Mathematical Finance, Vol. 7, No. 1 (January 1997), 1–71

BACKWARD STOCHASTIC DIFFERENTIAL EQUATIONS IN FINANCE

N. El Karoui

Laboratoire de Probabilités, CNRS-URA 224, Université de Paris VI, Paris, France

S. Peng

Institute of Mathematics, Shandong University, Jinan, China

M. C. Quenez[1]

Equipe de Mathématiques, Université de Marne la Vallée, Noisy-Le-Grand, France

We are concerned with different properties of backward stochastic differential equations and their applications to finance. These equations, first introduced by Pardoux and Peng (1990), are useful for the theory of contingent claim valuation, especially cases with constraints and for the theory of recursive utilities, introduced by Duffie and Epstein (1992a, 1992b).

Key Words: backward stochastic equation, mathematical finance, pricing, hedging portfolios, incomplete market, constrained portfolio, recursive utility, stochastic control, viscosity solution of PDE, Malliavin derivative

0. INTRODUCTION

We are concerned with backward stochastic differential equations (BSDE) and with their applications to finance. These equations were introduced by Bismut (1973) for the linear case and by Pardoux and Peng (1990) in the general case. According to these authors, the solution of a BSDE consists of a pair of adapted processes (Y, Z) satisfying

$$-dY_t = f(t, Y_t, Z_t)\,dt - Z_t^* dW_t; \qquad Y_T = \xi, \tag{0.1}$$

where f is the generator and ξ is the terminal condition.

Actually, this type of equation appears in numerous problems in finance (as pointed out in Quenez's doctorate 1993). First, the theory of contingent claim valuation in a complete market studied by Black and Scholes (1973), Merton (1973, 1991), Harrison and Kreps (1979), Harrison and Pliska (1981), Duffie (1988), and Karatzas (1989), among others, can be expressed in terms of BSDEs. Indeed, the problem is to determine the price of a contingent claim $\xi \geq 0$ of maturity T, which is a contract that pays an amount ξ at time T. In a complete market it is possible to construct a portfolio which attains as final wealth the amount ξ. Thus, the dynamics of the value of the replicating portfolio Y are given by a BSDE with linear generator f, with Z corresponding to the hedging portfolio. Then the price at time t is associated naturally with the value at time t of the hedging portfolio. However, there exists an infinite number of replicating portfolios and consequently the

[1]The authors are grateful to Stanley Pliska for his fruitful comments about the revision of this paper and to the anonymous referees for their careful reading and numerous suggestions.

Manuscript received December 1994; final revision received August 1996.

price is not well defined. But arbitrage pricing theory imposes some restrictions on the integrability of the hedging portfolios. In general, these assumptions are related to a risk-adjusted probability measure. Using BSDE theory, we will show that the problem is well posed—that is, there exist a unique price and a unique hedging portfolio—by restricting admissible strategies to square-integrable ones under the primitive probability.

On the other hand, the pricing theory has been studied in the context of an incomplete market by Föllmer and Schweizer (1990) and El Karoui and Quenez (1995). In this situation it is not always possible to construct a portfolio which attains exactly as final wealth the amount ξ, and the price cannot be determined by no-arbitrage arguments. The replication error is called the tracking error. El Karoui and Quenez (1995) only considered superstrategies, which are strategies with a positive tracking error, and defined the upper price for each contingent claim ξ as the smallest investment which allows one to superhedge the contingent claim ξ. They showed that the upper price is equal to the value function of a control problem. Using this dual characterization, they stated that the upper price corresponds to a superstrategy. The upper price process is not a solution of a BSDE, but it can be written as the increasing limit of penalized price processes which are solutions of nonlinear BSDEs. We shall see that the duality between the hedging problem and the pricing one, emphasized by El Karoui and Quenez (1995), corresponds to a general duality between convex BSDEs and some control problems.

Concerning this problem of pricing a contingent claim ξ in an incomplete market, Föllmer and Schweizer (1990) introduced the notion of local risk-minimizing strategies, for which the tracking error is a square-integrable martingale orthogonal to the basic securities. We show that this pricing rule corresponds to a standard valuation in a market where only the traded securities have a return different from the spot rate. The price for ξ is still a solution of a linear BSDE.

Recall that the results stated by El Karoui and Quenez (1995) for the constrained case of an incomplete market were generalized to convex constraints on the portfolios by Cvitanic and Karatzas (1992). Other nonlinear backward equations were introduced by those authors for the hedging problem with a higher interest rate for borrowing. In this case, the dynamics of the wealth process are given by a nonlinear convex BSDE.

Duffie and Epstein (1992a, 1992b) presented a stochastic differential formulation of recursive utility in the case of information generated by Brownian motion. Recursive utility is an extension of the standard additive utility with the instantaneous utility depending not only on the instantaneous consumption rate c_t but also on the future utility. Actually, it corresponds to the solution of a particular BSDE associated with a generator f which does not depend on z. Duffie and Epstein showed that, under Lipschitz conditions, the recursive utility exists and satisfies the usual properties of standard utilities (concavity with respect to consumption if the BSDE is concave). The BSDE point of view gives a simple formulation of recursive utilities and their properties.

In this paper we summarize the results on existence and uniqueness by Pardoux and Peng (1990) and give new (shorter) proofs. We state some a priori estimates of the difference between the solutions of two BSDEs; the existence and uniqueness of the solution of a BSDE follow from a fixed-point theorem. Also, we recall that one of the most important properties of BSDEs is a comparison theorem which can be obtained under quite general conditions. For example, this theorem gives a sufficient condition for the wealth process to be nonnegative and yields the classical properties of utilities.

Also, BSDEs with concave (or convex) generators are associated by duality with the value function of a control problem. Thus, the duality introduced by El Karoui and Quenez

(1991, 1995) in the hedging and pricing problem in a constrained case can be generalized to other cases such as recursive utilities. This result gives an interesting interpretation of recursive utilities of consumption: the utility can be expressed as the minimum of ex post utilities over all possible future price deflators. Actually, this variational formulation of recursive utilities had been introduced by Geoffard (1995) in a deterministic framework. Moreover, this result can be applied to European option pricing. In some constrained cases the price of a contingent claim is given by the solution of a nonlinear convex BSDE. As the utility function, the price can be written as the maximum of “ex post” prices taken over all changes of “numeraire” feasible for the wealth; also the set of controls is bounded and the maximum is attained for an optimal change of numeraire.

Furthermore, we are concerned with the solution of BSDEs associated with a state process satisfying some forward classical SDEs. The main property is that the solution is Markovian in the sense that it can be written as a function of time and a state process. Important results concerning the link between those BSDEs and PDEs have been stated by Pardoux and Peng (1992) and Peng (1991, 1992a, 1992b, 1992c) in the Markovian case; these Markovian BSDEs give a Feynman-Kac representation of some nonlinear parabolic partial differential equations. Conversely, under smoothness assumptions the solution of the BSDE corresponds to the solution of a system of quasilinear parabolic PDEs. These results can be applied to European option pricing in the constrained Markovian case and give a generalization of the Black and Scholes formula.

