(t, Y_t^n, Z_t^n)\, dt - Z_t^{n+1}\, dW_t, & t \in [0, T], \\ Y_T^{n+1} = \xi_T. \end{cases}$$

Then the sequence $(Y^n, Z^n)_{n=0}^{\infty}$ converges in $\mathcal{H}_m^2(0,T) \times \mathcal{H}_{m\times d}^2(0,T)$ and $d\mathbb{P}\otimes dt-$a.e. to the unique solution (Y, Z) of BSDE (1.1). Moreover, the sequence $(Y^n)_{n=0}^{\infty}$ of stochastic processes converges uniformly to Y, $\mathbb{P}-$a.s.

1.6 A Priori Estimates

The following lemma furnishes useful *a priori* estimates for solutions to BSDEs with standard parameters. We take this result for granted (for the proof, see Proposition 2.1 in El Karoui et al. [30]).

Lemma 1.3. *Let (Y^i, Z^i), $i = 1, 2$ be any solutions in $\mathcal{H}_m^2(0,T) \times \mathcal{H}_{m\times d}^2(0,T)$ to BSDEs with the standard parameters (g^i, ξ_T^i), $i = 1, 2$. Let $L > 0$ be the Lipschitz constant for g^1 and let us denote $\delta_2 g_t := g^1(t, Y_t^2, Z_t^2) - g^2(t, Y_t^2, Z_t^2)$. Then*

$$\mathbb{E}\Big(\sup_{t\in[0,T]} \|Y_t^i\|^2 + \int_0^T \|Z_t^i\|^2\, dt\Big) \leq L\mathbb{E}\left(\|\xi_T^i\|^2 + \int_0^T \|g^i(t,0,0)\|^2\, dt\right). \tag{1.16}$$

For any (λ, μ, α) such that $\lambda^2 > L$, $\mu > 0$ and $\alpha \geq L(2 + \lambda^2) + \mu^2$

$$\|Y^1 - Y^2\|_{2,\alpha}^2 \leq T\left(e^{\alpha T}\, \mathbb{E}\big(\|\xi_T^1 - \xi_T^2\|^2\big) + \frac{1}{\mu^2}\, \|\delta_2 g\|_\alpha^2\right) \tag{1.17}$$

and

$$\|Z^1 - Z^2\|_{2,\alpha}^2 \leq \frac{\lambda^2}{\lambda^2 - L}\left(e^{\alpha T}\, \mathbb{E}\big(\|\xi_T^1 - \xi_T^2\|^2\big) + \frac{1}{\mu^2}\, \|\delta_2 g\|_\alpha^2\right). \tag{1.18}$$

Then

$$\mathbb{E}\Big(\sup_{t\in[0,T]} \|Y_t^1 - Y_t^2\|^2\Big) \leq C\, \mathbb{E}\left(\|\xi_T^1 - \xi_T^2\|^2 + \int_0^T \|\delta_2 g_t\|^2\, dt\right) \tag{1.19}$$

where C is a positive constant, which only depends on T.

1.7 Forward Dynamics for a Solution to BSDEs

Assume that we have already established the existence and uniqueness of a solution to the BSDE

$$Y_t = \xi_T + \int_t^T g(s, Y_s, Z_s)\, ds - \int_t^T Z_s\, dW_s \tag{1.20}$$

in a suitable space of stochastic processes. Then, by applying (1.20) with $t = 0$, we obtain

$$Y_0 = \xi_T + \int_0^T g(s, Y_s, Z_s)\, ds - \int_0^T Z_s\, dW_s. \tag{1.21}$$

By subtracting equality (1.20) from (1.21), we get

$$Y_0 - Y_t = \int_0^t g(s, Y_s, Z_s)\, ds - \int_0^t Z_s\, dW_s, \tag{1.22}$$

which means that the forward dynamics of a state process Y are given by

$$Y_t = Y_0 - \int_0^t g(s, Y_s, Z_s)\, ds + \int_0^t Z_s\, dW_s \tag{1.23}$$

where we may assume that the process Z is given since we already know that a solution (Y, Z) to the BSDE (1.20) is unique. For any fixed process Z, equation (1.23) can also be seen as a (forward) SDE where Y is an unknown process and, if needed, we may apply the theory of SDEs to study some properties of our interest for the process Y.

Formally, we may denote $\sigma_t = Z_t$ and $\mu(t, Y_t) = g(\omega, t, Y_t, Z_t)$ so that (1.23) becomes

$$Y_t = Y_0 - \int_0^t \mu(s, Y_s)\, ds + \int_0^t \sigma_s\, dW_s,$$

which is a (forward) SDE but, of course, neither σ nor Y_0 are known *a priori*, that is, before the BSDE (1.20) is solved (usually only numerically).

1.8 Nonlinear Evaluation Associated with a BSDE

Notice that Theorem 1.1 can be easily extended to the case of a BSDE on $[0, \tau]$, rather than on $[0, T]$, where $\tau \in \mathcal{T}_{[0,T]}$ is arbitrary and the terminal condition ζ_τ is an $\mathcal{F}_\tau$-measurable and square-integrable random variable

$$\begin{cases} -dY'_t = g(t, Y'_t, Z'_t)\, dt - Z'_t\, dW_t, \quad t \in [0, \tau], \\ Y'_\tau = \zeta_\tau. \end{cases}$$

This means that we deal with the BSDE on $[0, \tau]$ given by

$$Y'_t = \zeta_\tau + \int_t^\tau g(s, Y'_s, Z'_s)\, ds - \int_t^\tau Z'_s\, dW_s \tag{1.24}$$

where a stopping time τ and a random variable $\zeta_\tau \in L^2(\mathcal{F}_\tau)$ are arbitrary. Notice that the definitions of spaces $\mathcal{H}^2_m(0, T), \mathcal{H}^2_{m\times d}(0, T)$ and $\mathcal{S}^{2;c}_m(0, T)$ can be extended to the stochastic interval $[0, \tau]$ for any fixed $\tau \in \mathcal{T}_{[0,T]}$.

Proposition 1.2. *Under the assumptions of Theorem 1.1, the BSDE 1.24 has a unique solution on $[0, \tau]$ in the space $\mathcal{H}^2_m(0, \tau) \times \mathcal{H}^2_{m\times d}(0, \tau)$ and Y belongs to $\mathcal{S}^{2;c}_m(0, \tau)$.*

Proof. To show that the BSDE (1.24) has a unique solution under the assumptions of Theorem 1.1, we argue as follows. We introduce the modified generator $\widetilde{g}$, which is given by $\widetilde{g}(t,x,y) := g(t,x,y)\mathbb{1}_{\{t\leq\tau\}}$. Then the generator $\widetilde{g}$ is uniformly Lipschitz continuous and it is easy to see that a unique solution (Y,Z) to the BSDE on $[0,T]$ with the generator $\widetilde{g}$ is stopped at τ, that is, $(Y,Z)=(Y^\tau,Z^\tau)$ where, for any process X and any stopping time τ, the *stopped process* X^τ is defined as $X^\tau_t := X_{t\wedge\tau}$ for every $t\in[0,T]$. Consequently, a unique solution (Y,Z) to the BSDE on $[0,T]$ with the terminal condition $\xi_T := \zeta_\tau$ is also a unique solution to the BSDE on $[0,\tau]$ with the terminal condition ζ_τ, that is, the equality $(Y,Z)=(Y',Z')$ holds on the stochastic interval $[0,\tau]$. □

It follows from the above considerations that one may introduce the notion of a *time-consistent* family of solutions to the BSDE (1.20), where *time-consistency* of solutions means that if (Y',Z') is a unique solution with the terminal value ζ_τ at time $\tau\in\mathcal{T}_{[0,T]}$ and $\sigma\in\mathcal{T}_{[0,T]}$ is any stopping time such that $\sigma\leq\tau$, then on the random interval $[0,\sigma]$ we have the equality $(Y',Z')=(Y'',Z'')$ where (Y'',Z'') is a unique solution with the terminal value Y'_σ at time σ. Of course, here $\sigma\in\mathcal{T}_{[0,T]}$ and $\tau\in\mathcal{T}_{[0,T]}$ are arbitrary stopping times such that $\sigma\leq\tau$ and ζ_τ is an arbitrary $\mathcal{F}_\tau$-measurable and square-integrable random variable.

Definition 1.2. Let $\tau\in\mathcal{T}_{[0,T]}$ and ζ_τ be an $\mathcal{F}_\tau$-measurable and square-integrable random variable. The *g-evaluation* of ζ_τ is given by the process Y defined on $[0,\tau]$ by

$$\mathcal{E}^g_{\cdot,\tau}(\zeta_\tau) := Y.$$

where (Y,Z) is a unique solution of the BSDE with the generator g and terminal condition ζ_τ at an $\mathbb{F}$-stopping time τ.

Then the time-consistency of solutions to the BSDE implies that $Y_\sigma = \mathcal{E}^g_{\sigma,\rho}(Y_\rho)$ when $\sigma\leq\rho$ and thus

$$Y_\sigma = \mathcal{E}^g_{\sigma,\tau}(\zeta_\tau) = \mathcal{E}^g_{\sigma,\rho}(\mathcal{E}^g_{\rho,\tau}(\zeta_\tau)) = \mathcal{E}^g_{\sigma,\rho}(Y_\rho)$$

for every $\sigma,\rho,\tau\in\mathcal{T}_{[0,T]}$ such that $\sigma\leq\rho\leq\tau$ and any $\zeta_\tau\in L^2(\mathcal{F}_\tau)$.

Observe also that Definition 1.2 can be used to define the mapping

$$\mathcal{E}^g = \{\mathcal{E}^g_{\sigma,\tau} : L^2(\mathcal{F}_\tau)\to L^2(\mathcal{F}_\sigma)\,|\,\sigma,\tau\in\mathcal{T}_{[0,T]},\ \sigma\leq\tau\}$$

by setting $\mathcal{E}^g_{\sigma,\tau}(\zeta_\tau)=Y_\sigma$ for every $\sigma\leq\tau$ and any $\zeta_\tau\in L^2(\mathcal{F}_\tau)$. The mapping $\mathcal{E}^g$ defined in that way is called the *g-evaluation*.

The g-evaluation can be seen as a special case of a *nonlinear evaluation*, which is defined through an axiomatic approach (see, e.g., Peng [80]). It is then typically postulated that a BSDE is such that the comparison property for BSDEs holds, so that the g-evaluation is monotonic (see condition (A.3)).

1.9 Abstract Nonlinear Evaluation

It order to develop a more comprehensive nonlinear probability theory, it is convenient to focus on general nonlinear mappings satisfying certain fundamental properties of conditional expectation with respect to a reference filtration $\mathbb{F}$. The relevant concepts of a *nonlinear expectation* and a *nonlinear evaluation* were introduced through an axiomatic approach and studied by several authors in early 2000s. In particular, Peng [80] (see Definition 2.1 therein) introduced the concept of an $\mathbb{F}$-*consistent nonlinear evaluation* on $[0,T]$ as a system of mappings

$$\mathcal{E} = \{\mathcal{E}_{\sigma,\tau} : L^2(\mathcal{F}_\tau) \to L^2(\mathcal{F}_\sigma) \,|\, \sigma, \tau \in \mathcal{T}_{[0,T]},\ \sigma \le \tau\}$$

satisfying the following conditions, for every $\sigma, \tau, \rho \in \mathcal{T}_{[0,T]}$ and $\zeta_\tau, \eta_\tau \in L^2(\mathcal{F}_\tau)$:
(A.1) $\mathcal{E}_{\tau,\tau}(\zeta_\tau) = \zeta_\tau$ (identity);
(A.2) $\mathcal{E}_{\sigma,\tau}(\zeta_\tau) = \mathcal{E}_{\sigma,\rho}(\mathcal{E}_{\rho,\tau}(\zeta_\tau))$ for every $\sigma \le \rho \le \tau$ (time-consistency);
(A.3) if $\zeta_\tau \ge \eta_\tau$, then $\mathcal{E}_{\sigma,\tau}(\zeta_\tau) \ge \mathcal{E}_{\sigma,\tau}(\eta_\tau)$ (monotonicity);
(A.4) $\mathbb{1}_E\, \mathcal{E}_{\sigma,\tau}(\zeta_\tau) = \mathbb{1}_E\, \mathcal{E}_{\sigma,\tau}(\mathbb{1}_E \zeta_\tau)$ for every $E \in \mathcal{F}_\sigma$ (local property).

Furthermore, one can also postulate that $\mathcal{E}$ has the stability property, which reads:
(A.5) there exists a constant $C > 0$ such that for every $\varepsilon > 0$, every $\sigma, \tau \in \mathcal{T}_{[0,T]}$ such that $\sigma \le \tau$, and every $\xi_\tau, \eta_\tau \in L^2(\mathcal{F}_\tau)$

$$|\xi_\tau - \eta_\tau| \le \varepsilon \quad \Rightarrow \quad |\mathcal{E}_{\sigma,\tau}(\xi_\tau) - \mathcal{E}_{\sigma,\tau}(\eta_\tau)| \le C\varepsilon.$$

Note that the properties (A.1)–(A.5) of the nonlinear evaluation $\mathcal{E}^g$ associated with solutions to the BSDE (1.1) can be proven when the assumptions of Theorem 1.1 are satisfied. In fact, properties (A.1)–(A.4) of $\mathcal{E}^g$ are easy to check and property (A.5) is an immediate consequence of inequality (1.19) in Lemma 1.3.

The following definition extends the concepts of a martingale and a supermartingale to the case of a nonlinear evaluation $\mathcal{E}$.

Definition 1.3. (i) An $\mathbb{F}$-optional process Y is called an *$\mathcal{E}$-martingale* (resp., an *$\mathcal{E}$-supermartingale*) if the equality $Y_t = \mathcal{E}_{t,s}(Y_s)$ holds (resp., $Y_t \ge \mathcal{E}_{t,s}(Y_s)$) for all $t, s \in [0,T]$ such that $t \le s$.
(ii) An $\mathbb{F}$-optional process Y is called a *strong $\mathcal{E}$-martingale* (resp., a *strong $\mathcal{E}$-supermartingale*) if $Y_\sigma = \mathcal{E}_{\sigma,\tau}(Y_\tau)$ (resp., $Y_\sigma \ge \mathcal{E}_{\sigma,\tau}(Y_\tau)$) for every $\sigma, \tau \in \mathcal{T}_{[0,T]}$ such that $\sigma \le \tau$.

Notice that a strong $\mathcal{E}$-martingale (resp., a strong $\mathcal{E}$-supermartingale) is also an $\mathcal{E}$-martingale (resp., an $\mathcal{E}$-supermartingale). It is clear that Definition 1.3 can be applied to g-evaluation $\mathcal{E}^g$ and then an $\mathcal{E}^g$-martingale (resp., an $\mathcal{E}^g$-supermartingale) is also called a *g-martingale* (resp., a *g-supermartingale*).

Under the postulate of *$\mathcal{E}^{g_\mu}$-domination*, Peng [81] showed (see Theorem 10.19 in [81]) that the Optional Stopping Theorem holds and thus a càdlàg $\mathcal{E}$-martingale (resp., a càdlàg $\mathcal{E}$-supermartingale) is also a strong $\mathcal{E}$-martingale (resp., a strong $\mathcal{E}$-supermartingale). These properties are well known to hold for classical strong martingales and strong supermartingales.

It should be recalled that, by definition, the classical martingales and supermartingales are càdlàg since one may show that any martingale or supermartingale has a càdlàg modification. This is due to the fact that the expected value of a martingale or a supermartingale is a càdlàg function and thus the Doob theorem on a càdlàg modification of a stochastic process can be applied.

2 Linear BSDEs

Let W be a d-dimensional Wiener process on $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P})$ where $\mathbb{F} = \mathbb{F}^W$ is the filtration generated by W. We will now study BSDEs with a linear generator.

Definition 2.1. The *linear BSDE* with $m = 1$ has the form

$$Y_t = \xi_T + \int_t^T (\varphi_s + \mu_s Y_s + \sigma_s Z_s)\, ds - \int_t^T Z_s\, dW_s \tag{2.1}$$

or, equivalently,

$$-dY_t = (\varphi_t + \mu_t Y_t + \sigma_t Z_t)\, dt - Z_t\, dW_t, \quad Y_T = \xi_T. \tag{2.2}$$

Assumption 2.1. We assume that:
(i) the real-valued process μ is $\mathbb{F}$-progressively measurable and bounded;
(ii) the $\mathbb{R}^d$-valued process σ is $\mathbb{F}$-progressively measurable and bounded;
(iii) the real-valued, $\mathbb{F}$-progressively measurable process φ satisfies

$$\mathbb{P}\Big(\int_0^T |\varphi_s|\, ds < \infty \Big) = 1. \tag{2.3}$$

To analyse (2.1), we first consider the following (adjoint) linear SDE on $[t, T]$, for any fixed $t \in [0, T]$,

$$d\Gamma_{t,s} = \Gamma_{t,s}(\mu_s\, ds + \sigma_s\, dW_s), \quad \Gamma_{t,t} = 1. \tag{2.4}$$

Lemma 2.1. *If Assumption 2.1 is valid, then:*
(a) for any fixed $t \in [0, T]$, the SDE (2.4) has a unique solution $(\Gamma_{t,s})_{s\in[t,T]}$;
(b) the stochastic flow property holds: $\Gamma_{t,s}\Gamma_{s,u} = \Gamma_{t,u}$ for all $0 \le t \le s \le u \le T$;
(c) the process $(\Gamma_{0,t})_{t\in[0,T]}$ satisfies $\sup_{t\in[0,T]} |\Gamma_{0,t}| \in L^2_1(\mathcal{F}_T)$.

Proof. The existence of a unique solution is a consequence of the classical Itô theorem for SDEs with Lipschitz continuous coefficients (see Theorem 12.3). Furthermore, by applying the Itô lemma to the function $f(x) = e^x$, one can check that the unique solution to the SDE (2.4) is given by

$$\Gamma_{t,s} = \exp\Big(\int_t^s \mu_u\, du + \int_t^s \sigma_u\, dW_u - \frac{1}{2} \int_t^s \|\sigma_u\|^2\, du \Big). \tag{2.5}$$

Part (b) is immediate from (2.3). Part (c) is a consequence of (2.3) and standard inequalities for semimartingales. □

Lemma 2.2. *Let Assumption 2.1 hold and (Y, Z) be any solution to the linear BSDE (2.1). Then*

$$\Gamma_{0,t} Y_t = Y_0 - \int_0^t \Gamma_{0,s}\varphi_s\, ds + \int_0^t \Gamma_{0,s}\big(Z_s + Y_s\sigma_s\big)\, dW_s \tag{2.6}$$

and thus the process $M = (M_t)_{t\in[0,T]}$ given by

$$M_t := \Gamma_{0,t} Y_t + \int_0^t \Gamma_{0,s}\varphi_s\, ds \tag{2.7}$$

is a local martingale.

Proof. It suffices to apply the Itô integration by parts formula to (2.2) and (2.4). Notice that the integrals in (2.6) and (2.7) are well defined since the processes Y and $\Gamma_{0,\cdot}$ are continuous. □

2.1 Square-Integrable Solutions to a Linear BSDE

We first examine the existence, uniqueness and closed-form expression for a solution to the linear BSDE in the space $\mathcal{H}^2_1(0,T) \times \mathcal{H}^2_d(0,T)$. For brevity, they are sometimes referred to as *square-integrable solutions*. We also establish some elementary versions of the *comparison property* (in part (b)) and the *strict comparison property* (in part (c)) for solutions to linear BSDEs. For an important extension to solutions of nonlinear BSDEs, see Theorem 3.1.

Theorem 2.1. *Let Assumption 2.1 be satisfied and:*
(i) the random variable ξ_T belong to $L^2_1(\mathcal{F}_T)$;
(ii) the process φ belong to $\mathcal{H}^2_1(0,T)$.
Then the following assertions are valid.
(a) The linear BSDE (2.1) *has a unique solution (Y,Z) in the space $\mathcal{H}^2_1(0,T) \times \mathcal{H}^2_d(0,T)$. Moreover, the process Y belongs to $\mathcal{S}^{2,c}_1(0,T)$ and is given by the following expression*

$$Y_t = \mathbb{E}\Big(\Gamma_{t,T}\xi_T + \int_t^T \Gamma_{t,s}\varphi_s\,ds \,\Big|\, \mathcal{F}_t\Big) \tag{2.8}$$

where $\Gamma_{t,s}$ is given by (2.3)*.*
(b) If ξ_T and φ are non-negative, then Y is non-negative.
(c) Assume that ξ_T and φ are non-negative. For a fixed $t \in [0,T]$, let B belong to the σ-field $\mathcal{F}_t$. Assume that $Y_t = 0$ on B or, equivalently, that $\mathbb{1}_B Y_t = 0$ where $\mathbb{1}_B$ is the indicator function of B. Then the following equalities are valid: $\mathbb{1}_B\xi_T = 0$, $\mathbb{P}$–a.s. and $\varphi_s = Y_s = Z_s = 0$, $d\mathbb{P} \otimes ds$–a.e. on $B \times [t,T]$.

Proof. To establish part (a), we note that the assumptions of Theorem 1.1 are satisfied and thus the linear BSDE (2.1) has a unique solution $(Y,Z) \in \mathcal{H}^2_1(0,T) \times \mathcal{H}^2_d(0,T)$ and Y belongs to $\mathcal{S}^{2,c}_1(0,T)$. If (Y,Z) is any solution to the linear BSDE (2.1), then, by Lemma 2.2, the process M given by (2.7) is a local martingale. We claim that, under the present assumptions, M is a martingale. Indeed, it is not hard to show that the random variables

$$Y^*_T := \sup_{t\in[0,T]} |Y_t|, \quad \Gamma^*_T := \sup_{t\in[0,T]} |\Gamma_{0,t}|$$

are square-integrable. Hence $M^*_T := \sup_{t\in[0,T]} |M_t|$ is integrable and thus, by Lemma 12.3, M is a uniformly integrable martingale. The martingale property of M yields

$$\mathbb{E}(M_T \,|\, \mathcal{F}_t) = \mathbb{E}\Big(\Gamma_{0,T}\xi_T + \int_0^T \Gamma_{0,s}\varphi_s\,ds \,\Big|\, \mathcal{F}_t\Big) = M_t = \Gamma_{0,t}Y_t + \int_0^t \Gamma_{0,s}\varphi_s\,ds$$

which proves (2.8) since $\Gamma_{t,s} = \Gamma^{-1}_{0,t}\Gamma_{0,s}$ for all $t \le s \le T$. Part (b) is an immediate consequence of equation (2.8).

To prove (c), we multiply both sides in (2.8) by the indicator function $\mathbb{1}_B$ to obtain

$$\mathbb{E}\Big(\Gamma_{t,T}\mathbb{1}_B\xi_T + \int_t^T \Gamma_{t,s}\mathbb{1}_B\varphi_s\,ds \,\Big|\, \mathcal{F}_t\Big) = \mathbb{1}_B Y_t = 0. \tag{2.9}$$

Since $\Gamma_{t,s} > 0$, ξ_T and φ are non-negative, we deduce that $\mathbb{1}_B\xi_T$ and $\int_t^T \mathbb{1}_B\varphi_s\,ds = 0$. Hence $Y_s = \varphi_s = 0$, $d\mathbb{P} \otimes ds$–a.e. on $B \times [t,T]$. This in turn implies that $Z_s = 0$, $d\mathbb{P} \otimes ds$–a.e. on $B \times [t,T]$. $\square$

2.2 Non-Uniqueness of a Solution to a BSDE with Null Generator

One may ask whether it is necessary to search for a solution the BSDE (1.1) in the space $\mathcal{H}^2_m(0,T) \times \mathcal{H}^2_{m\times d}(0,T)$ or some other Banach space of stochastic processes. Obviously, the answer to that question is negative but we will argue that a suitable integrability of solutions is needed to ensure the uniqueness of a solution to a BSDE even in the simplest case of the null generator. To this end, we will use a generalised version of the predictable representation property of the Brownian filtration. Let W be a one-dimensional Brownian motion. In was shown in Dudley [22] that for any time $T > 0$ there exists a real-valued, $\mathbb{F}^W$-progressively measurable process γ such that

$$\mathbb{P}\Big(\int_0^T |\gamma_t|^2 \, dt < \infty \Big) < +\infty \tag{2.10}$$

so that the stochastic integral

$$I_t := \int_0^t \gamma_s \, dW_s$$

is well defined and $I_T = 1$. The following more general result was established in [22].

Proposition 2.1 (Dudley [22]). *For any $\mathcal{F}^W_T$-measurable random variable ξ there exists an $\mathbb{F}^W$-progressively measurable process γ such that condition* (2.10) *is satisfied and $\xi = \int_0^T \gamma_t \, dW_t$.*

Observe that if ξ is integrable, but $\mathbb{E}(\xi) \neq 0$, then the process $M_t := \int_0^t \gamma_s \, dW_s$ is a continuous local martingale, but it cannot be a martingale (this is obvious, since $M_0 = 0$ and $\mathbb{E}(M_T) = \mathbb{E}(\xi) \neq 0$).

Example 2.1. We give here an example of non-uniqueness of a solution to a BSDE with null generator and the terminal value $\xi_T \in L^2(\mathcal{F}_T)$. From Lemma 1.1, we know that the BSDE

$$Y_t = \xi_T - \int_t^T Z_s \, dW_s \tag{2.11}$$

has a unique solution (Y, Z) in the space $\mathcal{H}^2_m(0,T) \times \mathcal{H}^2_{m\times d}(0,T)$ and, in fact, Y belongs to $\mathcal{S}^{2,c}_m(0,T)$. We will check that for any real number λ, the pair $(\widetilde{Y}, \widetilde{Z}) = (Y + \lambda(I-1), Z + \lambda\gamma)$ is also a solution of BSDE with parameters $(0, \xi_T)$.

It is clear that $\widetilde{Y}$ is an $\mathbb{F}$-adapted, continuous process satisfying $\widetilde{Y}_T = Y_T = \xi_T$. Moreover, for every $t \in [0,T]$

$$I_t = \int_0^t \gamma_s \, dW_s$$

and $I_T = 1$, so that

$$I_t - 1 = I_t - I_T = -\int_t^T \gamma_s \, dW_s.$$

Consequently, for every $t \in [0,T]$ we have

$$\begin{aligned} \widetilde{Y}_t = Y_t + \lambda(I_t - 1) &= \xi_T - \int_t^T Z_s \, dW_s - \lambda \int_t^T \gamma_s \, dW_s \\ &= \xi_T - \int_t^T (Z_s + \lambda\gamma_s) \, dW_s = \xi_T - \int_t^T \widetilde{Z}_s \, dW_s. \end{aligned}$$

Of course, the solution $(\widetilde{Y}, \widetilde{Z})$ to (2.11) does not belong to the space $\mathcal{H}^2_1(0,T) \times \mathcal{H}^2_1(0,T)$.

2.3 Non-uniqueness of Square-integrable Solutions to a Linear BSDE

It is easy to see that if (Y^1, Z^1) and (Y^2, Z^2) are any two solutions to the linear BSDE (2.1), then the pair $(\widehat{Y}, \widehat{Z}) := (Y^1 - Y^2, Z^1 - Z^2)$ is a solution to the following linear BSDE

$$-d\widehat{Y}_t = (\mu_t \widehat{Y}_t + \sigma_t \widehat{Z}_t)\, dt - \widehat{Z}_t\, dW_t, \quad \widehat{Y}_T = 0. \tag{2.12}$$

As was discussed in Section 2.2, this does not imply that $(Y^1, Z^1) = (Y^2, Z^2)$ since the linear BSDE (2.12) may have several solutions even when the processes μ and σ are assumed to be bounded. We have, however the following corollary to Theorem 1.1.

Corollary 2.1. *If Assumption 2.1 is valid, the process φ satisfies* (2.3) *and the random variable ξ_T is $\mathcal{F}_T$-measurable, then the uniqueness of solutions to the linear BSDE* (2.1) *holds in $\mathcal{H}^2_1(0,T) \times \mathcal{H}^2_d(0,T)$.*

Proof. From Theorem 1.1, if μ and σ are bounded processes, then $(\widehat{Y}, \widehat{Z}) = (0,0)$ is the unique solution to (2.12) in $\mathcal{H}^2_1(0,T) \times \mathcal{H}^2_d(0,T)$. Therefore, if (Y^1, Z^1) and (Y^2, Z^2) are two solutions to the linear BSDE (2.1) and they both belong to $\mathcal{H}^2_1(0,T) \times \mathcal{H}^2_d(0,T)$, then necessarily $(Y^1, Z^1) = (Y^2, Z^2)$. □

The following example of non-uniqueness of square-integrable solutions to a linear BSDE was given by El Karoui [26] (see Jeanblanc and Réveillac [43] for more results and examples in the same vein).

We consider the linear BSDE with $\xi_T = 0$

$$Y_t = \int_t^T \big[\big(\mu + f(s)\big) Y_s + \sigma Z_s\big]\, ds - \int_t^T Z_s\, dW_s \tag{2.13}$$

where $\mu, \sigma \in \mathbb{R}$ and the (unbounded, continuous) function f is given by $f(t) = \lambda\big(e^{-\lambda(T-t)} - 1\big)^{-1}$ for every $t \in [0,T)$. We will show that for every fixed $\beta > 0$ and $\lambda := (\beta+1)(\mu - \frac{1}{2}\sigma^2) + \frac{1}{2}\sigma^2(1+\beta)^2$ we have an infinite number of square-integrable solutions (Y^α, Z^α) to the linear BSDE (2.13) where $\alpha \in \mathbb{R}_+$ is a parameter. For any fixed $t \in [0,T]$ and every $s \in [t,T]$, we define

$$d\Gamma_{t,s} = \Gamma_{t,s}(\mu\, ds + \sigma\, dW_s), \quad \Gamma_{t,t} = 1,$$

which means that, for every $0 \le t \le s \le T$,

$$\Gamma_{t,s} = \exp\Big(\big(\mu - \tfrac{1}{2}\sigma^2\big)(s-t) + \sigma(W_s - W_t)\Big).$$

Step 1. We fix $\beta > 0$ and, for every $\alpha \in \mathbb{R}_+$, we define the pair (Y^α, Z^α) where the process Y^α equals, for every $t \in [0,T]$,

$$Y^\alpha_t := \alpha\, \mathbb{E}\Big(\int_t^T \Gamma_{t,s} (\Gamma_{0,s})^\beta\, ds \,\Big|\, \mathcal{F}_t \Big) = \mathbb{E}\Big(\int_t^T \Gamma_{t,s} \varphi^\alpha_s\, ds \,\Big|\, \mathcal{F}_t \Big)$$

where the continuous process $\varphi^\alpha_s := \alpha(\Gamma_{0,s})^\beta$ satisfies (2.3) and the process Z^α is obtained from the integral representation of the square-integrable martingale M^α, which is given by

$$\begin{aligned} M^\alpha_t &:= \Gamma_{0,t} Y^\alpha_t + \int_0^t \Gamma_{0,s} \varphi^\alpha_s\, ds = \Gamma_{0,t} Y^\alpha_t + \alpha \int_0^t \Gamma_{0,s} (\Gamma_{0,s})^\beta\, ds \\ &= \alpha\, \mathbb{E}\Big(\int_0^T (\Gamma_{0,s})^{\beta+1}\, ds \,\Big|\, \mathcal{F}_t \Big), \end{aligned}$$

that is,

$$M_t^\alpha = M_0^\alpha + \int_0^t Z_s^\alpha \, dW_s.$$

It is now easy to deduce from Theorem 1.1 that, for every $\alpha \in \mathbb{R}_+$, the pair (Y^α, Z^α) is a unique square-integrable solution to the linear BSDE

$$Y_t = \int_t^T \big(\varphi_s^\alpha + \mu Y_s + \sigma Z_s\big)\, ds - \int_t^T Z_s \, dW_s. \tag{2.14}$$

Step 2. We note that Y_t^α can be computed explicitly

$$\begin{aligned} Y_t^\alpha &= \alpha(\Gamma_{0,t})^\beta \, \mathbb{E}\Big(\int_t^T (\Gamma_{t,s})^{\beta+1} \, ds \,\big|\, \mathcal{F}_t \Big) = \alpha(\Gamma_{0,t})^\beta \int_t^T \mathbb{E}\big[(\Gamma_{t,s})^{\beta+1}\big]\, ds \\ &= \alpha(\Gamma_{0,t})^\beta \int_t^T e^{\left((\beta+1)\left(\mu - \frac{1}{2}\sigma^2\right) + \frac{1}{2}\sigma^2(\beta+1)^2\right)(s-t)}\, ds = \alpha(\Gamma_{0,t})^\beta \lambda^{-1}\big(e^{\lambda(T-t)} - 1\big) \\ &= \alpha(\Gamma_{0,t})^\beta (f(t))^{-1} = \varphi_t^\alpha (f(t))^{-1} \end{aligned}$$

where $\lambda = (\beta+1)(\mu - \frac{1}{2}\sigma^2) + \frac{1}{2}\sigma^2(\beta+1)^2$ (notice that λ and hence the function f do not depend on α). We thus see that there is a one-to-one correspondence between Y^α and $\varphi_t^\alpha(f(t))^{-1}$ and thus we deduce from (2.14) that (Y^α, Z^α) is a solution to the BSDE

$$Y_t = \int_t^T \big[\big(\mu + f(s)\big)Y_s + \sigma Z_s\big]\, ds - \int_t^T Z_s \, dW_s \tag{2.15}$$

We conclude that for every $\alpha \in \mathbb{R}_+$, the pair (Y^α, Z^α) where $Y^\alpha = \varphi_t^\alpha(f(t))^{-1}$ is a square-integrable solution to the linear BSDE (2.15) where the function $\mu + f(t)$, $t \in [0,T)$ is continuous but unbounded.

2.4 Comparison Properties of Solutions to Linear BSDEs

In this subsection, we relax the assumptions (iii) and (iv) in Theorem 2.1 and we consider general solutions to the linear BSDE (2.1).

Proposition 2.2. *Let Assumption 2.1 be valid and $\zeta \in L_1^2(\mathcal{F}_T)$ be non-negative. If φ and ξ_T are non-negative and (Y,Z) is any solution to the linear BSDE* (2.1)*, then:*
(a) if $Y \geq -\zeta$, then Y is non-negative;
(b) if $Y \geq -\zeta$ and $\mathbb{1}_B Y_t = 0$ for some $t \in [0,T]$ and $B \in \mathcal{F}_t$, then $\mathbb{1}_B \xi_T = 0$ and $\varphi_s = Y_s = Z_s = 0$, $d\mathbb{P} \otimes ds$–a.e. on $B \times [t,T]$.

Proof. Since μ and σ are bounded, it can be shown that $\Gamma_T^* := \sup_{t\in[0,T]} |\Gamma_{0,t}| \in L_1^2(\mathcal{F}_T)$ and thus the random variable $\Gamma_T^*\zeta$ is integrable. The local martingale M satisfies

$$M_t = \Gamma_{0,t} Y_t + \int_0^t \Gamma_{0,s}\varphi_s \, ds \geq \Gamma_{0,t} Y_t \geq -\Gamma_T^*\zeta$$

and thus it is bounded from below by an integrable random variable. By Lemma 12.2, M is a supermartingale, so that $M_t \geq \mathbb{E}(M_T \,|\, \mathcal{F}_t)$, which in turn implies that

$$Y_t \geq \mathbb{E}\Big(\Gamma_{t,T}\xi_T + \int_t^T \Gamma_{t,s}\varphi_s \, ds \,\Big|\, \mathcal{F}_t\Big) \geq 0. \tag{2.16}$$

It is now enough to argue as in the proof of Theorem 2.1. □

We now consider linear BSDEs with different generators and terminal conditions.

Corollary 2.2. *Let Assumption 2.1 be valid and a non-negative random variable ζ^1 belong to $L^2_1(\mathcal{F}_T)$. Let (Y^1, Z^1) and (Y^2, Z^2) be solutions to the linear BSDE, for $i = 1, 2$,*

$$-dY^i_t = (\varphi^i_t + \mu_t Y^i_t + \sigma_t Z^i_t)\, dt - Z^i_t\, dW_t, \quad Y^i_T = \xi^i_T,$$

where $\varphi^1 \geq \varphi^2$ and $\xi^1_T \geq \xi^2_T$. If Y^1 is bounded from below by $-\zeta^1$ and Y^2 belongs to $\mathcal{S}^{2,c}_1(0,T)$, then $Y^1 \geq Y^2$. If, in addition, $\mathbb{1}_B Y^1_t = \mathbb{1}_B Y^2_t$ for some $t \in [0,T]$ and $B \in \mathcal{F}_t$, then $\mathbb{1}_B \xi^1_T = \mathbb{1}_B \xi^1_T$ and $\varphi^1 = \varphi^2, Y^1 = Y^2, Z^1 = Z^2$, $d\mathbb{P} \otimes ds$–a.e. on $B \times [t,T]$.

Proof. We note that the process $\widehat{Y} := Y^1 - Y^2$ is bounded from below by $-\zeta$ where the non-negative, square-integrable random variable ζ is given by

$$\zeta := \zeta^1 + \zeta^2, \quad \zeta^2 = \sup_{t\in[0,T]} |Y^2_t|.$$

Moreover, $(\widehat{Y}, \widehat{Z}) := (Y^1 - Y^2, Z^1 - Z^2)$ is a solution to the linear BSDE (2.1) with $\varphi = \varphi^1 - \varphi^2 \geq 0$ and $\xi_T = \xi^1_T - \xi^2_T \geq 0$, that is,

$$-d\widehat{Y}_t = \big(\varphi^1_t - \varphi^2_t + \mu_t \widehat{Y}_t + \sigma_t \widehat{Z}_t\big)\, dt - \widehat{Z}_t\, dW_t, \quad \widehat{Y}_T = \xi^1_T - \xi^2_T.$$

It now suffices to apply Proposition 2.2 to obtain the inequality $\widehat{Y} \geq 0$ and thus also $Y^1 \geq Y^2$. The second assertion is also a consequence of Proposition 2.2. □

Corollary 2.3. *If a solution (Y^2, Z^2) to* (2.1) *is such that Y^2 belongs to the space $\mathcal{S}^{2,c}_1(0,T)$, then Y^2 is dominated by Y^1 where (Y^1, Z^1) is any solution to* (2.1) *such that Y^1 is bounded from below by a square-integrable random variable.*

3 Comparison Theorems for BSDEs

Recall that Theorem 1.1 was established under the assumption that (g, ξ_T) are *standard parameters*, in the sense of the following definition (see Assumption 1.1 and Definition 1.1).

Definition 3.1. We say that (g, ξ_T) are *standard parameters* if:
(i) the mapping $g : \Omega \times [0,T] \times \mathbb{R}^m \times \mathbb{R}^{m\times d} \to \mathbb{R}^m$ is $\mathcal{P} \otimes \mathcal{B}(\mathbb{R}^m) \otimes \mathcal{B}(\mathbb{R}^{m\times d})$-measurable and $g(\cdot, 0, 0) \in \mathcal{H}^2_m$;
(ii) g is uniformly Lipschitz continuous: there exists a constant L such that $d\mathbb{P} \otimes dt$–a.e.

$$\|g(\omega, t, y_1, z_1) - g(\omega, t, y_2, z_2)\| \le L\big(\|y_1 - y_2\| + \|z_1 - z_2\|\big),\ \forall\, y_1, y_2 \in \mathbb{R}^m,\ z_1, z_2 \in \mathbb{R}^{m\times d};$$

(iii) $\xi_T \in L^2_m(\mathcal{F}_T)$.

From Theorem 1.1, we know that if (g, ξ_T) is a standard parameter then the BSDE has a unique solution in $\mathcal{S}^2_m(0,T) \times \mathcal{H}^2_{m\times d}(0,T)$. We take $m = 1$, so that the processes Y^1 and Y^2 are real-valued, although the Brownian motion W is d-dimensional.

Theorem 3.1 (Peng [77])**.** *Let (g^1, ξ^1_T) and (g^2, ξ^2_T) be standard parameters and let (Y^1, Z^1) and (Y^2, Z^2) be the unique solutions in $\mathcal{S}^2_1(0,T) \times \mathcal{H}^2_d(0,T)$ to the associated BSDEs for $i = 1, 2$*

$$-dY^i_t = g^i(t, Y^i_t, Z^i_t)\, dt - Z^i_t\, dW_t, \quad Y^i_T = \xi^i_T. \tag{3.1}$$

We assume that $\xi^1_T \ge \xi^2_T$ and

$$\delta_2 g_t := g^1(t, Y^2_t, Z^2_t) - g^2(t, Y^2_t, Z^2_t) \ge 0, \quad d\mathbb{P} \otimes dt - a.e.$$

Then the following assertions are valid:
(a) the comparison property holds: $Y^1_t \ge Y^2_t$ for all $t \in [0,T]$;
(b) the strict comparison property holds: if $\mathbb{1}_B Y^1_t = \mathbb{1}_B Y^2_t$ for some event $B \in \mathcal{F}_t$, then $\mathbb{1}_B \xi^1_T = \mathbb{1}_B \xi^2_T$, $\delta_2 g = 0$ on $B \times [t,T]$, $d\mathbb{P} \otimes dt$–a.e., and $\mathbb{1}_B Y^1_s = \mathbb{1}_B Y^2_s$ for all $s \in [t,T]$.

Remark 3.1. A comparison theorem for multidimensional BSDEs was established by Hu and Peng [41] who used the results on the *viability property* of solutions to multidimensional BSDEs established by Buckdahn et al. [12]. For comparison theorems for solutions to SDEs, see Yamada [95, 96] or Yamada and Ogura [97].

Lemma 3.1. *For all $y \in \mathbb{R}$ and $z \in \mathbb{R}^d$*

$$-\mathbb{1}_{\{y>0\}}\|z\|^2 + Ly^+(|y| + \|z\|) \le (L + L^2)(y^+)^2. \tag{3.2}$$

Proof. If $y \le 0$, then both sides are manifestly equal to 0. If $y > 0$, then we need to check that

$$-\|z\|^2 + Ly(y + \|z\|)) \le (L + L^2)y^2.$$

This is equivalent to

$$L^2y^2 - Ly\|z\| + \|z\|^2 \ge L^2y^2 - 2Ly\|z\| + \|z\|^2 = (Ly - \|z\|)^2 \ge 0$$

and thus the lemma is established. □

Lemma 3.2 (Gronwall [39], Bellman [2]). *Let $u : \mathbb{R}_+ \to \mathbb{R}_+$ be a non-negative, locally integrable function. If*

$$u(t) \le a + b \int_0^t u(s)\,ds$$

where a, b are two constants, then $u(t) \le ae^{bt}$ for every $t \in \mathbb{R}_+$.

Proof. Manifestly,

$$u(t) \le a + b\Big(\int_0^t \Big(a + b \int_0^v u(s)\,ds \Big) dv \Big) = a + abt + \cdots + a\frac{(bt)^n}{n!} + \frac{b^{n+1}t^n}{n!} \int_0^t u(s)\,ds.$$

Since u is locally bounded, the last term tends to zero as n goes to infinity. Hence the desired inequality follows. □

First proof of Theorem 3.1. The proof hinges on the extended Itô formula and Lemma 3.2. We denote $Y := Y^2 - Y^1$, $Z := Z^2 - Z^1$, $\xi_T := \xi_T^2 - \xi_T^1$ so that (3.1) yields

$$dY_t = d(Y_t^2 - Y_t^1) = \big(g^1(t, Y_t^1, Z_t^1) - g^2(t, Y_t^2, Z_t^2)\,dt - (Z_t^1 - Z_t^2)\,dW_t \tag{3.3}$$

with the terminal condition $Y_T = \xi_T \le 0$. Let us consider the process $f(Y_t) = (Y_t^+)^2$. Since $f(y) = (y^+)^2$, $f'(y) = 2y^+$ and $f''(y) = 2\mathbb{1}_{\{y>0\}}$, it is clear that $f \in C^1(\mathbb{R})$ and f'' is locally integrable. Hence an application of the extended Itô formula (see Corollary 12.2)

$$f(Y_T) = f(Y_t) + \int_t^T f'(Y_s)\,dY_s + \frac{1}{2}\int_0^t f''(Y_s)\,d\langle Y\rangle_s$$

yields

$$\begin{aligned}(Y_t^+)^2 =& (\xi_T^+)^2 - 2\int_t^T Y_s^+ Z_s\,dW_s + 2\int_t^T Y_s^+\big(g^2(s, Y_s^2, Z_s^2) - g^1(s, Y_s^1, Z_s^1)\big)\,ds \\ & - \int_t^T \mathbb{1}_{\{Y_s>0\}} \|Z_s\|^2\,ds \end{aligned} \tag{3.4}$$

since $d\langle Y\rangle_s = \|Z_s\|^2\,ds$. Since $\xi_T^1 \ge \xi_T^2$, we have $\xi_T \le 0$ and thus $\xi_T^+ = 0$. Observe that

$$g^2(s, Y_s^2, Z_s^2) - g^1(s, Y_s^1, Z_s^1) = -\delta_2 g_s + g^1(s, Y_s^2, Z_s^2) - g^1(s, Y_s^1, Z_s^1)$$

where, by assumption, $-\delta_2 g_s = g^2(s, Y_s^2, Z_s^2) - g^1(s, Y_s^2, Z_s^2) \le 0$. We note that the integral

$$\int_t^T Y_s^+ g^2(s, Y_s^2, Z_s^2)\,ds$$

is finite, $\mathbb{P}-$a.s., since the process Y (and thus also the process Y^+) belongs to $\mathcal{S}_1^{2,c}(0,T)$ and the process $g^2(\cdot, Y^2, Z^2)$ belongs to $\mathcal{H}_1^2(0,T)$. Therefore, using also the uniform Lipschitz condition satisfied by g^1, we obtain

$$\begin{aligned}
&\int_t^T Y_s^+(g^2(s,Y_s^2,Z_s^2) - g^1(s,Y_s^1,Z_s^1))\,ds \\
&= \int_t^T Y_s^+\big(-\delta_2 g_s + g^1(s,Y_s^2,Z_s^2) - g^1(s,Y_s^1,Z_s^1)\big)\,ds \\
&\le \int_t^T Y_s^+\big(g^1(s,Y_s^2,Z_s^2) - g^1(s,Y_s^1,Z_s^1)\big)\,ds \\
&\le \int_t^T Y_s^+|g^1(s,Y_s^2,Z_s^2) - g^1(s,Y_s^1,Z_s^1)|\,ds \\
&\le L\int_t^T Y_s^+(|Y_s| + \|Z_s\|)\,ds.
\end{aligned} \tag{3.5}$$

Moreover, as in the proof of Theorem 1.1, we obtain

$$\begin{aligned}
\mathbb{E}\Big[\Big(\int_0^T (Y_s^+)^2\|Z_s\|^2\,ds\Big)^{1/2}\Big] &\le \mathbb{E}\Big[\sup_{t\in[0,T]}|Y_t|\Big(\int_0^T \|Z_s\|^2\,ds\Big)^{1/2}\Big] \\
&\le \frac{1}{2}\,\mathbb{E}\Big[\sup_{t\in[0,T]}|Y_t|^2 + \int_0^T \|Z_s\|^2\,ds\Big] < \infty
\end{aligned}$$

where the last inequality holds since (Y,Z) belongs to $\mathcal{S}_1^2(0,T)\times\mathcal{H}_d^2(0,T)$.

Hence the expectation of the Itô stochastic integral (3.4) equals zero. By combining inequalities (3.2), (3.4) and (3.5), we obtain

$$\mathbb{E}\big((Y_t^+)^2\big) \le (L+L^2)\int_t^T \mathbb{E}(Y_s^+)^2\,ds$$

and thus the Gronwall–Bellman lemma gives $\mathbb{E}\big((Y_t^+)^2\big) \le 0$ for every $t\in[0,T]$. We conclude that the equality $Y_t^+ = 0$ is valid for every $t\in[0,T]$, which immediately implies that the inequality $Y_t^1 \ge Y_t^2$ is satisfied for every $t\in[0,T]$ and thus part (a) in Theorem 3.1 is established. □

Second proof of Theorem 3.1. We define

$$\Delta_Y g_t^1 := \frac{g^1(t,Y_t^1,Z_{t,1}^1,Z_{t,2}^1,\dots,Z_{t,d}^1) - g^1(t,Y_t^2,Z_{t,1}^1,Z_{t,2}^1,\dots,Z_{t,d}^1)}{Y_t^1 - Y_t^2}\mathbb{1}_A$$

where $A := \{Y_t^1 \ne Y_t^2\}$ and $\Delta_Z g_t^1 := (\Delta_{Z_1}g_t^1,\dots,\Delta_{Z_d}g_t^1)$ where for $j=1,\dots,d$

$$\Delta_{Z_j} g_t^1 := \frac{g^1(t,Y_t^2,\dots,Z_{t,j-1}^2,Z_{t,j}^2,Z_{t,j+1}^1,\dots,Z_{t,d}^1) - g^1(t,Y_t^2,\dots,Z_{t,j-1}^2,Z_{t,j}^1,Z_{t,j+1}^1,\dots,Z_{t,d}^1)}{Z_{t,j}^1 - Z_{t,j}^2}\mathbb{1}_{A_j}$$

where $A_j := \{Z_{t,j}^1 \ne Z_{t,j}^2\}$. Using (3.1), we obtain

$$-d(Y_t^1 - Y_t^2) = \big(g^1(t,Y_t^1,Z_t^1) - g^2(t,Y_t^2,Z_t^2)\,dt - (Z_t^1 - Z_t^2)\,dW_t \tag{3.6}$$

with the terminal condition $Y^1_T - Y^2_T = \xi^1_T - \xi^2_T$. Note that when $d = 1$

$$
\begin{aligned}
g^1(t, Y^1_t, Z^1_t) - g^2(t, Y^2_t, Z^2_t) &= g^1(t, Y^1_t, Z^1_t) - g^1(t, Y^2_t, Z^1_t) \\
&\quad + g^1(t, Y^2_t, Z^1_t) - g^1(t, Y^2_t, Z^2_t) + g^1(t, Y^2_t, Z^2_t) - g^2(t, Y^2_t, Z^2_t) \\
&= \frac{g^1(t, Y^1_t, Z^1_t) - g^1(t, Y^2_t, Z^1_t)}{Y^1_t - Y^2_t} (Y^1_t - Y^2_t) \mathbb{1}_{\{Y^1_t \neq Y^2_t\}} \\
&\quad + \frac{g^1(t, Y^2_t, Z^1_t) - g^1(t, Y^2_t, Z^2_t)}{Z^1_t - Z^2_t} (Z^1_t - Z^2_t) \mathbb{1}_{\{Z^1_t \neq Z^2_t\}} + \delta_2 g_t \\
&= \delta_2 g_t + \mu_t (Y^1_t - Y^2_t) + \sigma_t (Z^1_t - Z^2_t)
\end{aligned}
$$

where we write

$$
\mu_t = \Delta_Y g^1_t := \frac{g^1(t, Y^1_t, Z^1_t) - g^1(t, Y^2_t, Z^1_t)}{Y^1_t - Y^2_t}
$$

and

$$
\sigma_t = \Delta_Z g^1_t := \frac{g^1(t, Y^2_t, Z^1_t) - g^1(t, Y^2_t, Z^2_t)}{Z^1_t - Z^2_t}.
$$

Since the mapping g^1 is uniformly Lipschitz continuous with respect to (y, z) the $\mathbb{F}$-progressively measurable processes β and μ are bounded. Notice that an analogous decomposition can be derived for any $d \geq 1$.

Let $Y := Y^1 - Y^2$, $Z := Z^1 - Z^2$, $\varphi_t := \delta_2 g_t$ and $\xi_T := \xi^1_T - \xi^2_T$. Then (Y, Z) solves

$$
-dY_t = \big(\varphi_t + \mu_t Y_t + \sigma_t Z_t\big)\, dt - Z_t\, dW_t, \quad Y_T = \xi_T.
$$

Using Theorem 2.1, we obtain

$$
Y_t = \mathbb{E}\Big(\xi_T \Gamma_{t,T} + \int_t^T \Gamma_{t,s} \varphi_s \, ds \,\Big|\, \mathcal{F}_t\Big), \quad \mathbb{P} - \text{a.s.},
$$

where $\Gamma_{t,s},\ s \in [t, T]$ is the unique solution to the SDE

$$
d\Gamma_{t,s} = \Gamma_{t,s}(\mu_s\, ds + \sigma_s\, dW_s), \qquad \Gamma_{t,t} = 1.
$$

Since the random variable ξ_T and processes $\varphi = \delta_2 g$ and $\Gamma_{t,\cdot}$ are non-negative it is clear that the inequality $Y_t \geq 0$ is satisfied for every $t \in [0, T]$. Moreover, the strict comparison property holds as well due to Theorem 2.1. □

An inspection of the second proof of Theorem 3.1 shows that the following extension of Theorem 3.1 is valid.

Proposition 3.1. *Assume that:*
(i) (Y^1, Z^1) *and* (Y^2, Z^2) *are (not necessarily unique) solutions to BSDEs with parameters* (g^1, ξ^1_T) *and* (g^2, ξ^2_T) *where* $\xi^1_T, \xi^2_T \in L^2_1(\mathcal{F}_T)$ *are such that* $\xi^1_T \geq \xi^2_T$*;*
(ii) the generator g^1 *is uniformly Lipschitz continuous;*
(iii) the process $\delta_2 g$ *belongs to the space* $\mathcal{H}^2_1(0, T)$ *and satisfies*

$$
\delta_2 g_t := g^1(t, Y^2_t, Z^2_t) - g^2(t, Y^2_t, Z^2_t) \geq 0, \quad d\mathbb{P} \otimes dt - a.e.
$$

Then the assertions (a) and (b) in Theorem 3.1 are valid.

To establish Proposition 3.1, it suffices to check that the second proof of Theorem 3.1 is still valid when the assumptions of Theorem 3.1 are replaced by assumptions (i)–(iii). Note that the statement is still valid if (ii) and (iii) are replaced by: g^2 is uniformly Lipschitz continuous and the process $\delta_1 g$ belongs to the space $\mathcal{H}^2_1(0,T)$ and satisfies

$$\delta_1 g_t := g^1(t, Y^1_t, Z^1_t) - g^2(t, Y^1_t, Z^1_t) \geq 0, \quad d\mathbb{P} \otimes dt - \text{a.e.}$$

We now extend Corollary 2.2 to nonlinear BSDEs. In the next result, we do not assume that $\delta_2 g$ belongs to $\mathcal{H}^2_1(0,T)$ or that the terminal conditions are square-integrable.

Corollary 3.1. *Assume that:*
(i) (Y^1, Z^1) *and* (Y^2, Z^2) *are (not necessarily unique) solutions to BSDEs with parameters* (g^1, ξ^1_T) *and* (g^2, ξ^2_T) *where* $\xi^1_T \geq \xi^2_T$*;*
(ii) the generator g^1 *is uniformly Lipschitz continuous;*
(iii) the following inequality holds

$$\delta_2 g_t := g^1(t, Y^2_t, Z^2_t) - g^2(t, Y^2_t, Z^2_t) \geq 0, \quad d\mathbb{P} \otimes dt - \textit{a.e.}$$

(iv) $Y^1 \geq -\zeta$ *for some non-negative random variable* $\zeta \in L^2_1(\mathcal{F}_T)$*;*
(v) Y^2 *belongs to* $\mathcal{S}^{2,c}_1(0,T)$*.*
Then the inequality $Y^1_t \geq Y^2_t$ *holds for all* $t \in [0,T]$*.*

Proof. Let $Y := Y^1 - Y^2$, $Z := Z^1 - Z^2$ and $\xi_T := \xi^1_T - \xi^2_T$. Then (Y, Z) solves the linear BSDE

$$-dY_t = \big(\delta_2 g_t + \mu_t Y_t + \sigma_t Z_t\big)\, dt - Z_t \, dW_t, \quad Y_T = \xi_T.$$

Since $\delta_2 g$ and ξ_T are non-negative, the assertion follows from Corollary 2.2. □

4 BSDE with a Continuous Generator

In this section, we will relax the assumption that the generator g is uniformly Lipschitz continuous and we will postulate instead that g is continuous and satisfies the linear growth condition in (y, z). As in the preceding section, we assume that $m = 1$, that is, the process Y in BSDEs under consideration is one-dimensional.

4.1 Continuous Generators

We will now work under the following assumption (H.2), which is manifestly weaker than assumption (H.1).

Assumption 4.1. We say that g satisfies *assumption (H.2)* if:
(i) the mapping $g : \Omega \times [0,T] \times \mathbb{R}^{d+1} \to \mathbb{R}$ is $\mathcal{P} \otimes \mathcal{B}(\mathbb{R}^{d+1})$-measurable;
(ii) the process $g(\cdot, 0, 0)$ belongs to the space $\mathcal{H}^2_1(0,T)$;
(iii) the function $(y, z) \mapsto g(\omega, t, y, z)$ is continuous, $d\mathbb{P} \otimes dt-$a.e.;
(iv) g satisfies the linear growth condition: there exists a constant $k > 0$ such that

$$|g(\omega, t, y, z)| \leq k(1 + |y| + \|z\|), \quad \forall\, (y, z) \in \mathbb{R}^{d+1},\ d\mathbb{P} \otimes dt - \text{a.e.}$$

Remark 4.1. Condition (iv) in Assumption 4.1 may be weakened as follows: there exists a constant $k > 0$ such that

$$|g(\omega, t, y, z) - g(\omega, t, 0, 0)| \leq k(1 + |y| + \|z\|), \quad \forall\, (y, z) \in \mathbb{R}^{d+1},\ d\mathbb{P} \otimes dt - \text{a.e.} \tag{4.1}$$

4.2 Definition of the Minimal Solution

The definition of the *minimal solution* to the BSDE is fairly natural. It is clear that it can be complemented by an analogous definition of the *maximal solution* and the proof of Theorem 4.1 can be modified to cover also the latter case.

Definition 4.1. We say that (Y, Z) is a *minimal solution* to BSDE (1.1) if for any solution (Y', Z') to BSDE (1.1) we have $Y_t \leq Y'_t$ for every $t \in [0, T]$.

4.3 Existence of a Minimal Solution

The next result establishes the existence of a minimal solution to BSDE (1.1). The uniqueness of a minimal (or a maximal) solution is obvious. However, the uniqueness of a solution under the assumptions of Theorem 4.1 does not hold, in general.

Theorem 4.1 (Lepeltier and San Martín [55])**.** *Assume that the generator g satisfies assumption (H.2) and the terminal condition ξ_T belongs to $L^2(\mathcal{F}_T)$. Then BSDE* (1.1) *with parameters (g, ξ_T) has a minimal solution (Y, Z). Moreover, the minimal solution (Y, Z) belongs to $\mathcal{S}^2_1(0, T) \times \mathcal{H}^2_d(0, T)$.*

In the proof of Theorem 4.1, we will employ Lemma 4.1 in which the symbol $\Box$ stands for the *infimal convolution* (see, e.g., [89]). The infimal convolution appears to be a convenient tool to construct an increasing sequence of Lipschitz continuous functions converging to a given continuous function. Obviously, the Lipschitz constant depends here on n, unless the function g is already Lipschitz continuous with $L \leq n$.

Lemma 4.1 (Infimal convolution). *Let $g : \mathbb{R}^d \to \mathbb{R}$ be a continuous function satisfying the linear growth condition for a fixed $k > 0$, that is,*

$$|g(x)| \leq k(1 + \|x\|), \quad \forall\, x \in \mathbb{R}^d.$$

For any natural number n, we set $b_n^0(x) := n\|x\|$ and we define the function $g_n : \mathbb{R}^d \to \mathbb{R}$ by the formula

$$g_n(x) := \inf_{y \in \mathbb{Q}^d} \big(g(y) + n\|x - y\| \big) = (g \,\Box\, b_n^0)(x).$$

Then the following properties are valid:
(i) the sequence $g_n : \mathbb{R}^d \to \mathbb{R}$ is non-decreasing, that is, $g_{n+1}(x) \geq g_n(x)$ for every $x \in \mathbb{R}^d$ and $n \in \mathbb{N}$;
(ii) for any $n \geq k$, we have $|g_n(x)| \leq k(1 + \|x\|)$ for all $x \in \mathbb{R}^d$;
(iii) the function g_n is Lipschitz continuous with constant n, that is,

$$|g_n(x) - g_n(y)| \leq n\|x - y\|, \quad \forall\, x, y \in \mathbb{R}^d,$$

(iv) the convergence of g_n to g is uniform on compact sets;
(v) if $\lim_{n\to\infty} x_n = x$, then $\lim_{n\to\infty} g_n(x_n) = g(x)$.

Proof of Theorem 4.1. In what follows, we fix $k > 0$ and we consider $n \geq k$.

Step 1. Let g be a generator satisfying assumption (H.2) and let $b_n^0(x) := n\|x\|$ for all $x \in \mathbb{R}^{d+1}$. In view of Lemma 4.1, the generator $g_n = g \,\Box\, b_n^0$ satisfies assumption (H.1), that is, (g_n, ξ_T) are standard parameters. Due to Theorem 1.1, the BSDE

$$\begin{cases} -dY_t^n = g_n(t, Y_t^n, Z_t^n)\, dt - Z_t^n\, dW_t, \quad t \in [0, T], \\ Y_T^n = \xi_T, \end{cases} \tag{4.2}$$

has a unique solution $(Y^n, Z^n) \in \mathcal{S}_1^2(0,T) \times \mathcal{H}_d^2(0,T)$ for every n. Let us denote by (U^1, V^1) and (U^2, V^2) the unique solutions in $\mathcal{S}_1^2(0,T) \times \mathcal{H}_d^2(0,T)$ to BSDEs with standard parameters $(h_1 = g^0 - b_k, \xi_T)$ and $(h_2 = g^0 + b_k, \xi_T)$ where $g_t^0 := g(\omega, t, 0, 0)$ for all $t \in [0,T]$ and $b_k(y,z) := k(1 + |y| + \|z\|)$ for every $(y,z) \in \mathbb{R}^{d+1}$. The linear growth condition (4.1) yields

$$-k(1 + |Y_t^n| + \|Z_t^n\|) \leq g_n(t, Y_t^n, Z_t^n) - g_t^0 \leq k(1 + |Y_t^n| + \|Z_t^n\|)$$

and thus

$$\begin{aligned} g_n(t, Y_t^n, Z_t^n) - h_1(t, Y_t^n, Z_t^n) &= (g_n(t, Y_t^n, Z_t^n) - g_t^0) + b_k(t, Y_t^n, Z_t^n) \\ &\geq b_k(t, Y_t^n, Z_t^n) - k(1 + |Y_t^n| + \|Z_t^n\|) = 0 \end{aligned}$$

and

$$\begin{aligned} h_2(t, Y_t^n, Z_t^n) - g_n(t, Y_t^n, Z_t^n) &= b_k(t, Y_t^n, Z_t^n) - (g_n(t, Y_t^n, Z_t^n) - g_t^0) \\ &\geq b_k(t, Y_t^n, Z_t^n) - k(1 + |Y_t^n| + \|Z_t^n\|) = 0. \end{aligned}$$

By an application of the comparison theorem for BSDEs with standard parameters (see Theorem 3.1), we obtain for every n

$$U_t^1 \leq Y_t^n \leq Y_t^{n+1} \leq U_t^2, \quad \forall\, t \in [0, T].$$

We also observe that the increasing sequence $I_n := \sup_{0\le t\le T}(Y_t^n)^2$ is bounded in $L^1(\mathcal{F}_T)$, since for every n

$$\mathbb{E}\Big(\sup_{0\le t\le T}(Y_t^n)^2\Big) \le \mathbb{E}\Big(\sup_{0\le t\le T}\big(|U_t^1|+|U_t^2|\big)^2\Big) =: \gamma_1 < \infty. \tag{4.3}$$

Hence the increasing sequence of continuous processes $(Y^n)_{n=k}^\infty$ converges $\mathbb{P}-$a.s. and in $\mathcal{H}_1^2(0,T)$ to a lower semi-continuous process, which is denoted as Y. Note that

$$\mathbb{P}\big(\lim_{n\to\infty} Y_t^n = Y_t,\ \forall\, t\in[0,T]\big) = 1. \tag{4.4}$$

It is clear that $\mathbb{E}\big(\sup_{0\le t\le T}(Y_t)^2\big) < \infty$, that is, Y belongs to $\mathcal{S}_1^2(0,T)$.

Step 2. We will now establish some estimates for the sequence of random variables

$$J_n := \int_0^T \|Z_t^n\|^2\, dt.$$

By applying the Itô formula to $(Y^n)^2$, we obtain

$$d(Y_t^n)^2 = 2Y_t^n\, dY_t^n + d\langle Y^n\rangle_t = 2Y_t^n\big(-g_n(t,Y_t^n,Z_t^n)\,dt + Z_t^n\,dW_t\big) + \|Z_t^n\|^2\,dt$$

so that

$$\xi_T^2 = (Y_0^n)^2 - \int_0^T 2Y_t^n g_n(t,Y_t^n,Z_t^n)\,dt + \int_0^T 2Y_t^n Z_t^n\,dW_t + \int_0^T \|Z_t^n\|^2\,dt$$

or, equivalently,

$$(Y_0^n)^2 + \int_0^T \|Z_t^n\|^2\,dt = \xi_T^2 + \int_0^T 2Y_t^n g_n(t,Y_t^n,Z_t^n)\,dt + \int_0^T 2Y_t^n Z_t^n\,dW_t$$

where the expectation of the last integral is zero. The following inequality is valid for any $k>0$ and arbitrary real numbers a,b

$$2k|a||b| \le 2k^2|a|^2 + \frac{1}{2}\,|b|^2. \tag{4.5}$$

Therefore, using also the postulated linear growth condition satisfied by g (and thus also by g_n for $n\ge k$) and the assumption that the process $g(\cdot,0,0)$ belongs to the space $\mathcal{H}_1^2(0,T)$, we obtain the following estimates

$$\begin{aligned}
&\mathbb{E}\left((Y_0^n)^2 + \int_0^T \|Z_t^n\|^2\,dt\right)\\
&\le \mathbb{E}(\xi_T^2) + \mathbb{E}\left(\int_0^T |Y_t^n|\Big(|g(t,0,0)| + k\big(1+|Y_t^n|+\|Z_t^n\|\big)\Big)\,dt\right)\\
&\le c_1 + c_2\,\mathbb{E}\left(\int_0^T |Y_t^n|^2\,dt\right) + 2k\,\mathbb{E}\left(\int_0^T |Y_t^n|\|Z_t^n\|\,dt\right)\\
&\le c_1 + c_2\,\mathbb{E}\left(\int_0^T (|U_t^1|+|U_t^2|)^2\,dt\right) + \frac{1}{2}\,\mathbb{E}\left(\int_0^T \|Z_t^n\|^2\,dt\right)
\end{aligned}$$

where c_1 and c_2 are positive constants, which may change from place to place but are independent of n. Consequently,

$$\mathbb{E}(J_n) = \mathbb{E}\left(\int_0^T \|Z_t^n\|^2\,dt\right) \le 2c_1 + 2c_2\,\mathbb{E}\left(\int_0^T (|U_t^1|+|U_t^2|)^2\,dt\right) =: \gamma_2 < \infty. \tag{4.6}$$

We conclude that the sequence J_n is bounded in $L^1(\mathcal{F}_T)$. Since g_n satisfies the linear growth condition and the process $g(\cdot,0,0)$ belongs to the space $\mathcal{H}_1^2(0,T)$, the sequence K_n is also bounded in $L^1(\mathcal{F}_T)$ where

$$K_n := \int_0^T g_n^2(t, Y_t^n, Z_t^n)\,dt.$$

Step 3. In this step, we aim to show that $(Z^n)_{n=k}^{\infty}$ is the Cauchy sequence in the space $\mathcal{H}_d^2(0,T)$. Since the sequences I_n and J_n are bounded in $L^1(\mathcal{F}_T)$, there exists a constant β such that for all m, n

$$\begin{aligned}
&\mathbb{E}\left(\int_0^T (g_n(t,Y_t^n,Z_t^n) - g_m(t,Y_t^m,Z_t^m))^2\,dt\right)\\
&\le 2\,\mathbb{E}\left(\int_0^T \Big(\big(g_n(t,Y_t^n,Z_t^n) - g(t,0,0)\big)^2 + \big(g_m(t,Y_t^m,Z_t^m) - g(t,0,0)\big)^2\Big)\,dt\right)\\
&\le 2k^2\,\mathbb{E}\left(\int_0^T \big((1+|Y_t^n| + \|Z_t^n\|)^2 + (1+|Y_t^m| + \|Z_t^m\|)^2\big)\,dt\right) \le \beta.
\end{aligned}$$

Let us denote $c_{m,n} := (Y_0^n - Y_0^m)^2$. Using first the Itô formula and subsequently the following versions of the Cauchy-Schwarz inequality:

$$\Big(\int_0^T f(t)g(t)\,dt\Big)^2 \le \Big(\int_0^T f^2(t)\,dt\Big)\Big(\int_0^T g^2(t)\,dt\Big)$$

and

$$(\mathbb{E}(XY))^2 \le \mathbb{E}(X^2)\mathbb{E}(Y^2),$$

we obtain

$$\begin{aligned}
&c_{m,n} + \mathbb{E}\left(\int_0^T \|Z_t^n - Z_t^m\|^2\,dt\right)\\
&\le 2\,\mathbb{E}\left(\int_0^T (Y_t^n - Y_t^m)\big(g_n(t,Y_t^n,Z_t^n) - g_m(t,Y_t^m,Z_t^m)\big)\,dt\right)\\
&\le 2\,\mathbb{E}\left[\left(\int_0^T (Y_t^n - Y_t^m)^2\,dt\right)^{1/2}\left(\int_0^T \big(g_n(t,Y_t^n,Z_t^n) - g_m(t,Y_t^m,Z_t^m)\big)^2\,dt\right)^{1/2}\right]\\
&\le 2\left[\mathbb{E}\left(\int_0^T (Y_t^n - Y_t^m)^2\,dt\right)\right]^{1/2}\left[\mathbb{E}\left(\int_0^T \big(g_n(t,Y_t^n,Z_t^n) - g_m(t,Y_t^m,Z_t^m)\big)^2\,dt\right)\right]^{1/2}\\
&\le 2\sqrt{\beta}\left[\mathbb{E}\left(\int_0^T (Y_t^n - Y_t^m)^2\,dt\right)\right]^{1/2}
\end{aligned}$$

for arbitrary natural numbers $m, n \ge k$. We conclude that $(Z^n)_{n=k}^{\infty}$ is a Cauchy sequence in the space $\mathcal{H}_d^2(0,T)$, since we have shown in Step 1 that the sequence $(Y^n)_{n=k}^{\infty}$ converges (hence is a Cauchy sequence) in $\mathcal{H}_1^2(0,T)$. Therefore, there exists a process $Z \in \mathcal{H}_d^2(0,T)$ such that $(Z^n)_{n=k}^{\infty}$ converges to Z in $\mathcal{H}_d^2(0,T)$, that is,

$$\lim_{n\to\infty} \mathbb{E}\left(\int_0^T \|Z_t^n - Z_t\|^2\,dt\right) = 0. \tag{4.7}$$

Step 4. Let τ be an arbitrary stopping time taking values in $[0,T]$. Then from (4.7)

$$\lim_{n\to\infty} \mathbb{E}\left(\left(\int_0^\tau (Z_t^n - Z_t)\,dW_t\right)^2\right) \le \lim_{n\to\infty} \mathbb{E}\left(\int_0^T \|Z_t^n - Z_t\|^2\,dt\right) = 0.$$

In view of (4.3) and (4.4), we also have

$$\lim_{n\to\infty} \mathbb{E}\left((Y_\tau^n - Y_\tau)^2\right) = 0.$$

Our next goal is to show that $\lim_{n\to\infty} L_n = 0$ where

$$L_n := \mathbb{E}\left(\Big|\int_0^\tau \big(g_n(t, Y_t^n, Z_t^n) - g(t, Y_t, Z_t)\big)\,dt\Big|\right). \tag{4.8}$$

Observe that

$$\begin{aligned} L_n &\le \mathbb{E}\left(\int_0^\tau \big|g_n(t, Y_t^n, Z_t^n) - g(t, Y_t, Z_t)\big|\,dt\right) \\ &\le \mathbb{E}\left(\int_0^T \big|g_n(t, Y_t^n, Z_t^n) - g(t, Y_t, Z_t)\big|\,dt\right). \end{aligned}$$

We observe that the linear growth property of g_n yields

$$\big|g_n(t, Y_t^n, Z_t^n) - g(t, Y_t, Z_t)\big| \le 2k\big(1 + |Y_t^n| + \|Z_t^n\| + |\Psi_t|\big)$$

where $\Psi := |Y| + \|Z\|$ belongs to $\mathcal{H}_1^2(0,T)$. Moreover, it is easy to deduce from (4.3) and (4.6) that the sequence $(|Y^n| + \|Z^n\|)$ is bounded in $\mathcal{H}_1^2(0,T)$. Using part (v) in Lemma 4.1, (4.4) and (4.7), we obtain

$$\lim_{n\to\infty} \big|g_n(t, Y_t^n, Z_t^n) - g(t, Y_t, Z_t)\big| = 0.$$

We conclude that $\lim_{n\to\infty} L_n = 0$.

Step 5. We are now ready to show that the limit $(Y,Z) = \lim_{n\to\infty}(Y^n, Z^n)$ is a solution to BSDE (1.1). From the previous steps, we deduce that for an arbitrary stopping time τ taking values in $[0,T]$ we have

$$Y_\tau = Y_0 - \int_0^\tau g(t, Y_t, Z_t)\,dt + \int_0^\tau Z_t\,dW_t.$$

Using the optional section theorem, we conclude that Y and the right-hand side in the formula above are indistinguishable, which means, in particular, that Y is a continuous process and thus (Y,Z) is indeed a solution to the BSDE with parameters (g, ξ_T). We have also shown that the solution (Y,Z) belongs to $\mathcal{S}_1^2(0,T) \times \mathcal{H}_d^2(0,T)$.

Step 6. It remains to show that (Y,Z) is the minimal solution. Let (Y', Z') be any solution to BSDE (1.1). Since the generator g_n is Lipschitz continuous and $g_n \le g$, we may apply Proposition 3.1 to deduce that $Y^n \le Y'$ for every n. By taking the limit, we obtain the desired inequality $Y \le Y'$. This completes the proof of the theorem. □

Some recent results on the uniqueness of solutions to BSDEs:

- Fang and Jiang [32] showed that if the generator g is continuous, monotonic and has a general growth in y, g is uniformly continuous in z, and the process g^0 is square-integrable, then the one-dimensional BSDE with the generator g has a unique solution for any square-integrable terminal condition ξ_T.