The outline of the paper is as follows. In the first section we give some examples of BSDEs which appear naturally in the problem of pricing and hedging contingent claims. Actually, the dynamics of wealth processes can be written as a BSDE; these equations are linear in the classical case and nonlinear (but convex) in the case of constraints on the portfolio. Another example is given by the stochastic differential formulation of recursive utilities introduced by Duffie and Epstein (1992a, 1992b).

In Section 2 we present some important results for BSDEs: a priori estimates of the difference between two solutions, existence and uniqueness, a comparison theorem, and supersolutions. Also, we study the properties of continuity and differentiability of the solutions of BSDEs with respect to parameters, properties which follow essentially from the a priori estimates. Finally, we give the flow properties of BSDEs.

In Section 3 we study the properties of concave (or convex) BSDEs. We show that, in this case, the solution of the BSDE can be written as the value function of a control problem. Then we give some applications of these results to finance.

In Section 4 we study different properties of the solution of a BSDE associated with some forward SDEs (regularity, measurability), and, in particular, we prove under very weak assumptions that the solution only depends on time and the state process. Then we recall different results concerning the link between these solutions and some PDEs, and we give a simple application to European option pricing in the constrained case.

In Section 5 we give some complementary results on BSDEs. First, we solve the BSDE in the case of a non-Brownian filtration and under p-integrability assumptions ($p > 1$). Second, we study in detail the properties of differentiation on Wiener space of the solution of a BSDE in the spirit of the work of Pardoux and Peng (1992). Applying these results to finance, we show that the portfolio process of a hedging strategy corresponds to the Malliavin derivative of the price process. This important property was first emphasized by Karatzas and Ocone (1992) in the nonconstrained case (i.e., the linear case).

1. BACKWARD DIFFERENTIAL EQUATIONS IN FINANCE

1.1. The Model

We begin with the typical setup for continuous-time asset pricing: the basic securities consist of $n+1$ assets. One of them is a locally riskless asset (the money market instrument or bond) with price per unit P^0 governed by the equation

$$dP_t^0 = P_t^0 r_t \, dt, \tag{1.1}$$

where r_t is the short rate. In addition to the bond, n risky securities (the stocks) are continuously traded. The price process P^i for one share of ith stock is modeled by the linear stochastic differential equation

$$dP_t^i = P_t^i \left[b_t^i \, dt + \sum_{j=1}^{n} \sigma_t^{i,j} \, dW_t^j \right], \tag{1.2}$$

where $W = (W^1, \ldots, W^n)^*$ is a standard Brownian motion on $\mathbb{R}^n$, defined on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$. $\mathbb{P}$ is said to be the "objective" probability measure. The information structure is given by a right-continuous filtration $(\mathcal{F}_t\,;\ 0 \le t \le T)$. Usually, $(\mathcal{F}_t)$ is the σ-algebra generated by the Brownian motion $W = (W^1, \ldots, W^n)^*$ and augmented.[2] We make the following standard assumptions.

HYPOTHESIS 1.1.

- The short rate r is a predictable and bounded process. It is generally nonnegative.
- The column vector of the stock appreciation rates $b = (b^1, \ldots, b^n)^*$ is a predictable and bounded process.
- The volatility matrix $\sigma = (\sigma^{i,j})$ is a predictable and bounded process. σ_t has full rank a.s. for all $t \in [0, T]$ and the inverse matrix σ^{-1} is a bounded process.
- There exists a predictable and bounded-valued process vector θ, called a risk premium, such that

  $$b_t - r_t \mathbf{1} = \sigma_t \, \theta_t, \qquad d\mathbb{P} \otimes dt \text{ a.s.},$$

  where $\mathbf{1}$ is the vector whose every component is 1.

Under these assumptions the market is dynamically complete.

[2] The augmented Brownian filtration $\mathbf{F} = \{(\mathcal{F}_t), t \in [0, T]\}$ is defined by $\mathcal{F}_t = \sigma(\mathcal{F}_t^W \cup \mathcal{Z})$, where $\mathcal{F}_t^W = \sigma(W_s; s \in [0, t])$ is the smallest σ-field with respect to which W_s is measurable for every $s \in [0, t]$ and $\mathcal{Z} = \{E \subseteq \Omega; \exists G \in \mathcal{F}, E \subseteq G, \mathbb{P}(G) = 0\}$ denotes the set of $\mathbb{P}$-null events. It is well known that the augmented filtration is continuous and that W is still a Brownian motion with respect to it (Karatzas and Shreve 1987, Corollary 2.7.8. and Proposition 2.7.9).

All the stochastic processes to appear in the sequel are progressively measurable with respect to $\mathbf{F}$, all the equalities involving random variables are understood to hold $\mathbb{P}$ a.s., and the equalities involving stochastic processes are understood to hold $d\mathbb{P} \otimes dt$ a.s. Sometimes we shall refer to the following notion of equality between two processes: two processes X and Y are said to be indistinguishable if $\{\omega, \exists t \in [0, T]\ X_t(\omega) \neq Y_t(\omega)\}$ is a $\mathbb{P}$-null set. The same definitions hold for the inequalities.

Let us consider a small investor whose actions cannot affect market prices and who can decide at time $t \in [0, T]$ what amount π_t^i of the wealth V_t to invest in the ith stock, $i = 1, \dots, n$. In the Merton model (1971), he also chooses his consumption c_t (actually c_t is the positive rate of consumption at time t). Of course, his decisions can only be based on the current information $(\mathcal{F}_t)$; i.e., the processes $\pi = (\pi^1, \pi^2, \dots, \pi^n)^*$, $\pi^0 = V - \sum_{i=1}^n \pi^i$, and c are predictable. Following Harrison and Pliska (1981), we say a strategy is self-financing if the wealth process $V = \sum_{i=0}^n \pi^i$ satisfies the equality

$$V_t = V_0 + \int_0^t \sum_{i=0}^n \pi_t^i \frac{dP_t^i}{P_t^i}$$

or, equivalently, if the wealth process satisfies the linear stochastic differential equation (LSDE)

$$\begin{aligned} dV_t &= r_t V_t\, dt + \pi_t^*(b_t - r_t \mathbf{1})\, dt + \pi_t^* \sigma_t\, dW_t \\ &= r_t V_t\, dt + \pi_t^* \sigma_t [dW_t + \theta_t\, dt]. \end{aligned}$$

In the Merton model the equation becomes

$$dV_t = r_t V_t\, dt - c_t\, dt + \pi_t^* \sigma_t [dW_t + \theta_t\, dt].$$

More precisely,

DEFINITION 1.1. A self-financing *trading* strategy is a pair (V, π), where V is the market value and $\pi = (\pi^1, \dots, \pi^n)^*$ is the portfolio process, such that (V, π) satisfies

$$(1.3) \qquad dV_t = r_t V_t\, dt + \pi_t^* \sigma_t [dW_t + \theta_t\, dt], \qquad \int_0^T |\sigma_t^* \pi_t|^2\, dt < +\infty, \ \mathbb{P} \text{ a.s.}$$

The strategy is called *feasible* if the constraint of nonnegative wealth holds:

$$V_t \geq 0, \qquad t \in [0, T], \ \mathbb{P} \text{ a.s.}$$

REMARK. Generally, the initial wealth x is taken as a primitive, and for an initial endowment x and portfolio process π there exists a unique (continuous) wealth process that is a solution of (1.3) with initial value $V_0 = x$, since the process r is bounded. There exists a useful one-to-one correspondence between the pair $(V_0 = x, \pi)$ and the trading strategy (V, π).