- Jia [37] proved that if g is uniformly continuous in z, uniformly with respect to (t, ω) and independent of y, then the solution to the BSDE with generator g is unique.
- Ma et al. [58] established an existence and uniqueness result for L^p $(p > 1)$ solutions to one-dimensional BSDEs where the generator g is monotonic in y and uniformly continuous in z.

5 BSDE Approach to Financial Derivatives

Suppose that we are given a family $(X^u,\, u \in \mathcal{U})$ of stochastic processes where the dynamics of X^u depend on a *control process* $u = (u_t)_{t\in[0,T]}$ taking values in the set U.

Definition 5.1. In the *stochastic target problem*, we search for a control process $u^* \in \mathcal{U}$ such that $X_T^{u^*} = \bar{X}$ where the target random variable $\bar{X}$ is predetermined and $X_0^{u^*}$ attains the minimum in the set $\{X_0^u : u \in \mathcal{U} \text{ and } X_T^u = \bar{X}\}$ where $\mathcal{U}$ is the class of all *admissible* control processes.

Alternatively, we may search for a control process u^* such that $X_T^{u^*} \geq \bar{X}$ and $X_0^{u^*}$ is the minimum of the set $\{X_0^u : u \in \mathcal{U} \text{ and } X_T^u \geq \bar{X}\}$. Other related problems can also be studied; for instance, one may search for a control u such that $\mathbb{E}_{\mathbb{P}}(X_T^u) \geq m$ for some fixed m or $\mathbb{P}(X_T^u \geq \bar{X}) \geq \alpha$ for a predetermined level of $\alpha > 0$. The concept of *admissibility* of a control process depends on a particular problem under study.

5.1 Dynamics of the Wealth Process

We focus on a particular instance of a stochastic target problem, which arises in financial mathematics and is called the *replication* problem (also known as the *hedging* problem). More generally, one may study the so-called *superhedging* problem and show that the two problems have the same solution.

Definition 5.2. Let $C_T^i := (S_T^i - K)^+$ (resp., $P_T^i := (K - S_T^i)^+$) be the payoff at T of the *call* (resp., *put*) *option* on the risky asset $(S_t^i)_{t\in[0,T]}$ where $i = 1, 2, \ldots, d$ and the strike K is an arbitrary positive number. More generally, let $\bar{X} = h(S_T^1, \ldots, S_T^d)$ be a (path-independent) *contingent claim* with maturity T where $h : \mathbb{R}^d \to \mathbb{R}_+$ is a non-negative payoff function.

Of course, in practice the payoff $\bar{X}$ is not necessarily nonnegative but that assumption is made here for mathematical convenience and hence it can be weakened. In the *replication problem*, we search for a portfolio π such that the *wealth process* (see Definition 5.4) satisfies $V_T^\pi = \bar{X}$ has the minimal initial wealth V_0^π among all portfolios π satisfying that equality. Of course, we need to describe the problem more precisely by analyzing first the dynamics of the wealth process for a trading portfolio and imposing the condition of admissibility of a portfolio.

Recall that according to the classical Black-Scholes model the stock price process is governed by the stochastic differential equation of the form

$$dS_t = S_t\big(\mu\, dt + \sigma\, dW_t\big), \quad S_0^i > 0, \tag{5.1}$$

for constant parameters $\mu \in \mathbb{R}, \sigma > 0$ and a one-dimensional standard Brownian motion W. Hence

$$S_t = S_0 \exp\left(\mu t + \sigma W_t - \frac{1}{2}\sigma^2 t\right).$$

It can be easily extended to the case of several stocks and stochastic parameters through the following definition.

Definition 5.3. Let W be a d-dimensional standard Brownian motion on the underlying probability space $(\Omega, \mathbb{F}, \mathbb{P})$. The *extended Black-Scholes model* postulates that the price S^0 of a non-risky asset satisfies

$$dS_t^0 = r_t S_t^0 \, dt, \quad S_0^0 = 1, \tag{5.2}$$

and the prices $S^1, \dots, S^d$ of risky assets are given as unique solution to SDEs

$$dS_t^i = S_t^i \big(\mu_t^i \, dt + \sigma_t^i \, dW_t \big), \quad S_0^i > 0, \tag{5.3}$$

where:
(i) the $\mathbb{F}$-adapted real-valued processes r, μ^i, and the $\mathbb{R}^d$-valued process σ^i are bounded;
(ii) the d-dimensional matrix $\sigma_t = [\sigma_t^{ij}]$ is invertible for every $t \in [0, T]$, so that the process σ^{-1} is well defined.

For any $i = 0, 1, \dots, d$, let φ_t^i stand for the number of shares of the ith asset held at time t. Then $\pi_t^i := \varphi_t^i S_t^i$ is the cash value of our investment in the ith asset at time t. We may thus identify a *portfolio* with a $(d+1)$-dimensional, $\mathbb{F}$-progressively measurable process $\pi = (\pi^0, \pi^1, \dots, \pi^d)$.

Definition 5.4. The *wealth process* V^π of π is given as $V_t^\pi = \sum_{i=0}^d \pi_t^i$ for every $t \in [0, T]$. We say that a portfolio π is *admissible* if π is *self-financing*, meaning that for all $t \in [0, T]$

$$V_t^\pi = V_0^\pi + \sum_{i=0}^d \int_0^t \varphi_s^i \, dS_s^i = V_0^\pi + \int_0^t \pi_s^0 r_s \, ds + \sum_{i=1}^d \int_0^t \pi_s^i \big(\mu_s^i \, ds + \sigma_s^i \, dW_s \big)$$

and the wealth process is non-negative, that is, $V_t^\pi \geq 0$ for all $t \in [0, T]$. We denote by $\mathcal{A}$ the class of all admissible portfolios.

The next lemma is an immediate consequence of Definition 5.4, equations (8.6)–(8.7) and the Itô formula.

Lemma 5.1. *For any admissible portfolio π, the wealth process V^π satisfies*

$$dV_t^\pi = r_t V_t^\pi \, dt + \pi_t (\mu_t - r_t \mathbb{1}) \, dt + \pi_t \sigma_t \, dW_t$$

where $\mathbb{1} = (1, 1, \dots, 1) \in \mathbb{R}^d$. More explicitly,

$$V_t^\pi = V_0^\pi + \int_0^t r_s V_s^\pi \, ds + \sum_{i=1}^d \int_0^t \pi_s^i (\mu_s^i - r_s) \, ds + \sum_{i=1}^d \sum_{j=1}^d \int_0^t \pi_s^i \sigma_s^{ij} \, dW_s^j. \tag{5.4}$$

The discounted wealth process $\widetilde{V}^\pi := (S^0)^{-1} V^\pi$ satisfies

$$d\widetilde{V}_t^\pi = \pi_t (\mu_t - r_t \mathbb{1}) \, dt + \pi_t \sigma_t \, dW_t,$$

that is,

$$\widetilde{V}_t^\pi = \widetilde{V}_0^\pi + \sum_{i=1}^d \int_0^t \pi_s^i (\mu_s^i - r_s) \, ds + \sum_{i=1}^d \sum_{j=1}^d \int_0^t \pi_s^i \sigma_s^{ij} \, dW_s^j. \tag{5.5}$$

It is important to observe that π^0 does not appear in dynamics (5.4) and (5.5). Consequently, we may formulate the following lemma.

Lemma 5.2. *A portfolio process π is admissible if the process $(\pi^1, \ldots, \pi^d)$ is $\mathbb{F}$-progressively measurable, the integrals in* (5.4) *are well defined and the wealth process V^π (or, equivalently, the process $\widetilde{V}^\pi$) is non-negative.*

From now on, we consider *admissible portfolios* $\pi = (\pi^1, \ldots, \pi^d)$. We are in a position to formulate a particular version of the stochastic target problem, which is called the *replication problem* or the *hedging problem*. Note that here an $\mathbb{R}^d$-valued process $u = \pi$ is a control process and $\mathcal{U} = \mathcal{A}$ is the set of all admissible control processes. Finally, ξ_T is the payoff at time T of a generic contingent claim (financial derivative).

5.2 Pricing and Superhedging via BSDE

Let ξ_T be an arbitrary $\mathcal{F}_T$-measurable random variable such that $\mathbb{E}_{\mathbb{P}}(\xi_T^2) < \infty$. We wish to find a pair (V^{π^*}, π^*) where π^* is an admissible portfolio such that $V_T^{\pi^*} = \xi_T$ and $V_0^{\pi^*} \leq V_0^\pi$ for any admissible portfolio π such that $V_T^\pi = \xi_T$.

It is clear that our goal is to solve the following minimisation problem

$$V_0^{\pi^*} = \min_{\pi \in \mathcal{A}} \{ V_0^\pi : V_T^\pi = \xi_T \}.$$

In fact, instead of searching for an admissible portfolio π^* such that $V_0^{\pi^*} \leq V_0^\pi$, we will show that the inequality $V_t^{\pi^*} \leq V_t^\pi$ holds for every $t \in [0, T]$.

The following result shows that in the extended Black-Scholes model a solution to the replication problem for a nonnegative random variable from the class $L_1^2(\mathcal{F}_T)$ (a nonnegative *contingent claim*) can be solved using a BSDE approach. It is worth noting that it can be easily extended to a market model with frictions by first identifying nonlinear dynamics of the wealth process and solving an appropriate nonlinear BSDE.

Theorem 5.1 (El Karoui et al. [30]). *Assume that the process γ given by*

$$\gamma_t := \sigma_t^{-1}(\mu_t - r_t \mathbb{1}) \tag{5.6}$$

is bounded. Let a random variable $\xi_T \geq 0$ belong to $L_1^2(\mathcal{F}_T)$. Then $(V^{\pi^*}, \pi^*) = (\bar{Y}, \bar{Z}\sigma^{-1})$ where the pair $(\bar{Y}, \bar{Z})$ is the unique solution in the space $\mathcal{S}_1^2(0,T) \times \mathcal{H}_d^2(0,T)$ of the linear BSDE

$$dY_t = (r_t Y_t \, dt + Z_t \gamma_t) \, dt + Z_t \, dW_t, \quad Y_T = \xi_T. \tag{5.7}$$

Proof. **Step 1.** In the first part of the proof, we search for a candidate for an optimal admissible replicating portfolio π^*. Upon denoting $Z = \pi\sigma$, we see that the linear BSDE (5.7) is equivalent to the following equation

$$dY_t = (r_t Y_t + \pi_t \sigma_t \gamma_t) \, dt + \pi_t \sigma_t \, dW_t, \quad Y_T = \xi_T, \tag{5.8}$$

which is consistent with dynamics (5.4) of the wealth process V^π. Since the processes r and γ are assumed to be bounded, we deduce from Theorem 2.1 that the linear BSDE (5.7) has a unique solution $(\bar{Y}, \bar{Z})$ in the space $\mathcal{S}_1^2(0,T) \times \mathcal{H}_d^2(0,T)$ and the process $\bar{Y}$ equals

$$\bar{Y}_t = \mathbb{E}_{\mathbb{P}}\big(\Gamma_{t,T} \xi_T \,|\, \mathcal{F}_t\big) \tag{5.9}$$

where $\Gamma_{t,T}$ is given by

$$\Gamma_{t,T} = \exp\left(-\int_t^T r_u\,du - \int_t^T \gamma_u\,dW_u - \frac{1}{2}\int_t^T \|\gamma_u\|^2\,du\right).$$

We claim that $\pi^* := \bar{Z}\sigma^{-1}$ solves the replication problem. We note that $V^{\pi^*} = \bar{Y}$ and thus π^* is an admissible portfolio, since Y is non-negative (this is an immediate consequence of (5.9) and the assumption that $\xi_T \geq 0$). We conclude that in the space $\mathcal{S}_1^2(0,T)\times\mathcal{H}_d^2(0,T)$ there exists a unique pair $(\bar{Y},\bar{Z})$ such that $\pi^* := \bar{Z}\sigma^{-1}$ is an admissible replicating portfolio and $\bar{Y}$ is equal to the wealth process V^{π^*}.

Step 2. In the second part of the proof, we aim to show that $V_0^\pi \geq \bar{Y}_0$ for any admissible portfolio π such that $V_T^\pi = \xi_T$. In fact, we will prove that the inequality $V_t^\pi \geq \bar{Y}_t$ is satisfied for every $t \in [0,T]$ if $V_T^\pi = \xi_T$.

To this end, it suffices to show that:
(a) for the portfolio π^*, the process $(\Gamma_{0,t}\bar{Y}_t)_{t\in[0,T]}$ is a uniformly integrable martingale,
(b) for any admissible portfolio π, the process $(\Gamma_{0,t}V_t^\pi)_{t\in[0,T]}$ is a non-negative local martingale and thus, in view of Lemma 12.2, it is also a supermartingale.

For statement (a), we note that the process $\Gamma_{0,\cdot}\bar{Y}$ coincides with the process M given by equation (2.7) with $\varphi = 0$. The fact that M is uniformly integrable martingale was established in the proof of Theorem 2.1.

For statement (b), we observe that $d\Gamma_{0,t} = \Gamma_{0,t}(-r_t\,dt - \gamma_t\,dW_t)$ and (see Lemma 5.1)

$$dV_t^\pi = r_tV_t^\pi\,dt + \pi_t(\mu_t - r_t\mathbb{1})\,dt + \pi_t\sigma_t\,dW_t.$$

An application of the Itô integration by parts formula gives

$$\begin{aligned}
d(\Gamma_{0,t}V_t^\pi) &= \Gamma_{0,t}\,dV_t^\pi + V_t^\pi\,d\Gamma_{0,t} + \langle\Gamma_{0,\cdot},V^\pi\rangle_t \\
&= \Gamma_{0,t}\Big(r_tV_t^\pi\,dt + \pi_t(\mu_t - r_t\mathbb{1})\,dt + \pi_t\sigma_t\,dW_t\Big) \\
&\quad + V_t^\pi\Gamma_{0,t}(-r_t\,dt - \gamma_t\,dW_t) - \Gamma_{0,t}\gamma_t\pi_t\sigma_t\,dt \\
&= \Gamma_{0,t}(\pi_t\sigma_t - V_t^\pi\gamma_t)\,dW_t
\end{aligned}$$

where we used the equality $\gamma = \sigma^{-1}(\mu - r\mathbb{1})$ (see (5.6)). It is thus clear that the process $(\Gamma_{0,t}V_t^\pi)_{t\in[0,T]}$ is a local martingale and its non-negativity is the consequence of the postulated admissibility of π. We conclude that for any admissible portfolio π replicating ξ_T, the process $\Gamma_{0,\cdot}V^\pi$ is a supermartingale

$$\mathbb{E}_{\mathbb{P}}(\Gamma_{0,T}\xi_T\,|\,\mathcal{F}_t) = \mathbb{E}_{\mathbb{P}}(\Gamma_{0,T}V_T^\pi\,|\,\mathcal{F}_t) \leq \Gamma_{0,t}V_t^\pi$$

whereas for π^* the process $\Gamma_{0,\cdot}V^{\pi^*}$ is a martingale (see (5.9) or (2.8) with $\varphi = 0$)

$$\mathbb{E}_{\mathbb{P}}(\Gamma_{0,T}\xi_T\,|\,\mathcal{F}_t) = \mathbb{E}_{\mathbb{P}}(\Gamma_{0,T}V_T^{\pi^*}\,|\,\mathcal{F}_t) = \Gamma_{0,t}V_t^{\pi^*}.$$

It is now clear that the inequality $V_t^{\pi^*} \leq V_t^\pi$ is satisfied for all $t \in [0,T]$ and every $\pi \in \mathcal{A}$ such that $V_T^\pi = \xi_T$. We conclude that π^* is the replicating strategy for ξ_T with the minimal wealth process among all admissible replicating strategies for ξ_T. □

Remark 5.1. From the comparison property of solutions to linear BSDEs (see Corollary 3.1), it follows that π^* is also a solution to the *super-hedging* problem, meaning that

$$V_0^{\pi^*} = \min_{\pi\in\mathcal{A}}\{\,V_0^\pi : V_T^\pi \geq \xi_T\}.$$

5.3 Pricing via Equivalent Local Martingale Measure

Since the extended Black-Scholes model is inherently linear, it is also possible to use an alternative approach to the target problem, which is based on the concept of an equivalent martingale measure. Of course, the method of an equivalent change of a probability measure is widely used in financial mathematics to deal with pricing of financial derivatives and portfolio optimisation problem.

From (5.9) and the flow property of $\Gamma_{t,s}$, we deduce that

$$Y_t = V_t^{\pi^*} = \mathbb{E}_{\mathbb{P}}\big(\Gamma_{t,T}\xi_T \,|\, \mathcal{F}_t\big) = (\Gamma_{0,t})^{-1}\mathbb{E}_{\mathbb{P}}\big(\Gamma_{0,T}\xi_T \,|\, \mathcal{F}_t\big).$$

This representation for Y can be simplified by an equivalent change of a probability measure. One may check that

$$Y_t = V_t^{\pi^*} = S_t^0\,\mathbb{E}_{\widetilde{\mathbb{P}}}\big((S_T^0)^{-1}\xi_T \,|\, \mathcal{F}_t\big)$$

where

$$S_t^0 = \exp\Big(\int_0^t r_s\,ds\Big)$$

and $\widetilde{\mathbb{P}}$ is a probability measure equivalent to $\mathbb{P}$ on $(\Omega, \mathcal{F}_T)$ and such that the process $\widetilde{W}_t := W_t + \int_0^t \gamma_u\,du$, $t \in [0,T]$ is a d-dimensional standard Brownian motion under $\widetilde{\mathbb{P}}$.

From the Girsanov theorem (see Theorem 12.4), it is known that the probability measure $\widetilde{\mathbb{P}}$ on $(\Omega, \mathcal{F}_T)$ is given by the Radon-Nikodým density

$$\frac{d\widetilde{\mathbb{P}}}{d\mathbb{P}} = \mathcal{E}_T\Big(-\int_0^{\cdot}\gamma_t\,dW_t\Big) = \exp\left(-\int_0^T \gamma_t\,dW_t - \frac{1}{2}\int_0^T \|\gamma_t\|^2\,dt\right), \quad \mathbb{P}-\text{a.s.}$$

where $\mathcal{E}$ denotes the *Doléans–Dade exponential* (see Definition 12.4). The probability measure $\widetilde{\mathbb{P}}$ is called the *equivalent local martingale measure* since the discounted prices $\widetilde{S}^i := (S^0)^{-1}S^i$, $i = 1, 2, \ldots, d$ are local martingales under $\widetilde{\mathbb{P}}$ and, for any self-financing portfolio π, the discounted wealth process $\widetilde{V}^\pi := (S^0)^{-1}V^\pi$ is a local martingale under $\widetilde{\mathbb{P}}$.

6 Optimal Stopping Problem and Reflected BSDEs

The goal of this section is to study reflected BSDEs (RBSDEs) and their relationship to either the classical (linear) or nonlinear optimal stopping problem.

6.1 Optimal Stopping Problem

We first give a very brief survey of results on the classical optimal stopping problem. Let W be a d-dimensional Brownian motion on $(\Omega, \mathcal{F}, \mathbb{P})$ and let $\mathbb{F}$ be the filtration generated by W. Let X be a real-valued, càdlàg, $\mathbb{F}$-adapted (or $\mathbb{F}$-optional) stochastic process, which is interpreted as the *reward process*.

Definition 6.1. We consider the following *optimal stopping problem*

$$v_0(X) = \sup_{\tau \in \mathcal{T}_{[0,T]}} \mathbb{E}\big(X_\tau\big) \tag{6.1}$$

where $\mathcal{T}_{[0,T]}$ is the set of all $\mathbb{F}$-stopping times with values in $[0,T]$. The goal is to find the *value* $v_0(X)$ and a *maximiser*, that is, a stopping time $\widehat{\tau} \in \mathcal{T}_{[0,T]}$ satisfying

$$\mathbb{E}\big(X_{\widehat{\tau}}\big) = \sup_{\tau \in \mathcal{T}_{[0,T]}} \mathbb{E}\big(X_\tau\big). \tag{6.2}$$

Notice that the unique value $v_0(X)$ is always well defined provided that the process X satisfies appropriate integrability conditions (or is simply bounded) but the existence of a maximizer is not obvious when X is not necessarily a continuous process. Furthermore, the uniqueness of a maximizer may fail to hold, in general. Finally, one may introduce the concept of an ε-optimal stopping time, which always exists but is not necessarily unique.

6.2 Snell Envelope

A standard approach to the classical optimal stopping problem hinges on the concept of the *Snell envelope*, which was introduced by Snell [87]. For a given real-valued stochastic process X, we define the Snell envelope $\widehat{X}$ of X as the smallest supermartingale of class (D) that dominates X, in the sense that $\widehat{X}_t \geq X_t$ for all $t \in [0,T]$. We say that a process $(X_t)_{t\in[0,T]}$ is of *class* (D) if the family of random variables $\{X_\tau : \tau \in \mathcal{T}_{[0,T]}\}$ is uniformly integrable.

Definition 6.2. We say that a process $\widehat{X}$ is the *Snell envelope* of X if:
(i) $\widehat{X}$ dominates X, that is, $\widehat{X}_t \geq X_t$ for all $t \in [0,T]$;
(ii) $\widehat{X}$ is a supermartingale of class (D);
(iii) $\widehat{X}$ is minimal, in the sense that if $\widetilde{X}$ is any supermartingale of class (D) that dominates X, then $\widetilde{X}$ dominates $\widehat{X}$.

The uniqueness of the Snell envelope is an immediate consequence of Definition 6.2. We henceforth assume that X satisfies the following condition

$$\mathbb{E}\Big(\sup_{t\in[0,T]} (X_t^+)^2\Big) < \infty, \tag{6.3}$$

that is, X^+ belongs to $\mathcal{S}^2_1(0,T)$. It is known that if condition (6.3) is satisfied, then the Snell envelope is well defined. Furthermore, it is closely related to a dynamic version of the optimal stopping problem, as the following classical result shows. Notice that the continuity of X is not postulated in Lemma 6.1 but if X is a continuous process, then the Snell envelope $\widehat{X}$ inherits that property.

Lemma 6.1. *If an $\mathbb{F}$-adapted, càdlàg process X satisfies condition* (6.3)*, then the Snell envelope $\widehat{X}$ belongs to $\mathcal{S}^2_1(0,T)$ and satisfies, for every $t \in [0,T]$,*

$$\widehat{X}_t = \operatorname*{ess\,sup}_{\tau \in \mathcal{T}_{[t,T]}} \mathbb{E}(X_\tau \,|\, \mathcal{F}_t) =: v_t(X) \tag{6.4}$$

where $v_t(X)$ is the value of the optimal stopping problem on $[t,T]$ for any fixed $t \in [0,T]$.

The random variable $\widehat{X}_t$ is equal to the essential supremum of the $\mathcal{F}_t$-conditional expectation of X_τ over all choices of a stopping time $\tau \in \mathcal{T}_{[t,T]}$. It is thus clear that it represents the value of the optimal stopping problem for the conditional expectation of X with optional stopping on the interval $[t,T]$. Note that when dealing with the stopping problem on $[t,T]$, we implicitly assume that the process X has not been stopped before t. The following lemma is an immediate consequence of the *Doob-Meyer decomposition* of a (resp., continuous) supermartingale of class (D).

Lemma 6.2. *There exists a unique $\mathbb{F}$-predictable (resp., $\mathbb{F}$-adapted and continuous), non-decreasing process K with $K_0 = 0$ such that $\mathbb{E}(K_T^2) < \infty$ and $\widehat{X} = \widehat{X}_0 + M - K$ where M with $M_0 = 0$ is a (resp., continuous) uniformly integrable martingale belonging to $\mathcal{S}^2_1(0,T)$ (resp., $\mathcal{S}^{2,c}_1(0,T)$).*

6.3 Optimal Stopping Time

We now turn to the problem of finding a maximiser $\widehat{\tau}$, which is also called an *optimal stopping time.* Recall that for any process X and any stopping time τ the *stopped process* X^τ is defined as $X^\tau_t = X_{t\wedge\tau}$ for every $t \in [0,T]$. We first give a general characterization of all optimal stopping times.

Lemma 6.3. *A stopping time $\widehat{\tau} \in \mathcal{T}_{[0,T]}$ is a maximiser for the optimal stopping problem* (6.1) *if and only if the stopped process $\widehat{X}^{\widehat{\tau}}$ is an $\mathbb{F}$-martingale and the equality $\widehat{X}_{\widehat{\tau}} = X_{\widehat{\tau}}$ holds.*

Proof. If $\widehat{X}^{\widehat{\tau}}$ is an $\mathbb{F}$-martingale and $\widehat{X}_{\widehat{\tau}} = X_{\widehat{\tau}}$, then using (6.4) with $t=0$ we obtain

$$\sup_{\tau \in \mathcal{T}_{[0,T]}} \mathbb{E}(X_\tau) = v_0(X) = \widehat{X}_0 = \mathbb{E}(\widehat{X}_{\widehat{\tau}}) = \mathbb{E}(X_{\widehat{\tau}}) \tag{6.5}$$

and thus $\widehat{\tau}$ is a (not necessarily unique) optimal stopping time for the problem (6.1).

Conversely, if $\widehat{\tau}$ is an optimal stopping time, then $v_0(X) = \widehat{X}_0 = \mathbb{E}(X_{\widehat{\tau}})$. Since $\widehat{X}_{\widehat{\tau}} \geq X_{\widehat{\tau}}$ and $\widehat{X}$ is an $\mathbb{F}$-supermartingale, we obtain $\widehat{X}_0 \geq \mathbb{E}(\widehat{X}_{\widehat{\tau}}) \geq \mathbb{E}(X_{\widehat{\tau}}) = \widehat{X}_0$ and thus the equalities $\widehat{X}_{\widehat{\tau}} = X_{\widehat{\tau}}$ and $\widehat{X}_0 = \mathbb{E}(\widehat{X}_{\widehat{\tau}})$ hold. The latter equality implies that $\widehat{X}^{\widehat{\tau}}$ is an $\mathbb{F}$-martingale by the optional sampling theorem. $\square$

Assume now that X is a continuous process and thus also the Snell envelope $\widehat{X}$ and the processes M and K from its Doob-Meyer decomposition are continuous. Then it is known that the process K increases only on the set $\{t \in [0,T] : \widehat{X}_t = X_t\}$, which means that, for every $t \in [0,T]$,

$$K_t = \int_0^t \mathbb{1}_{\{\widehat{X}_s = X_s\}} \, dK_s. \tag{6.6}$$

Since K is a non-decreasing process, equality (6.6) is equivalent to either of the following conditions

$$\int_0^T \mathbb{1}_{\{\widehat{X}_s = X_s\}} \, dK_s = 0, \tag{6.7}$$

as well as to

$$\int_0^T (\widehat{X}_s - X_s) \, dK_s = 0. \tag{6.8}$$

Any of the above conditions can be used as the *minimality condition* for K (also known as the *Skorokhod condition*).

In fact, if a continuous supermartingale $\widetilde{X}$ dominates X and the process K from its Doob-Meyer decomposition satisfies the Skorokhod condition, then $\widetilde{X}$ is the Snell envelope of X, that is, the equality $\widetilde{X} = \widehat{X}$ is satisfied. That property explains why the Skorokhod condition is also called the minimality condition. Notice that for a càdlàg (or, more generally, làdlàg, i.e., RLLL process) reward process we need to impose more minimality conditions to effectively handle the jumps of X and $\widehat{X}$ (see Section 6.10 for more details).

The next result is an easy consequence of Lemma 6.3.

Lemma 6.4. *Assume that X (and thus also $\widehat{X}$) is a continuous processes such that X^+ belongs to $\mathcal{S}_1^{2,c}(0,T)$. Let stopping time $\tau^* = \tau_0^*$ be given by*

$$\tau^* := \inf \big\{ t \in [0,T] : \widehat{X}_t = X_t \big\}.$$

Then stopped process $\widehat{X}^{\tau^*}$ is an $\mathbb{F}$-martingale and thus the equality $\widehat{X}^{\tau^*} = \widehat{X}_0 + M^{\tau^*}$ holds or, equivalently, $K_{\tau^*} = 0$. Moreover, τ^* is the minimal optimal stopping time. More generally, for the conditional optimal stopping problem on $[t,T]$, the minimal optimal stopping time τ_t^* is given by

$$\tau_t^* := \inf \big\{ s \in [t,T] : \widehat{X}_s = X_s \big\}$$

and the equality $K_{\tau_t^*} = K_t$ holds for every $t \in [0,T]$.

If X is not assumed to be continuous but $K_{\tau^*} = 0$, then τ^* is the minimal optimal stopping time (once again it suffice to apply Lemma 6.3). Furthermore, the latest moment when it is optimal to stop is given by

$$\bar{\tau} := \inf \big\{ t \in [0,T] : K_t > 0 \big\}$$

where $\inf \emptyset = T$ so that all optimal stoping times satisfy $\tau^* \leq \widehat{\tau} \leq \bar{\tau}$.

Finally, since the equality $\widehat{X}_T = X_T$ holds, we obtain the following equality, for all $t \in [0,T]$

$$\widehat{X}_t = X_T + (K_T - K_t) - (M_T - M_t),$$

which leads to the *reflected backward equation*

$$Y_t = X_T + (K_T - K_t) - (M_T - M_t)$$

where we search for a triplet (Y, M, K) such that $Y \geq X$, the process M is a continuous, $\mathbb{F}$-martingale, and the $\mathbb{F}$-adapted, non-decreasing, continuous process K satisfies the Skorokhod condition. This can be seen as a special case of the reflected BSDE (6.19) with $g = 0$.

6.4 Reflected BSDE

The notion of the Snell envelope can be extended to a nonlinear setup by introducing the concept of a reflected BSDE. Let W be a d-dimensional Brownian motion on $(\Omega, \mathcal{F}, \mathbb{P})$ and let $\mathbb{F}$ be the filtration generated by W. We consider the *reflected BSDE* with data (g, ξ_T, L) represented as

$$\begin{cases} Y_t = \xi_T + \int_t^T g(s, Y_s, Z_s)\, ds + (K_T - K_t) - \int_t^T Z_s\, dW_s, \quad t \in [0, T], \\ Y_t \geq L_t, \end{cases} \tag{6.9}$$

where $\xi_T \in L^2(\mathcal{F}_T)$ is such that $\xi_T \geq L_T$ and L is an $\mathbb{F}$-adapted, continuous process such that L^+ belongs to $\mathcal{S}_1^{2,c}(0, T)$.

Definition 6.3. Let (Y, K) be a pair of $\mathbb{F}$-adapted, continuous processes and let Z be $\mathbb{F}$-progressively measurable. Then (Y, Z, K) is a *solution* to the reflected BSDE with data (g, ξ_T, L) if $(Y, Z, K) \in \mathcal{H}_1^2(0, T) \times \mathcal{H}_d^2(0, T) \times \mathcal{A}_1^2(0, T)$ and:
(i) the following equality is satisfied for all $t \in [0, T]$

$$Y_t = \xi_T + \int_t^T g(s, Y_s, Z_s)\, ds + (K_T - K_t) - \int_t^T Z_s\, dW_s,$$

(ii) the inequality $Y_t \geq L_t$ holds for all $t \in [0, T]$;
(iii) the $\mathbb{F}$-adapted, continuous process K is non-decreasing, with $K_0 = 0$ and it can increase only on the random set $\{t \in [0, T] : Y_t = L_t\}$, so that for all $t \in [0, T]$

$$K_t = \int_0^t \mathbb{1}_{\{Y_s = L_s\}}\, dK_s. \tag{6.10}$$

Since $Y \geq L$ and K is non-decreasing, equality (6.10) is equivalent to the following condition

$$\int_0^T \mathbb{1}_{\{Y_s > L_s\}}\, dK_s = 0, \tag{6.11}$$

as well as to

$$\int_0^T (Y_s - L_s)\, dK_s = 0. \tag{6.12}$$

The proof of Proposition 6.1 can be found in El Karoui et al. [27].

Proposition 6.1. *Under assumptions (H.1) if $L^+ \in \mathcal{S}_1^{2,c}(0, T)$, then the reflected BSDE* (6.19) *has at most one solution and $(Y, Z, K) \in \mathcal{S}_1^{2,c}(0, T) \times \mathcal{H}_d^2(0, T) \times \mathcal{A}_1^{2,c}(0, T)$*

6.5 Optimal Stopping via Reflected BSDE

We first examine the relationship between a reflected BSDE and a particular linear optimal stopping problem.

Theorem 6.1. *Assume that a triplet* $(Y,Z,K) \in \mathcal{S}_1^{2,c}(0,T) \times \mathcal{H}_d^2(0,T) \times \mathcal{A}_1^{2,c}(0,T)$ *is a unique solution to the reflected BSDE* (6.19)*. Then*

$$Y_t = \operatorname*{ess\,sup}_{\tau \in \mathcal{T}_{[t,T]}} \mathbb{E}\left(\int_t^\tau g(s,Y_s,Z_s)\,ds + L_\tau \mathbb{1}_{\{\tau<T\}} + \xi_T \mathbb{1}_{\{\tau=T\}} \,\Big|\, \mathcal{F}_t \right). \tag{6.13}$$

Proof. First, we observe that for every $\tau \in \mathcal{T}_{[t,T]}$ by taking the conditional expectation with respect to $\mathcal{F}_t$ we obtain

$$\begin{aligned} Y_t &= \mathbb{E}\left(\int_t^\tau g(s,Y_s,Z_s)\,ds + Y_\tau + K_\tau - K_t \,\Big|\, \mathcal{F}_t \right) \\ &\geq \mathbb{E}\left(\int_t^\tau g(s,Y_s,Z_s)\,ds + L_\tau \mathbb{1}_{\{\tau<T\}} + \xi_T \mathbb{1}_{\{\tau=T\}} \,\Big|\, \mathcal{F}_t \right) \end{aligned}$$

since $Y_\tau \geq L_\tau \mathbb{1}_{\{\tau<T\}} + \xi_T \mathbb{1}_{\{\tau=T\}}$ and $K_\tau - K_t \geq 0$. We conclude that $Y_t \geq v_t(X)$ for every $t \in [0,T]$.

For the converse inequality, we fix t and we define

$$\tau_t := \inf\{s \in [t,T] : Y_s = L_s\}$$

where $\inf \emptyset = T$. The continuity of K and condition (6.11) imply that $K_{\tau_t} = K_t$ and thus

$$Y_t = \mathbb{E}\left(\int_t^{\tau_t} g(s,Y_s,Z_s)\,ds + L_{\tau_t} \mathbb{1}_{\{\tau_t<T\}} + \xi_T \mathbb{1}_{\{\tau_t=T\}} \,\Big|\, \mathcal{F}_t \right),$$

which entails that $Y_t \leq v_t(X)$ for every $t \in [0,T]$. This shows that $Y_t = v_t(X)$ and thus the proof is completed. □

If we take a fixed process g in (6.19) and (6.13), then we obtain a relationship between the classical optimal stopping problem with reward

$$X_t = \int_0^t g_s\,ds + L_t \mathbb{1}_{\{t<T\}} + \xi_T \mathbb{1}_{\{t=T\}}$$

and the reflected BSDE with null generator

$$\begin{cases} Y_t = \xi_T + \int_t^T g_s\,ds + (K_T - K_t) - \int_t^T Z_s\,dW_s, & t \in [0,T], \\ Y_t \geq L_t. \end{cases}$$

Finally, if we set $g = 0$ in (6.19) and (6.13), then we obtain a reflected BSDE with null generator, that is,

$$\begin{cases} Y_t = \xi_T + (K_T - K_t) - \int_t^T Z_s\,dW_s, & t \in [0,T], \\ Y_t \geq L_t. \end{cases}$$

which corresponds to the optimal stopping problem with reward $X_t = L_t \mathbb{1}_{\{t<T\}} + \xi_T \mathbb{1}_{\{t=T\}}$.

6.6 Solution to a Reflected BSDE via Penalisation Method

Theorem 6.2. *Under assumptions (H.1) and $L^+ \in \mathcal{S}_1^{2,c}(0,T)$, the reflected BSDE* (6.19) *has a unique solution $(Y,Z,K) \in \mathcal{H}_1^2(0,T) \times \mathcal{H}_d^2(0,T) \times \mathcal{A}_1^2(0,T)$ and (Y,Z,K) belongs to the space $\mathcal{S}_1^{2,c}(0,T) \times \mathcal{H}_d^2(0,T) \times \mathcal{A}_1^{2,c}(0,T)$.*

Proof. To prove the existence of a solution, we use the *penalisation method*. To this end, we introduce the following sequence of BSDEs with parameters (g_n, ξ_T)

$$Y_t^n = \xi_T + \int_t^T g(s, Y_s^n, Z_s)\, du + \int_t^T n(L_s - Y_s^n)^+\, ds - \int_t^T Z_s\, dW_s. \tag{6.14}$$

Note that the generator g_n of BSDE (6.14) equals

$$g_n(\omega,t,y,z) = g(\omega,t,y,z) + n(L_t(\omega) - y)^+$$

and thus it satisfies condition (H.1). Therefore, by virtue of Theorem 1.1, BSDE (6.14) has a unique solution (Y^n, Z^n) in the space $\mathcal{S}_1^{2,c}(0,T) \times \mathcal{H}_d^2(0,T)$ for every natural n. For every n, we define the process K^n by

$$K_t^n := \int_0^t n(L_s - Y_s^n)^+\, ds, \quad \forall\, t \in [0,T],$$

so that K^n is manifestly an $\mathbb{F}$-adapted, continuous and non-decreasing process. To complete the proof, one may proceed as follows (for a detailed proof, we refer to El Karoui et al. [27, 29]).