To extend this formulation to the Merton model, we choose to introduce the cumulative amount of consumption between 0 and t, namely $C_t = \int_0^t c_s\, ds$, and still refer to the Merton model even if the adapted increasing consumption process C is not absolutely continuous. Instead of a consumption process, this process can sometimes be interpreted as the liquidity necessary to satisfy some constraints.

DEFINITION 1.2. A self-financing *superstrategy* is a vector process (V, π, C), where V is the market value (or wealth process), π is the portfolio process, and C is the cumulative consumption process, such that

$$dV_t = r_t V_t \, dt - dC_t + \pi_t^* \sigma_t [dW_t + \theta_t \, dt] \,, \qquad \int_0^T |\sigma_t^* \pi_t|^2 \, dt < +\infty \,, \ \mathbb{P} \text{ a.s.,} \tag{1.4}$$

and C is an increasing, right-continuous, adapted process with $C_0 = 0$. The superstrategy is called *feasible* if the constraint of nonnegative wealth holds:

$$V_t \geq 0, \qquad t \in [0, T], \mathbb{P} \text{ a.s.}$$

1.2. Pricing and Hedging Positive Contingent Claims

Fair Price of Positive Contingent Claims. A European contingent claim ξ settled at time T is an $\mathcal{F}_T$-measurable random variable. It can be thought of as a contract which pays ξ at maturity T. The arbitrage-free pricing of a positive contingent claim is based on the following principle: if we start with the price of the claim as initial endowment and invest it in the $n + 1$ assets, the value of the portfolio at time T must be just enough to guarantee ξ. We follow the presentation of Karatzas and Shreve (1987).

DEFINITION 1.3. Let $\xi \geq 0$ be a positive contingent claim.

(1) A *hedging strategy* against ξ (resp. a superhedging strategy) is a feasible self-financing strategy (V, π) (resp. (V, π, C)) such that $V_T = \xi$.
We denote by $\mathcal{H}(\xi)$ (resp. $\mathcal{H}'(\xi)$) the class of hedging strategies (resp. superhedging strategies) against ξ. If $\mathcal{H}(\xi)$ (resp. $\mathcal{H}'(\xi)$) is nonempty, ξ is called *hedgeable* (resp. *superhedgeable*).

(2) The *fair price* X_0 (resp. upper price X_0') at time 0 of the hedgeable (resp. superhedgeable) claim ξ is the smallest initial endowment needed to hedge ξ; i.e.,

$$X_0 = \inf\{x \geq 0; \ \exists \, (V, \pi) \in \mathcal{H}(\xi) \text{ such that } V_0 = x\},$$

$$X_0' = \inf\{x \geq 0; \exists \, (V, \pi, C) \in \mathcal{H}'(\xi) \text{ such that } V_0 = x\}.$$

These definitions hold for X_t at any time t.

If Hypothesis 1.1 is satisfied, for any square-integrable nonnegative claim ξ, $\mathcal{H}(\xi)$ is nonempty, and the market is said to be *complete*.[3] Moreover, the fair price is the market value of a hedging strategy in $\mathcal{H}(\xi)$ (Karatzas and Shreve 1987), as proved in the following theorem.

[3]In the case of non-Brownian filtration, the market is incomplete; that is, there exists some contingent claim ξ for which $\mathcal{H}(\xi)$ is empty. For such a contingent claim the fair price is not defined. However, the set $\mathcal{H}'(\xi)$ is nonempty and the upper price is well defined (El Karoui and Quenez 1991,1995).

THEOREM 1.1. *Assume H1.1. Let ξ be a positive square-integrable contingent claim. There exists a hedging strategy (X,π) against ξ such that*

$$dX_t = r_t X_t\,dt + {\pi_t}^*\sigma_t\theta_t\,dt + \pi_t^*\sigma_t\,dW_t, \qquad X_T = \xi, \tag{1.5}$$

and such that the market value X is the fair price and the upper price of the claim.

Let $(H_s^t;\, s \geq t)$ be the deflator started at time t; that is,

$$dH_s^t = -H_s^t[r_s\,ds + \theta_s^*\,dW_s], \qquad H_t^t = 1. \tag{1.6}$$

Then

$$X_t = \mathbb{E}[H_T^t\xi|\mathcal{F}_t], \text{ a.s.} \tag{1.7}$$

Proof. Following Duffie and Epstein (1992a, 1992b), the process

$$H_t = \exp-[\int_0^t r_s\,ds + \int_0^t \theta_s^*\,dW_s + \frac{1}{2}\int_0^t |\theta_s|^2\,ds]$$

is said to be a deflator; it is also the solution of (1.6) started at time 0. Since r and θ are bounded processes, it follows from Novikov's theorem (Karatzas and Shreve 1987, p. 198) that $\mathbb{E}(H_T^2) < +\infty$ and $E(H_T\xi) < +\infty$ for any square-integrable contingent claim. Define the continuous adapted process X from

$$H_t X_t = \mathbb{E}[H_T\xi|\mathcal{F}_t] = M_t,$$

where M is the continuous version of the uniformly integrable nonnegative martingale $\mathbb{E}[H_T^t\xi|\mathcal{F}_t]$. From the martingale representation for the Brownian motion (Karatzas and Shreve 1987, p. 185), M can be represented as a stochastic integral; i.e., there exists a predictable process (U_t) such that

$$H_t X_t = E(H_T\xi) + \int_0^t U_s^*\,dW_s, \qquad \int_0^T |U_t|^2\,dt < +\infty \text{ a.s.}$$

Put $\pi_t = (\sigma_t^*)^{-1}[H_t^{-1}U_t + X_t\,\theta_t]$. Then $H_tX_t = E(H_T\xi) + \int_0^t H_s(\pi_s^*\sigma_s - X_s\theta_s^*)\,dW_s$. By Itô's lemma, (X,π) satisfies the linear BSDE (1.5). Thanks to the continuity of the processes H and X and to the boundedness of θ, we can show that $\int_0^T |\sigma_t^*\pi_t|^2\,dt < +\infty$ a.s. So (X,π) is a hedging strategy against ξ with $X_0 = E(H_T\xi)$.