Step 1. Show that the inequality $Y^n \le Y^{n+1}$ is satisfied for every n.

Step 2. Show that the following convergence is valid

$$\lim_{n\to\infty} \mathbb{E}\big(\sup_{t\in[0,T]} [(L_t - Y_t^n)^+]^2 \big) = 0.$$

Step 3. Show that the following convergence holds

$$\lim_{n\to\infty} \mathbb{E}\big(\sup_{t\in[0,T]} [Y_t^n - Y_t]^2 \big) = 0.$$

Step 4. Let us denote $K_t^n := \int_0^t n(L_u - Y_u^n)^+\, du$. Show that there exists an $\mathbb{F}$-adapted, continuous, non-decreasing process $K \in \mathcal{A}_1^{2,c}(0,T)$ such that

$$\lim_{n\to\infty} \mathbb{E}\big(\sup_{t\in[0,T]} [K_t^n - K_t]^2 \big) = 0.$$

Step 5. Finally, show that the triplet $(Y,Z,K) := \lim_{n\to\infty}(Y^n, Z^n, K^n)$ is a solution to the reflected BSDE (6.19). □

6.7 Solution to a Reflected BSDE via Snell Envelope

Let W be a d-dimensional Brownian motion on $(\Omega, \mathcal{F}, \mathbb{P})$ and let $\mathbb{F}$ be the filtration generated by W. We consider the optimal stopping problem introduced in Definition 6.1. The next result shows that the existence of a solution to reflected BSDE (6.19) can be deduced from results on the classical optimal stopping problem. Hence Theorem 6.3 can be seen as a converse to Theorem 6.1 where the value of the optimal stopping problem was obtained from a solution to a reflected BSDE. Recall that the uniqueness of a solution to reflected BSDE (6.19) with a Lipschitz continuous generator holds in view of Proposition 6.1 and thus it suffices to focus on the existence of a solution.

Theorem 6.3. *(i) Let the reward process X in the optimal stopping problem from Definition 6.1 be given by*

$$X_t := L_t \mathbb{1}_{\{t<T\}} + \xi_T \mathbb{1}_{\{t=T\}} + \int_0^t g_s\, ds. \tag{6.15}$$

If $\xi_T \in L^2(\mathcal{F}_T)$ and $L^+ \in \mathcal{S}_1^{2,c}$, then X satisfies (6.3) *and thus the Snell envelope $\widehat{X}$ is well defined and belongs to the class (D).*
(ii) Let $\widehat{X} = \widehat{X}_0 + M - K$ be the Doob-Meyer decomposition of the Snell envelope $\widehat{X}$ of X and let the process Z from $\mathcal{H}_d^2(0,T)$ be such that

$$M_t = M_0 + \int_0^t Z_s\, dW_s, \quad \forall\, t \in [0,T], \tag{6.16}$$

and

$$Y_t := \widehat{X}_t - \int_0^t g_s\, ds, \quad \forall\, t \in [0,T]. \tag{6.17}$$

Then the triplet (Y,Z,K) is a unique solution to the reflected BSDE

$$\begin{cases} Y_t = \xi_T + \int_t^T g_s\, ds + (K_T - K_t) - \int_t^T Z_s\, dW_s, \quad t \in [0,T], \\ Y_t \geq L_t, \end{cases} \tag{6.18}$$

where $\xi_T \in L^2(\mathcal{F}_T)$ and g is a fixed generator from $\mathcal{H}_1^2(0,T)$.
(iii) If (H.1) is valid and $L^+ \in \mathcal{S}_1^{2,c}$, then the reflected BSDE

$$\begin{cases} Y_t = \xi_T + \int_t^T g(s,Y_s,Z_s)\, ds + (K_T - K_t) - \int_t^T Z_s\, dW_s, \quad t \in [0,T], \\ Y_t \geq L_t, \end{cases} \tag{6.19}$$

has a unique solution (Y,Z,K) in the space $\mathcal{S}_1^{2,c}(0,T) \times \mathcal{H}_d^2(0,T) \times \mathcal{A}_1^2(0,T)$.

Proof. **Step 1.** One can check that (i) is valid and thus the Snell envelope $\widehat{X}$ of X is well defined and is given by

$$\widehat{X}_t = \operatorname*{ess\,sup}_{\tau \in \mathcal{T}_{[t,T]}} \mathbb{E}\left(L_\tau \mathbb{1}_{\{\tau<T\}} + \xi_T \mathbb{1}_{\{\tau=T\}} + \int_0^\tau g_s\, ds \,\Big|\, \mathcal{F}_t \right).$$

Hence the process Y given by (6.17) satisfies

$$Y_t = \operatorname*{ess\,sup}_{\tau \in \mathcal{T}_{[t,T]}} \mathbb{E}\left(L_\tau \mathbb{1}_{\{\tau<T\}} + \xi_T \mathbb{1}_{\{\tau=T\}} + \int_t^\tau g_s\, ds \,\Big|\, \mathcal{F}_t \right),$$

so that $Y_T = \xi_T$. Furthermore, for every $t \in [0,T]$,

$$Y_t = \widehat{X}_t - \int_0^t g_s\, ds = \widehat{X}_0 + M_t - K_t - \int_0^t g_s\, ds$$

where $M_t = \int_0^t Z_u \, dW_u$ for some process $Z \in \mathcal{H}_d^2(0,T)$ and thus

$$\xi_T - Y_t = Y_T - Y_t = M_T - M_t - (K_T - K_t) - \int_t^T g_s \, ds,$$

which gives

$$Y_t = \xi_T + \int_t^T g_s \, ds + (K_T - K_t) - \int_t^T Z_s \, dW_s.$$

Note that the existence of Z is a consequence of the predictable representation property of the Brownian motion W. Moreover, it can be checked that the process Y given by (6.17) belongs to $\mathcal{S}_1^2(0,T)$.

We have thus shown that for every $t \in [0,T]$

$$\begin{aligned} Y_t &= \xi_T + \int_t^T g_s \, ds + (K_T - K_t) - \int_t^T Z_s \, dW_s \\ &= \operatorname*{ess\,sup}_{\tau \in \mathcal{T}_{[t,T]}} \mathbb{E}\left(L_\tau \mathbb{1}_{\{\tau < T\}} + \xi_T \mathbb{1}_{\{\tau = T\}} + \int_t^\tau g_s \, ds \,\Big|\, \mathcal{F}_t \right) \end{aligned}$$

where the second equality implies that $Y_t \geq L_t$ for all $t \in [0,T]$ (it suffices to take $\tau = t$).

To complete the proof of existence of a solution to the reflected BSDE (6.18), we need to show that (see (6.11))

$$\int_0^T \mathbb{1}_{\{Y_s > L_s\}} \, dK_s = 0. \tag{6.20}$$

Let us fix $t \in [0,T)$ and let us consider the event $\{Y_t > L_t\}$. From the martingale property of the Snell envelope stopped at τ_t^*, we obtain $K_{\tau_t^*} - K_t = 0$ on $\{Y_t > L_t\}$ (see Lemma 6.4) and, from the continuity of the processes Y and L, we deduce that the inequality $\tau_t^* > t$ holds on the event $\{Y_t > L_t\}$. This leads to equality (6.20), although that argument is not completely trivial. We need to consider all rational numbers $t \in [0,T)$ and use also the continuity of processes Y and L.

Step 2. In the second step, we consider the general reflected BSDE

$$\begin{cases} Y_t = \xi_T + \int_t^T g(s, Y_s, Z_s) \, ds + (K_T - K_t) - \int_t^T Z_s \, dW_s, & t \in [0,T], \\ Y_t \geq L_t. \end{cases} \tag{6.21}$$

The arguments used in that step are similar as in the proof of Theorem 1.1. Since the generator g is assumed to satisfy condition (H.1), it is possible to show that the mapping obtained from reflected BSDE (6.21) is a contraction in an appropriate Banach space of stochastic processes and thus a unique solution to reflected BSDE (6.21) exists. We thus obtain here an alternative proof of Theorem 6.1 as well as the convergence of Picard's sequence for a solution to the reflected BSDE. □

6.8 Comparison Theorem for Reflected BSDEs

Let us mention that one can also establish the comparison theorem for reflected BSDEs analogous to Theorem 3.1. For the proof of Theorem 6.4, we refer to El Karoui et al. [29].

Theorem 6.4. *Let (g^1, ξ^1_T, L^1) and (g^2, ξ^2_T, L^2) be standard parameters and let (Y^1, Z^1, K^1) and (Y^2, Z^2, K^2) be the unique solutions in $\mathcal{S}^2_1(0,T) \times \mathcal{H}^2_d(0,T) \times \mathcal{A}^2_1(0,T)$ to the associated reflected BSDEs for $i = 1, 2$*

$$-dY^i_t = g^i(t, Y^i_t, Z^i_t)\, dt + (K^i_T - K^i_t) - Z^i_t\, dW_t, \quad Y^i_T = \xi^i_T. \tag{6.22}$$

We assume that $L^1, L^2 \in \mathcal{A}^{2,c}_1(0,T)$ are such that $L^1_t \geq L^2_t$ for all $t \in [0,T]$ and $\xi^1_T \geq \xi^2_T$. Moreover

$$\delta_2 g_t := g^1(t, Y^2_t, Z^2_t) - g^2(t, Y^2_t, Z^2_t) \geq 0, \quad d\mathbb{P} \otimes dt - a.e.$$

(i) The inequality $Y^1_t \geq Y^2_t$ holds for all $t \in [0,T]$.
(ii) If the equality $L^1_t = L^2_t$ holds for all $t \in [0,T]$, then $K^1_t - K^1_s \geq K^2_t - K^2_s$ for every $0 \leq s \leq t \leq T$.

We will also need the following result.

Theorem 6.5. *Let (g, ξ_T, L) and (g, ξ'_T) be standard parameters and let $(Y, Z, K) \in \mathcal{S}^2_1(0,T) \times \mathcal{H}^2_d(0,T) \times \mathcal{A}^2_1(0,T)$ and $(Y', Z') \in \mathcal{S}^2_1(0,T) \times \mathcal{H}^2_d(0,T)$ be unique solutions*

$$-dY_t = g(t, Y_t, Z_t)\, dt + (K_T - K_t) - Z_t\, dW_t, \quad Y \geq L, \ Y_T = \xi_T,$$

and

$$-dY'_t = g(t, Y'_t, Z'_t)\, dt - Z'_t\, dW_t, \quad Y'_T = \xi'_T.$$

If $\xi_T \geq \xi'_T$, then the inequality $Y_t \geq Y'_t$ holds for all $t \in [0,T]$.

6.9 Nonlinear Optimal Stopping

Our next goal is to analyze an extension of the classical optimal stopping problem to a nonlinear setup where the linear operator $\mathbb{E}$ is replaced by a nonlinear g-evaluation denoted as $\mathcal{E}^g$, which was introduced in Section 1.9.

Definition 6.4. The *value* of the nonlinear optimal stopping problem with reward ζ is given by

$$v^g_0(\zeta) = \sup_{\tau \in \mathcal{T}_{[0,T]}} \mathcal{E}^g_{0,\tau}(\zeta_\tau) \tag{6.23}$$

and an $\mathbb{F}$-stopping time $\widehat{\tau} \in \mathcal{T}_{[0,T]}$ is called a *maximiser* if

$$v^g_0(\zeta) = \mathcal{E}^g_{0,\widehat{\tau}}(\zeta_{\widehat{\tau}}) = \max_{\tau \in \mathcal{T}_{[0,T]}} \mathcal{E}^g_{0,\tau}(\zeta_\tau).$$

In the next result, we consider a nonlinear optimal stopping problem based on $\mathcal{E}^g$ defined by the BSDE (1.1) with a Lipschitz continuous generator. Then it is known that the stability property (A.5) is satisfied by $\mathcal{E}^g$. Recall that the stability property of solutions to the BSDE (1.1) can be proven when g is a Lipschitz continuous generator as a consequence of inequality (1.19) in Lemma 1.3. Recall that we write $\zeta_t = L_t \mathbb{1}_{\{t<T\}} + \xi_T \mathbb{1}_{\{t=T\}}$.

Notice that Theorem 6.6 is very similar to results from Section 6.3 on the classical optimal stopping problem and, in fact, the classical case is obtained if we set $g = 0$. To stress similarities, we recall that a process Y from a solution (Y, Z) to the BSDE (1.1) is an *g-martingale* and a process Y from a solution (Y, Z, K) to the reflected BSDE (6.19) is an *g-supermartingale*.

Theorem 6.6. *Let $\zeta \in \mathcal{S}^2_1(0,T)$ and let the triplet (Y,Z,K) be a unique solution to the reflected BSDE* (6.19)*. If $\mathcal{E}^g$ has the strict comparison property, then the following assertions are valid:*
(i) Y_0 is the value of the nonlinear optimal stopping problem with the reward process ζ, that is, $Y_0 = v^g_0(\zeta)$;
(ii) an $\mathbb{F}$-stopping time $\widehat{\tau}$ is a maximiser if and only if the stopped process $Y^{\widehat{\tau}}$ is a g-martingale and $Y_{\widehat{\tau}} = \zeta_{\widehat{\tau}}$ (i.e., $Y = Y'$ on $[0,\widehat{\tau}]$ where (Y',Z') is a solution to the BSDE (1.1) *with $Y'_{\widehat{\tau}} = \zeta_{\widehat{\tau}}$);*
(iii) the $\mathbb{F}$-stopping time $\tau^* := \inf\{t \in [0,T] \,|\, Y_t = \zeta_t\}$ is a (earliest) maximiser provided that the equality $K_{\tau^*} = 0$ holds so that $Y^{\tau^*} = \zeta_{\tau^*}$.

Proof. (i) The inequality $Y_0 \geq v^g_0(\zeta)$ is a consequence of (6.23), the definition of the nonlinear evaluation $\mathcal{E}^g$ and the comparison property of from Theorem 6.5. We obtain the inequality $Y_0 \geq Y'_0 = \mathcal{E}^g_{0,\tau}(\zeta_\tau)$ for every $\tau \in \mathcal{T}_{[0,T]}$ and thus

$$Y_0 \geq \sup_{\tau \in \mathcal{T}_{[0,T]}} \mathcal{E}^g_{0,\tau}(\zeta_\tau) = v^g_0(\zeta).$$

For the converse inequality, we set for any fixed $\varepsilon > 0$

$$\tau_\varepsilon := \inf\{t \in [0,T] \,|\, Y_t \leq \zeta_t + \varepsilon\}.$$

Then, by noticing the right-continuity of Y and ζ, we have that $Y_{\tau_\varepsilon} \leq \zeta_{\tau_\varepsilon} + \varepsilon$ and $Y_t = \mathcal{E}^g_{t,\tau_\varepsilon}(Y_{\tau_\varepsilon})$ for $0 \leq t \leq \tau_\varepsilon$. The monotonicity property of $\mathcal{E}^g$ yields

$$Y_0 = \mathcal{E}^g_{0,\tau_\varepsilon}(Y_{\tau_\varepsilon}) \leq \mathcal{E}^g_{0,\tau_\varepsilon}(\zeta_{\tau_\varepsilon} + \varepsilon).$$

In view of the stability property of BSDE (1.1), there exists a positive constant C such that

$$\left|\mathcal{E}^g_{0,\tau_\varepsilon}(\zeta_{\tau_\varepsilon} + \varepsilon) - \mathcal{E}^g_{0,\tau_\varepsilon}(\zeta_{\tau_\varepsilon})\right| \leq C\varepsilon.$$

Consequently, $Y_0 \leq \mathcal{E}^g_{0,\tau_\varepsilon}(\zeta_{\tau_\varepsilon}) + C\varepsilon$, meaning that τ_ε is a $(C\varepsilon)$-optimal stopping time for the nonlinear optimal stopping problem (6.23). Since ε is arbitrary, it is now easy to conclude that $Y_0 \leq v^g_0(\zeta)$, which ends the proof of the equality $Y_0 = v^g_0(\zeta)$.

(ii) If $\widehat{X}^{\widehat{\tau}}$ is a g-martingale and $\widehat{X}_{\widehat{\tau}} = X_{\widehat{\tau}}$, then

$$\sup_{\tau \in \mathcal{T}_{[0,T]}} \mathcal{E}^g_{0,\widehat{\tau}}(\zeta_\tau) = v^g_0(\zeta) = Y_0 = \mathcal{E}^g_{0,\widehat{\tau}}(\widehat{X}_{\widehat{\tau}}) = \mathcal{E}^g_{0,\widehat{\tau}}(X_{\widehat{\tau}}) \tag{6.24}$$

and thus $\widehat{\tau}$ is an optimal stopping time for the problem (6.1).

Conversely, if $\widehat{\tau}$ is an optimal stopping time, then

$$v_0(X) = Y_0 = \mathcal{E}^g_{0,\widehat{\tau}}(\zeta_{\widehat{\tau}}).$$

Since $Y_{\widehat{\tau}} \geq \zeta_{\widehat{\tau}}$ and Y is a g-supermartingale, we obtain

$$Y_0 \geq \mathcal{E}^g_{0,\widehat{\tau}}(Y_{\widehat{\tau}}) \geq \mathcal{E}^g_{0,\widehat{\tau}}(\zeta_{\widehat{\tau}}) = Y_0$$

so that $Y_0 = \mathcal{E}^g_{0,\widehat{\tau}}(Y_{\widehat{\tau}})$ and, by the domination $Y_{\widehat{\tau}} \geq \zeta_{\widehat{\tau}}$ and the strict comparison property of $\mathcal{E}^g$, the equality $Y_{\widehat{\tau}} = \zeta_{\widehat{\tau}}$ holds. The former equality implies that $Y^{\widehat{\tau}}$ is a g-martingale by the nonlinear version of the classis optional sampling theorem.

(iii) It suffices to notice that, under the assumptions of part (iii), we immediately obtain the equality $Y_0 = \mathcal{E}^g_{0,\tau^*}(\zeta_{\tau^*})$ and thus, in view of part (i) in the theorem, we also have that $v_0(\zeta) = \mathcal{E}^g_{0,\tau^*}(\zeta_{\tau^*})$, which means that the stopping time

$$\tau^* := \inf\{t \in [0,T] \,|\, Y_t = \zeta_t\}$$

is indeed a maximiser for the nonlinear optimal stopping problem (6.23). □

Using Theorem 10 in Chapter VII of [19] (see also Theorem 2.6 in [83] or Proposition B.10 in [51] for a general case), it can be shown that the increasing process K in the solution (Y,Z,K) to the reflected BSDE (6.19) is continuous if the reward process ζ is left-upper-semicontinuous along stopping times.

In that case, the equality $K_{\tau^*} = 0$ is indeed satisfied and thus the nonlinear optimal stopping problem has at least one maximiser. Obviously, this is manifestly true under the postulate that the reward process ζ is continuous, which was our usual assumption when dealing with reflected BSDEs.

6.10 Optimal Stopping with Discontinuous Reward

If the reward process ζ is $\mathbb{F}$-optional but possibly discontinuous, then the Skorokhod condition should be modified. We denote by $\mathcal{K}$ (resp., $\overline{\mathcal{K}}$) the class of all càdlàg, nondecreasing, $\mathbb{F}$-predictable (resp., làdlàg, nondecreasing, $\mathbb{F}$-predictable) processes. The following lemma is elementary.

Lemma 6.5. *Any process $K \in \mathcal{K}$ with $K_0 = 0$ has a unique decomposition $K = K^c + K^d$ where $K^c_0 = K^d_0 = 0$, K^c is an $\mathbb{F}$-adapted, continuous, nondecreasing process and K^d is an $\mathbb{F}$-predictable, càdlàg, purely discontinuous, nondecreasing process. More generally, if K belongs to $\overline{\mathcal{K}}$ and $K_0 = 0$ then the decomposition becomes $K = K^c + K^d + K^g$ where K^g with $K^g_0 = 0$ is an $\mathbb{F}$-adapted, càglàd, purely discontinuous, nondecreasing process.*

If X and Y are arbitrary $\mathbb{F}$-optional processes, then the inequality $Y \geq X$ means that $Y_\tau \geq X_\tau$ for every $\tau \in \mathcal{T}_{[0,T]}$. We denote by $\mathcal{T}^p_{[0,T]}$ the class of all $\mathbb{F}$-predictable stopping times (but any $\mathbb{F}$-stopping time is an $\mathbb{F}$-predictable stopping time when $\mathbb{F}$ is the Brownian filtration).

Definition 6.5. The *left-upper-semicontinuous envelope* of an $\mathbb{F}$-optional process ζ is given by $\bar{\zeta}_t := \limsup_{s\to t,\, s<t} \zeta_s$ for every $t \in [0,T]$.

It can be checked that $\bar{\zeta}$ is an $\mathbb{F}$-predictable process, in general. Furthermore, if the left-hand limits of ζ exist (for instance, when ζ is a càdlàg process), then we have that $\bar{\zeta} = \zeta_-$, that is, the left-upper-semicontinuous envelope of ζ coincides with the process of left-hand limits of ζ (of course, ζ_- is a càglàd version of ζ if the latter is a càdlàg process).

Definition 6.6. A triplet $(Y,Z,K) \in \mathcal{H}^2_1(0,T) \times \mathcal{H}^2_d(0,T) \times \mathcal{H}^2_1(0,T)$ is a solution to the *reflected BSDE* with the lower obstacle ζ if $Y \geq \zeta$ and the following equality holds, for every $\tau \in \mathcal{T}_{[0,T]}$,

$$Y_\tau = \zeta_T + \int_\tau^T g(s, Y_s, Z_s)\, dW_s - \int_\tau^T Z_s\, dW_s + K_T - K_\tau \tag{6.25}$$

where $K \in \overline{\mathcal{K}}$ satisfies $K_0 = 0$ and the Skorokhod (minimality) conditions

$$
\begin{aligned}
&\int_0^T \mathbb{1}_{\{Y_{s-} > \bar{\zeta}_s\}} \, dK^c_s = 0, \\
&(Y_{\tau-} - \bar{\zeta}_\tau)\Delta K^d_\tau = 0, \quad \forall\, \tau \in \mathcal{T}^p_{[0,T]}, \\
&(Y_\tau - \zeta_\tau)\Delta^+ K^g_\tau = 0, \quad \forall\, \tau \in \mathcal{T}_{[0,T]},
\end{aligned}
$$

where $\Delta K^d_\tau := K^d_\tau - K^d_{\tau-}$ and $\Delta^+ K^g_\tau := K^g_{\tau+} - K^g_\tau$.

We refer to Kobylanski and Quenez [51] for a thorough study of the classical optimal stopping and to Grigorova et al. [33] and [35] for the case of nonlinear optimal stopping with irregular reward under a nonlinear evaluation $\mathcal{E}^g$. Let us only make some comments on Skorokhod conditions related to left and right jumps of ζ and K.

Let us first analyze condition

$$
(Y_\tau - \zeta_\tau)\Delta^+ K^g_\tau = 0, \quad \forall\, \tau \in \mathcal{T}_{[0,T]}.
$$

It is clear from (6.25) that $\Delta^+ Y_\tau = -\Delta^+ K^g_\tau$ for every $\tau \in \mathcal{T}_{[0,T]}$. The conditions $Y \geq \zeta$ and $(Y_\tau - \zeta_\tau)\Delta^+ K^g_\tau = 0$ for all $\tau \in \mathcal{T}_{[0,T]}$ are known to be equivalent to the equality $Y_\tau = \zeta_\tau \vee Y_{\tau+}$ for all $\tau \in \mathcal{T}_{[0,T]}$ and thus they are also equivalent to the equality $\Delta^+ K^g_\tau = (\zeta_\tau - Y_{\tau+})^+$ for all $\tau \in \mathcal{T}_{[0,T]}$. Intuitively, this means that the upward push represented by K^g is only needed on the event $\{\zeta_\tau > Y_{\tau+}\}$ where it is not possible to set $Y_\tau = Y_{\tau+}$. Since the (backward) upward push should be minimal, the postulated inequality $Y \geq \zeta$ becomes the equality $Y_\tau = \zeta_\tau$ on that event.

An analysis of condition

$$
(Y_{\tau-} - \bar{\zeta}_\tau)\Delta K^d_\tau = 0, \quad \forall\, \tau \in \mathcal{T}^p_{[0,T]}
$$

is less transparent since it involves the process ζ before the moment of jump. We note that $\Delta Y_\tau = \Delta M_\tau - \Delta K^d_\tau$ and thus for every $\tau \in \mathcal{T}^p_{[0,T]}$ we have that $\mathbb{E}(\Delta Y_\tau \,|\, \mathcal{F}_{\tau-}) = -\Delta K^d_\tau$ since $\mathbb{E}(\Delta M_\tau \,|\, \mathcal{F}_{\tau-}) = 0$ and $\mathbb{E}(\Delta K^d_\tau \,|\, \mathcal{F}_{\tau-}) = \Delta K^d_\tau$. The conditions $Y \geq \zeta$ (and hence $Y_- \geq \bar{\zeta}$) and $(Y_{\tau-} - \bar{\zeta}_\tau)\Delta K^d_\tau = 0$ for all $\tau \in \mathcal{T}^p_{[0,T]}$ are equivalent to the equality $\Delta Y_\tau = -(\bar{\zeta}_\tau - Y_{\tau-})^+$ on the event $\{\Delta K^d_\tau > 0\}$ for every $\tau \in \mathcal{T}^p_{[0,T]}$ and thus they are also equivalent to the equality $\Delta K^d_\tau = \mathbb{E}((\bar{\zeta}_\tau - Y_\tau)^+ \,|\, \mathcal{F}_{\tau-})$ on the event $\{\Delta K^d_\tau > 0\}$ for every $\tau \in \mathcal{T}^p_{[0,T]}$.

If the process ζ is left-upper-semicontinuous along stopping times, then the process K^d is continuous and if ζ is right-lower-semicontinuous along stopping times, then the process K^g is continuous.

7 Dynkin Games and Doubly Reflected BSDEs

We first give a brief summary of concepts and results for multi-person games and Dynkin games.

7.1 Nash and Optimal Equilibria

We summarise here the basic results for Nash and optimal equilibria in a multi-player game.

Let $\mathcal{M} = \{1, \dots, m\}$ be the set of players and let $S = S^1 \times \cdots \times S^m$ stand for the class of all strategy profiles $s = (s^1, \dots, s^m)$. For each $s \in S$, we denote by $V^k(s) = V^k(s^1, \dots, s^m)$ a (possibly random) payoff of the kth player. Every player aims to maximise his payoff. It is convenient to write $s = (s^k, s^{-k})$ and $\mathcal{M}^{-k} = \mathcal{M} \setminus \{k\}$.

Nash equilibrium. Let us first recall the concept of Nash equilibrium (see Nash [61]).

Definition 7.1. A strategy profile $\sigma = (\sigma^1, \sigma^2, \dots, \sigma^m) \in S$ is called a *Nash equilibrium* if the inequality $V^k(\sigma^k, \sigma^{-k}) \geq V^k(s^k, \sigma^{-k})$ holds for each k and all $s^k \in S^k$. In other words, for each $k \in \mathcal{M}$,

$$V^k(\sigma^k, \sigma^{-k}) = \operatorname*{ess\,sup}_{s^k \in S^k} V^k(s^k, \sigma^{-k}). \tag{7.1}$$

Optimal equilibrium. The notion of an optimal equilibrium is stronger than a by far more widely used concept of a Nash equilibrium (although the two concepts coincide in the case of a two-person zero-sum game).

Definition 7.2. A Nash equilibrium $\sigma = (\sigma^1, \sigma^2, \dots, \sigma^m) \in S$ is called an *optimal equilibrium* if the inequality $V^k(\sigma^k, \sigma^{-k}) \leq V^k(\sigma^k, s^{-k})$ holds for each $k \in \mathcal{M}$ and all $s^{-k} \in S^{-k}$ or, equivalently, for each $k \in \mathcal{M}$,

$$V^k(\sigma^k, \sigma^{-k}) = \operatorname*{ess\,inf}_{s^{-k} \in S^{-k}} V^k(\sigma^k, s^{-k}). \tag{7.2}$$

In view of condition (7.1) of a Nash equilibrium, an optimal equilibrium σ satisfies

$$V^k(\sigma^k, \sigma^{-k}) = \operatorname*{ess\,inf}_{s^{-k} \in S^{-k}} V^k(\sigma^k, s^{-k}) = \operatorname*{ess\,sup}_{s^k \in S^k} V^k(s^k, \sigma^{-k})$$

or, equivalently, the following *saddle point* property is valid

$$V^k(s^k, \sigma^{-k}) \leq V^k(\sigma^k, \sigma^{-k}) \leq V^k(\sigma^k, s^{-k}), \quad \forall\, s^k \in S^k, s^{-k} \in S^{-k}, \tag{7.3}$$

for every $k \in \mathcal{M}$.

Remark 7.1. For the zero-sum two-player game, we have $m = 2$ and $V^1(s) + V^2(s) = 0$ for every strategy profile s. Then it is easy to check that any Nash equilibrium (τ^*, σ^*) has the saddle point property (7.3), that is,

$$V(\tau^*, \sigma) \leq V(\tau^*, \sigma^*) \leq V(\tau, \sigma^*), \quad \forall\, \tau \in S^1,\ \sigma \in S^2,$$

where $V = V^1$. Therefore, any Nash equilibrium is also an optimal equilibrium.

Maxmin and minmax values. It is clear that an optimal equilibrium is essentially a saddle point. In addition to the properties of a Nash equilibrium, each player can guarantee his optimal equilibrium payoff without knowing the actions of other players. Furthermore, as shown in Corollary 7.1 below, all optimal equilibria achieve the same value. The concept of an optimal equilibrium is also closely related to the notions of the maxmin and minmax values of the game.

The *maxmin value* is the maximum payoff player k can guarantee.

Definition 7.3. The *maxmin value* $\underline{V}^k$ equals

$$\underline{V}^k := \operatorname*{ess\,sup}_{s^k \in S^k} \operatorname*{ess\,inf}_{s^{-k} \in S^{-k}} V^k(s^k, s^{-k}). \tag{7.4}$$

A *maxmin strategy* for player k is any strategy $\widehat{s}^k$ belonging to the class S^k such that

$$\operatorname*{ess\,inf}_{s^{-k} \in S^{-k}} V^k(\widehat{s}^k, s^{-k}) = \underline{V}^k,$$

so that $\widehat{s}^k$ realises the supremum in (7.4).

The *minmax value* is the lowest payoff that the other players can force upon player k.

Definition 7.4. The *minmax value* $\overline{V}^k$ equals

$$\overline{V}^k := \operatorname*{ess\,inf}_{s^{-k} \in S^{-k}} \operatorname*{ess\,sup}_{s^k \in S^k} V^k(s^k, s^{-k}). \tag{7.5}$$

A *minmax strategy profile* for the player set $\mathcal{M}^{-k}$ is any strategy profile $\widehat{s}^{-k} \in S^{-k}$ such that

$$\operatorname*{ess\,sup}_{s^k \in S^k} V^k(s^k, \widehat{s}^{-k}) = \overline{V}^k,$$

so that $\widehat{s}^{-k}$ realises the infimum in (7.5).

Value of a game. In general, the maxmin value is never higher than the minmax value, since the players from the set $\mathcal{M}^{-k}$ cannot force the payoff of player k to be lower than an amount that can be guaranteed by him. This explains why the maxmin value (resp., minmax value) is also known as the *lower value* (resp., *upper value*).

Proposition 7.1. *(i) The inequality $\overline{V}^k \geq \underline{V}^k$ is valid for all k.*
(ii) If σ is a Nash equilibrium, then $V^k(\sigma) \geq \overline{V}^k$ for all k.
(iii) If σ satisfies (7.2)*, then $V^k(\sigma) \leq \underline{V}^k$ for all k.*
(iv) If σ is an optimal equilibrium, then $V^k(\sigma) = \overline{V}^k = \underline{V}^k$ for all k.
(v) If σ is an optimal equilibrium, then for all k the strategy σ^k (resp., strategy profile σ^{-k}) is a maxmin strategy for player k (resp., a minmax strategy profile the players set $\mathcal{M}^{-k}$).

Proof. (i) For every $\widehat{s}^k \in S^k$ and $\widehat{s}^{-k} \in S^{-k}$, we have that

$$G(\widehat{s}^{-k}) := \operatorname*{ess\,sup}_{s^k \in S^k} V^k(s^k, \widehat{s}^{-k}) \geq V^k(\widehat{s}^k, \widehat{s}^{-k}) \geq \operatorname*{ess\,inf}_{s^{-k} \in S^{-k}} V^k(\widehat{s}^k, s^{-k}) =: H(\widehat{s}^k)$$

and thus $G(s^{-k}) \geq H(s^k)$ for every s^k and s^{-k}. Consequently,

$$\begin{aligned}\overline{V}^k &= \operatorname*{ess\,inf}_{s^{-k}\in S^{-k}} \operatorname*{ess\,sup}_{s^k\in S^k} V^k(s^k, s^{-k}) = \operatorname*{ess\,inf}_{s^{-k}\in S^{-k}} G(s^{-k}) \\ &\geq \operatorname*{ess\,sup}_{s^k\in S^k} H(s^k) = \operatorname*{ess\,sup}_{s^k\in S^k} \operatorname*{ess\,inf}_{s^{-k}\in S^{-k}} V^k(s^k, s^{-k}) = \underline{V}^k.\end{aligned}$$

(ii) If condition (7.1) holds then

$$V^k(\sigma^k, \sigma^{-k}) = \operatorname*{ess\,sup}_{s^k\in S^k} V^k(s^k, \sigma^{-k}) \geq \operatorname*{ess\,inf}_{s^{-k}\in S^{-k}} \operatorname*{ess\,sup}_{s^k\in S^k} V^k(s^k, s^{-k}) = \overline{V}^k.$$

(iii) If condition (7.2) holds then

$$V^k(\sigma^k, \sigma^{-k}) = \operatorname*{ess\,inf}_{s^{-k}\in S^{-k}} V^k(\sigma^k, s^{-k}) \leq \operatorname*{ess\,sup}_{s^k\in S^k} \operatorname*{ess\,inf}_{s^{-k}\in S^{-k}} V^k(s^k, s^{-k}) = \underline{V}^k.$$

(iv) If both conditions (7.1) and (7.2) hold, then, by (ii) and (iii), we get $\underline{V}^k \geq V^k(\sigma) \geq \overline{V}^k$. In view of (i), we thus obtain the equality $V^k(\sigma) = \overline{V}^k = \underline{V}^k$.

(v) By combining (iv) with condition (7.3), we obtain

$$\begin{aligned}V^k(\sigma^k, s^{-k}) &\geq V^k(\sigma^k, \sigma^{-k}) = \underline{V}^k, \quad \forall\, s^{-k} \in S^{-k}, \\ V^k(s^k, \sigma^{-k}) &\leq V^k(\sigma^k, \sigma^{-k}) = \overline{V}^k, \quad \forall\, s^k \in S^k.\end{aligned}$$

Hence σ^k and σ^{-k} are maxmin and minmax strategies, respectively. □

Definition 7.5. If the equality $\overline{V}^k = \underline{V}^k$ holds, then $V^{*k} := \overline{V}^k = \underline{V}^k$ is the *value for player* k. The *value of the game* V^* is the vector $(V^{*1}, \dots, V^{*m})$ of values for all players.

Since the equality is not necessarily achieved in part (i) in Proposition 7.1, the existence of the value is not guaranteed. However, by part (v) in this proposition, the existence of an optimal equilibrium implies the existence of the value of the game.

Corollary 7.1. *(i) If the value V^* of a game exists, then it is unique.*
(ii) If there exists an optimal equilibrium σ, then the value exists for every player and

$$V(\sigma) = (V^{*1}, \dots, V^{*m}) = V^*.$$

(iii) Every optimal equilibrium σ achieves the same value.

Proof. Statement (i) is implicit in Definition 7.5 and (ii) follows immediately from Proposition 7.1 (iv). Finally (iii) follows immediately from (i) and (ii). □

We recall the definitions of an ε*-optimal strategy* and ε*-equilibrium* of a multi-player game.

Definition 7.6. Suppose that a game has a value V^{k*} for player k. For $\varepsilon \geq 0$, an ε*-optimal strategy* $\tau^\varepsilon \in S^k$ for player k guarantees the payoff to within ε of the value. In other words

$$\operatorname*{ess\,inf}_{s^{-k}\in S^{-k}} V^k(\tau^\varepsilon, s^{-k}) \geq V^{k*} - \varepsilon.$$

Definition 7.7. For $\varepsilon \geq 0$, a strategy profile $\sigma = (\sigma^1, \ldots, \sigma^m) \in S^1 \times \cdots \times S^m$ is called an *ε-equilibrium* of a game if for each $k \in \mathcal{M}$

$$\operatorname*{ess\,sup}_{s^k \in S^k} V^k(s^k, \sigma^{-k}) \leq V^k(\sigma^k, \sigma^{-k}) + \varepsilon.$$

Note that a Nash equilibrium is a 0-equilibrium, while an optimal equilibrium is a 0-equilibrium with 0-optimal strategies. The following result is easy to prove and thus its proof is omitted.