It remains to show that X_0 (resp. X_t) is the upper price (resp. the smallest superhedging strategy). Let (V,φ,C) be a superhedging strategy against ξ. Again using Itô's lemma for the product of the RCLL semimartingale V and the continuous semimartingale H and using (1.4), we have that $(H_tV_t)_{t\in[0,T]}$ is a positive local supermartingale with decomposition

$d H_t V_t = -H_t \, dC_t + (U_t^V)^* \, dW_t$, where $U_t^V = H_t[-V_t \, \theta_t + (\sigma_t^*)\varphi_t]$. Hence, by Fatou's lemma, $(H_t V_t)_{t \in [0,T]}$ is a positive supermartingale and

$$H_t V_t \geq E[H_T V_T | \mathcal{F}_t] = H_t X_t \,, \qquad V_0 \geq \mathbb{E}(H_T \xi) = X_0. \qquad \square \tag{1.8}$$

REMARK. MARTINGALE MEASURE. Recall that (1.7) corresponds to the well-known property that the fair price of ξ can be evaluated as the expectation of the discounted value of the claim under the so-called risk-neutral probability measure or martingale measure (Harrison and Pliska 1981); that is,

$$X_t = \mathbb{E}_{\mathbb{Q}}[e^{-\int_t^T r_s \, ds} \xi | \mathcal{F}_t],$$

where $\mathbb{Q}$ is the risk-neutral probability measure with Radon-Nikodym derivative with respect to $\mathbb{P}$ on $\mathcal{F}_T$, given by

$$\frac{d\,\mathbb{Q}}{d\mathbb{P}} = \exp - \left[\int_0^T \theta_s^* \, dW_s \; + \frac{1}{2} \int_0^T |\theta_s|^2 \, ds \right].$$

Notice that $\mathbb{Q}$ is well defined as a probability measure since, by assumption, θ is bounded. Moreover, $\mathbb{Q}$ is a martingale measure; that is, the discounted wealth processes are $\mathbb{Q}$-local martingales.

Arbitrage Opportunity and Uniqueness of the Hedging Strategies. We have defined a hedging strategy against a positive contingent claim ξ as a positive solution (V, π) $(V_t \geq 0)$ of the equation

$$dV_t = r_t V_t \, dt + \pi_t{}^* \sigma_t \theta_t \, dt + \pi_t^* \sigma_t \, dW_t, \qquad V_T = \xi. \tag{1.9}$$

Recall that, by definition, an *arbitrage opportunity* is a self-financing strategy (V, π) with freelunch, that is, $V_0 = 0$, $V_T \geq 0$, and $\mathbb{P}(V_T > 0) > 0$. Notice that a hedging strategy against $\xi \geq 0$ and, more generally, any feasible (positive) self-financing strategy cannot be an arbitrage opportunity since if $V_T \geq 0$ and $\mathbb{P}(V_T > 0) > 0$, then $V_0 \geq \mathbb{E}(H_T V_T) > 0$ by (1.8). If the positivity assumption on $(V_t, 0 \leq t < T)$ is relaxed, the self-financing strategy can be an arbitrage opportunity. An example of such a strategy is given below.

EXAMPLE. By using one of Dudley's results (Dudley 1977 or Karatzas and Shreve 1987, p. 189), we can construct a predictable process ψ such that

$$\int_0^T \psi_t^* \, dW_t \; = 1, \qquad 0 < \int_0^T |\psi_t|^2 \, dt < +\infty \text{ a.s.}$$

Put $H_t Y_t = \int_0^t \psi_s \, dW_s$ and $\phi_t = (\sigma_t^*)^{-1}[H_t^{-1} \psi_t + Y_t \, \theta_t]$. The strategy (Y, ϕ) is a self-financing strategy with initial endowment 0 and final value H_T^{-1}. It is an arbitrage opportunity.

Using that arbitrage opportunity, we build an infinite number of self-financing strategies (not positive) which hedge ξ at time T, that is, an infinite number of solutions of the linear BSDE (1.9). From the pricing formula (1.7) we derive that H_t^{-1} is the fair price at time t of the contingent claim H_T^{-1} with hedging portfolio $H_t^{-1}(\sigma_t^*)^{-1}\theta_t$. So the pair $(X^0, \pi^0) = (H^{-1} - Y, H^{-1}(\sigma^*)^{-1}\theta - \phi)$ is a self-financing strategy with initial endowment 1 and terminal value 0 (such a strategy is called a "suicide strategy" by Harrison and Pliska 1981). This pair satisfies the following BSDE:

$$dX_t^0 = r_t X_t^0\, dt + \pi_t^{0*}\sigma_t\theta_t\, dt + (\pi_t^0)^*\sigma_t\, dW_t\,, \qquad X_T^0 = 0.$$

Let (X, π) be a solution of the LBSDE (1.9). Then for each λ the pair $(X + \lambda X^0, \pi + \lambda\pi^0)$ is also a solution of (1.9). So there exists an infinite number of solutions of LBSDE (1.9). Nonuniqueness for the LBSDE holds in general. Furthermore, notice that the value at 0 of the strategy $(X + \lambda X^0, \pi + \lambda\pi^0)$ is equal to $X_0 + \lambda$. Hence, if such strategies are allowed to hedge against ξ, then the fair price for ξ cannot be defined.

Instead of introducing a positivity assumption on the present value of the self-financing strategies, an alternative is to impose some integrability constraints on the strategies. In Section 2, according to the results of Pardoux and Peng (1990), we prove that, under assumption H1.1, (1.9) has a unique solution (X, π) such that $\mathbb{E}\int_0^T |\sigma_t^*\pi_t|^2\, dt < +\infty$. If we consider only square-integrable, self-financing strategies, there exists no arbitrage opportunity. Moreover, for a square-integrable contingent claim, there exists a unique square-integrable hedging strategy.

For a complete study of the different formulations of arbitrage opportunities, see the seminal papers of Harrison and Kreps (1979) and Harrison and Pliska (1981, 1983), and the papers of Delbaen and Schachermayer (1994a, 1994b). Important contributions to this problem can be found in Stricker (1990), Back and Pliska (1987, 1991), Delbaen (1992), Jacka (1994), and many others listed in the references.

1.3. Constrained Portfolios

Recently, in studying the pricing of contingent claims with constraints on the wealth or portfolio processes, many authors have introduced some nonlinear backward equations for the fair price of claims. We present a few examples.

EXAMPLE 1.1. HEDGING CLAIMS WITH HIGHER INTEREST RATE FOR BORROWING. Bergman (1991), Korn (1992), and Cvitanic and Karatzas (1993) consider the following problem: the investor is allowed to borrow money at time t at an interest rate $R_t > r_t$, where r_t is the bond rate (R is assumed to be predictable and bounded). It is not reasonable to borrow money and to invest money in the bond at the same time. Therefore, we restrict ourselves to policies for which the amount borrowed at time t is equal to $(V_t - \sum_{i=1}^n \pi_t^i)^-$. Then the strategy (wealth, portfolio) (V, π) satisfies

$$(1.10) \quad dV_t = r_t\, V_t\, dt + \pi_t^*\sigma_t\theta_t\, dt\ +\ \pi_t^*\sigma_t\, dW_t - (R_t - r_t)\left(V_t - \sum_{i=1}^n \pi_t^i\right)^- dt.$$

Given an inital investment $V_0 = x$ and a risky portfolio π, there exists a unique solution to this forward stochastic differential equation with Lipschitz coefficients. The fair price (resp. upper price) of a claim is still defined as the minimal endowment to finance a strategy which guarantees ξ at time T. According to the results of Pardoux and Peng 1990 (Section 2), there exists a unique square-integrable strategy (X, π) which is a solution of the nonlinear backward stochastic differential equation, where the nonlinear term depends on both wealth and portfolio:

$$
\begin{aligned}
(1.11) \qquad dX_t \;&=\; r_t\, X_t\, dt + \pi_t^* \sigma_t \theta_t\, dt \\
&\quad -(R_t - r_t)\Big(X_t - \sum_{i=1}^{n} \pi_t^i\Big)^- dt + \pi_t^* \sigma_t\, dW_t, \qquad X_T = \xi,
\end{aligned}
$$

with $\mathbb{E}\int_0^T |\pi_t^* \sigma_t|^2\, dt < +\infty$. Furthemore X_t is the fair price and the upper price of ξ at time t.

Similar equations appear in continuous trading with short-sales constraints with different risk premia for long and short positions (Jouini and Kallal 1995a, 1995b; He and Pearson 1991). Let $\theta^1 - \theta^2$ be the difference in excess return between long and short positions in the stocks. Then the present value V of the portfolio strategy π must satisfy

$$
(1.12) \qquad dV_t = r_t\, V_t\, dt + \pi_t^* \sigma_t \theta_t^1\, dt + [\pi_t^*]^- \sigma_t [\theta_t^1 - \theta_t^2]\, dt + \pi_t^* \sigma_t\, dW_t.
$$

Let us suppose that θ^1 and θ^2 are bounded, predictable processes. Given an initial endowment x and a portfolio strategy π there exists a unique solution to this forward equation. Conversely, from the results stated in Section 2, we have that, given a square-integrable contingent claim, there exists a unique square-integrable solution $(X, \sigma^* \pi)$ of the BSDE

$$
dX_t = r_t\, X_t\, dt + \pi_t^* \sigma_t \theta_t^1\, dt + [\pi_t^*]^- \sigma_t [\theta_t^1 - \theta_t^2]\, dt + \pi_t^* \sigma_t\, dW_t, \qquad X_T = \xi.
$$

Here X_t is the fair price and the upper price of the claim ξ at time t. In Section 3.3 we develop a general pricing theory for contingent claims with respect to convex constrained portfolios.

EXAMPLE 1.2. HEDGING CLAIMS IN INCOMPLETE MARKETS. In this section we suppose that only some securities can be traded, and the hedging portfolios can be built by using only these primary securities. In this case the market is incomplete; that is, it is not always possible to replicate a payoff by a controlled portfolio of the basic securities. The assumption that the set of superhedging strategies is nonempty is much milder. Under this assumption, El Karoui and Quenez (1991,1995) showed that there exists an upper price process $(X_t)_{t\in[0,T]}$ for a contingent claim ξ and gave a characterization. This process does not correspond exactly to the solution of a BSDE, but it can be obtained as the increasing limit of a sequence of processes $(X_t^k)_{t\in[0,T]}$ associated with the solutions of nonlinear BSDEs. In that paper a general filtration was considered, but here, for expository simplicity, we restrict our presentation to a Brownian filtration.

More precisely, in this market only some securities, the first j ones ($j \le n$), can be traded, and the hedging portfolios can be built by using only these primary securities. The

$j \times n$ volatility matrix of the primary securities is denoted by $\sigma^1 = (\sigma^{i,k})_{1\le i\le j, 1\le k\le n}$, and the volatility matrix of the others is denoted by σ^2. Hence the volatility matrix σ of the set of market securities can be decomposed at time t as

$$\sigma_t = \begin{pmatrix} \sigma_t^1 \\ \sigma_t^2 \end{pmatrix}.$$

Notice that the matrix $(\sigma_t^1)^*$ has full rank because the global matrix $(\sigma_t)^*$ has full rank, so the matrix $\sigma_t^1(\sigma_t^1)^*$ is invertible. The amount of a general portfolio π_t invested at time t in the primary securities is denoted by ${}^1\pi_t$ and the amount in the others is denoted by ${}^2\pi_t$ so that

$$\pi_t^*\sigma_t = ({}^1\pi_t)^*\sigma_t^1 + ({}^2\pi_t)^*\sigma_t^2.$$

So an admissible hedging portfolio is to be constrained to $({}^2\pi_t) = 0$. Hence it satisfies $\pi_t^*\sigma_t = ({}^1\pi_t)^*\sigma_t^1$, and the admissible wealth V is modeled by

$$dV_t = r_t V_t\,dt + ({}^1\pi_t)^*\sigma_t^1(dW_t + \theta_t\,dt).$$

This equation is unchanged if θ is replaced by the "minimal risk premium" θ^1 defined at time t as the orthogonal projection of θ_t onto the range of $(\sigma_t^1)^*$ since $\theta_t - \theta_t^1$ belongs to the kernel of σ_t^1. Notice that classical results from linear algebra allow us to give a closed formula for θ^1, namely

$$\theta_t^1 = (\sigma_t^1)^*[(\sigma_t^1)(\sigma_t^1)^*]^{-1}\sigma_t^1\theta_t. \tag{1.13}$$

In what follows we suppose the process θ^1 to be bounded.

Given a square-integrable contingent claim ξ, there does not necessarily exist a hedging portfolio built on the primary securities to finance ξ. In others words, the BSDE

$$dX_t = r_t X_t\,dt + ({}^1\pi_t)^*\sigma_t^1[dW_t + \theta_t^1\,dt], \qquad X_T = \xi,$$

can have no solution. Consequently, it is interesting to study the upper price for ξ given by

$$X_0 = \inf\{x\,;\,\exists(V,\pi,C)\;\in\mathcal{H}_1'(\xi);\,V_0 = x\},$$

where $\mathcal{H}_1'(\xi)$ is the set of the superstrategies which only depend on the primary securities; that is,

$$\mathcal{H}_1'(\xi) = \{(V,\pi,C);\,V_T = \xi, dV_t = r_t V_t\,dt + ({}^1\pi_t)^*\sigma_t^1[dW_t + \theta_t^1\,dt] - dC_t;\,V_t \ge 0\}.$$

Suppose that $\mathcal{H}_1'(\xi)$ is nonempty. Then the upper price is well defined and is achieved by a superhedging strategy. This property can be proved by a duality argument.

Let K^1 be the subspace of bounded and predictable processes which take values in the kernel of σ^1; that is,

$$K^1 = \left\{\beta;\ \sigma_t^1 \beta_t = 0,\ \text{a.s. } t \in [0,T];\ \exists B > 0, |\beta_t| \le B\right\}.$$

We strongly penalize the presence of assets $(j+1, \dots, n)$ in the general replicating strategies by introducing a risk premium $\beta \in K^1$. The corresponding β-hedging strategy (V^β, π^β) is the solution of the BSDE

$$dV_t^\beta = r_t V_t^\beta\, dt + (\pi_t^\beta)^* \sigma_t [dW_t + \theta_t^1\, dt + \beta_t\, dt], \qquad V_T = \xi.$$

In the terminology of Karatzas et al. (1989), (V^β, π^β) is the fair price for ξ in a fictitious market, which completes the initial market.