Proposition 7.2. *In a two-player, zero-sum game the following are equivalent:*
(i) The game has a value for both players.
(ii) For all $\varepsilon > 0$, there exist ε-optimal strategies for both players.
(iii) For all $\varepsilon > 0$, there exists an ε-equilibrium.
(iv) For all $\varepsilon > 0$, there exists a real number v^ε and a strategy profile $(\sigma^\varepsilon, \tau^\varepsilon)$ such that

$$\operatorname*{ess\,sup}_{\tau \in S^1} V^1(\tau, \sigma^\varepsilon) \leq v^\varepsilon \leq \operatorname*{ess\,inf}_{\sigma \in S^2} V^1(\tau^\varepsilon, \sigma).$$

7.2 Ordered Dynkin Games

Let X, Y and Z be $\mathbb{F}$-adapted, càdlàg processes satisfying the usual integrability condition and such that $X \leq Z \leq Y$. Consider the *Dynkin game* with the payoff to the maximiser given by

$$R(\tau, \sigma) = \mathbb{1}_{\{\tau < \sigma\}} X_\tau + \mathbb{1}_{\{\sigma < \tau\}} Y_\sigma + \mathbb{1}_{\{\sigma = \tau\}} Z_\sigma \tag{7.6}$$

where τ, σ are $\mathbb{F}$-stopping times. The zero-sum property means that the payoff to the minimizer equals $-R(\tau, \sigma)$. Without loss of generality, we set $X_T = Y_T = Z_T$. In fact, one can show that the process Z plays no role in an ordered Dynkin game.

The case of a continuous-time Dynkin game was studied by several authors (see, for instance, Kobylanski et al. [52], Laraki and Solan [53], Lepeltier and Maingueneau [54], Neveu [62] and Stettner [88]). The following result summarises some of their findings.

Theorem 7.1. *Consider the Dynkin game with the payoff R given by* (7.6).
(i) For any $t \in [0, T]$, the Dynkin game on $[t, T]$ has the value V_t^* satisfying

$$V_t^* = \operatorname*{ess\,inf}_{\sigma_t \in \mathcal{T}_{[t,T]}} \operatorname*{ess\,sup}_{\tau_t \in \mathcal{T}_{[t,T]}} \mathbb{E}_{\mathbb{P}}\big(R(\tau_t, \sigma_t) \,|\, \mathcal{F}_t\big) = \operatorname*{ess\,sup}_{\tau_t \in \mathcal{T}_{[t,T]}} \operatorname*{ess\,inf}_{\sigma_t \in \mathcal{T}_{[t,T]}} \mathbb{E}_{\mathbb{P}}\big(R(\tau_t, \sigma_t) \,|\, \mathcal{F}_t\big).$$

The value process V^* of the Dynkin game can be chosen to be right-continuous.
(ii) For any $t \in [0, T]$ and $\varepsilon > 0$, the $\mathbb{F}$-stopping times $(\tau_t^\varepsilon, \sigma_t^\varepsilon) \in \mathcal{T}_{[t,T]} \times \mathcal{T}_{[t,T]}$ defined by

$$\tau_t^\varepsilon := \inf\{u \geq t : X_u \geq V_u^* - \varepsilon\}, \quad \sigma_t^\varepsilon := \inf\{u \geq t : Y_u \leq V_u^* + \varepsilon\} \tag{7.7}$$

are ε-optimal strategies

$$\operatorname*{ess\,sup}_{\tau_t \in \mathcal{T}_{[t,T]}} \mathbb{E}_{\mathbb{P}}\big(R(\tau_t, \sigma_t^\varepsilon) \,|\, \mathcal{F}_t\big) - \varepsilon \leq V_t^* \leq \operatorname*{ess\,inf}_{\sigma_t \in \mathcal{T}_{[t,T]}} \mathbb{E}_{\mathbb{P}}\big(R(\tau_t^\varepsilon, \sigma_t) \,|\, \mathcal{F}_t\big) + \varepsilon.$$

(iii) If we further assume that X and $-Y$ are left upper semi-continuous (only have positive jumps), then the Dynkin game on $[t, T]$ has a Nash equilibrium $(\tau_t^*, \sigma_t^*) \in \mathcal{T}_{[t,T]} \times \mathcal{T}_{[t,T]}$ satisfying

$$\sigma_t^* := \lim_{\varepsilon \to 0} \sigma_t^\varepsilon, \quad \tau_t^* := \lim_{\varepsilon \to 0} \tau_t^\varepsilon \tag{7.8}$$

and for every $\tau_t, \sigma_t \in \mathcal{T}_{[t,T]}$

$$\mathbb{E}_{\mathbb{P}}\big(R(\tau_t, \sigma_t^*)|\mathcal{F}_t\big) \le V_t^* = \mathbb{E}_{\mathbb{P}}\big(R(\tau_t^*, \sigma_t^*)|\mathcal{F}_t\big) \le \mathbb{E}_{\mathbb{P}}\big(R(\tau_t^*, \sigma_t)|\mathcal{F}_t\big).$$

Observe that there may be other ε-optimal strategy pairs (resp., Nash equilibria) than the ones specified by (7.7) (resp., (7.8)). Also σ_t^*, τ_t^* do not necessarily coincide with stopping times σ_t^0, τ_t^0, which are defined by setting $\varepsilon = 0$ in (7.7), that is,

$$\tau_t^0 := \inf\{u \ge t : X_u \ge V_u^*\}, \quad \sigma_t^0 := \inf\{u \ge t : Y_u \le V_u^*\}.$$

In general, we have that $\tau_t^* \le \tau_t^0$ and $\sigma_t^* \le \sigma_t^0$.

7.3 General Dynkin Games

We now assume that X, Y and Z arbitrary $\mathbb{F}$-adapted, càdlàg processes satisfying the usual integrability condition. To study the existence of the value process V^*, we define the *modified payoff* $\widetilde{R}(\sigma,\tau)$ by (see Guo [36])

$$\widetilde{R}(\sigma,\tau) := (X_\tau \wedge Z_\tau)\mathbb{1}_{\{\tau<\sigma\}} + (Y_\sigma \vee Z_\sigma)\mathbb{1}_{\{\sigma<\tau\}} + Z_\sigma \mathbb{1}_{\{\tau=\sigma\}}. \tag{7.9}$$

If we set $L := X \wedge Z$ and $U := Y \vee Z$, then $\widetilde{R}(\sigma,\tau)$ can be represented as follows

$$\widetilde{R}(\sigma,\tau) := L_\tau \mathbb{1}_{\{\tau<\sigma\}} + U_\sigma \mathbb{1}_{\{\sigma<\tau\}} + Z_\sigma \mathbb{1}_{\{\tau=\sigma\}}.$$

Since $L \le Z \le U$ (the ordered case), it is well known that the Dynkin game with the payoff $\widetilde{R}(\sigma,\tau)$ has the value process, which is henceforth denoted as $\widetilde{V}$. Furthermore, it can be shown that for every $\sigma_t, \tau_t \in \mathcal{T}_{[t,T]}$ there exist $\widehat{\sigma}_t \in \mathcal{T}_{[t,T]}$ and $\widehat{\tau}_t \in \mathcal{T}_{[t,T]}$ such that

$$R(\widehat{\sigma}_t, \tau_t) \le \widetilde{R}(\sigma_t, \tau_t) \le R(\sigma_t, \widehat{\tau}_t)$$

so that

$$\operatorname*{ess\,sup}_{\tau\in\mathcal{T}_{[t,T]}} \operatorname*{ess\,inf}_{\sigma\in\mathcal{T}_{[t,T]}} \mathbb{E}(R(\sigma,\tau)\,|\,\mathcal{F}_t) \le \widetilde{V}_t \le \operatorname*{ess\,inf}_{\sigma\in\mathcal{T}_{[t,T]}} \operatorname*{ess\,sup}_{\tau\in\mathcal{T}_{[t,T]}} \mathbb{E}(R(\sigma,\tau)\,|\,\mathcal{F}_t).$$

Hence if the value process V^* is well defined, then necessarily $V^* = \widetilde{V}$.

We are ready to state necessary and sufficient for the existence of the value of a general Dynkin game with arbitrary processes X, Y, Z.

Assumption 7.1. Assume that $X_t \wedge Y_t \le \widetilde{V}_t \le X_t \vee Y_t$ for every $t \in [0,T]$.

Proposition 7.3. *The value process V^* of a general Dynkin game with the payoff R given by* (7.6) *is well defined if and only if Assumption* (7.1) *is valid. Then the equality $V_t^* = \widetilde{V}_t$ is satisfied for every $t \in [0,T]$ where $\widetilde{V}$ is the value process of the standard Dynkin game with the payoff $\widetilde{R}$ given by* (7.9).

If $X \le Z \le Y$ and, more generally, if $X \wedge Y \le Z \le X \vee Y$ (see Ohtsubo [70]), then Assumption 7.1 is satisfied. In particular, if either $Z = X$ or $Z = Y$, then Assumption 7.1 holds.

Case $Z = X$**.** Assume first that the equality $Z = X$ holds (as in Szimayer [90]). Then we obtain

$$R(\sigma,\tau) = R_1(\sigma,\tau) := X_\tau \mathbb{1}_{\{\tau\le\sigma\}} + Y_\sigma \mathbb{1}_{\{\sigma<\tau\}} \tag{7.10}$$

and

$$\begin{aligned}\widetilde{R}(\sigma,\tau) = \widetilde{R}_1(\sigma,\tau) &:= X_\tau \mathbb{1}_{\{\tau<\sigma\}} + (X_\sigma \vee Y_\sigma)\mathbb{1}_{\{\sigma<\tau\}} + X_\sigma \mathbb{1}_{\{\tau=\sigma\}} \\ &= X_\tau \mathbb{1}_{\{\tau\le\sigma\}} + (X_\sigma \vee Y_\sigma)\mathbb{1}_{\{\sigma<\tau\}}.\end{aligned}$$

Corollary 7.2. *If $Z = X$, then the value process V^* of the general Dynkin game is well defined and the equality $V^* = \widetilde{V}$ holds.*

Proof. Recall that $L \le Z \le U$ and thus also $X \wedge Z = L \le \widetilde{V} \le U = Y \vee Z$ from the standard case. Therefore, if we assume that $Z = X$, then we obtain $X \wedge Y \le X \le \widetilde{V} \le X \vee Y$ and thus Assumption 7.1 is satisfied. Hence, by Proposition 7.3, the value process V^* is well defined and $V^* = \widetilde{V}$. □

Let us consider some special cases:
(a) If $Z = X \le Y$ then $R_1(\sigma,\tau) = \widetilde{R}_1(\sigma,\tau)$ and we deal with the standard Dynkin game.
(b) If $Z = X \ge Y$ then $\widetilde{R}_1(\sigma,\tau) = X_{\sigma\wedge\tau}$.

Case $Z = Y$**.** We now postulate that $Z = Y$ (see, in particular, Ohtsubo [70]). Then

$$R(\sigma,\tau) = R_2(\sigma,\tau) := X_\tau \mathbb{1}_{\{\tau<\sigma\}} + Y_\sigma \mathbb{1}_{\{\sigma\le\tau\}} \tag{7.11}$$

and

$$\begin{aligned}\widetilde{R}(\sigma,\tau) = \widetilde{R}_2(\sigma,\tau) &:= (X_\tau \wedge Y_\tau)\mathbb{1}_{\{\tau<\sigma\}} + Y_\sigma \mathbb{1}_{\{\sigma<\tau\}} + Y_\sigma \mathbb{1}_{\{\tau=\sigma\}} \\ &= (X_\tau \wedge Y_\tau)\mathbb{1}_{\{\tau<\sigma\}} + Y_\sigma \mathbb{1}_{\{\sigma\le\tau\}}.\end{aligned}$$

Corollary 7.3. *If $Z = Y$, then the value process V^* of the general Dynkin game is well defined and the equality $V^* = \widetilde{V}$ holds.*

Proof. Recall that $L \le Z \le U$ and thus also $X \wedge Z = L \le \widetilde{V} \le U = Y \vee Z$ from the standard case. Therefore, if we assume that $Z = Y$, then we obtain $X \wedge Y \le \widetilde{V} \le Y \le X \vee Y$ and thus Assumption 7.1 is satisfied. Hence, by Proposition 7.3, the value process V^* is well defined and $V^* = \widetilde{V}$. □

Let us consider some special cases:
(a) If $X \le Y = Z$ then $R_2(\sigma,\tau) = \widetilde{R}_2(\sigma,\tau)$ and we deal with the standard Dynkin game.
(b) If $X \ge Y = Z$ then $\widetilde{R}_2(\sigma,\tau) = Y_{\sigma\wedge\tau}$.

Remark 7.2. If $Z = Y$, then in a discrete time set-up, we have $V_T^* = Y_T = Z_T$ and for $t = 0, 1, \ldots, T-1$

$$V_t^* = Y_t \mathbb{1}_{\{Y_t\le X_t\}} + \min\Big(Y_t, \max\big(X_t,\, \mathbb{E}(V_{t+1}^* \,|\, \mathcal{F}_t)\big)\Big)\mathbb{1}_{\{Y_t>X_t\}}$$

whereas $\widetilde{V}_T = Y_T = Z_T$ and for $t = 0, 1, \ldots, T-1$

$$\widetilde{V}_t = \min\Big(Y_t, \max\big(X_t \wedge Y_t,\, \mathbb{E}(\widetilde{V}_{t+1} \,|\, \mathcal{F}_t)\big)\Big).$$

Hence it is easy to check that the value process V^* coincides with $\widetilde{V}$.

7.4 Doubly Reflected BSDE

Let W be a d-dimensional Brownian motion on $(\Omega, \mathcal{F}, \mathbb{P})$ and let $\mathbb{F}$ be the filtration generated by W.

Definition 7.8. We say that (Y, Z, K^+, K^-) is a *solution to the doubly reflected BSDE* (DRBSDE) (g, ξ_T, L, U) if

$$(Y, Z, K^+, K^-) \in \mathcal{H}^2_T(\mathbb{R}) \times \mathcal{H}^2_T(\mathbb{R}^d) \times \mathcal{H}^2_T(\mathbb{R}) \times \mathcal{H}^2_T(\mathbb{R}).$$

and the following conditions are satisfied:
(i) for all $t \in [0, T]$

$$Y_t = \xi_T + \int_t^T g(s, Y_s, Z_s)\, ds + (K^+_T - K^+_t) - (K^-_T - K^-_t) - \int_t^T Z_s\, dW_s,$$

(ii) $\mathbb{P}(L_t \leq Y_t \leq U_t,\ \forall\, t \in [0, T]) = 1$;
(iii) K^+ and K^- are $\mathbb{F}$-adapted, continuous, and non-decreasing processes such that $K^+_0 = K^-_0 = 0$ and

$$\int_0^T (Y_t - L_t)\, dK^+_t = \int_0^T (U_t - Y_t)\, dK^-_t = 0.$$

We will study doubly reflected BSDEs with Lipschitz generators. Hence we make the following *standard assumptions* for doubly reflected BSDE (g, ξ_T, L, U):
(i) g and ξ_T are standard parameters, in the sense of Definition 1.1;
(ii) L and U are $\mathbb{F}$-adapted, continuous, real-valued processes satisfying

$$\mathbb{E}\big[\sup_{0 \leq t \leq T} (L^+_t)^2\big] < +\infty, \quad \mathbb{E}\big[\sup_{0 \leq t \leq T} (U^+_t)^2\big] < +\infty.$$

The following condition is frequently used to ensure that the doubly reflected BSDE is well posed.

Definition 7.9. We say that a pair (L, U) satisfies the *Mokobodzki condition* if there exist two non-negative, càdlàg supermartingales $(A_t)_{t \in [0,T]}$ and $(B_t)_{t \in [0,T]}$, which are uniformly square-integrable and such that

$$\mathbb{P}(L_t \leq A_t - B_t \leq U_t,\ \forall t \in [0, T]) = 1.$$

The following existence and uniqueness result for doubly reflected BSDEs was established by Cvitanić and Karatzas [15].

Theorem 7.2. *(i) If the standard doubly reflected BSDE problem has a solution, then it is unique and the solution (Y, Z, K^+, K^-) belongs to the space $\mathcal{S}^2_1(0, T) \times \mathcal{H}^2_d(0, T) \times \mathcal{S}^2_1(0, T) \times \mathcal{S}^2_1(0, T)$.*
(ii) The Mokobodzki condition is a sufficient condition for the existence of a solution to the standard doubly reflected BSDE.

7.5 Non-linear Dynkin Games

We are in a position to formally define a nonlinear Dynkin game (g, L, U, ν).

Definition 7.10. Assume that $L, U \in \mathcal{S}_1^2(0,T)$ are such that $L_t \le U_t$ for all $t \in [0,T]$. The *reward* for the max-player is given by

$$I(\tau,\sigma) := L_\tau \mathbb{1}_{\{\tau \le \sigma\}} + U_\sigma \mathbb{1}_{\{\sigma < \tau\}}$$

where $\tau, \sigma \in \mathcal{T}_{[0,T]}$. In the *nonlinear Dynkin game* (g, L, U, ν) started at $\nu \in \mathcal{T}_{[0,T]}$

- the max-player chooses $\tau \in \mathcal{T}_{[\nu,T]}$ to maximise $\mathcal{E}^g_{\nu,\tau\wedge\sigma}(I(\tau,\sigma))$,
- the min-player chooses $\sigma \in \mathcal{T}_{[\nu,T]}$ to minimise $\mathcal{E}^g_{\nu,\tau\wedge\sigma}(I(\tau,\sigma))$.

Note that $\mathcal{E}^g_{\nu,\tau\wedge\sigma}(I(\tau,\sigma))$ represents the g-conditional expected reward at ν.

We define the value of a nonlinear Dynkin game.

Definition 7.11. The *upper* and *lower values* of a nonlinear Dynkin game (g, L, U, ν) are given by

$$\overline{V}^g_\nu(L,U) := \operatorname*{ess\,inf}_\sigma \operatorname*{ess\,sup}_\tau \mathcal{E}^g_{\nu,\tau\wedge\sigma}\big(I(\tau,\sigma)\big),$$
$$\underline{V}^g_\nu(L,U) := \operatorname*{ess\,sup}_\tau \operatorname*{ess\,inf}_\sigma \mathcal{E}^g_{\nu,\tau\wedge\sigma}\big(I(\tau,\sigma)\big),$$

so that $\overline{V}^g_\nu(L,U) \ge \underline{V}^g_\nu(L,U)$. If the equality $\underline{V}^g_\nu(L,U) = \overline{V}^g_\nu(L,U) =: V^g_\nu(L,U)$ holds, then we say that $V^g_\nu(L,U)$ is the *value at time* ν of a nonlinear Dynkin game.

As in the linear case, the concept of a saddle point (also known as a Nash equilibrium) is a convenient tool to analyze a nonlinear Dynkin game.

Definition 7.12. A pair $(\tau^*, \sigma^*) \in \mathcal{T}_{[\nu,T]} \times \mathcal{T}_{[\nu,T]}$ is an ν*-saddle point* if for all pairs $(\tau,\sigma) \in \mathcal{T}_{[\nu,T]} \times \mathcal{T}_{[\nu,T]}$

$$\mathcal{E}^g_{\nu,\tau\wedge\sigma^*}\big(I(\tau,\sigma^*)\big) \le \mathcal{E}^g_{\nu,\tau^*\wedge\sigma^*}\big(I(\tau^*,\sigma^*)\big) \le \mathcal{E}^g_{\nu,\tau^*\wedge\sigma}\big(I(\tau^*,\sigma)\big).$$

It is clear that the existence of a ν-saddle point implies the existence of the value $V^g_\nu(L,U)$ for the nonlinear Dynkin game.

7.6 Equivalence of DRBSDEs and Nonlinear Dynkin Games

The following result is due to Dumitrescu et al. [23].

Theorem 7.3. *Let (Y, Z, K^+, K^-) be the unique solution to the standard doubly reflected BSDE $(g, \xi_T, L, U,)$ with the barriers L, U satisfying the Mokobodzki condition and the terminal condition $\xi_T = L_T$. Then:*
(i) the value of the nonlinear Dynkin game (g, L, U, ν) exists and it satisfies the equality $V^g_\nu(L,U) = Y_\nu$;
(ii) the pair $(\tau^*_\nu, \sigma^*_\nu) \in \mathcal{T}_{[\nu,T]} \times \mathcal{T}_{[\nu,T]}$ given by

$$\tau^*_\nu := \inf\{t \ge \nu \,|\, Y_t = L_t\}, \quad \sigma^*_\nu := \inf\{t \ge \nu \,|\, Y_t = U_t\},$$

is a ν-saddle point.

Proof. The proof extends the proof of equivalence of classical Dynkin games and (linear) reflected BSDEs from Cvitanić and Karatzas [15] and it is omitted. □

From Theorem 7.3, Dumitrescu et al. [23] deduced the following comparison result for the doubly reflected BSDE with a Lipschitz continuous generators.

Proposition 7.4. *For $i = 1, 2$, consider the standard doubly reflected BSDE with barriers (L^i, U^i) satisfying the Mokobodzki condition, the generator g^i and the terminal condition $\xi^i_T = L^i_T$. Assume that:*
(i) $L^2_t \leq L^1_t$ and $U^2_t \leq U^1_t$ for all $t \in [0, T]$;
(ii) $g^2(t, Y^i_t, Z^i_t) \leq g^1(t, Y^i_t, Z^i_t)$ for all $t \in [0, T]$ for some $i = 1, 2$.
Then the inequality $Y^2_t \leq Y^1_t$ holds for all $t \in [0, T]$.

Proof. We have $g^2(t, Y^1_t, Z^1_t) \leq g^1(t, Y^1_t, Z^1_t)$ and for all $\tau, \sigma \in \mathcal{T}_{[t,T]}$

$$I^2(\tau, \sigma) := L^2_\tau \mathbb{1}_{\{\tau \leq \sigma\}} + U^2_\sigma \mathbb{1}_{\{\sigma < \tau\}} \leq L^1_\tau \mathbb{1}_{\{\tau \leq \sigma\}} + U^1_\sigma \mathbb{1}_{\{\sigma < \tau\}} =: I^1(\tau, \sigma).$$

From the comparison theorem for standard BSDEs, we obtain for all $t \in [0, T]$

$$\mathcal{E}^{g_2}_{t,\tau\wedge\sigma}\big(I^2(\tau, \sigma)\big) \leq \mathcal{E}^{g_1}_{t,\tau\wedge\sigma}\big(I^1(\tau, \sigma)\big).$$

By considering the Dynkin games (g^i, L^i, U^i, t) for $i = 1, 2$, we obtain

$$\begin{aligned} Y^2_t = V^{g_2}_t(L, U) &= \operatorname*{ess\,inf}_{\sigma \in \mathcal{T}_{[t,T]}} \operatorname*{ess\,sup}_{\tau \in \mathcal{T}_{[t,T]}} \mathcal{E}^{g_2}_{t,\tau\wedge\sigma}\big(I^2(\tau, \sigma)\big) \\ &\leq \operatorname*{ess\,inf}_{\sigma \in \mathcal{T}_{[t,T]}} \operatorname*{ess\,sup}_{\tau \in \mathcal{T}_{[t,T]}} \mathcal{E}^{g_1}_{t,\tau\wedge\sigma}\big(I^1(\tau, \sigma)\big) = V^{g_1}_t(L, U) = Y^1_t \end{aligned}$$

which concludes the proof. □

8 Stochastic Optimal Control

We will study two classes of optimal control problems for diffusion processes: (a) controlling of the drift term and (b) controlling of the drift and diffusion terms. For examples of explicitly solvable singular optimal control problems, see Beneš et al. [5].

Let $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P})$ be a filtered probability space satisfying the usual hypotheses of right-continuity and completeness. Let U be a compact metric space. We first define the class of all admissible control processes.

Definition 8.1. Let U be a compact metric space. We say that a control process u is *admissible* if (i) u takes values in U and (ii) u is $\mathbb{F}$-adapted. We denote by $\mathcal{U}$ the space of all admissible control processes.

8.1 Optimal Control via Hamilton–Jacobi–Bellman Equation

We summarise here the dynamic programming approach to optimal control problem in a Markovian setting (for a more detailed exposition, see Pham [74]). Let $(W_t)_{t\in[0,T]}$ be an $\mathbb{R}^d$-valued standard Brownian motion on $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P})$. We assume that $\mathbb{F} = (\mathcal{F}_t)_{t\in[0,T]}$ is the filtration generated by W and augmented by the class of $\mathbb{P}$-null sets of $\mathcal{F}_T$. For any process $u \in \mathcal{U}$, we consider the following d-dimensional SDE

$$dX_t^u = \mu(t, X_t^u, u_t)\,dt + \sigma(t, X_t^u)\,dW_t, \quad X_0^u = x_0 \in \mathbb{R}^d, \tag{8.1}$$

where $\mu : [0,T]\times\mathbb{R}^d\times\mathcal{U} \to \mathbb{R}^d$ and $\sigma : [0,T]\times\mathbb{R}^d \to \mathbb{R}^{d\times d}$ satisfy conditions of Theorem 12.3 with respect to x and u. Hence for any fixed process u, SDE (8.1) admits a unique strong solution. It is worth stressing that the solution X^u to (8.1) is not necessarily Markovian with respect to $\mathbb{F}$, unless the process u is given as $u_t = u(t, X_t^u)$ for some function $u(t,x)$. Obviously, there is no reason to make such an assumption, but it is natural to conjecture that an optimal control process u^* for the optimisation problem (8.2) can be found in the following form: $u_t^* = u^*(t, X_t^{u*})$ for some function $u^*(t,x)$. Then one may use Markovian techniques to search for a candidate for an optimal control u^* and subsequently verify that it is indeed optimal in the class $\mathcal{U}$ of all $\mathbb{F}$-adapted control processes taking values in U.

We are ready to formulate the optimisation problem studied in this section. The goal of the controller is to find a process $u^* \in \mathcal{U}$ such that $J_0(u^*) \geq J_0(u)$ for every process $u \in \mathcal{U}$ where the real-valued functional J_0 is given by

$$J_0(u) := \mathbb{E}_{\mathbb{P}}\bigg(\int_0^T h(s, X_s^u, u_s)\,ds + \Psi(X_T^u)\bigg).$$

If an optimal control u^* exists, then

$$J_0(u^*) = \max_{u\in\mathcal{U}} \mathbb{E}_{\mathbb{P}}\bigg(\int_0^T h(s, X_s^u, u_s)\,ds + \Psi(X_T^u)\bigg). \tag{8.2}$$

More generally, for any control process $u \in \mathcal{U}$, we denote by $J(u)$ the real-valued process given by the following expression

$$J_t(u) := \mathbb{E}_{\mathbb{P}}\bigg(\int_t^T h(s, X_s^u, u_s)\,ds + \Psi(X_T^u)\,\Big|\,\mathcal{F}_t\bigg).$$

Then the optimisation problem can also be formulated on the interval $[t,T]$ for any fixed $t \in [0,T]$. If we postulate, in addition, that $u_t = u(t, X_t^u)$, then the solution X^u to (8.1) is Markovian with respect to the filtration $\mathbb{F}$ and thus

$$\begin{aligned} J_t(u) &= \mathbb{E}_{\mathbb{P}}\bigg(\int_t^T h\big(s, X_s^u, u(s, X_s^u)\big)\, ds + \Psi(X_T^u) \,\Big|\, \mathcal{F}_t \bigg) \\ &= \mathbb{E}_{\mathbb{P}}\bigg(\int_t^T h\big(s, X_s^u, u(s, X_s^u)\big)\, ds + \Psi(X_T^u) \,\Big|\, X_t^u \bigg) = f(t, X_t^u) \end{aligned}$$

where $f(t,\cdot)$ is a Borel measurable function on $\mathbb{R}^d$. Hence we may also write, on the event $\{X_t^u = x\}$,

$$J_t(u) = \mathbb{E}_{\mathbb{P}}\bigg(\int_t^T h\big(s, X_s^u, u(s, X_s^u)\big)\, ds + \Psi(X_T^u) \,\Big|\, X_t^u = x \bigg) = f(t,x).$$

Of course, we have that $J_T(u) = \Psi(X_T^u)$ for every $u \in \mathcal{U}$, which in fact means that, independently of u, the terminal condition reads $J_T(u) = \Psi(x)$ provided that $X_T^u = x$.

Since the optimisation problem is Markovian, we may address it by considering a family of control problems indexed by (t,x) where x is the value of the controlled process at time t. To this end, for any admissible control process $u \in \mathcal{U}$ and for any fixed $(t,x) \in [0,T] \times \mathbb{R}^d$, let $X^{t,x,u}$ be the unique strong solution to the SDE

$$dX_s^{t,x,u} = \mu(s, X_s^{t,x,u}, u_s)\, ds + \sigma(s, X_s^{t,x,u})\, dW_s \tag{8.3}$$

with the initial condition $X_t^{t,x,u} = x$. We define the function $J : [0,T] \times \mathbb{R}^d \times \mathcal{U} \to \mathbb{R}$

$$J(t,x,u) := \mathbb{E}_{\mathbb{P}}\bigg(\int_t^T h(s, X_s^{t,x,u}, u_s)\, ds + \Psi(X_T^{t,x,u}) \bigg).$$

Although we do not assume in the definition of the function J that $u_t = u(t, X_t^u)$, it is rather clear that

Definition 8.2. The function $v : [0,T] \times \mathbb{R}^d \to \mathbb{R}$ given by $v(t,x) = \sup_{u \in \mathcal{U}} J(t,x,u)$ is called the *value function* for the maximisation problem. If an optimal control u^* exists, then the process $J_t(u^*) = v(t, X_t^{u^*})$ is called the *value process*.

In view of the shape of our optimisation problem, it is natural to conjecture that the *dynamic programming principle* (DPP), which is also known as the *Bellman principle* (see Bellman [3]) should be satisfied. In the present setup, the DPP states that

$$v(t,x) = \sup_{u \in \mathcal{U}} \mathbb{E}_{\mathbb{P}}\bigg(\int_t^\tau h(s, X_s^{t,x,u}, u_s)\, ds + v(\tau, X_\tau^{t,x,u}) \bigg) \tag{8.4}$$

for an arbitrary stopping time $\tau \in \mathcal{T}_{[t,T]}$. Using (8.1), (8.4) and the Itô formula, it is possible to derive the Hamilton–Jacobi–Bellman (HJB) equation.

Definition 8.3. Let $w : [0,T] \times \mathbb{R}^d \to \mathbb{R}$. The *Hamilton–Jacobi–Bellman equation* for the problem (8.2) has the following form

$$\frac{\partial w}{\partial t}(t,x) + \frac{1}{2} \sum_{i,j=1}^d a_{ij}(t,x) \frac{\partial^2 w}{\partial x_i \partial x_j}(t,x) + \sup_{u \in U} \bigg(\sum_{i=1}^d \mu_i(s,x,u) \frac{\partial w}{\partial x_i}(t,x) + h(t,x,u) \bigg) = 0$$

with the terminal $w(T,x) = \Psi(x)$.

Let us sketch the standard derivation of the HJB equation from the DPP. We fix (t,x) and we consider the process $X = X^{t,x,u}$. We assume that a (yet unknown) value function v is sufficiently regular, say of class $C^{1,2}([0,T]\times\mathbb{R}^d)$. Then, by an application of the Itô formula, we obtain for all $s\in[t,T]$

$$\mathbb{E}_{\mathbb{P}}(v(s,X_s)) = v(t,x) + \mathbb{E}_{\mathbb{P}}\bigg(\int_t^s \frac{\partial v}{\partial t}\,(z,X_z)\,dz\bigg)$$
$$+\,\mathbb{E}_{\mathbb{P}}\bigg(\int_t^s \Big(\frac{1}{2}\sum_{i,j=1}^d a_{ij}(z,X_z)\frac{\partial^2 v}{\partial x_i\partial x_j}\,(z,X_z) + \sum_{i=1}^d \mu_i(z,X_z,u_z)\frac{\partial v}{\partial x_i}\,(z,X_z)\Big)\,dz\bigg).$$

By combining this equality with the Bellman principle with $\tau = s$, we obtain

$$\mathbb{E}_{\mathbb{P}}\bigg(\int_t^s \Big(\frac{\partial v}{\partial t}\,(z,X_z) + \frac{1}{2}\sum_{i,j=1}^d a_{ij}(z,X_z)\frac{\partial^2 v}{\partial x_i\partial x_j}\,(z,X_z)\Big)\,dz\bigg)$$
$$+\,\mathbb{E}_{\mathbb{P}}\bigg(\sup_{u\in\mathcal{U}}\Big(\int_t^s\Big(\sum_{i=1}^d \mu_i(z,X_z,u_z)\frac{\partial v}{\partial x_i}\,(z,X_z) + h(z,X_z,u_z)\Big)\,dz\bigg) = 0$$

By dividing by $s-t$, assuming that s tends to t and replacing $\sup_{u\in\mathcal{U}}$ by $\sup_{u\in U}$ and X_z by x, we obtain the HJB equation, meaning that the HJB equation is simply an infinitesimal version of the DPP. It is clear that the derivation of the HJB equation was informal, but this is not a serious issue, since any result obtained using this equation needs to be validated by a direct proof. Also, in some models, it can be formally proven that the HJB equation is necessarily satisfied by the value function.

Before we formulate the main result in this section, we find it convenient to introduce the concept of the Hamiltonian function.

Definition 8.4. The *Hamiltonian function* $H:[0,T]\times\mathbb{R}^d\times\mathbb{R}^d\to\mathbb{R}$ is defined by

$$H(t,x,p) = \sup_{u\in U}\big(\mu(t,x,u)p + h(t,x,u)\big)$$

and we set $\Gamma(t,x,p) = \arg\max_{u\in U} H(t,x,p)$ or, more explicitly,

$$\Gamma(t,x,p) = \big\{u\in U : \mu(t,x,u)p + h(t,x,u) = H(t,x,p)\big\}.$$

Note that the HJB can be represented as follows

$$\frac{\partial v}{\partial t}(t,x) + \frac{1}{2}\,\mathbf{Trace}\big(a(t,x)\Delta_x v(t,x)\big) + H\big(t,x,\nabla_x v(t,x)\big) = 0.$$

Lemma 8.1. *Assume that the HJB equation has a unique solution, which is sufficiently regular. Then:*
(i) for every control process $u\in\mathcal{U}$ *we have* $v(t,x)\geq J(t,x,u)$*,*
(ii) the equality $v(t,x) = J(t,x,u)$ *is valid if* u *satisfies,* $\mathbb{P}$*-a.s. for almost all* $s\in[t,T]$*,*

$$u_s \in \Gamma\Big(s, X_s^{t,x,u}, \frac{\partial v}{\partial x}\,\big(s, X_s^{t,x,u}\big)\Big).$$

Consequently, if a process $u_t^* = u^*(t,X_t^{u^*})$ *is such that,* $\mathbb{P}$*-a.s. for almost all* $t\in[0,T]$*,*

$$u_t^* \in \Gamma\Big(t, X_t^{u^*}, \frac{\partial v}{\partial x}\,\big(t, X_t^{u^*}\big)\Big),$$

then $J_0(u^*) \geq J_0(u)$ *for every* $u \in \mathcal{U}$ *and thus* u^* *is optimal and* $J_0(u^*) = v(0, x_0)$ *where* x_0 *is the initial condition in the SDE* (8.1).

In practice, the HJB equation should first be solved for v in order to find an explicit representation for the closed-loop Markovian optimal control $u^*(t,x)$ or, at least, one has to prove a suitable regularity of a solution to the HJB equation. Subsequently, one has to prove the verification theorem, that is, the result showing that the process u^* is indeed optimal in the class $\mathcal{U}$ of all admissible control processes.

Theorem 8.1 (Verification Theorem)**.** *Assume that:*
(i) there exists a function $w \in C^{1,2}([0,T] \times \mathbb{R}^d)$ *satisfying a suitable growth condition, which is a solution to the HJB equation;*
(ii) there exists a Borel measurable mapping $u^* : [0,T] \times \mathbb{R}^d \to \mathbb{R}$ *such that the following equality holds*

$$\frac{\partial w}{\partial t}(t,x) + \frac{1}{2}\sum_{i,j=1}^{d} a_{ij}(t,x)\frac{\partial^2 w}{\partial x_i \partial x_j}(t,x) + \sum_{i=1}^{d} \mu_i(s,x,u^*(t,x))\frac{\partial w}{\partial x_i}(t,x) + h(t,x,u^*(t,x)) = 0,$$

(iii) for every (t,x)*, the SDE on* $[t,T]$

$$dX_s = \mu(s, X_s, u^*(s,X_s))\, ds + \sigma(s, X_s)\, dW_s, \quad X_t = x, \tag{8.5}$$

has a unique strong solution, denoted as X^{t,x,u^*}*. Then the following statements hold:*
(i) $w = v$*, that is,* w *is the value function;*
(ii) for every (t,x)*, the process* $u^*(s, X_s^{t,x,u^*}),\ s \in [t,T]$ *is an optimal control on* $[t,T]$.