Define X as the right-continuous, left-limited (RCLL) process which satisfies

$$X_t = \operatorname{ess\,sup}\{V_t^\beta;\ \beta \in K^1\}, \qquad \mathbb{P} \text{ a.s.} \tag{1.14}$$

El Karoui and Quenez (1995) showed that the process X is the upper price and that X is the market value of an admissible superhedging strategy; i.e., there exist a portfolio process ${}^1\pi$ and a consumption process C such that

$$dX_t = r_t X_t\, dt + ({}^1\pi_t)^* \sigma_t^1 [dW_t + \theta_t^1\, dt] - dC_t, \qquad X_T = \xi. \tag{1.15}$$

The process X can be approximated by the continuous processes (X^k) satisfying, for $k \in \mathbb{N}$,

$$X_t^k = \operatorname{ess\,sup}\{V_t^\beta;\ |\beta| \le k,\ \beta \in K^1\}, \qquad \mathbb{P} \text{ a.s.} \tag{1.16}$$

We prove (see Section 3) that, since

$$\begin{aligned}
dV_t^\beta &= r_t V_t^\beta\, dt + (\pi_t^\beta)^* \sigma_t \beta_t\, dt + (\pi_t^\beta)^* \sigma_t [dW_t + \theta_t^1\, dt], \qquad V_T^\beta = \xi,\\
dX_t^k &= r_t X_t^k - \sup_{\beta \in K^1;\ |\beta| \le k} [(-\pi_t^k)^* \sigma_t \beta]\, dt + (\pi_t^k)^* \sigma_t [dW_t + \theta_t^1\, dt], \qquad X_T^k = \xi.
\end{aligned}$$

Because $\sup_{\beta \in K^1;\ |\beta| \le k} (-\pi_t^* \sigma_t \beta) = k \|\operatorname{Proj}_t(\sigma_t^* \pi_t)\|$, where Proj_t denotes the orthogonal projection onto the kernel of σ_t^1,

$$dX_t^k = r_t X_t^k - k \|\operatorname{Proj}_t(\sigma_t^* \pi_t^k)\|\, dt + (\pi_t^k)^* \sigma_t [dW_t + \theta_t^1\, dt], \qquad X_T^k = \xi.$$

The strategies (X^k, π^k) can also be considered as penalized strategies. The penalizing process given by $k \int_0^t \|\operatorname{Proj}_s(\sigma_s^* \pi_s^k)\|\, ds$ has an intensity proportional to the length of the nonadmissible part of $(\sigma_s^* \pi_s^k)$. The larger this length, the more the local "variance" of the nonadmissible part of $\int_0^t (\pi_s^k)^* \sigma_s \, dW_s$ is large and the more the penalty is expensive. Notice that the penalizing process only depends on $\mathrm{Proj}_t((\sigma_t^2)^*(^2\pi_t)) = \mathrm{Proj}_t((\sigma_t)^*(\pi_t))$.

By Pardoux and Peng's results on nonlinear BSDEs that we recall in Section 2, these equations admit a unique square-integrable solution (X^k, π^k). By the comparison theorem for BSDEs, we can show that the sequence of the processes X^k is increasing, and by using their interpretation as a value function of a control problem, it follows that their limit is equal to X.

Notice that these solutions are obtained as the value functions of a control problem associated with the fair prices for the claim in fictitious markets. We shall see in Section 3 that this property is very general and results from the convexity of the function $\mathrm{Proj}_t(\cdot)$.

Cvitanic and Karatzas (1993) studied this methodology for very general constraints on the portfolio in a Brownian market and obtained similar approximations for the upper price of contingent claim ξ. Bardhan (1993) also studied some problems in this area. More recently, Kramkov (1994) and Föllmer and Kabanov (1995) extended this methodology to a general arbitrage-free asset pricing setting in incomplete markets without restrictions on the filtration.

EXAMPLE 1.3. FÖLLMER-SCHWEIZER HHEDGING STRATEGY IN INCOMPLETE MARKETS. The model and the notation are the same as in Example 1.2. Let ξ be a square-integrable contingent claim. In this context a strategy $(V,{}^1\pi,\phi)$, is called a *nonadjusted hedging strategy* against ξ when

$$dV_t = r_t V_t \, dt + (^1\pi_t)^* \sigma_t^1 [dW_t + \theta_t \, dt] + d\phi_t \,, \qquad V_T = \xi, \tag{1.17}$$

where the process ϕ is a RCLL semimartingale satisfying $\phi_0 = 0$. The process $(-\phi_t)$ is called the *tracking error*. In particular, at the terminal time the tracking error measures the spread between the contingent claim ξ and the portfolio value, and ϕ corresponds to the cost process introduced by Föllmer and Schweizer (1990). Notice that a self-financing hedging strategy corresponds to a tracking error equal to zero and that a superhedging strategy corresponds to an increasing tracking error. If ϕ is a martingale orthogonal to $\int_0^\cdot \sigma_s^1 \, dW_s$, it will be called a Föllmer-Schweizer hedging strategy. Again, we remark that this equation is unchanged if θ is replaced by the "minimal risk premium" θ^1. More precisely,

DEFINITION 1.5. A strategy $(X,{}^1\pi, M)$ is called a *Föllmer-Schweizer hedging strategy* (or FS-strategy) if ${}^1\pi$ is a j-dimensional predictable process such that $\mathbb{E}(\int_0^T |(\sigma_s^1)^*(^1\pi_t)|^2 dt) < +\infty$ and if M is a square-integrable martingale orthogonal to $\int_0^\cdot \sigma_s^1 \, dW_s$ such that

$$dX_t = r_t X_t \, dt + (^1\pi_t)^* \sigma_t^1 [dW_t + \theta_t \, dt] + dM_t \,, \qquad X_T = \xi.$$

REMARK. Such strategies were first introduced by Föllmer and Sondermann (1986). Initially, the problem is to find a strategy with minimal variance for the tracking error, but it is rather complicated. In the case $r = 0$ and $\theta = 0$ (that is, $\mathbb{P}$ is a martingale measure) considered by Föllmer and Sondermann, it is easy to see that the tracking error of the minimal variance strategy is a martingale (such a strategy is said to be mean-self-financing); actually, the minimal variance tracking error is characterized to be a martingale process orthogonal to $\int_0^{\cdot} \sigma_s^1\, dW_s$. If $\theta \neq 0$ the situation is more subtle. Schweizer (1991) introduced a criterion of local risk minimization and showed that a nonadjusted hedging strategy is locally risk minimizing if the associated tracking error is a martingale orthogonal to $\int_0^t \sigma_s^1\, dW_s$. New developments in this area include Ansel and Stricker (1992a,1992b), Schweizer (1992, 1994a, 1994b), Delbaen et al. (1994), and Monat and Stricker (1995).

Recall that in their paper, Föllmer and Schweizer (1990) characterized the FS-strategy in a different framework: there exists only one primary risky security, and the price process of this unique basic security is supposed to be any continuous square-integrable semimartingale. Actually, the arbitrage-free hypothesis implies that it can be written $N_t + \int_0^t \alpha_s\, d\langle N\rangle_s$ for some continuous square-integrable martingale N with quadratic variation $\langle N\rangle$ and some predictable process α.

Indeed, the FS-strategy is simply given by the solution of a linear BSDE.