Proof. Let us fix (t,x). From (a) and (b), we deduce that for any process $u \in \mathcal{U}$ we have

$$\frac{\partial w}{\partial t}(t,x) + \frac{1}{2}\sum_{i,j=1}^{d} a_{ij}(t,x)\frac{\partial^2 w}{\partial x_i \partial x_j}(t,x) + \sum_{i=1}^{d} \mu_i(s,x,u_t)\frac{\partial w}{\partial x_i}(t,x) + h(t,x,u_t) \leq 0,$$

and for the mapping $u^*(t,x)$ we get

$$\frac{\partial w}{\partial t}(t,x) + \frac{1}{2}\sum_{i,j=1}^{d} a_{ij}(t,x)\frac{\partial^2 w}{\partial x_i \partial x_j}(t,x) + \sum_{i=1}^{d} \mu_i(s,x,u^*(t,x))\frac{\partial w}{\partial x_i}(t,x) + h(t,x,u^*(t,x)) = 0$$

From the Itô formula on $[t,T]$ applied to the solution to SDEs (8.3) and (8.5), for any process $u \in \mathcal{U}$, we obtain for the process $w(s,X_s),\ s \in [t,T]$ where $X = X^{t,x,u}$

$$\begin{aligned}
&\mathbb{E}_{\mathbb{P}}(\Psi(X_T)) = \mathbb{E}_{\mathbb{P}}(w(T,X_T)) = w(t,x) + \mathbb{E}_{\mathbb{P}}\bigg(\int_t^T \Big(\frac{\partial w}{\partial t}(s,X_s)\, ds\bigg) \\
&+ \mathbb{E}_{\mathbb{P}}\bigg(\int_t^T \Big(\frac{1}{2}\sum_{i,j=1}^{d} a_{ij}(s,X_s)\frac{\partial^2 w}{\partial x_i \partial x_j}(s,X_s) + \sum_{i=1}^{d} \mu_i(z,X_s,u_s)\frac{\partial w}{\partial x_i}(s,X_s)\Big)\, ds\bigg) \\
&\leq w(t,x) - \mathbb{E}_{\mathbb{P}}\bigg(\int_t^T h(s,X_s,u_s)\, ds\bigg)
\end{aligned}$$

and for the processes $X = X^{t,x,u^*}$ and $u^*_s = u^*(s, X_s^{t,x,u^*})$ we get equality. This corresponds to the supermartingale and the martingale property of $w(t, X_t) - \int_0^t h(s, X_s, u_s)\, ds$ for u and u^*, respectively. We conclude that

$$w(t,x) = \mathbb{E}_{\mathbb{P}}\bigg(\int_t^T h(s, X_s^{t,x,u^*}, u^*_s)\, ds + \Psi(X_T) \bigg) \geq \mathbb{E}_{\mathbb{P}}\bigg(\int_t^T h(s, X_s^{t,x,u}, u_s)\, ds + \Psi(X_T) \bigg)$$

for every process $u \in \mathcal{U}$. Hence $w(t,x) = v(t,x)$ and u^* is optimal on $[t,T]$. □

8.2 Merton's Portfolio Selection via HJB Equation

We consider the classical Black-Scholes model where W be a one-dimensional standard Brownian motion on the underlying probability space $(\Omega, \mathbb{F}, \mathbb{P})$. We postulate that the price S^0 of a non-risky asset satisfies

$$dS_t^0 = rS_t^0\, dt, \quad S_0^0 = 1, \tag{8.6}$$

for some constant r and the stock price $S = S^1$ is given as a unique solution to the SDE

$$dS_t^1 = S_t^1\big(\mu\, dt + \sigma\, dW_t\big), \quad S_0^1 > 0, \tag{8.7}$$

for some constants μ and $\sigma > 0$. Let φ_t^i stand for the number of shares of the ith asset held at time t. Then $\pi_t^i := \varphi_t^i S_t^i$ is the cash value of our investment in the ith asset at time t. We may thus identify a *portfolio* with a two-dimensional, $\mathbb{F}$-progressively measurable process $\pi = (\pi^0, \pi^1)$. As in Definition 5.4, the wealth process V^π is given as $V_t^\pi = \sum_{i=0}^1 \pi_t^i$ and we say that a portfolio π is self-financing if for all $t \in [0,T]$

$$V_t^\pi = V_0^\pi + \sum_{i=0}^1 \int_0^t \varphi_s^i\, dS_s^i = V_0^\pi + \int_0^t \pi_s^0 r_s\, ds + \int_0^t \pi_s^1 \big(\mu\, ds + \sigma\, dW_s\big).$$

It is easy to check that the wealth process V^π satisfies

$$V_t^\pi = V_0^\pi + \int_0^t rV_s^\pi\, ds + \int_0^t \pi_s^1(\mu - r)\, ds + \int_0^t \pi_s^1 \sigma\, dW_s \tag{8.8}$$

and hence the dynamics of the wealth process depends on a choice of π^1.

If we denote by α_t the proportion of wealth invested in the risky asset S at time t and by $1 - \alpha_t$ the proportion of wealth invested in the savings account S^0 at time t, then we obtain an equivalent representation

$$V_t^\alpha = V_0^\alpha + \int_0^t rV_s^\alpha\, ds + \int_0^t \alpha_s(\mu - r)V_s^\alpha\, ds + \int_0^t \alpha_s V_s^\alpha \sigma\, dW_s \tag{8.9}$$

for every $\mathbb{F}$-progressively measurable process α belonging to the class $\mathcal{A}$ of all admissible control processes.

We assume that the preferences of an agent are described by a strictly increasing and strictly concave utility function and the performance of a portfolio is measured by the expected utility of terminal wealth so that the value function is given by

$$v(t,x) = \sup_{\alpha \in \mathcal{A}} \mathbb{E}_{\mathbb{P}}\big[U(V_T^\alpha)\big], \quad \forall (t,x) \in [0,T] \times (0,\infty),$$

where $\mathcal{A}$ is the set of all admissible controls.

Lemma 8.2. *The HJB equation for Merton's portfolio optimisation problem reads*

$$w_t + rxw_x + \sup_{\alpha\in A}\left[\alpha(\mu-r)xw_x + \frac{1}{2}\alpha^2\sigma^2x^2w_{xx}\right] = 0, \quad \forall\,(t,x)\in[0,T)\times(0,\infty),$$

with the terminal condition $w(T,x)=U(x)$ *for all* $x>0$.

We will find an explicit solution for the power utility function

$$U(x) = \frac{x^\gamma}{\gamma}, \quad \gamma<1,\ \gamma\neq 0.$$

which has the property of Constant Relative Risk Aversion (CRRA), in the sense that

$$-\frac{xU''(x)}{U'(x)} = 1-\gamma.$$

We search for a solution $w(t,x)$ in the form $w(t,x)=\varphi(t)U(x)$. Plugging into HJB, we see that the unknown function φ should satisfy the ODE

$$\varphi'(t)+\beta\varphi(t)=0, \quad \varphi(T)=1,$$

where

$$\beta = r\gamma + \gamma\sup_{\alpha\in A}\left[\alpha(\mu-r)-\frac{1}{2}\alpha^2(1-\gamma)\sigma^2\right].$$

Then $\varphi(t)=\exp(\beta(T-t))$ and thus a candidate for the value function is $v(t,x)=\exp(\beta(T-t))U(x)$. Furthermore, the optimal control is a constant proportion of the wealth invested in risky asset

$$\alpha^* = \arg\max_{\alpha\in A}\left[\alpha(\mu-r)-\frac{1}{2}\alpha^2(1-\gamma)\sigma^2\right].$$

Proposition 8.1. *In the unconstrained case where* $A=\mathbb{R}$*, the solution to Merton's portfolio optimisation problem with power utility function is given by* $v(t,x)=\varphi(t)U(x)$ *where* $\varphi(t)=\exp(\beta(T-t))$ *with*

$$\beta = \frac{(\mu-r)^2}{2\sigma^2}\frac{\gamma}{1-\gamma}+\gamma r$$

and the optimal proportion process α^* *is constant*

$$\alpha_t^* = \frac{\mu-r}{\sigma^2(1-\gamma)}, \quad \forall\, t\in[0,T].$$

It is worth noting that for $U(x)=\log x$ the optimal proportion is also constant and equals

$$\alpha_t^* = \frac{\mu-r}{\sigma^2}, \quad \forall\, t\in[0,T].$$

This can be shown directly without using the HJB equation.

8.3 Optimal Control via BSDE Approach

Let $\mu : [0,T] \times \mathbb{R}^d \times U \to \mathbb{R}^d$, $h : [0,T] \times \mathbb{R}^d \times U \to \mathbb{R}$ and $\Psi : \mathbb{R}^d \to \mathbb{R}$ be given bounded functions, which are continuous in u. Let us first consider the strong formulation of the *stochastic optimal control* problem. It is worth stressing that in the strong formulation we need to impose additional regularity conditions on μ and σ to ensure the existence and uniqueness of a strong solution to SDE (8.10).

Strong formulation. Let $(W_t)_{t\in[0,T]}$ be an $\mathbb{R}^d$-valued standard Brownian motion on $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P})$. We assume that $\mathbb{F} = (\mathcal{F}_t)_{t\in[0,T]}$ is the filtration generated by W and augmented by the class of $\mathbb{P}$-null sets of $\mathcal{F}_T$. For any process $u \in \mathcal{U}$, we consider the following d-dimensional SDE

$$dX_t^u = \mu(t, X_t^u, u_t)\,dt + \sigma(t, X_t^u)\,dW_t, \quad X_0^u = x_0 \in \mathbb{R}^d, \tag{8.10}$$

and we assume that it admits a unique strong solution. The goal is to find a process $u^* \in \mathcal{U}$ such that $J_0(u^*) \geq J_0(u)$ for every process $u \in \mathcal{U}$ where

$$J_0(u) := \mathbb{E}_{\mathbb{P}}\bigg(\int_0^T h(s, X_s^u, u_s)\,ds + \Psi(X_T^u) \bigg).$$

More generally, we denote by $J(u)$ the process given by the following expression

$$J_t(u) := \mathbb{E}_{\mathbb{P}}\bigg(\int_t^T h(s, X_s^u, u_s)\,ds + \Psi(X_T^u) \,\Big|\, \mathcal{F}_t \bigg).$$

Weak formulation. It is no longer necessary to assume that $\mathbb{F} = (\mathcal{F}_t)_{t\in[0,T]}$ is generated by the Brownian motion W. We instead assume that W is a Brownian motion with respect to $\mathbb{F}$. We start by considering the following SDE

$$dX_t = \sigma(t, X_t)\,dW_t, \quad X_0 = x_0 \in \mathbb{R}^d, \tag{8.11}$$

and we assume that $\sigma : [0,T] \times \mathbb{R}^d \to \mathbb{R}^d$ is such that this equation admits a unique strong solution. We postulate, in addition, that the process $\sigma^{-1}(t, X_t)$ is bounded and, for an arbitrary process $u \in \mathcal{U}$, we define the following bounded process

$$\gamma_t^u := \sigma^{-1}(t, X_t)\mu(t, X_t, u_t).$$

We define the probability measure $\mathbb{P}^u$ on $(\Omega, \mathcal{F}_T)$ by postulating that the Radon-Nikodým density of $\mathbb{P}^u$ with respect to $\mathbb{P}$ equals

$$\eta_T^u = \frac{d\mathbb{P}^u}{d\mathbb{P}} := \mathcal{E}_T\Big(- \int_0^{\cdot} \gamma_t^u\,dW_t \Big) = \exp\,\Big(\int_0^T \gamma_t^u\,dW_t - \frac{1}{2}\int_0^T \|\gamma_t^u\|^2\,dt \Big), \quad \mathbb{P}-\text{a.s.}$$

Since the process γ^u is bounded, Novikov's condition is satisfied and thus $\mathbb{E}_{\mathbb{P}}(\eta_T^u) =$ for an arbitrary process $u \in \mathcal{U}$ (see Proposition 12.5). Therefore, the probability measure $\mathbb{P}^u$ is well defined. Moreover, from Girsanov's theorem (see Theorem 12.4), we know that the process $W_t^u := W_t - \int_0^t \gamma_s^u\,ds$ is a d-dimensional Brownian motion under $\mathbb{P}^u$. The following lemma is obvious.

Lemma 8.3. *For each process $u \in \mathcal{U}$, the process X given by SDE* (8.11) *under $\mathbb{P}$ is also a solution to the following SDE under $\mathbb{P}^u$*

$$dX_t = \mu(t, X_t, u_t)\,dt + \sigma(t, X_t)\,dW_t^u, \quad X_0 = x_0 \in \mathbb{R}^d. \tag{8.12}$$

For $t = 0$, we define the functional $J_0 : \mathcal{U} \to \mathbb{R}$

$$J_0(u) := \mathbb{E}_{\mathbb{P}^u}\bigg(\int_0^T h(s, X_s, u_s)\, ds + \Psi(X_T) \bigg) \tag{8.13}$$

and, more generally, for every $t \in [0, T]$ we set

$$J_t(u) := \mathbb{E}_{\mathbb{P}^u}\bigg(\int_t^T h(s, X_s, u_s)\, ds + \Psi(X_T) \,\bigg|\, \mathcal{F}_t \bigg). \tag{8.14}$$

The goal in the weak formulation is to find an optimal control process $u^* \in \mathcal{U}$ such that $J_0(u^*) \geq J_0(u)$ or, more generally, $J_t(u^*) \geq J_t(u)$ for every $t \in [0, T]$ and every process $u \in \mathcal{U}$. The following definition will be useful.

Definition 8.5. The *value process* J^* is given by $J^*_t := \operatorname{ess\,sup}_{u \in \mathcal{U}} J_t(u)$ for every $t \in [0, T]$. In particular, $J^*_0 := \sup_{u \in \mathcal{U}} J_0(u)$.

It is clear that necessarily $J^*_0 = J_0(u^*)$ and $J^*_t = J_t(u^*)$, provided that an optimal control process u^* exists. Note, however, that neither the existence nor the uniqueness of an optimal control process are ensured. By contrast, the process X is fixed in what follows as a unique solution to (8.11). The mapping H introduced in the next lemma is called the *Hamiltonian process.* For the proof of the lemma, see Exercise 2(a) in Assignment 3.

Lemma 8.4. *For any process $u \in \mathcal{U}$, let (Y^u, Z^u) be the unique solution in $\mathcal{S}^2_1(0, T) \times \mathcal{H}^2_d(0, T)$ to the linear BSDE with the generator $H : [0, T] \times \Omega \times \mathbb{R}^d \to \mathbb{R}$ given by*

$$H(t, X_t, z, u_t) := z\sigma^{-1}(t, X_t)\mu(t, X_t, u_t) + h(t, X_t, u_t)$$

and the terminal condition $\Psi(X_T)$ or, more explicitly,

$$-dY^u_t = H(t, X_t, Z^u_t, u_t)\, dt - Z^u_t\, dW_t, \quad Y^u_T = \Psi(X_T), \tag{8.15}$$

where W is a d-dimensional Brownian motion under $\mathbb{P}$. Then we have $J(u) = Y^u$.

Let us define the mapping $H^* : [0, T] \times \Omega \times \mathbb{R}^d \to \mathbb{R}$ by

$$H^*(t, X_t, z) := \operatorname*{ess\,sup}_{u \in U} H(t, X_t, z, u), \quad \forall\, (t, z) \in [0, T] \times \mathbb{R}^d. \tag{8.16}$$

Lemma 8.5. *There exists a measurable mapping $u^* : [0, T] \times \mathbb{R}^d \times \mathbb{R}^d \to U$, denoted as $u^*(t, X_t, z)$, such that for every $t \in [0, T]$ and $z \in \mathbb{R}^d$*

$$H^*(t, X_t, z) := \operatorname*{ess\,sup}_{u \in U} H(t, X_t, z, u) = H(t, X_t, z, u^*(t, X_t, z)).$$

Proof. Since the function $u \to H(t, X_t, Z^*_t, u)$ is continuous on a compact set U, the existence of u^* follows from the *measurable selection theorem* due to Beneš [4]. The proof of the measurable selection theorem is technical and thus it is omitted. It relies on proving that there exists a measurable version of $\arg\max_{u \in U} H(t, X_t, z, u)$ when U is a compact set and $H(t, X_t, Z^*_t, u)$ is continuous in u. $\square$

Lemma 8.6. *The generator H^* satisfies assumption (H.1) and is a convex function with respect to z. Hence the BSDE*

$$-dY_t^* = H^*(t, X_t, Z_t^*)\, dt - Z_t^*\, dW_t, \quad Y_T^* = \Psi(X_T), \tag{8.17}$$

has a unique solution (Y^*, Z^*) in $\mathcal{S}_1^2(0,T) \times \mathcal{H}_d^2(0,T)$. Moreover, $Y_t^* = Y_t^{u^*} = J_t(u^*)$ for every $t \in [0,T]$ where $u_t^* = u^*(t, X_t, Z_t^*)$.

Proof. The proof of the lemma hinges on Theorem 1.1 for BSDEs with Lipschitz continuous generators. See Exercise 2(b) in Assignment 3. □

Lemma 8.7. *For any process $u \in \mathcal{U}$, let (Y^u, Z^u) be the unique solution in $\mathcal{S}_1^2(0,T) \times \mathcal{H}_d^2(0,T)$ to BSDE (8.15). If (Y^*, Z^*) is the unique solution in $\mathcal{S}_1^2(0,T) \times \mathcal{H}_d^2(0,T)$ to BSDE (8.17), then $Y_t^* \geq Y_t^u$ for every $t \in [0,T]$.*

Proof. It suffices to check that the assumptions of Proposition 3.1 are satisfied. The validity of assumption (i) is clear. For (ii), we note that H^* is a Lipschitz continuous function with respect to z, since it satisfies (H.1) (see Lemma 8.6). For (iii), we observe that

$$H^*(t, X_t, Z_t^u) - H(t, X_t, Z_t^u, u_t) \geq 0, \quad d\mathbb{P} \otimes dt - \text{a.e.},$$

and the process given above belongs to the space $\mathcal{H}_1^2(0,T)$ since $Z^u \in \mathcal{H}_1^2(0,T)$, the mappings H^* and H are Lipschitz continuous with respect to z, and the processes $H^*(t, X_t, 0)$ and $H(t, X_t, 0, u_t)$ are bounded (hence they belong to $\mathcal{H}_1^2(0,T)$). From Proposition 3.1, we deduce that $Y_t^* \geq Y_t^u$ for every $t \in [0,T]$. □

Lemma 8.8. *The equality $Y_t^* = J_t^*$ holds for every $t \in [0,T]$.*

Proof. From Lemmas 8.4 and 8.7, it is clear that $Y_t^* \geq Y_t^u = J_t(u)$ and thus $Y_t^* \geq J_t^*$ for all $t \in [0,T]$. To show that $Y_t^* \leq J_t^*$, we note that, from Lemma 8.6, there exists an admissible control process u^* such that $Y_t^* = J_t(u^*) = Y_t^{u^*} \leq J_t^*$ for all $t \in [0,T]$. □

Theorem 8.2. *Let U be a compact metric space and let $\mathcal{U}$ be the class of all $\mathbb{F}$-adapted processes with values in U. Assume that the functions $h : [0,T] \times \mathbb{R}^d \times U \to \mathbb{R}$ and $\Psi : \mathbb{R}^d \to \mathbb{R}$ are bounded and continuous in u. Let (Y^*, Z^*) be the unique solution to BSDE (8.17) with the generator H^* given by (8.16) and the terminal condition $\xi_T = \Psi(X_T)$. Then an optimal control process u^* exists and $Y_t^* = J_t(u^*) = J_t^*$ for every $t \in [0,T]$ where*

$$J_t^* := \operatorname{ess\,sup}_{u \in \mathcal{U}} \mathbb{E}_{\mathbb{P}^u}\bigg(\int_t^T h(s, X_s, u_s)\, ds + \Psi(X_T) \,\bigg|\, \mathcal{F}_t \bigg).$$

In particular,

$$Y_0^* = J_0(u^*) = J_0^* := \sup_{u \in \mathcal{U}} J_0(u)$$

where $J_0(u)$ is given by

$$J_0(u) := \mathbb{E}_{\mathbb{P}^u}\bigg(\int_0^T h(s, X_s, u_s)\, ds + \Psi(X_T) \bigg).$$

It is fair to acknowledge that the BSDE approach is clearly more general than the HJB method but a solution to a BSDE gives less explicit results than a solution to the HJB equation.

8.4 Stochastic Pontryagin's Maximum Principle

Let $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P})$ be a filtered probability space satisfying the usual hypotheses of right-continuity and completeness. Let $(W_t)_{t\in[0,T]}$ be an $\mathbb{R}^d$-valued standard Brownian motion on $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P})$. We assume that $\mathbb{F} = (\mathcal{F}_t)_{t\in[0,T]}$ is the filtration generated by W and augmented by the class of $\mathbb{P}$-null sets of $\mathcal{F}_T$. In this section, a real-valued process u is said to be an *admissible control* if it is $\mathbb{F}$-progressively measurable. We denote by $\mathcal{U}$ the class of all admissible controls. Let $\mu, \sigma : [0,T] \times \mathbb{R}^2 \to \mathbb{R}$ be such that for every $u \in \mathcal{U}$ the one-dimensional SDE

$$dX_t^u = \mu(t, X_t^u, u_t)\, dt + \sigma(t, X_t^u, u_t)\, dW_t, \quad X_0^u = x_0 \in \mathbb{R}, \tag{8.18}$$

admits a unique strong solution X^u in the space $\mathcal{S}_1^2(0,T)$. For more results on the approach presented in this section, we refer to Peng [75, 77, 78, 79].

Remark 8.1. For concreteness, one may postulate that μ and σ are of class C^1 with respect to (x, u) and that μ and σ as well as their derivatives are uniformly Lipschitz continuous functions. Then SDE (8.18) and BSDE (8.20) will be covered by Theorems 1.1 and 12.4, respectively.

Let $\Psi : \mathbb{R} \to \mathbb{R}$ be a given bounded function, which is of class $C^1(\mathbb{R})$ with a bounded derivative. The goal is to find a process $u^* \in \mathcal{U}$ such that $J_0(u^*) \geq J_0(u)$ for every process $u \in \mathcal{U}$ where $J_0(u) := \mathbb{E}_{\mathbb{P}}\big(\Psi(X_T^u)\big)$. In other words, the optimal control $u^* \in \mathcal{U}$ satisfies $J_0(u^*) = \max_{u\in\mathcal{U}} \mathbb{E}_{\mathbb{P}}\big(\Psi(X_T^u)\big)$.

The *Hamiltonian* H is defined as follows

$$H(t, x, u, p, q) := \mu(t, x, u)p + \sigma(t, x, u)q, \quad \forall\, (t, x, u, p, q) \in [0,T] \times \mathbb{R}^4. \tag{8.19}$$

The following result gives a sufficient condition for optimality of an admissible control u^*. In fact, under mild technical assumptions, it is also possible to show that condition (8.21) is a necessary condition for the optimality of a control process u^*. Equality (8.21) is referred to as the *stochastic Pontryagin maximum principle* since it gives conditions that an optimal stochastic trajectory must satisfy.

Theorem 8.3. *Assume that:*
(i) the function $\Psi : \mathbb{R} \to \mathbb{R}$ *is concave;*
(ii) $u^* \in \mathcal{U}$ *is such that there exists a pair* $(p, q) \in \mathcal{S}_1^2(0,T) \times \mathcal{H}_1^2(0,T)$ *solving the following BSDE*

$$p_t = \partial_x \Psi(X_T^{u^*}) + \int_t^T \partial_x H(s, X_s^{u^*}, u_s^*, p_s, q_s)\, ds - \int_t^T q_s\, dW_s, \tag{8.20}$$

and for all $t \in [0,T]$

$$H(t, X_t^{u^*}, u_t^*, p_t, q_t) = \max_{u\in\mathbb{R}} H(t, X_t^{u^*}, u, p_t, q_t) = H^*(t, X_t^{u^*}, p_t, q_t), \tag{8.21}$$

where

$$H^*(t, x, p_t, q_t) := \sup_{u\in\mathbb{R}} H(t, x, u, p_t, q_t), \tag{8.22}$$

(iii) for almost all $t \in [0,T]$*, the mapping* $x \mapsto H^*(t, x, p_t, q_t)$ *is concave.*
Then an admissible control $u^* \in \mathcal{U}$ *is optimal:* $J_0(u^*) = \max_{u\in\mathcal{U}} J_0(u)$.

Proof. **Step 1.** By assumption (i), the function Ψ is concave and thus for all $x, y \in \mathbb{R}$

$$\Psi(y) - \Psi(x) \geq \partial_x \Psi(y)(y - x).$$

Hence for arbitrary processes $u, u^* \in \mathcal{U}$

$$J_0(u^*) - J_0(u) = \mathbb{E}_{\mathbb{P}}\big(\Psi(X_T^{u^*}) - \Psi(X_T^u)\big) \geq \mathbb{E}_{\mathbb{P}}\big(\partial_x(\Psi(X_T^{u^*}))(X_T^{u^*} - X_T^u)\big).$$

Step 2. For a process $u^* \in \mathcal{U}$, let (p, q) be a solution to the *adjoint equation* (8.20). By applying the Itô integration by parts formula to the product $p_t(X_t^{u^*} - X_t^u)$ and noting that $X_0^{u^*} = X_0^u = x_0$, we obtain

$$\begin{aligned}
p_T(X_T^{u^*} - X_T^u) &= \int_0^T \big(\mu(t, X_t^{u^*}, u_t^*) - \mu(t, X_t^u, u_t)\big)p_t \, dt \\
&+ \int_0^T \big(\sigma(t, X_t^{u^*}, u_t^*) - \sigma(t, X_t^u, u_t)\big)q_t \, dt - \int_0^T \partial_x H(t, X_t^{u^*}, u_t^*, p_t, q_t)(X_t^{u^*} - X_t^u) \, dt \\
&+ \int_0^T \Big(\big(\sigma(t, X_t^{u^*}, u_t^*) - \sigma(t, X_t^u, u_t)\big)p_t + \big(X_t^{u^*} - X_t^u\big)q_t\Big) \, dW_t.
\end{aligned}$$

Under the present assumptions, we also have that

$$\mathbb{E}_{\mathbb{P}}\bigg(\int_0^T \Big(\big(\sigma(t, X_t^{u^*}, u_t^*) - \sigma(t, X_t^u, u_t)\big)p_t + \big(X_t^{u^*} - X_t^u\big)q_t\Big) \, dW_t\bigg) = 0.$$

Step 3. Since $p_T = \partial_x \Psi(X_T^{u^*})$, from Steps 1 and 2 and (8.19), we get

$$\begin{aligned}
&J_0(u^*) - J_0(u) \geq \mathbb{E}_{\mathbb{P}}\big(p_T(X_T^{u^*} - X_T^u)\big) \\
&= \mathbb{E}_{\mathbb{P}}\bigg(\int_0^T \Big(\big(\mu(t, X_t^{u^*}, u_t^*) - \mu(t, X_t^u, u_t)\big)p_t + \big(\sigma(t, X_t^{u^*}, u_t^*) - \sigma(t, X_t^u, u_t)\big)q_t\Big) dt\bigg) \\
&\quad - \mathbb{E}_{\mathbb{P}}\bigg(\int_0^T \partial_x H(t, X_t^{u^*}, u_t^*, p_t, q_t)(X_t^{u^*} - X_t^u) \, dt\bigg) \\
&= \mathbb{E}_{\mathbb{P}}\bigg(\int_0^T \big(H(t, X_t^{u^*}, u_t^*, p_t, q_t) - H(t, X_t^u, u_t, p_t, q_t)\big) \, dt\bigg) \\
&\quad - \mathbb{E}_{\mathbb{P}}\bigg(\int_0^T \partial_x H(t, X_t^{u^*}, u_t^*, p_t, q_t)(X_t^{u^*} - X_t^u) \, dt\bigg).
\end{aligned}$$

Step 4. Using (8.21) for u^* and (8.22) for u, we conclude that

$$\begin{aligned}
J_0(u^*) - J_0(u) &\geq \mathbb{E}_{\mathbb{P}}\bigg(\int_0^T \big(H^*(t, X_t^{u^*}, p_t, q_t) - H^*(t, X_t^u, p_t, q_t)\big) \, dt\bigg) \\
&\quad - \mathbb{E}_{\mathbb{P}}\bigg(\int_0^T \partial_x H^*(t, X_t^{u^*}, p_t, q_t)(X_t^{u^*} - X_t^u) \, dt\bigg) \geq 0
\end{aligned}$$

where the second inequality is a consequence of assumption (iii). □

Remark 8.2. Note that it suffices to assume that the function Ψ is concave and such that $\partial_x \Psi(X_T^{u^*})$ (i.e., the terminal condition in BSDE (8.20)) is a well-defined, square-integrable random variable.

Remark 8.3. It is possible to extend the method presented here to the case of the functional

$$J_0(u) := \mathbb{E}_{\mathbb{P}}\bigg(\int_0^T h(s, X_s^u, u_s)\, ds + \Psi(X_T^u) \bigg).$$

To this end, one should modify the definition of the Hamiltonian as follows

$$H(t,x,u,p,q) := \mu(t,x,u)p + \sigma(t,x,u)q + h(t,x,u), \quad \forall\, (t,x,u,p,q) \in [0,T] \times \mathbb{R}^4,$$

and make suitable adjustments in the proof of Theorem 8.3.

9 Stochastic Differential Games

Let $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P})$ be a filtered probability space satisfying the usual hypotheses of right-continuity and completeness. Let U and V be compact metric spaces. We define the classes of all admissible control processes for the two players.

Definition 9.1. We say that control processes u and v are *admissible* if the following conditions are satisfied:
(i) the process u takes values in U,
(ii) the process v takes values in V,
(iii) the processes u and v are $\mathbb{F}$-adapted.

We denote by $\mathcal{U}$ (resp., $\mathcal{V}$) the space of all admissible control processes for the first (resp., the second) player. Let

$$f : [0,T] \times \mathbb{R}^d \times U \times V \to \mathbb{R}^d, \quad h : [0,T] \times \mathbb{R}^d \times U \times V \to \mathbb{R}$$

be given bounded functions, which are continuous in (u, v). Also, let $\Psi : \mathbb{R}^d \to \mathbb{R}$ and $h : [0,T] \times \mathbb{R}^d \times U \times V \to \mathbb{R}^d$ be bounded, continuous functions.

Let us first consider the strong formulation of the *stochastic differential game*. It is worth stressing that in the strong formulation we need to impose additional regularity conditions on the functions f and σ in order to ensure the existence and uniqueness of a strong solution to SDE (9.1).

Strong formulation. Let $(W_t)_{t\in[0,T]}$ be an $\mathbb{R}^d$-valued standard Brownian motion on $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P})$. We assume that $\mathbb{F} = (\mathcal{F}_t)_{t\in[0,T]}$ is the filtration generated by W and augmented by the class of $\mathbb{P}$-null sets of $\mathcal{F}_T$. For any process $u \in \mathcal{U}$ and $v \in \mathcal{V}$, we consider the following d-dimensional SDE

$$dX_t^{u,v} = f(t, X_t^{u,v}, u_t, v_t)\, dt + \sigma(t, X_t^{u,v})\, dW_t, \quad X_0^{u,v} = x_0 \in \mathbb{R}^d, \tag{9.1}$$

and we assume that the coefficients

$$f : [0,T] \times \mathbb{R}^d \times U \times V \to \mathbb{R}^d, \quad \sigma : [0,T] \times \mathbb{R}^d \to \mathbb{R}^d$$

are such that SDE (9.1) has a unique strong solution. We define the functional $J_0 : \mathcal{U} \times \mathcal{V} \to \mathbb{R}$

$$J_0(u,v) := \mathbb{E}_{\mathbb{P}}\bigg(\int_0^T h(s, X_s^{u,v}, u_s, v_s)\, ds + \Psi(X_T^{u,v}) \bigg).$$

More generally, we denote by $J(u,v)$ the process given by the following expression

$$J_t(u,v) := \mathbb{E}_{\mathbb{P}}\bigg(\int_t^T h(s, X_s^{u,v}, u_s, v_s)\, ds + \Psi(X_T^{u,v}) \,\Big|\, \mathcal{F}_t \bigg).$$

We consider a zero-sum two-player differential game in which the first (resp., the second) player aims to maximise (resp., to minimise) the expected payoff $J_0(u,v)$. Our goal is to find a pair of process $u^* \in \mathcal{U}$ and $u^* \in \mathcal{U}$ such that (u^*, v^*) is an *optimal equilibrium* (see Definition 7.2), meaning that

$$J_0(u, v^*) \le J_0(u^*, v^*) \le J_0(u^*, v), \quad \forall\, u \in \mathcal{U},\ v \in \mathcal{V}, \tag{9.2}$$

and to compute the value of the game, that is, the process $J_t(u^*, v^*)$, $t \in [0,T]$.

Weak formulation. We no longer postulate that $\mathbb{F} = (\mathcal{F}_t)_{t\in[0,T]}$ is generated by the Brownian motion W, but we maintain the assumption that W is a Brownian motion with respect to $\mathbb{F}$. As in Section 8, we postulate that X is the unique (strong or weak) solution to SDE

$$dX_t = \sigma(t, X_t)\, dW_t, \quad X_0 = x_0 \in \mathbb{R}^d, \tag{9.3}$$

and we assume that $\sigma : [0,T] \times \mathbb{R}^d \to \mathbb{R}^d$ is such that this equation admits a unique (either strong or weak) solution. We postulate, in addition, that the process $\sigma^{-1}(t, X_t)$ is bounded and, for arbitrary processes $u \in \mathcal{U}$ and $v \in \mathcal{V}$, we define the following bounded process

$$\gamma_t^{u,v} := \sigma^{-1}(t, X_t) f(t, X_t, u_t, v_t).$$

The probability measure $\mathbb{P}^{u,v}$ on $(\Omega, \mathcal{F}_T)$ is given by the Radon-Nikodým density with respect to $\mathbb{P}$

$$\eta_T^{u,v} = \frac{d\mathbb{P}^{u,v}}{d\mathbb{P}} := \mathcal{E}_T\Big(- \int_0^\cdot \gamma_t^{u,v}\, dW_t \Big) = \exp\,\Big(\int_0^T \gamma_t^{u,v}\, dW_t - \frac{1}{2} \int_0^T \|\gamma_t^{u,v}\|^2\, dt \Big), \quad \mathbb{P}-\text{a.s.}$$

Since the process $\gamma^{u,v}$ is bounded, Novikov's condition is satisfied and thus $\mathbb{E}_{\mathbb{P}}(\eta_T^{u,v}) = 1$ for arbitrary processes $u \in \mathcal{U}$ and $v \in \mathcal{V}$ (see Proposition 12.5). Therefore, the probability measure $\mathbb{P}^{u,v}$ is well defined. Moreover, from Girsanov's theorem (see Theorem 12.4), we know that the process $W^{u,v}$ given by

$$W_t^{u,v} := W_t - \int_0^t \gamma_s^{u,v}\, ds, \quad \forall\, t \in [0,T],$$

is a d-dimensional Brownian motion under $\mathbb{P}^{u,v}$. The following lemma is an immediate consequence of Girsanov's theorem.

Lemma 9.1. *For each process $u \in \mathcal{U}$, the process X given by SDE* (8.1) *under $\mathbb{P}$ is also a solution to the following SDE under $\mathbb{P}^u$*

$$dX_t = f(t, X_t, u_t, v_t)\, dt + \sigma(t, X_t)\, dW_t^{u,v}, \quad X_0 = x_0 \in \mathbb{R}^d. \tag{9.4}$$

We now define for $t = 0$

$$J_0(u, v) := \mathbb{E}_{\mathbb{P}^{u,v}} \bigg(\int_0^T h(s, X_s, u_s, v_s)\, ds + \Psi(X_T) \bigg) \tag{9.5}$$

and for every $t \in [0,T]$

$$J_t(u, v) := \mathbb{E}_{\mathbb{P}^{u,v}} \bigg(\int_t^T h(s, X_s, u_s, v_s)\, ds + \Psi(X_T) \,\Big|\, \mathcal{F}_t \bigg). \tag{9.6}$$

Definition 9.2. The *value process* J^* is given by

$$J_t^* := \operatorname*{ess\,sup}_{u\in\mathcal{U}} \operatorname*{ess\,inf}_{v\in\mathcal{V}} J_t(u,v) = \operatorname*{ess\,inf}_{v\in\mathcal{V}} \operatorname*{ess\,sup}_{u\in\mathcal{U}} J_t(u,v),$$

provided that the second equality holds. In particular, the *value* of the game is given by

$$J_0^* := \sup_{u\in\mathcal{U}} \inf_{v\in\mathcal{V}} J_0(u,v) = \inf_{v\in\mathcal{V}} \sup_{u\in\mathcal{U}} J_0(u,v),$$

provided that the second equality is valid.

Our goal in the weak formulation can be stated as follows.

Differential game problem. Find an optimal equilibrium for the diferential game, that is, a pair $(u^*, V^*) \in \mathcal{U} \times \mathcal{V}$ such that

$$J_0(u, v^*) \leq J_0(u^*, v^*) \leq J_0(u^*, v), \quad \forall\, u \in \mathcal{U},\, v \in \mathcal{V}.$$

It is clear that necessarily $J_0^* = J_0(u^*, v^*)$ and $J_t^* = J_t(u^*, v^*)$ for every $t \in [0, T]$, provided that an optimal equilibrium (u^*, v^*) exists. The mapping H introduced in the next lemma extends the concept of the Hamiltonian process introduced in Section 8.

Lemma 9.2. *For arbitrary processes $u \in \mathcal{U}$ and $v \in \mathcal{V}$, we denote by $(Y^{u,v}, Z^{u,v})$ the unique solution to the linear BSDE with the generator $H : [0, T] \times \Omega \times \mathbb{R}^d \to \mathbb{R}$ given by*

$$H(t, X_t, z, u_t, v_t) := z\sigma^{-1}(t, X_t) f(t, X_t, u_t, v_t) + h(t, X_t, u_t, v_t)$$

and the terminal condition $\Psi(X_T)$ or, more explicitly,

$$-dY_t^{u,v} = H(t, X_t, Z_t^{u,v}, u_t, v_t)\, dt - Z_t^{u,v}\, dW_t, \quad Y_T^{u,v} = \Psi(X_T), \tag{9.7}$$

where W is a d-dimensional Brownian motion under $\mathbb{P}$. Then we have $J(u, v) = Y^{u,v}$.

Proof. The proof of the lemma is left as an easy exercise. □

We henceforth assume that the following *Isaacs condition* is satisfied.

Definition 9.3. We say that the *Isaacs condition* is satisfied if for every $t \in [0, T]$

$$\operatorname*{ess\,sup}_{u \in U} \operatorname*{ess\,inf}_{v \in V} H(t, X_t, z, u, v) = \operatorname*{ess\,inf}_{v \in V} \operatorname*{ess\,sup}_{u \in U} H(t, X_t, z, u, v). \tag{9.8}$$

Let us define the mapping $H^* : [0, T] \times \Omega \times \mathbb{R}^d \to \mathbb{R}$ by

$$H^*(t, X_t, z) := \operatorname*{ess\,inf}_{v \in V} \operatorname*{ess\,sup}_{u \in U} H(t, X_t, z, u, v), \quad \forall\, (t, z) \in [0, T] \times \mathbb{R}^d. \tag{9.9}$$

The next result is technical, so we take it for granted without a proof.

Lemma 9.3. *There exist two measurable functions $u^*(t, X_t, \cdot)$ and $v^*(t, X_t, \cdot)$ with values in U and V, respectively such that*

$$H^*(t, X_t, z) = H(t, X_t, z, u^*(t, X_t, z), v^*(t, X_t, z)) \tag{9.10}$$

and for any $u \in U, v \in V, z \in \mathbb{R}^d$ and $t \in [0, T]$

$$H(t, X_t, z, u, v^*(t, X_t, z)) \leq H(t, X_t, z, u^*(t, X_t, z), v^*(t, X_t, z)) \leq H(t, X_t, z, u^*(t, X_t, z), v),$$

that is,

$$H(t, X_t, z, u, v^*(t, X_t, z)) \leq H^*(t, X_t, z) \leq H(t, X_t, z, u^*(t, X_t, z), v). \tag{9.11}$$

The proof of the next lemma hinges on Theorem 1.1 for BSDEs with Lipschitz continuous generators and the comparison property of solutions to BSDEs.

Lemma 9.4. *The generator H^* satisfies assumption (H.1) and is an affine function with respect to z. Hence the BSDE*

$$-dY_t^* = H^*(t, X_t, Z_t^*)\,dt - Z_t^*\,dW_t, \quad Y_T^* = \Psi(X_T), \tag{9.12}$$

has a unique solution $(Y^*, Z^*) \in \mathcal{S}_1^2(0,T) \times \mathcal{H}_d^2(0,T)$. Moreover, for every $u \in \mathcal{U}, v \in \mathcal{V}$ and $t \in [0,T]$

$$J_t(u, v^*) \le Y_t^* = J_t(u^*, v^*) \le J_t(u^*, v)$$

so that (u^*, v^*) is a saddle point for the game.

Proof. We first check that the assumptions of Theorem 1.1 and Proposition 3.1 are satisfied. The validity of assumption (i) in Proposition 3.1 is clear, since $\xi_T^1 = \xi_T^2 = \Psi(X_T)$. For assumption (ii), we note that H^* is a Lipschitz continuous function with respect to z and thus it satisfies condition (H.1). Finally, to check assumption (iii), we note that for any admissible control u

$$H^*(t, X_t, Z_t^{u,v^*}) - H(t, X_t, Z_t^{u,v^*}, u_t, v_t^*) \ge 0, \quad d\mathbb{P} \otimes dt - \text{a.e.}$$

Moreover, the process given above is bounded and thus it belongs to the space $\mathcal{H}_1^2(0,T)$. From Proposition 3.1, we deduce that $Y_t^* \ge Y_t^{u,v^*}$ for every $t \in [0,T]$. Similarly, for any admissible control v, we have

$$H^*(t, X_t, Z_t^{u^*,v}) - H(t, X_t, Z_t^{u^*,v}, u_t^*, v_t) \le 0, \quad d\mathbb{P} \otimes dt - \text{a.e.}$$

and thus, again from Proposition 3.1, we deduce that $Y_t^* \le Y_t^{u^*,v}$ for every $t \in [0,T]$. Note that the strict comparison property is also valid. $\square$

Theorem 9.1. *Let U and V be compact metric spaces and let $\mathcal{U}$ and $\mathcal{V}$ be the classes of all $\mathbb{F}$-adapted processes with values in U and V, respectively. Assume that the functions*

$$f : [0,T] \times \mathbb{R}^d \times U \times V \to \mathbb{R}^d, \quad h : [0,T] \times \mathbb{R}^d \times U \times V \to \mathbb{R}$$

are bounded and continuous in (u,v) and $\Psi : \mathbb{R}^d \to \mathbb{R}$ is bounded and continuous. Let (Y^*, Z^*) be the unique solution to BSDE (9.12) *with the generator H^* given by* (8.16) *and the terminal condition $\xi_T = \Psi(X_T)$. Then an optimal equilibrium (u^*, v^*) exists and $Y_t^* = J_t(u^*, v^*) = J_t^*$ for every $t \in [0,T]$ where*

$$J_t^* := \operatorname*{ess\,inf}_{v \in \mathcal{V}} \operatorname*{ess\,sup}_{u \in \mathcal{U}} \mathbb{E}_{\mathbb{P}^{u,v}} \left(\int_t^T h(s, X_s, u_s)\,ds + \Psi(X_T) \,\Big|\, \mathcal{F}_t \right).$$

In particular,

$$Y_0^* = J_0(u^*, v^*) = J_0^* := \inf_{v \in \mathcal{V}} \sup_{u \in \mathcal{U}} J_0(u,v) = \sup_{u \in \mathcal{U}} \inf_{v \in \mathcal{V}} J_0(u,v)$$

where $J_0(u,v)$ is given by (9.5)*.*

Proof. From Lemma 9.4, we know that there exists an optimal equilibrium (u^*, v^*) and $Y^* = J(u^*, v^*)$. Hence the equality $J(u^*, v^*) = J^*$ is clear (see, for instance, part (iv) in Proposition 7.1 or Corollary 7.1). $\square$

10 Markovian Forward–Backward SDEs

Let $(W_t)_{t\in[0,T]}$ be an $\mathbb{R}^d$-valued Brownian motion on a probability space $(\Omega,\mathcal{F},\mathbb{F},\mathbb{P})$ where the filtration $\mathbb{F}=(\mathcal{F}_t)_{t\in[0,T]}$ is generated by the Brownian motion W.

We consider the $\mathbb{R}^k$-valued diffusion process $(X^{t,x})$ where, for any fixed $(t,x)\in[0,T]\times\mathbb{R}^k$, the process $(X^{t,x}_s,\ s\in[t,T])$ is the unique strong solution to the following SDE

$$\begin{cases} dX^{t,x}_s=\mu(s,X^{t,x}_s)\,ds+\sigma(s,X^{t,x}_s)\,dW_s, & s\in[t,T],\\ X^{t,x}_t=x. \end{cases} \tag{10.1}$$

It is henceforth assumed that the functions μ and σ satisfy the assumptions of Theorem 12.3 and thus the uniqueness of a strong solution to SDE (10.1) is ensured.

For every $(t,x)\in[0,T]\times\mathbb{R}^k$, we denote by $(Y^{t,x},Z^{t,x})$ the pair of processes taking values in $\mathbb{R}^m$ and $\mathbb{R}^{m\times d}$, respectively, which solve the following Markovian BSDE

$$\begin{cases} -dY_s=g(s,X^{t,x}_s,Y_s,Z_s)\,dt-Z_s\,dW_s, & s\in[t,T],\\ Y_T=\Psi(X^{t,x}_T), \end{cases} \tag{10.2}$$

where $\Psi:\mathbb{R}^k\to\mathbb{R}^m$ and $g:[0,T]\times\mathbb{R}^k\times\mathbb{R}^m\times\mathbb{R}^{m\times d}\to\mathbb{R}^m$ are given functions.

Definition 10.1. The pair (10.1)–(10.2) is called the **Markovian forward–backward stochastic differential equation** (FBSDE).

We introduce the following Markovian counterparts of assumptions (H.1) and (H.2).

Assumption 10.1. We say that the generator g satisfies assumption (MH.1) if:
(i) $g:[0,T]\times\mathbb{R}^k\times\mathbb{R}^m\times\mathbb{R}^{m\times d}\to\mathbb{R}^m$ is $\mathcal{B}([0,T])\otimes\mathcal{B}(\mathbb{R}^k)\otimes\mathcal{B}(\mathbb{R}^m)\otimes\mathcal{B}(\mathbb{R}^{m\times d})$-measurable;
(ii) g is uniformly Lipschitz continuous: there exists a constant $L>0$ such that for all $y_1,y_2\in\mathbb{R}^m$ and $z_1,z_2\in\mathbb{R}^{m\times d}$

$$\|g(t,y_1,z_1)-g(t,y_2,z_2)\|\le L\big(\|y_1-y_2\|+\|z_1-z_2\|\big),$$

(iii) there exists a constant $c>0$ such that for a real constant $p\ge 1/2$

$$\|g(t,x,0,0)\|+\|\Psi(x)\|\le c(1+\|x\|^p),\quad \forall\,(t,x)\in[0,T]\times\mathbb{R}^k.$$

Assumption 10.2. A generator g satisfies assumption (MH.2) if:
(i) $g:[0,T]\times\mathbb{R}^k\times\mathbb{R}\times\mathbb{R}^d\to\mathbb{R}$ is $\mathcal{B}([0,T])\otimes\mathcal{B}(\mathbb{R}^k)\otimes\mathcal{B}(\mathbb{R})\otimes\mathcal{B}(\mathbb{R}^d)$-measurable;
(ii) the function $(y,z)\mapsto g(t,x,y,z)$ is continuous;
(iii) g satisfies the linear growth condition: there exists a constant $k>0$ such that

$$|g(t,x,y,z)|\le k(1+|y|+\|z\|),\quad \forall\,(y,z)\in\mathbb{R}^{d+1},$$

(iv) there exists a constant $c>0$ such that for a real constant $p\ge 1/2$

$$|g(t,x,0,0)|+|\Psi(x)|\le c(1+\|x\|^p),\quad \forall\,(t,x)\in[0,T]\times\mathbb{R}^k.$$

We take for granted the following result.

Theorem 10.1. *If (MH1) or (MH.2) holds, then there exist measurable functions u and w such that the unique (minimal) solution $(Y^{t,x},Z^{t,x})$ to BSDE* (10.2) *is given by*

$$Y^{t,x}=u(s,X^{t,x}_s),\quad Z^{t,x}_s=w(s,X^{t,x}_s)\sigma(s,X^{t,x}_s),\quad d\mathbb{P}\otimes ds-\textit{a.e.} \tag{10.3}$$

11 Feynman–Kac Formula for Quasi-Linear Parabolic PDEs

Let the k-dimensional matrix $a(t,x) = [a_{ij}(t,x)]$ be equal to $\sigma(t,x)\sigma^{\perp}(t,x)$ where $\sigma^{\perp}(t,x)$ is the transpose $\sigma(t,x)$. We define the second order differential operator $\mathcal{A}$

$$\mathcal{A} := \frac{1}{2}\sum_{i,j=1}^{k} a_{ij}(t,x)\frac{\partial^2}{\partial x_i \partial x_j} + \sum_{i=1}^{k} \mu_i(t,x)\frac{\partial}{\partial x_i}. \tag{11.1}$$

Let u be an $\mathbb{R}^m$-valued function defined on $[0,T] \times \mathbb{R}^k$ satisfying the following *quasi-linear parabolic PDE*

$$\begin{cases} -\frac{\partial u}{\partial t}(t,x) = \mathcal{A}u(t,x) + g(t,x,u(t,x),D_\sigma u(t,x)), \\ u(T,x) = \Psi(x), \end{cases} \tag{11.2}$$

where $D_\sigma u(t,x) := \nabla u(t,x)\sigma(t,x)$. The goal is to give a probabilistic representation for the function u similar to the classical Feynman–Kac formula, which is recalled in Section 12.7.

Lemma 11.1. *Let Assumption 10.1 hold. For any fixed* $(t,x) \in [0,T) \times \mathbb{R}^k$ *and any* $h > 0$ *such that* $t+h < T$ *we have*

$$v(t+h, X^{t,x}_{t+h}) = Y^{t+h,\, X^{t,x}_{t+h}}_{t+h} = Y^{t,x}_{t+h} \tag{11.3}$$

where $v(t+h,\widetilde{x}) := Y^{t+h,\widetilde{x}}_{t+h}$ *for every* $\widetilde{x} \in \mathbb{R}^k$.

Proof. The flow property of the unique solution to SDE (10.1) gives for all $s \in [t+h,T]$

$$X^{t,x}_s = X^{t+h,\, X^{t,x}_{t+h}}_s. \tag{11.4}$$

Observe that $\widetilde{Y} := Y^{t+h,\widetilde{x}}$ is a unique solution to the following FBSDE

$$\begin{cases} -d\widetilde{Y}_s = g(s, X^{t+h,\widetilde{x}}_s, \widetilde{Y}_s, \widetilde{Z}_s)\,dt - \widetilde{Z}_s\,dW_s, \quad s \in [t+h,T], \\ \widetilde{Y}_T = \Psi(X^{t+h,\widetilde{x}}_T), \end{cases}$$

whereas $\widehat{Y} := Y^{t,x}$ is a unique solution to the following FBSDE

$$\begin{cases} -d\widehat{Y}_r = g(r, X^{t,x}_r, \widehat{Y}_r, \widehat{Z}_r)\,dt - \widehat{Z}_r\,dW_r, \quad r \in [t,T], \\ \widehat{Y}_T = \Psi(X^{t,x}_T). \end{cases}$$

From the flow property (11.4) and the uniqueness of solutions to FBSDE (10.1)-(10.2), it is now easy to deduce that (11.3) is valid. □

Proposition 11.1. *Let Assumption 10.1 hold and let* $X^{t,x}$ *be given by* (10.1)*. If a function* u *belongs to the class* $C^{1,2}([0,T] \times \mathbb{R}^d, \mathbb{R}^m)$ *and is a solution to PDE* (11.2)*, then the pair*

$$(Y^{t,x}, Z^{t,x}) := (u(s, X^{t,x}_s), D_\sigma u(s, X^{t,x}_s)) \tag{11.5}$$

is the unique solution to BSDE (10.2) *on* $[t,T]$*. Moreover, for every* $t \in [0,T]$ *and* $x \in \mathbb{R}^k$*, the equality* $u(t,x) = Y^{t,x}_t$ *is valid.*

Proof. The proof is based on an application of the Itô formula and thus it is left as an exercise. □

Assumption 11.1. The functions μ, σ, g and Ψ belong to the class C^3 and have bounded derivatives.

The following theorem is due to Pardoux and Peng [73].

Theorem 11.1. *Let Assumptions 10.1, 10.2 and 11.1 be satisfied and let the pair* $(Y_s^{t,x}, Z_s^{t,x}),\ s \in [t,T]$ *be a unique solution to FBSDE* (10.1)–(10.2). *Then the function* $v(t,x) := Y_t^{t,x}$ *belongs to the class* $C^{1,2}([0,T] \times \mathbb{R}^k, \mathbb{R}^m)$ *and is a solution to quasi-linear parabolic PDE* (11.2).

Proof. The proof of Theorem 11.1 is quite technical and thus we only sketch the arguments used in the proof. In particular, we take for granted that the function $v(t,x) := Y_t^{t,x}$ belongs to the class $C^{1,2}([0,T] \times \mathbb{R}^k, \mathbb{R}^m)$ and we aim to show that the function $v(t,x) := Y_t^{t,x}$ is a solution to PDE (11.2). For this purpose, we consider an arbitrary partition $t = t_0 < t_1 < \cdots < t_n = T$ of the interval $[0,T]$. It is clear that

$$\begin{aligned} v(t,x) - \Psi(x) &= \sum_{i=0}^{n-1} \big(v(t_{i+1}, x) - v(t_i, x) \big) \\ &= \sum_{i=0}^{n-1} \Big(v(t_{i+1}, x) - v(t_{i+1}, X_{t_{i+1}}^{t,x}) \Big) + \sum_{i=0}^{n-1} \Big(v(t_{i+1}, X_{t_{i+1}}^{t,x}) - v(t_i, x) \Big). \end{aligned}$$

Since the function v is assumed to belong to the class $C^{1,2}([0,T] \times \mathbb{R}^k, \mathbb{R}^m)$, an application of the Itô formula yields

$$v(t_{i+1}, x) - v(t_{i+1}, X_{t_{i+1}}^{t,x}) = \int_{t_i}^{t_{i+1}} \mathcal{A} v(t_{i+1}, X_r^{t_i,x})\, dr + \int_{t_i}^{t_{i+1}} D_\sigma v(t_{i+1}, X_r^{t_i,x})\, dW_r$$

where the differential operator $\mathcal{A}$ is given by (11.1). From Lemma 11.1, we obtain

$$v(t_{i+1}, X_{t_{i+1}}^{t,x}) := Y_{t_{i+1}}^{t_{i+1}, X_{t_{i+1}}^{t_i,x}} = Y_{t_{i+1}}^{t_i,x}$$

and thus, using also BSDE (10.2), we get

$$v(t_{i+1}, X_{t_{i+1}}^{t,x}) - v(t_i, x) = Y_{t_{i+1}}^{t_i,x} - Y_{t_i}^{t_i,x} = \int_{t_i}^{t_{i+1}} g(r, X_r^{t_i,x}, Y_r^{t_i,x}, Z_r^{t_i,x})\, dr - \int_{t_i}^{t_{i+1}} Z_r^{t_i,x}\, dW_r.$$

We thus see that

$$\begin{aligned} v(t,x) - \Psi(x) = \sum_{i=0}^{n-1} \Big[&\int_{t_i}^{t_{i+1}} \Big(\mathcal{A} v(t_{i+1}, X_r^{t_i,x}) + g(r, X_r^{t_i,x}, Y_r^{t_i,x}, Z_r^{t_i,x}) \Big) dr \\ &- \int_{t_i}^{t_{i+1}} \Big(Z_r^{t_i,x} - D_\sigma v(t_{i+1}, X_r^{t_i,x}) \Big) dW_r \Big]. \end{aligned}$$

Assume that $\lim_{n\to\infty} \sup_i (t_{i+1}^n - t_t^n) = 0$. Using the equality $Z_s^{t,x} = D_\sigma v(s, X_s^{t,x})$, it is possible to show that under Assumptions 10.1 and 11.1 for every $(t,x) \in [0,T] \times \mathbb{R}^k$

$$v(t,x) - \Psi(x) = \int_t^T \Big(\mathcal{A} v(r,x) + g(r, x, v(r,x), D_\sigma v(r,x)) \Big) dr,$$

which in turn implies that the function v solves PDE (11.2). $\square$

12 Appendix: Stochastic Calculus

12.1 Properties of Local Martingales

Let $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P})$ be a filtered probability space satisfying the usual hypotheses of right-continuity and completeness. We also assume that the σ-field $\mathcal{F}_0$ is trivial.

We first recall the conditional Fatou lemma.

Lemma 12.1. *Let $\xi_1, \xi_2, \ldots$ be a sequence of extended real-valued random variables defined on $(\Omega, \mathcal{F}, \mathbb{P})$. If there exists a non-negative integrable random variable ζ such that $\xi_n \geq -\zeta$ for all n, then*

$$\mathbb{E}_{\mathbb{P}}\Big(\liminf_{n\to\infty} \xi_n \,\Big|\, \mathcal{F}_t\Big) \leq \liminf_{n\to\infty} \mathbb{E}_{\mathbb{P}}\big(\xi_n \,|\, \mathcal{F}_t\big).$$

Definition 12.1. Let $M = (M_t)_{t\in[0,T]}$ be an $\mathbb{F}$-adapted process. We say that M is a *local martingale* if there exists a non-decreasing sequence τ_n of $\mathbb{F}$-stopping times with values in $[0,T]$ such that $\lim_{n\to\infty} \tau_n = T$ and for every n the stopped process M^{τ_n} defined by $M^{\tau_n}_t = M_{t\wedge\tau_n}$ is a uniformly integrable martingale.

In many instances, it is important to know that a non-negative local martingale is a supermartingale, as can be deduced from the following lemma.

Lemma 12.2. *Let $M = (M_t)_{t\in[0,T]}$ be a local martingale such that $M \geq -\zeta$ for some non-negative integrable random variable ζ. Then M is a supermartingale, that is, $\mathbb{E}_{\mathbb{P}}(M_t \,|\, \mathcal{F}_s) \leq M_s$ for all $0 \leq s \leq t \leq T$.*

The next result is useful when one needs to check if a local martingale is a (true) martingale.

Lemma 12.3. *If $M = (M_t)_{\,t\in[0,T]}$ is a real-valued local martingale and the random variable $M^*_T = \sup_{\,t\in[0,T]} |M_t|$ is integrable, then M is a uniformly integrable martingale.*

Proof. Let $(\tau_n)_{n\in\mathbb{N}}$ be any reducing sequence of stopping times for M. For any $n \in \mathbb{N}$ and $t \leq T$ the martingale property of the stopped process M^{τ_n} yields

$$\mathbb{E}_{\mathbb{P}}(M_{T\wedge\tau_n} \,|\, \mathcal{F}_t) = M_{t\wedge\tau_n}. \tag{12.1}$$

Since the random variable M^*_T dominates $M_{T\wedge\tau_n}$ and $X_{t\wedge\tau_n}$, using the dominated convergence theorem, we may pass to the limit in (12.1). We thus obtain $\mathbb{E}_{\mathbb{P}}(M_T \,|\, \mathcal{F}_t) = M_t$ for every $t \in [0,T]$, which is the desired martingale property of M. $\square$

12.1.1 Doob's L^p Inequality

For $p \geq 1$, we denote by $\|\cdot\|_p$ the usual L^p norm for random variables, that is, $\|\,\eta\,\|_p := [\mathbb{E}(|\eta|^p)]^{1/p}$. For an arbitrary process X, we write $X^*_t = \sup_{0\leq s\leq t} |X_s|$ and $X^*_\infty = \sup_{\,t\geq 0} |X_t|$. For the proof the next result, see Theorem 1.7 in Chapter II in Revuz and Yor [84] (or, in discrete time, the monograph by Doob [21]).

Theorem 12.1 (Doob's inequality). *If $X = (X_t)_{t\in[0,T]}$ is either a real-valued right-continuous martingale or a positive submartingale, then for $p \geq 1$*

$$\lambda^p \, \mathbb{P}(X^*_\infty \geq \lambda) \leq \sup_{t>0} \mathbb{E}(|X_t|^p)$$

and for $p > 1$

$$\|X^*_\infty\|_p \leq \frac{p}{p-1} \sup_{t>0} \|X_t\|_p.$$

Consequently, for any right-continuous local martingale

$$\mathbb{E}\big((M^*_T)^2\big) \leq 4 \sup_{t\in[0,T]} \mathbb{E}(M_t^2) = 4\,\mathbb{E}(M_T^2).$$

12.1.2 Burkholder–Davis–Gundy (BDG) Inequalities

Let M be a real-valued continuous local martingale and let $M^*_\infty = \sup_{t\geq 0} |M_t|$. Recall that $\langle M, M\rangle$ (also denoted as $\langle M\rangle$) is the predictable covariation of M. For the proof of the next result, see Theorem 4.1 in Chapter IV in Revuz and Yor [84].

Theorem 12.2 (BDG inequalities). *For every $p \in]0, \infty[$, there exists two constants c_p and C_p such that for all continuous local martingales M vanishing at 0 we have*

$$c_p \, \mathbb{E}\big[\langle M, M\rangle_\infty^{p/2}\big] \leq \mathbb{E}\big[(M^*_\infty)^p\big] \leq C_p \, \mathbb{E}\big[\langle M, M\rangle_\infty^{p/2}\big].$$

Hence the norms $\|M^*_\infty\|_p$ and $\|\langle M, M\rangle_\infty^{1/2}\|_p$ are equivalent.

12.2 Itô Processes

Let W be a d-dimensional standard Brownian motion defined on a filtered probability space $(\Omega, \mathbb{F}, \mathbb{P})$. We introduce a particular class of continuous semimartingales, which are known as *Itô processes*.

Definition 12.2. An $\mathbb{F}$-adapted, continuous process X is called an *Itô process* if it admits the integral representation

$$X_t = X_0 + \int_0^t \alpha_s \, ds + \int_0^t \beta_s \, dW_s \tag{12.2}$$

for some $\mathbb{F}$-progressively measurable processes α and β satisfying suitable integrability conditions.

It is customary to represent the integral formula (12.2) using the differential notation as

$$dX_t = \alpha_t \, dt + \beta_t \, dW_t.$$

It is clear that the Itô process X follows a continuous semimartingale and formula (12.2) gives the canonical decomposition of X. In the present set-up, decomposition (12.2) is unique in the following sense: if X satisfies (12.2) and simultaneously we have that

$$X_t = X_0 + \int_0^t \widetilde{\alpha}_s \, ds + \int_0^t \widetilde{\beta}_s \, dW_s$$

for some $\mathbb{F}$-progressively measurable processes $\widetilde{\alpha}$ and $\widetilde{\beta}$, then the following equalities hold for every $t \in [0,T]$

$$\int_0^t \widetilde{\alpha}_s \, du = \int_0^t \alpha_s \, ds, \quad \int_0^t \widetilde{\beta}_s \, dW_s = \int_0^t \beta_s \, dW_s.$$

Note that an Itô process X given by (12.2) is a continuous local martingale if and only if it can be represented as follows

$$X_t = X_0 + \int_0^t \beta_s \, dW_s.$$

12.3 Itô's Lemma

We now ask the following question: is the process $f(X_t)$ a semimartingale if X is a continuous semimartingale and f is a sufficiently regular function? It turns out that the class of continuous semimartingales is invariant with respect to compositions with C^2-functions (more general results are also available).

It is clear that if X is a continuous semimartingale and $f : \mathbb{R} \times [0,T] \to \mathbb{R}$ is a jointly continuous function, then the process $Y_t = f(X_t, t)$ is $\mathbb{F}$-adapted and has almost all sample paths continuous.

12.3.1 One-dimensional Itô's Formula

The next result, which is a special case of *Itô's lemma,* shows that Y is a semimartingale provided that the function g is sufficiently smooth.

Proposition 12.1. *Suppose that $f : \mathbb{R} \times [0,T] \to \mathbb{R}$ is a function of class $C^{2,1}(\mathbb{R} \times [0,T], \mathbb{R})$ and X is an Itô process. Then the process $Y_t := f(X_t, t)$, $t \in [0,T]$ is an Itô process and its canonical decomposition is given by the Itô formula*

$$dY_t = f_t(X_t, t) \, dt + f_x(X_t, t)\alpha_t \, dt + f_x(X_t, t)\beta_t \, dW_t + \tfrac{1}{2} f_{xx}(X_t, t)\beta_t^2 \, dt.$$

More generally, if $X = X_0 + M + A$ is a real-valued continuous semimartingale and f is a function of class $C^{2,1}(\mathbb{R} \times [0,T], \mathbb{R})$, then the process $Y_t := f(X_t, t)$ is a continuous semimartingale with the following canonical decomposition

$$dY_t = f_t(X_t, t) \, dt + f_x(X_t, t) \, dX_t + \tfrac{1}{2} f_{xx}(X_t, t) \, d\langle M \rangle_t. \tag{12.3}$$

12.3.2 Multidimensional Itô's Formula

Recall that a process $W = (W^1, \ldots, W^d)$ defined on a filtered probability space $(\Omega, \mathbb{F}, \mathbb{P})$ is called the *d-dimensional standard Brownian motion* if the components $W^1, \ldots, W^d$ are independent one-dimensional standard Brownian motions with respect to $\mathbb{F}$. We henceforth assume that W is a d-dimensional standard Brownian motion. Let γ be an $\mathbb{R}^d$-valued, $\mathbb{F}$-progressively measurable process satisfying

$$\mathbb{P}\left(\int_0^T \|\gamma_s\|^2 \, ds < \infty \right) = 1.$$

Then the Itô stochastic integral of γ with respect to W is well defined and we have for all $t \in [0,T]$

$$\int_0^t \gamma_s \, dW_s = \sum_{i=1}^d \int_0^t \gamma_s^i \, dW_s^i.$$

Let $X = (X^1, \ldots, X^k)$ be a k-dimensional Itô process given by

$$X_t^i = X_0^i + \int_0^t \alpha_s^i \, ds + \int_0^t \beta_s^i \, dW_s. \tag{12.4}$$

In formula (12.4), we assume that α^i is a real-valued process and $\beta^i = (\beta^{i1}, \ldots, \beta^{id})$ is an $\mathbb{R}^d$-valued process. Furthermore, it is implicitly assumed that the processes α^i, β^i, $i = 1, \ldots, k$ are integrable in a suitable sense.

Let us recall the notion of the *quadratic covariation*. If $X^i = X_0^i + M^i + A^i$, $i = 1, 2, \ldots, k$ are real-valued continuous semimartingales, then we define quadratic covariation by setting

$$\langle X^i, X^j \rangle := \langle M^i, M^j \rangle$$

where in turn $\langle M^i, M^j \rangle$ is given by the following *polarisation equality*

$$\langle M^i, M^j \rangle := \tfrac{1}{2} \left(\langle M^i + M^j \rangle - \langle M^i \rangle - \langle M^j \rangle \right) = \tfrac{1}{4} \left(\langle M^i + M^j \rangle - \langle M^i - M^j \rangle \right).$$

It is easily seen that $\langle X^i, X^i \rangle = \langle X^i \rangle$. If X^i and X^j are Itô processes given by (12.4), then we have for all $t \in [0,T]$

$$\langle X^i, X^j \rangle_t = \int_0^t \beta_s^i \beta_s^j \, ds = \int_0^t \sum_{l=1}^d \beta_s^{il} \beta_s^{jl} \, ds.$$

The following result extends Proposition 12.1.

Proposition 12.2. *Suppose that f is a function of class $C^2(\mathbb{R}^k, \mathbb{R})$. Then the Itô formula holds*

$$df(X_t) = \sum_{i=1}^k f_{x_i}(X_t) \alpha_t^i \, dt + \sum_{i=1}^k f_{x_i}(X_t) \beta_t^i \, dW_t + \frac{1}{2} \sum_{i,j=1}^k f_{x_i x_j}(X_t) \, \beta_t^i \, \beta_t^j \, dt.$$

More generally, if X^i is a continuous semimartingale for $i = 1, \ldots, k$, then

$$f(X_t) = f(X_0) + \sum_{i=1}^k \int_0^t f_{x_i}(X_s) \, dX_s^i + \frac{1}{2} \sum_{i,j=1}^k \int_0^t f_{x_i x_j}(X_s) \, d\langle X^i, X^j \rangle_s$$

or, equivalently,

$$df(X_t) = \sum_{i=1}^k f_{x_i}(X_t) \, dX_t^i + \frac{1}{2} \sum_{i,j=1}^k f_{x_i x_j}(X_t) \, d\langle X^i, X^j \rangle_t.$$

The Itô integration by parts formula is obtained by taking $f(x_1, x_2) = x_1 x_2$.

Corollary 12.1. *Let X^1 and X^2 be real-valued continuous semimartingales. Then the Itô integration by parts formula is valid*

$$X_t^1 X_t^2 = X_0^1 X_0^2 + \int_0^t X_s^1 \, dX_s^2 + \int_0^t X_s^2 \, dX_s^1 + \langle X^1, X^2 \rangle_t. \tag{12.5}$$

In particular, for any real-valued continuous semimartingale X

$$X_t^2 = X_0^2 + 2 \int_0^t X_s \, dX_s + \langle X \rangle_t. \tag{12.6}$$

12.4 Itô–Tanaka–Meyer Formula

The classical Itô formula is obtained under the assumption that the transformation is twice continuously differentiable in the space variable. It is noteworthy that the preservation of the semimartingale property of a real-valued semimartingale X holds under a much weaker assumption that a function $f : \mathbb{R} \to \mathbb{R}$ is a difference of two convex functions. Furthermore, if W is a one-dimensional Brownian motion and $f : \mathbb{R} \to \mathbb{R}$ is a functions such that the process $f(W)$ is a semimartingale, then necessarily $f = f_1 - f_2$ where f_1 and f_2 are convex functions (see Wang [93] and Brosamler [11]).

Let X be a real-valued continuous semimartingale with the canonical decomposition $X = X_0 + M + A$ where M is a continuous local martingale and A is a continuous process of finite variation with $M_0 = A_0 = 0$. Recall that $\langle X \rangle = \langle M \rangle$ stands for the quadratic variation of X, meaning that $M^2 - \langle M \rangle$ is a continuous local martingale. For instance, if

$$X_t = X_0 + \int_0^t \sigma_s \, dW_s + \int_0^t \mu_s \, ds = X_0 + M_t + A_t,$$

then

$$\langle X \rangle_t = \langle M \rangle_t = \int_0^t |\sigma_s|^2 \, ds.$$

Local time. For any fixed $a \in \mathbb{R}$, we denote by $L_t^a(X)$ the (right) semimartingale *local time* of X at the level a. It is defined by the formula

$$L_t^a(X) := |X_t - a| - |X_0 - a| - \int_0^t \operatorname{sgn}(X_s - a) \, dX_s$$

for every $t \in [0, T]$, where $\operatorname{sgn}(x) = 1$ for $x > 0$ and $\operatorname{sgn}(x) = -1$ for $x < 0$. By convention, we set $\operatorname{sgn}(0) = -1$ so that the function $\operatorname{sgn}(x - a)$ is the left derivative of the function $|x - a|$. It is well known that the local time $L^a(X)$ of a continuous semimartingale X is an adapted process whose sample paths are continuous, non-decreasing functions and it satisfies

$$L_t^a(X) = \int_0^t \mathbb{1}_{\{X_s = a\}} \, dL_s^a(X) = \lim_{\varepsilon \downarrow 0} \frac{1}{2\varepsilon} \int_0^t \mathbb{1}_{\{|X_u - a| \leq \varepsilon\}} \, d\langle X \rangle_u.$$

In addition, for any locally integrable (or non-negative) Borel measurable function $h : \mathbb{R} \to \mathbb{R}$ the following *density of occupation time formula* holds

$$\int_{\mathbb{R}} h(a) L_t^a(X) \, da = \int_0^t h(X_s) \, d\langle X \rangle_s.$$

We take for granted the following result, first obtained Tanaka [91] and Wang [93] for the one-dimensional Brownian motion and later extended by Meyer to real-valued semimartingales (see McKean [59] or Chapter VI in Revuz and Yor [84]).

Lemma 12.4 (Itô–Tanaka–Meyer formula). *For an arbitrary convex function $f : \mathbb{R} \to \mathbb{R}$ and a real-valued continuous semimartingale X, the following equality holds*

$$f(X_t) = f(X_0) + \int_0^t f_l'(X_s) \, dX_s + \frac{1}{2} \int_{\mathbb{R}} L_t^a(X) \, \mu(da)$$

where f_l' is the left derivative of f and the measure $\mu = f''$ represents the second order derivative of f in the sense of distributions.

It is worth recalling that the left- and right-hand derivatives of a convex function are non-decreasing functions. If the function f belongs to the class $C^2(\mathbb{R})$, then $\mu(da) = f''(a)\,da$ for a continuous (hence locally integrable) function f'' and the density of occupation time formula gives

$$\int_{\mathbb{R}} L_t^a(X)\,\mu(da) = \int_{\mathbb{R}} L_t^a(X)\,f''(a)\,da = \int_0^t f''(X_s)\,d\langle X\rangle_s.$$

Consequently, if $f \in C^2(\mathbb{R})$, then the Itô–Tanaka–Meyer formula reduces to the classical Itô formula. More generally, we have the following extension of the classical Itô formula to the case of functions belonging to the class $C^1(\mathbb{R})$, but not necessarily to the class $C^2(\mathbb{R})$.

Corollary 12.2. *Assume that f belongs to the class $C^1(\mathbb{R})$, the first order derivative f' is differentiable almost everywhere and f'' is locally integrable. Then for any real-valued continuous semimartingale X we have*

$$f(X_t) = f(X_0) + \int_0^t f'(X_s)\,dX_s + \frac{1}{2}\int_0^t f''(X_s)\,d\langle X\rangle_s.$$

12.5 Itô's Theorem for SDEs

Let $\mu : [0,T] \times \mathbb{R}^k \to \mathbb{R}^k$ and $\sigma : [0,T] \times \mathbb{R}^k \to \mathbb{R}^{k\times d}$ be given functions and let W be a d-dimensional Brownian motion defined on a filtered probability space $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P})$.

Definition 12.3. By a *solution* of the stochastic differential equation (SDE)

$$dX_t = \mu(t, X_t)\,dt + \sigma(t, X_t)\,dW_t \tag{12.7}$$

with the initial condition $X_0 \in \mathbb{R}^k$, we mean an $\mathbb{R}^k$-valued, $\mathbb{F}$-adapted, continuous stochastic process X defined on the probability space $(\Omega, \mathbb{F}, \mathbb{P})$ and such that for every $t \in [0,T]$

$$X_t = X_0 + \int_0^t \mu(s, X_s)\,ds + \int_0^t \sigma(s, X_s)\,dW_s.$$

Let us recall Itô's theorem, which gives sufficient conditions for the existence and uniqueness of solutions to the SDE (12.7).