PROPOSITION 1.1. *Let (X, ψ) be the hedging strategy against ξ in a market with the "minimal risk premium" θ^1; that is,*

$$dX_t = r_t X_t\, dt + \psi_t^* \sigma_t \theta_t^1\, dt + \psi_t^* \sigma_t\, dW_t\,, \qquad X_T = \xi. \tag{1.18}$$

Put $q_t = \sigma_t^* \psi_t$. Let q_t^1 (resp. q_t^2) be the orthogonal projection of q_t onto the range of $(\sigma_t^1)^*$ (resp. the kernel of σ_t^1). Let ${}^1\pi_t$ be the vector process such that $(\sigma_t^1)^*({}^1\pi_t) = q_t^1$, given in closed form by ${}^1\pi_t = [\sigma_t^1(\sigma_t^1)^*]^{-1}\sigma_t^1 q_t$. Put $M_t = \int_0^t (q_s^2)^* dW_s$. Then $(X, {}^1\pi, M)$ is the unique FS-strategy associated with ξ. In other words, the FS-strategy corresponds to the hedging strategy in a fictitious market with no penalizing risk premium ($\beta = 0$).

Proof. Let (X, ψ) be the square-integrable hedging strategy against ξ, in a market with the risk premium vector θ^1, and put $q_t = \sigma_t^* \psi_t$. Project the vector q_t orthogonally onto the range of $(\sigma_t^1)^*$, so

$$q_t = q_t^1 + q_t^2 \qquad \text{where } q_t^1 \in \operatorname{Range}((\sigma_t^1)^*) \text{ and } q_t^2 \in \operatorname{Ker}(\sigma_t^1).$$

Let ${}^1\pi_t$ be the vector process satisfying $q_t^1 = (\sigma_t^1)^*({}^1\pi_t)$ (notice that ${}^1\pi_t$ is unique since the matrix $(\sigma_t^1)^*$ is full rank). We have $q_t^* \theta_t^1 = (q_t^1)^* \theta_t^1 = (q_t^1)^* \theta_t$. Put $M_t = \int_0^t (q_t^2)^*\, dW_s$. M is a martingale orthogonal to $\int_0^{\cdot} \sigma_s^1\, dW_s$, and $(X, {}^1\pi, M)$ is an FS-strategy.

Conversely, let $(\widehat{X}, ({}^1\widehat{\pi}), \widehat{M})$ be an FS-strategy. Let $\widehat{q}_t^2$ be such that $\widehat{M}_t = \int_0^t (\widehat{q}_t^2)^*\, dW_s$ and put $\widehat{q}_t = (\sigma_t^1)^*({}^1\widehat{\pi}_t) + \widehat{q}_t^2$, and $\widehat{\psi}_t = (\sigma_t^*)^{-1}\widehat{q}_t$. Then $(\widehat{X}, \widehat{\psi})$ is a solution of BSDE (1.18), and it follows from the uniqueness of the solution of a BSDE that the FS-strategy is unique. □

REMARK 1.1. Suppose that the matrix $\sigma_t^1(\sigma_t^2)^*$ is the null matrix; that is, the nonprimary securities do not introduce supplementary risk on the admissible portfolios. Then the FS-strategy consists of holding the amount ${}^1\pi_t = [\sigma_t^1(\sigma_t^1)^*]^{-1}\sigma_t^1(\sigma_t^1)^*({}^1\psi)_t = ({}^1\psi)_t$ in the primary securities. Moreover, by uniqueness the FS-strategy does not depend on the matrix σ^2. Consequently, the simplest way to compute the FS-strategy is to complete the "primary" market by introducing other securities whose volatility matrix satisfies $\sigma_t^1(\sigma_t^2)^* = 0$, a.s.

REMARK 1.2. Let $\mathbb{Q}^1$ be the martingale measure associated with θ^1. Using the same arguments as in the proof of Theorem 1.1, it follows directly from Proposition 1.1 that X_t is the conditional expectation of the discounted contingent claim computed under the martingale measure $\mathbb{Q}^1$.

Let us show that $\mathbb{Q}^1$ is the minimal martingale measure (El Karoui and Quenez 1995, Proposition 1.8.2) first introduced by Föllmer and Schweizer (1990) as a martingale measure such that any bounded $\mathbb{P}$-martingale orthogonal to the martingale $\int_0^{\cdot} \sigma_s^1\, dW_s$ remains a bounded $\mathbb{Q}^1$-martingale. Indeed, using the same arguments as in the proof of Theorem 1.1, we derive that a bounded, continuous process X which is also a $\mathbb{Q}^1$-local martingale corresponds to the square-integrable solution (X, ψ) of BSDE (1.18) with $r = 0$ and $\xi = X_T$. Now suppose that X' is a $\mathbb{P}$-martingale orthogonal to the martingale $\int_0^{\cdot} \sigma_s^1\, dW_s$. Then X' is the stochastic integral of a process q' in the kernel of σ^1. So $(q_t')^*\theta_t^1 = 0$ and $(X', (\sigma^*)^{-1}q')$ is a solution of the BSDE (1.18) with $r_. = 0$, so X' is a $\mathbb{Q}^1$-local martingale.

1.4. Recursive Utility

In the continuous-time, deterministic case, recursive utilities were first (to our knowledge) introduced by Epstein and Zin (1989). Let us consider a small agent who can consume between time 0 and time T. Let c_t be the (positive) consumption rate at time t. We assume that there exists a terminal reward Y at time T. The utility at time t is a function of the instantaneous consumption rate c_t and of the future utility (corresponding to the future consumption). In fact, the recursive utility Y is assumed to satisfy the following differential equation:

$$-\, dY_t = f(c_t, Y_t)\, dt\,, \qquad Y_T = Y. \tag{1.19}$$

The function f is called *the generator*. Thus, at time 0 the utility of the consumption path $(c_t,\ 0 \le t \le T)$ is

$$Y_0 = Y + \int_0^T f(c_s, Y_s)\, ds.$$

Under uncertainty, Duffie and Epstein (1992a, 1992b) (see also Duffie, Geoffard, and Skiadias 1992) introduced the following class of recursive utilities:

$$-\, dY_t = \left[f(c_t, Y_t) - A(Y_t)\tfrac{1}{2} Z^* Z_t \right] dt - Z_t^*\, dW_t\,, \qquad Y_T = Y, \tag{1.20}$$

where A is the “variance multiplier.” We can give another representation of the utility at time t of the future consumption $(c_s;\ \ t \le s \le T)$:

$$Y_t = E\left[Y + \int_t^T [f(c_s, Y_s) - A(Y_s)\tfrac{1}{2} Z_s^*\, Z_s]\, ds \,|\mathcal{F}_t \right].$$

Because of their economic motivations, we provide the following examples of recursive utilities, following Duffie and Epstein (1992a).

Examples.
• *Standard additive utility*. The generator of the standard utility is

$$f(c, y) = u(c) - \beta\, y,$$

The recursive utility is

$$Y_t = E\left[Y e^{-\beta(T-t)} + \int_t^T u(c_s)\, e^{-\beta(s-t)}\, ds\, |\mathcal{F}_t\right].$$

• *Uzawa utility*. The generator has the same form as additive utility, but the discounting rate β depends on the consumption rate c_t:

$$f(c, y) = u(c) - \beta(c)y.$$

• *Kreps-Porteus utility*. Let $0 \neq \rho \leq 1$ and $0 \leq \beta$. The generator is defined by

$$f(c, y) = \frac{\beta}{\rho}\frac{c^\rho - y^\rho}{y^{\rho-1}}.$$

In general, utilities must satisfy the following classical properties:

- Monotonicity with respect to the terminal value and to the consumption.
- Concavity with respect to the consumption.
- Time consistency: this means that, for any two consumption processes c^1 and c^2 and any time t, if c^1 and c^2 are identical up to time t and if the continuation of c^1 is preferred to the continuation of c^2 at time t, then c^1 is preferred to c^2 at time 0.