Theorem 12.3. *Let $\mu : [0,T] \times \mathbb{R}^k \to \mathbb{R}^k$ and $\sigma : [0,T] \times \mathbb{R}^k \to \mathbb{R}^{k\times d}$ satisfy the following conditions:*
(i) *the functions μ and σ are Lipschitz continuous with respect to the variable x: there exists a constant $L > 0$ such that for every $x, y \in \mathbb{R}^k$ and $t \in \mathbb{R}_+$*

$$\|\mu(t,x) - \mu(t,y)\|^2 + \|\sigma(t,x) - \sigma(t,y)\|^2 \le L\|x - y\|^2, \tag{12.8}$$

(ii) *the functions μ and σ satisfy the linear growth condition: there exists a constant K such that, for any $x \in \mathbb{R}^k$ and $t \in \mathbb{R}_+$,*

$$\|\mu(t,x)\|^2 + \|\sigma(t,x)\|^2 \le K(1 + \|x\|^2). \tag{12.9}$$

Then SDE (12.7) *has the unique solution X.*

Proposition 12.3. *Let μ and σ be $\mathbb{F}$-adapted, bounded processes. Then the unique solution of the SDE*

$$dX_t = X_t\big(\mu_t\,dt + \sigma_t\,dW_t\big)$$

is given by

$$X_t = X_0 \exp\Big(\int_0^t \sigma_s\,dW_s + \int_0^t \Big(\mu_s - \frac{1}{2}\,\|\sigma_s\|^2\Big)\,ds\Big).$$

12.6 Girsanov's Theorem

Girsanov's theorem states that, under mild technical conditions, a Brownian motion with an absolutely continuous drift becomes a standard Brownian motion after an equivalent change of a probability measure. Before proceeding to the general result, let us first consider the simple case of a linear drift.

Proposition 12.4. *Let W be a one-dimensional standard Brownian motion on a probability space $(\Omega, \mathbb{F}, \mathbb{P})$. For a real number $\gamma \in \mathbb{R}$, we define the process $\widetilde{W}$ by setting*

$$\widetilde{W}_t := W_t - \gamma t, \quad \forall\, t \in [0, T].$$

Let the probability measure $\widetilde{\mathbb{P}}$, equivalent to $\mathbb{P}$ on $(\Omega, \mathcal{F}_T)$, be defined through the formula

$$\frac{d\widetilde{\mathbb{P}}}{d\mathbb{P}} := \exp\left(\gamma W_T - \tfrac{1}{2}\gamma^2 T\right), \quad \mathbb{P} - a.s.$$

Then $\widetilde{W}$ is a standard Brownian motion on $(\Omega, \mathbb{F}, \widetilde{\mathbb{P}})$.

Let γ be an $\mathbb{R}^d$-valued, $\mathbb{F}$-progressively measurable process such that

$$\mathbb{P}\Big(\int_0^T \|\gamma_s\|^2\, ds < \infty\Big) = 1. \tag{12.10}$$

Then the stochastic exponential is defined as follows.

Definition 12.4. The *stochastic exponential* (also known as the *Doléans–Dade exponential*) is given by the formula

$$\mathcal{E}_t\Big(\int_0^{\cdot} \gamma_s\, dW_s\Big) := \exp\,\Big(\int_0^t \gamma_s\, dW_s - \frac{1}{2}\int_0^t \|\gamma_s\|^2\, ds\Big), \quad \forall\, t \in [0, T].$$

Lemma 12.5. *The stochastic exponential is the unique solution Z to the SDE*

$$dZ_t = Z_t\, \gamma_t\, dW_t, \quad Z_0 = 1.$$

The following result is the classical variant of the celebrated Girsanov theorem established in 1960.

Theorem 12.4 (Girsanov [38])**.** *Let W be a d-dimensional standard Brownian motion on a filtered probability space $(\Omega, \mathbb{F}, \mathbb{P})$. Suppose that γ is an $\mathbb{R}^d$-valued, $\mathbb{F}$-progressively measurable process such that* (12.10) *holds and*

$$\mathbb{E}_{\mathbb{P}}\Big\{\mathcal{E}_T\Big(\int_0^{\cdot} \gamma_s\, dW_s\Big)\Big\} = 1. \tag{12.11}$$

Let a probability measure $\widetilde{\mathbb{P}}$, equivalent to $\mathbb{P}$ on $(\Omega, \mathcal{F}_T)$, be defined by means of the Radon-Nikodým derivative

$$\frac{d\widetilde{\mathbb{P}}}{d\mathbb{P}} := \mathcal{E}_T\Big(\int_0^{\cdot} \gamma_s\, dW_s\Big), \quad \mathbb{P} - a.s.$$

Then the process $\widetilde{W}$ given by the formula

$$\widetilde{W}_t = W_t - \int_0^t \gamma_s\, ds, \quad \forall\, t \in [0, T],$$

is a d-dimensional Brownian motion on $(\Omega, \mathbb{F}, \widetilde{\mathbb{P}})$.

For the proof of the next result, see Proposition 1.15 in Chapter VIII of Revuz and Yor [84].

Proposition 12.5 (Novikov [69]). *Assume that*

$$\mathbb{E}_{\mathbb{P}}\Big\{ \exp\Big(\frac{1}{2}\int_0^T \|\gamma_t\|^2\,dt\Big)\Big\} < \infty.$$

Then equality (12.11) *holds and thus the process* $\mathcal{E}(U)$ *where* $U_t = \int_0^t \gamma_s\,dW_s$ *is a strictly positive, uniformly integrable, continuous* $\mathbb{F}$*-martingale.*

In particular, if the process γ is bounded, then Novikov's criterion is satisfied. A weaker, but also sufficient for the validity of (12.11), is Kazamaki's condition

$$\mathbb{E}_{\mathbb{P}}\Big\{ \exp\Big(\frac{1}{2}\int_0^t \gamma_s\,dW_s\Big)\Big\} < \infty, \quad \forall\, t \in [0,T].$$

It should be noted that Kazamaki's criterion is not a necessary condition for the process $\mathcal{E}(U)$ to be a martingale. For the proof of the next result, see Proposition 1.14 in Chapter VIII of Revuz and Yor [84].

Proposition 12.6 (Kazamaki [48]). *If the process*

$$\exp\Big(\frac{1}{2}\int_0^t \gamma_s\,dW_s\Big)$$

is a uniformly integrable submartingale, then the Doléans-Dade exponential $\mathcal{E}(U)$ *where* $U_t = \int_0^t \gamma_s\,dW_s$ *is a strictly positive, uniformly integrable, continuous* $\mathbb{F}$*-martingale.*

For a survey of more general results and new proofs of classical results, see Ruf [86].

12.7 Feynman–Kac Theorem

Let us first consider a particular example of a general set-up. If $f : \mathbb{R} \to \mathbb{R}$ is a bounded function and $g : \mathbb{R} \to \mathbb{R}$ is a function bounded from below then the unique bounded solution $v(t,x)$ of the partial differential equation (PDE)

$$\frac{\partial v}{\partial t}(t,x) = \frac{1}{2}\frac{\partial^2 v}{\partial x^2}(t,x) - g(x)v(t,x)$$

with the initial condition $v(0,x) = f(x)$ can be represented by the following version of the *Feynman–Kac formula*

$$v(t,x) = \mathbb{E}_{\mathbb{P}}\Big(f(x+W_t)\exp\Big(-\int_0^t g(x+W_s)\,ds\Big)\Big).$$

It is clear that this correspondence between expected values of certain functionals of a Brownian motion and solutions to partial differential equations allows us to compute some probabilistic quantities using the PDE approach and, conversely, to solve initial value problems for certain PDEs through purely probabilistic techniques (for instance, through Monte Carlo simulation of sample paths of a stochastic process).

12.7.1 One-dimensional Feynman–Kac Theorem

Let us first examine the one-dimensional version of the Feynman–Kac theorem (see Kac [44]) named after Richard Feynman and Mark Kac.

Proposition 12.7. *Let $\mu, \sigma : \mathbb{R} \to \mathbb{R}$ satisfy the the Lipschitz condition (12.8) and the linear growth condition (12.9). Let $f : \mathbb{R} \to \mathbb{R}$ be a bounded function and $g : \mathbb{R} \to \mathbb{R}$ be a function bounded from below. Assume that $v : \mathbb{R} \times \mathbb{R}_+ \to \mathbb{R}$ is the unique bounded solution to the parabolic PDE*

$$\frac{\partial v}{\partial t}(t,x) = \frac{1}{2}\sigma^2(x)\frac{\partial^2 v}{\partial x^2}(t,x) + \mu(x)\frac{\partial v}{\partial x}(t,x) - g(x)v(t,x) \tag{12.12}$$

with the initial condition $v(0,x) = f(x)$. Then v has the following representation

$$v(t,x) = \mathbb{E}_{\mathbb{P}}\Big(f(X_t^x)\exp\Big(-\int_0^t g(X_s^x)\,ds\Big)\Big) \tag{12.13}$$

where X^x is a solution to the SDE

$$dX_t^x = \mu(X_t^x)\,dt + \sigma(X_t^x)\,dW_t$$

with the initial condition $X_0^x = x$.

The *infinitesimal generator* $\mathcal{A}$ of X is given by the formula

$$\mathcal{A}f(x) = \frac{1}{2}\sigma^2(x)\frac{\partial^2 f}{\partial x^2}(x) + \mu(x)\frac{\partial f}{\partial x}(x).$$

Hence equation (12.12) can be rewritten as follows

$$\frac{\partial v}{\partial t} = \mathcal{A}v - gv.$$

Proof of Proposition 12.7. The proof relies on finding, for a fixed $t \in [0,T]$, a martingale $M = (M_s)_{s\in[0,t]}$ such that $\mathbb{E}_{\mathbb{P}}(M_0) = v(t,x)$ and

$$\mathbb{E}_{\mathbb{P}}M_t = \mathbb{E}_{\mathbb{P}}\Big(f(X_t^x)\exp\Big(-\int_0^t g(X_u^x)\,du\Big)\Big).$$

Since for any martingale M we have $\mathbb{E}_{\mathbb{P}}M_0 = \mathbb{E}_{\mathbb{P}}M_t$, this implies (12.13). Let us fix $t > 0$ and let us define the process $M = (M_s)_{s\in[0,t]}$ by the formula

$$M_s = v(t-s, X_s^x)\exp\Big(-\int_0^s g(X_u^x)\,du\Big). \tag{12.14}$$

The Itô formula combined with the assumption that the function v solves (12.12), yield

$$dM_s = \exp\Big(-\int_0^s g(X_u^x)\,du\Big)\sigma(X_s^x)\frac{\partial v}{\partial x}(t-s, X_s^x)\,dW_s$$

and thus M is a local martingale. We assumed that v, g are bounded functions, so that $v(x) \le L_1$ and $g(x) \ge -L_2$ for some constants L_1 and L_2. From (12.14), we thus get

$$\sup_{u\le t}|M_u| \le L_1\exp(L_2 t) < \infty,$$

so that $\mathbb{E}_{\mathbb{P}}(M_t^*) < \infty$. From Lemma 12.3, it follows that M is a martingale. Hence

$$\begin{aligned} v(t,x) &= \mathbb{E}_{\mathbb{P}}(M_0) = \mathbb{E}_{\mathbb{P}}(M_t) = \mathbb{E}_{\mathbb{P}}\Big(v(0,X_t^x)\exp\Big(-\int_0^t g(X_s^x)\,ds\Big)\Big) \\ &= \mathbb{E}_{\mathbb{P}}\Big(f(X_t^x)\exp\Big(-\int_0^t g(X_s^x)\,ds\Big)\Big) \end{aligned}$$

since, obviously, $v(0,X_t^x) = f(X_t^x)$. We conclude that (12.13) holds. □

12.7.2 Multidimensional Feynman–Kac Formula

Let W be a d-dimensional standard Brownian motion on the underlying probability space $(\Omega, \mathbb{F}, \mathbb{P})$. We consider the k-dimensional diffusion process X given as

$$dX_t^x = \mu(X_t^x)\,dt + \sigma(X_t^x)\,dW_t, \quad X_0^x = x, \tag{12.15}$$

where the coefficients $\mu : \mathbb{R}^k \to \mathbb{R}^k$ and $\sigma : \mathbb{R}^k \to \mathbb{R}^{k\times d}$ satisfy the Lipschitz condition (12.8) and the linear growth condition (12.9). Then, by virtue of Theorem 12.3, the SDE (12.15) admits a unique solution X. Moreover, X is a continuous process and has the *Markov property* with respect to the filtration $\mathbb{F}$.

Let the k-dimensional matrix $a(x) = [a_{ij}(x)]$ be equal to $\sigma(x)\sigma^t(x)$, where $\sigma^t(x)$ is the transpose $\sigma(x)$. We associate with the Markov process X its *infinitesimal generator* $\mathcal{A}$

$$\mathcal{A}f(x) = \frac{1}{2}\sum_{i,j=1}^k a_{ij}(x)\frac{\partial^2 f}{\partial x_i \partial x_j}(x) + \sum_{i=1}^k \mu_i(x)\frac{\partial f}{\partial x_i}(x)$$

for any function $f : \mathbb{R}^k \to \mathbb{R}$ belonging to a suitably defined class of functions, referred to as the domain $D(\mathcal{A})$ of the generator $\mathcal{A}$.

Remark 12.1. Although it is rather difficult to find explicitly the domain $D(\mathcal{A})$ of the infinitesimal generator $\mathcal{A}$ of a diffusion process X, $\mathcal{A}$ is always well defined on the set $C_c^2(\mathbb{R}^k)$ of all functions of class $C^2(\mathbb{R}^k)$ and with a compact support (that is, functions that vanish outside a bounded interval), so that $D(\mathcal{A}) \subseteq C_c^2(\mathbb{R}^k)$.

The following result furnishes a multidimensional version of the Feynman–Kac theorem.

Proposition 12.8. *Let $g \in C(\mathbb{R}^k)$ be a function bounded from below and let f belong to the class $C_c^2(\mathbb{R}^k)$. Define the function $v : \mathbb{R}_+ \times \mathbb{R}^k \to \mathbb{R}$ by the formula*

$$v(t,x) = \mathbb{E}_{\mathbb{P}}\Big(f(X_t^x)\exp\Big(-\int_0^t g(X_s^x)\,ds\Big)\Big).$$

Then v satisfies the parabolic PDE

$$\frac{\partial v}{\partial t} = \mathcal{A}v - gv \tag{12.16}$$

with the initial condition $v(0,x) = f(x)$ for $x \in \mathbb{R}^k$.

Under the assumptions of Proposition 12.8, we also have the following converse result.

Proposition 12.9. *If $w \in C^{1,2}(\mathbb{R}\times\mathbb{R}^k)$ is bounded on $[-K,K]\times\mathbb{R}^d$ for any $K > 0$ (equivalently, the function w is bounded on $A\times\mathbb{R}^d$ for any compact subset $A \subset \mathbb{R}$) and w is a solution to the parabolic PDE* (12.16) *with the initial condition $w(0,x) = f(x)$ for $x \in \mathbb{R}^k$, then $w(t,x) = v(t,x)$.*

References

[1] Banach, S.: Sur les opérations dans les ensembles abstraits et leur application aux équations intégrales. *Fundamenta Mathematicae* 3 (1922), 133–181.

[2] Bellman, R. E.: The stability of solutions of linear differential equations. *Duke Mathematical Journal* 10 (1943), 643–647.

[3] Bellman, R. E.: *The Theory of Dynamic Programming*. Rand Corporation, 1954.

[4] Beneš, V. E.: Existence of optimal strategies based on specific information for a class of stochastic decision problems. *SIAM Journal on Control and Optimization* 8 (1970), 179–188.

[5] Beneš, V. E., Shepp, L. A., Witsenhausen, H. S.: Some solvable stochastic control problems. *Stochastics* 4 (1980), 39–83.

[6] Bichteler, K.: *Stochastic Integration and Stochastic Differential Equations.* Cambridge University Press, 2002.

[7] Bielecki, T. R., Rutkowski, M.: Valuation and hedging of contracts with funding costs and collateralization. *SIAM Journal on Financial Mathematics* 6 (2015), 594–655.

[8] Bielecki, T. R., Cialenco, I., Rutkowski, M.: Arbitrage-free pricing of derivatives in nonlinear market models. *Probability, Uncertainty and Quantitative Risk* 3 (2018), 1–56.

[9] Bismut, J. M.: Conjugate convex functions in optimal stochastic control. *Journal of Mathematical Analysis and Applications* 44 (1973), 384–404.

[10] Bismut, J. M.: Sur un problème de Dynkin. *Z. Wahrsch. Verw. Gebiete* 39 (1977), 31–53.

[11] Brosamler, G. A.: Quadratic variation of potentials and harmonic functions. *Transactions of the American Mathematical Society* 149 (1970) 243–257.

[12] Buckdahn, R., Quincampoix, M., Răşcanu, A.: Viability property for a backward stochastic differential equation and applications to partial differential equations. *Probability Theory and Related Fields* 116 (2000), 485–504.

[13] Chassagneux, J. F., Elie, R., Kharroubi, I.: A note on existence and uniqueness for solutions of multidimensional reflected BSDEs. *Electronic Communications in Probability* 16 (2011), 120–128.

[14] Chassagneux, J. F., Elie, R., Kharroubi, I.: Discrete-time approximation of multidimensional BSDEs with oblique reflections. *Annals of Applied Probability* 22 (2012), 971–1007.

[15] Cvitanić, J. and Karatzas, I.: Backward stochastic differential equations with reflection and Dynkin games. *Annals of Probability* 24 (1996), 2024–2056.

[16] Dellacherie, C.: *Capacités et processus stochastiques.* Springer, 1972.

[17] Dellacherie, C., Meyer, P. A.: *Probabilités et potentiel, Chapitres I à IV*. Hermann, Paris, 1975.

[18] Dellacherie, C., Meyer, P. A.: *Probabilités et potentiel, Chapitres V à VIII*. Hermann, Paris, 1980.

[19] Dellacherie, C., Meyer, P. A.: *Probabilities and Potential. B. Theory of Martingales. North-Holland Publishing Co.*, Amsterdam, 1982.

[20] Darling, R., Pardoux, E.: Backward SDE with random terminal time and applications to semilinear elliptic PDE. *Annals of Probability* 25 (1997), 1135–1159.

[21] Doob, J. L.: *Stochastic Processes*. John Wiley & Sons, 1953.

[22] Dudley, R. M.: Wiener functionals as Itô integrals. *Annals of Probability* 5 (1977), 140–141.

[23] Dumitrescu, R., Quenez, M. C., Sulem, A.: Generalized Dynkin games and doubly reflected BSDEs with jumps. *Electronic Journal of Probability* 21(64) (2016), 1–32.

[24] Dumitrescu, R., Quenez, M. C., Sulem, A.: Game options in an imperfect market with default. *SIAM Journal on Financial Mathematics* 8 (2017), 532–559.

[25] Dynkin, E. B.: A game-theoretic version of an optimal stopping problem. (In Russian) *Doklady Akademii Nauk SSSR* 185, (1969), 16–19.

[26] El Karoui, N.: Backward stochastic differential equations: a general introduction. In: *Backward Stochastic Differential Equations, Pitman Research Notes in Mathematics Series 364*, N. El Karoui and L. Mazliak, editors, Addison Wesley Longman, 1997, pp. 7–26.

[27] El Karoui, N., Hamadène, S., Matoussi, A.: Backward stochastic differential equations and applications. In: *Indifference Pricing: Theory and Applications*, R. Carmona, editor, Princeton University Press, 2009, pp. 267–320.

[28] El Karoui, N., Huang, S. J.: A general result of existence and uniqueness of backward stochastic differential equations. In: *Backward Stochastic Differential Equations, Pitman Research Notes in Mathematics Series 364*, N. El Karoui and L. Mazliak, editors, Addison Wesley Longman, 1997, pp. 27–36.

[29] El Karoui, N., Kapoudjian, C., Pardoux, E., Peng, S., Quenez, M. C.: Reflected solutions of backward SDE's, and related obstacle problems for SDE's. *Annals of Probability* 25 (1997), 702–737.

[30] El Karoui, N., Peng, S., Quenez, M. C.: Backward stochastic differential equations in finance. *Mathematical Finance* 7 (1997), 1–71.

[31] El Karoui, N., Quenez, M. C.: Non-linear pricing theory and backward stochastic differential equations. In: *Financial Mathematics, Lect. Notes Math. 1656*, B. Biais et al., editors, Springer, Berlin, 1997, pp. 191–246.

[32] Fan, S. J., Jiang, L.: Uniqueness result for the BSDE whose generator is monotonic in y and uniformly continuous in z. *Comptes Rendus Acad. Sci. Paris, Ser. I* 348 (2010), 89–92.

[33] Grigorova, M., Imkeller, P., Offen, E., Ouknine, Y., and Quenez, M. C.: Reflected BSDEs when the obstacle is not right-continuous and optimal stopping. *Annals of Applied Probability* 27 (2017), 3153–3188.

[34] Grigorova, M., Imkeller, P., Ouknine, Y., and Quenez, M. C.: Doubly reflected BSDEs and $\mathcal{E}^f$-Dynkin games: beyond the right-continuous case. *Electronic Journal of Probability* 23 (2018), 1-38.

[35] Grigorova, M., Imkeller, P., Ouknine, Y., and Quenez, M. C.: Optimal stopping with f-expectations: the irregular case. *Stochastic Processes and their Applications* 130 (2020), 1258–1288.

[36] Guo, I.: *Competitive Multi-Player Stochastic Games with Applications to Multi-Person Financial Contracts.* Doctoral dissertation, University of Sydney, 2013.

[37] Jia, G.: A uniqueness theorem for the solution of Backward Stochastic Differential Equations. *Comptes Rendus Acad. Sci. Paris, Ser. I* 346 (2008), 439–444.

[38] Girsanov, I. V.: On transforming a certain class of stochastic processes by absolutely continuous substitution of measures. *Theory of Probability and its Applications* 5 (1960), 285–301.

[39] Gronwall, T. H.: Notes on the derivatives with respect to a parameter of the solutions of a system of differential equations. *Annals of Mathematics* 20 (1919), 292–296.

[40] Hamadène, S., Lepeltier, J. P., Matoussi, A.: Double barrier reflected backward SDE's with continuous coefficient. In: *Backward Stochastic Differential Equations, Pitman Research Notes in Mathematics Series 364*, N. El Karoui and L. Mazliak, editors, Addison Wesley Longman, 1997, pp. 161–171.

[41] Hu, Y., Peng, S.: On the comparison theorem for multidimensional BSDEs. *Comptes Rendus Acad. Sci. Paris, Ser. I* 343 (2006), 135–140.

[42] Isaacs, R.: *Differential Games.* John Wiley and Sons, 1965.

[43] Jeanblanc, M., Réveillac, A.: A note on BSDEs with singular driver coefficients. Working paper, 2014.

[44] Kac, M.: On distributions of certain Wiener functionals. *Transactions of the American Mathematical Society* 65 (1949), 1–13.

[45] Karatzas, I.: Lecture notes on capacities and on debut, section and projection theorems. Working paper, 2015.

[46] Karatzas, I., Li, Q.: BSDE approach to non-zero-sum stochastic differential games of control and stopping. Working paper, 2011.

[47] Karatzas, I., Shreve, S.: *Brownian Motion and Stochastic Calculus.* 2nd ed., Springer, 1998.

[48] Kazamaki, N.: On a problem of Girsanov. *Tohoku Mathematical Journal* 29 (1977), 597–600.

[49] Kazamaki, N.: A sufficient condition for the uniform integrability of exponential martingales. *Mathematics Reports, Toyama University* 2 (1979), 1–11.

[50] Kim, E., Nie, T., Rutkowski, M.: American options in nonlinear markets. *Electronic Journal of Probability* 26 (2021), no. 90, 1–41.

[51] Kobylanski, M., Quenez, M. C.: Optimal stopping time problem in a general framework. *Electronic Journal of Probability* 17/72 (2012), 1–28.

[52] Kobylanski, M., Quenez, M. C., de Campagnolle, M. R.: Dynkin games in a general framework. *Stochastics* 86 (2014), 304–329.

[53] Laraki, R., Solan, E.: The value of zero-sum stopping games in continuous time. *SIAM Journal on Control and Optimization* 43 (2005), 1913–1922.

[54] Lepeltier, J. P., Maingueneau, M. A.: Le jeu de Dynkin en théorie générale sans l'hypothèse de Mokobodski. *Stochastics* 13 (1984), 25–44.

[55] Lepeltier, J. P., San Martín, J.: Backward stochastic differential equations with continuous coefficients. *Statistics and Probability Letters* 32 (1997), 425–430.

[56] Lepeltier, J. P., San Martín, J.: On the existence or non-existence of solutions for certain backward stochastic differential equations. *Bernoulli* 8 (2002), 123–137.

[57] Lin, Q.: A class of backward doubly stochastic differential equations with non-Lipschitz coefficients. *Statistics and Probability Letters* 79 (2009), 2223–2229.

[58] Ma, M., Fan, S. J., Song, X.: L^p $(p > 1)$ solutions of backward stochastic differential equations with monotonic and uniformly continuous generators. *Bulletin des Sciences Mathématiques* 137 (2013), 97–106.

[59] McKean, H. P.: A Hölder condition for Brownian local time. *Journal of Mathematics of Kyoto University* 1-2 (1962), 195–201.

[60] Mokobodzki, G.: Sur l'opérateur de réduite. Remarques sur un travail de J. M. Bismut: "Sur un problème de Dynkin." In *Séminaire de Théorie du Potentiel, No. 3 (Paris 1976/1977), Lecture Notes in Math., Vol. 681,* F. Hirsch and G. Mokobodzki (Eds.). Springer, Berlin, 1978, pp. 188–208.

[61] Nash, J. F.: Equilibrium points in N-person games. *Proceedings of the National Academy of Sciences* 36 (1950), 48–49.

[62] Neveu, J.: *Discrete-Parameter Martingales.* North–Holland, 1975.

[63] Nie, T., Rutkowski, M.: Fair bilateral prices in Bergman's model with exogenous collateralization. *International Journal of Theoretical and Applied Finance* 18 (2015), 1550048 (26 pages).

[64] Nie, T., Rutkowski, M.: A BSDE approach to fair bilateral pricing under endogenous collateralization. *Finance and Stochastics* 20 (2016), 855–900.

[65] Nie, T., Rutkowski, M.: Fair bilateral pricing under funding costs and exogenous collateralization. *Mathematical Finance* 28 (2018), 621–655.

[66] Nie, T., Rutkowski, M.: Existence, uniqueness and strict comparison theorems for backward stochastic differential equations driven by RCLL martingales. *Probability, Uncertainty and Quantitative Risk* 6(4) (2021), 319–342.

[67] Nie, T., Rutkowski, M.: Reflected BSDEs and doubly reflected BSDEs driven by RCLL martingales. *Stochastics and Dynamics* 22(5) (2022), 2250012 (34 pages).

[68] Nikeghbali, A.: An essay on the general theory of stochastic processes. *Probability Surveys* 3 (2006), 345–412.

[69] Novikov, A.: On an identity for stochastic integrals. *Theory of Probability and its Applications* 17 (1972), 717–720.

[70] Ohtsubo, Y.: Optimal stopping in sequential games with or without a constraint of always terminating. *Mathematics of Operations Research* 11 (1986), 591–607.

[71] Pardoux, E.: BSDEs, weak convergence and homogenization of semilinear PDEs in nonlinear analysis. In: *Differential Equations and Control Lecture*, Kluwer Academic Publishers, F. H. Clarke and R. J. Stern, editors, 1999, pp. 503–549.

[72] Pardoux, E., Peng, S.: Adapted solution of a backward stochastic differential equation. *Systems and Control Letters* 14 (1990), 55–61.

[73] Pardoux, E., Peng, S.: Backward stochastic differential equations and quasilinear parabolic partial differential equations. In: *Stochastic Partial Differential Equations and their Applications, Lecture Notes in Control and Information Sciences 176*, B. L. Rozovskii and R. B. Sowers, editors, 1992, pp. 200–217.

[74] Pham, H.: On some recent aspects of stochastic control and their applications. *Probability Surveys* 2 (2005), 506–549.

[75] Peng, S.: A general stochastic maximum principle for optimal control problems. *SIAM Journal on Control and Optimization* 28 (1990), 966–979.

[76] Peng, S.: Probabilistic interpretation for systems of quasilinear parabolic partial differential equations. *Stochastics* 37 (1991), 61–74.

[77] Peng, S.: Stochastic Hamilton-Jacobi-Bellman equations. *SIAM Journal on Control and Optimization* 28 (1990), 966–979.

[78] Peng, S.: A generalized dynamic programming principle and Hamilton-Jacobi-Bellman equation. *Stochastics* 38 (1992), 119–134.

[79] Peng, S.: Backward stochastic differential equations and its application in optimal control. *Applied Mathematics and Optimization* 27 (1993), 125–144.

[80] Peng, S.: Nonlinear expactations, nonlinear evaluations and risk measures. In: *Lecture Notes in Mathematics 1856*, M. Frittelli and W. Runggaldier, editors, Springer, 2004, pp. 165–253.

[81] Peng, S.: Dynamically consistent nonlinear evaluations and expectations. Working paper, 2004 (arXiv:0501415v1).

[82] Peng, S., Xu, X.: BSDEs with random default time and their applications to default risk. Working paper, 2009.

[83] Quenez, M. C. and Sulem, A.: Reflected BSDEs and robust optimal stopping time for dynamic risk measures with jumps. *Stochastic Processes and their Applications* 124 (2014), 3031–3054.

[84] Revuz, D., Yor, M.: *Continuous Martingales and Brownian Motion.* 3rd ed., Springer, 1999.

[85] Rogers, L. C. G.: Duality in constrained optimal investment and consumption problems: a synthesis. *Lecture Notes in Mathematics 1814*, R. Carmona et al., editors, Springer, pp. 95–131.

[86] Ruf, J.: A new proof for the conditions of Novikov and Kazamaki. Working paper, University of Oxford, 2012.

[87] Snell, J. L.: Applications of martingale system theorems. *Transactions of the American Mathematical Society* 73 (1952), 293–312.

[88] Stettner, Ł.: Zero-sum Markov games with stopping and impulsive strategies. *Journal of Applied Mathematics and Optimization* 9 (1982), 1–24.

[89] Strömberg, T.: The operation of infimal convolution. *Dissertationes Mathematicae* 352 (1996), 58 pp.

[90] Szimayer, A. (2005). Valuation of American options in the presence of event risk. *Finance and Stochastics* 9, 89–107.

[91] Tanaka, H.: Note on continous additive functionals of the 1-dimensional Brownian path. *Z. Wahrscheinlichkeitstheorie verw. Grebiete* 1 (1963), 251–257.

[92] Trotter, H. F.: A property of Brownian motion paths. *Illinois Journal of Mathematics* 2 (1958), 425–433.

[93] Wang, A. T.: Generalized Itô's formula and additive functionals of the Brownian path. *Z. Wahrscheinlichkeitstheorie verw. Grebiete* 41 (1977), 153–159.

[94] Xu, M.: Backward stochastic differential equations with reflection and weak assumptions on the coefficients. *Stochastic Processes and their Applications* 118 (2008) 968–980.

[95] Yamada, T.: On a comparison theorem for solutions of stochastic differential equations and its applications. *Journal of Mathematics of Kyoto University* 13 (1977), 497–512.

[96] Yamada, T.: On the non-confluent property of solutions of stochastic differential equations and its applications. *Stochastics* 17 (1986), 111–124.

[97] Yamada, T., Ogura, Y.: On the strong comparison theorems for solutions of stochastic differential equations. *Z. Wahrscheinlichkeitstheorie verw. Gebiete* 56 (1981), 3–19.

[98] Yin, J., Mao, X.: The adapted solution and comparison theorem for backward stochastic differential equations with Poisson jumps and applications. *Journal of Mathematical Analysis and Applications* 346 (2008), 345–358.

[99] Zheng, S., Zhou, S.: A generalized existence theorem of reflected BSDEs with double obstacles. *Statistics and Probability Letters* 78 (2008), 528–536.

Who's Who

- Robert Brown (1773–1858) was a Scottish botanist who made important contributions to botany, in particular, through his pioneering use of the microscope.
- Augustin-Louis Cauchy (1789–1857) was a French mathematician, who made essential contributions to calculus, complex analysis and algebra.
- Carl Gustav Jacob Jacobi (1804–1851) was a German mathematician, who made contributions to elliptic functions, differential equations and number theory.
- William Rowan Hamilton (1805–1865) was an Irish physicist and mathematician, who made contributions to mechanics, optics and algebra.
- Rudolf Lipschitz (1832–1903) was a German mathematician, who contributed to mathematical analysis, differential geometry, number theory and algebra.
- Karl Schwarz (1843–1921) was a German mathematician, known for his work in complex analysis.
- Henri Poincaré (1854–1912) was a French mathematician, theoretical physicist, engineer, and a philosopher of science.
- Andrei Andreyevich Markov (1856–1922) was a Russian mathematician best known for his work on stochastic processes, in particular, Markov chains.
- Émile Picard (1856–1941) was a French mathematician, who contributed to complex analysis, the theory of differential equations and algebra.
- David Hilbert (1862–1943) was a German mathematician, who discovered and developed a broad range of fundamental ideas in many areas, including invariant theory and the axiomatization of geometry.
- Émile Borel (1871–1956) was a French mathematician known for his founding work in the areas of measure theory and probability.
- Henri Lebesgue (1875–1941) was a French mathematician most famous for his contributions to the theory of integration.
- Thomas Hakon Grönwall (1877–1932) was a Swedish American mathematician working in the areas of physical chemistry and atomic physics.
- Pierre Fatou (1878–1929) was a French mathematician and astronomer. He is known for major contributions to several branches of analysis.
- Guido Fubini (1879–1943) was an Italian mathematician, who worked in mathematical analysis, functional analysis and complex analysis.
- Johann Radon (1887–1956) was an Austrian mathematician, who contributed to the development of the integration theory and integral geometry.
- Otto Nikodým (1887–1974) was a Polish mathematician, who contributed to the integration theory and abstract Boolean lattices.

- Stefan Banach (1892–1945) was a Polish mathematician, who was one of the founders of modern functional analysis.
- Norbert Wiener (1894–1964) was an American mathematician, who contributed to the theory of stochastic processes, electronic engineering, and control systems.
- Lev Semyonovich Pontryagin (1908–1988) was a Russian mathematician working in the areas of optimal control theory and algebraic topology.
- Joseph Leo Doob (1910–2004) was an American mathematician specialised in analysis and probability theory and the theory of martingales.
- Rufus Philip Isaacs (1914–1981) was an American mathematician, who developed the theory of differential games.
- Mark Kac (1914–1984) was a Polish American mathematician, who worked in mathematical analysis and probability theory.
- Kiyosi Itô (1915–2008) was a Japanese mathematician, who pioneered the theory of stochastic integration and stochastic differential equations.
- Richard Feynman (1918–1988) was an American physicist known for the path integral formulation of quantum mechanics and the theory of quantum electrodynamics.
- Richard Bellman (1920–1984) was an American mathematician, who introduced dynamic programming in 1953.
- Eugene Borisovich Dynkin (1924–2014) was a Russian mathematician, who made important contributions to the fields of probability and algebra, in particular, Markov processes and Lie groups.
- James Laurie Snell (1925–2011) was an American mathematician specialising in probability theory.
- Donald Burkholder (1927–2013) was an American mathematician known for his contributions to probability theory and the theory of martingales.
- John Forbes Nash (1928–2015) was an American mathematician with fundamental contributions in game theory, differential geometry and PDEs.
- Henry P. McKean, Jr. (born 1930) is an American mathematician at the Courant Institute in New York University. He works in various areas of analysis and probability theory (in particular, the theory of diffusion processes).
- Igor Vladimirovich Girsanov (1934–1967) was a Russian mathematician specialising in probability theory and control theory.
- Paul-André Meyer (1934–2003) was a French mathematician, who played a major role in the development of the general theory of stochastic processes by the famous Strasbourg School.
- Catherine Doléans-Dade (1942–2004) was a French mathematician, who made significant contributions to the Itô stochastic calculus including a general change of variables formula and exponential processes of semimartingales.

- Claude Dellacherie (born 1943) is a French mathematician, who made significant contributions to the general theory of stochastic processes.
- Nicole El Karoui (born 1944) is a French mathematician and pioneer in the development of optimal control and mathematical finance. She is considered one of the pioneers on the French school of mathematical finance.
- Mark Davis (1945–2020) was Professor of Mathematics at Imperial College London. He made fundamental contributions to the theory of stochastic processes, stochastic control and mathematical finance.
- Monique Jeanblanc (born 1947) is a French mathematician known for her work in mathematical finance and optimal control theory.
- Peng Shige (born 1947) is a Chinese mathematician noted for his contributions in stochastic analysis (the theory of BSDEs, G-Brownian motion) and mathematical finance (nonlinear market models).
- Marc Yor (1949–2014) was a French mathematician well known for his work on stochastic processes, especially properties of semimartingales, Brownian motion, Lévy processes and Bessel processes, and their applications to mathematical finance (CGMY model).
- Chris Rogers (born 1954) is a mathematician working in probability theory and quantitative finance. Rogers' specialist fields include stochastic analysis and applications to quantitative finance. With David Williams he has written two influential textbooks on diffusion processes.