Duffie and Epstein (1992a) showed that if f is Lipschitz with respect to y, then

$$-dY_t = f(t, c_t, Y_t)\, dt - Z_t^*\, dW_t\,, \qquad Y_T = Y,$$

has a unique solution. Also, they state that if f is concave with respect to (c, y) and increasing with respect to c, the above properties are satisfied.[4]

We will consider a more general class of recursive utilities, defined as associated with the solution of a general BSDE

$$(1.21) \qquad - dY_t = f(t, c_t, Y_t, Z_t)\, dt - Z_t^*\, dW_t\,, \qquad Y_T = Y,$$

with concave generator f. The existence and uniqueness of solutions of (1.21) are proved in Section 2. Also, the above properties are obtained as direct consequences of a comparison

[4]More recently, Duffie and Singleton (1994) and Duffie and Huang (1994) have used this type of BSDE to solve some pricing problems.

theorem. The main result stated in Section 3 is the interpretation of recursive utility as the value function of a control problem.

2. BACKWARD STOCHASTIC DIFFERENTIAL EQUATIONS

A linear backward stochastic differential equation was introduced by Bismut (1973) as the equation for the conjugate variable (or adjoint process) in the stochastic version of the Pontryagin maximum principle. Other works on the maximum principle were also done using linear BSDEs by Arkin and Saksonov (1979), Kabanov (1978), and Cadenillas and Karatzas (1995). Bismut (1978) introduced a nonlinear BSDE (a Riccati equation) for which he showed existence and uniqueness. Pardoux and Peng (1990) were the first to consider general BSDEs.

Several papers extended their results, particularly Antonelli (1993), Ma, Protter, and Yong (1994), Buckdahn (1993), Buckdahn and Pardoux (1994), and, of course, Pardoux and Peng (1990,1992,1993) and Peng (1990, 1991, 1992a, 1992b, 1992c, 1993).

In this section we present some important results for BSDEs. First we state some a priori estimates of the spread between the solutions of two BSDEs, from which we derive the results of existence and uniqueness. Then we give different properties concerning BSDEs. In particular, we study one-dimensional linear BSDEs which are classical in finance, for which we state a comparison theorem.

First we fix some notation. For $x \in \mathbb{R}^d$, $|x|$ denotes its Euclidian norm and $\langle x, y\rangle$ denotes the inner product. An $n \times d$ matrix will be considered as an element $y \in \mathbb{R}^{n\times d}$; note that its Euclidean norm is also given by $|y| = \sqrt{\text{trace}(y y^*)}$ and that $\langle y, z\rangle = \text{trace}(yz^*)$.

Given a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ and an $\mathbb{R}^n$-valued Brownian motion W, we consider

- $\{(\mathcal{F}_t);\, t \in [0, T]\}$, the filtration generated by the Brownian motion W and augmented, and $\mathcal{P}$ the σ-field of predictable sets of $\Omega \times [0, T]$.
- $\mathbb{L}^2_T(\mathbb{R}^d)$, the space of all $\mathcal{F}_T$-measurable random variables $X\colon \Omega \mapsto \mathbb{R}^d$ satisfying $||X||^2 = \mathbb{E}(|X|^2) < +\infty$.
- $\mathbb{H}^2_T(\mathbb{R}^d)$, the space of all predictable processes $\phi\colon \Omega\times[0, T] \mapsto \mathbb{R}^d$ such that $||\varphi||^2 = \mathbb{E}\int_0^T |\varphi_t|^2\,dt < +\infty$.
- $\mathbb{H}^1_T(\mathbb{R}^d)$, the space of all predictable processes $\phi\colon \Omega \times [0, T] \mapsto \mathbb{R}^d$ such that $\mathbb{E}\sqrt{\int_0^T |\varphi_t|^2\,dt} < +\infty$.
- For $\beta > 0$ and $\phi \in \mathbb{H}^2_T(\mathbb{R}^d)$, $||\phi||^2_\beta$ denotes $\mathbb{E}\int_0^T e^{\beta t}|\phi_t|^2\,dt$. $\mathbb{H}^2_{T,\beta}(\mathbb{R}^d)$ denotes the space $\mathbb{H}^2_T(\mathbb{R}^d)$ endowed with the norm $||\cdot||_\beta$.

For notational simplicity we sometimes use $\mathbb{L}^2_T(\mathbb{R}^d) = \mathbb{L}^{2,d}_T$, $\mathbb{H}^2_T(\mathbb{R}^d) = \mathbb{H}^{2,d}_T$, $\mathbb{H}^1_T(\mathbb{R}^d) = \mathbb{H}^{1,d}_T$, and $\mathbb{H}^2_{T,\beta}(\mathbb{R}^d) = \mathbb{H}^{2,d}_{T,\beta}$.

2.1. Existence and Uniqueness of Backward Stochastic Differential Equations

The Main Result. Consider the BSDE

$$-\,dY_t = f(t, Y_t, Z_t)\,dt - Z^*_t\,dW_t\,, \qquad Y_T = \xi, \tag{2.1}$$

or, equivalently,

$$Y_t = \xi + \int_t^T f(s, Y_s, Z_s)\, ds \;-\; \int_t^T Z_s^* dW_s, \tag{2.2}$$

where

- The terminal value is an $\mathcal{F}_T$-measurable random variable, $\xi : \Omega \mapsto \mathbb{R}^d$.
- The generator f maps $\Omega \times \mathbb{R}^+ \times \mathbb{R}^d \times \mathbb{R}^{n\times d}$ onto $\mathbb{R}^d$ and is $\mathcal{P} \otimes \mathcal{B}^d \otimes \mathcal{B}^{n\times d}$-measurable.

A *solution* is a pair (Y, Z) such that $\{Y_t;\, t \in [0, T]\}$ is a continuous $\mathbb{R}^d$-valued adapted process and $\{Z_t;\, t \in [0, T]\}$ is an $\mathbb{R}^{n\times d}$-valued predictable process satisfying $\int_0^T |Z_s|^2\, ds < +\infty$, $\mathbb{P}$ a.s.

Suppose that $\xi \in \mathbb{L}_T^2(\mathbb{R}^d)$, $f(\cdot, 0, 0) \in \mathbb{H}_T^2(\mathbb{R}^d)$, and f is uniformly Lipschitz; i.e., there exists $C > 0$ such that $d\mathbb{P} \otimes dt$ a.s.

$$|f(\omega, t, y_1, z_1) - f(\omega, t, y_2, z_2)| \le C\,(|y_1 - y_2| + |z_1 - z_2|) \quad \forall (y_1, z_1), \forall (y_2, z_2).$$

Then (f, ξ) are said to be *standard* parameters for the BSDE.

THEOREM 2.1 (Pardoux-Peng 1990). *Given standard parameters (f, ξ), there exists a unique pair $(Y, Z) \in \mathbb{H}_T^2(\mathbb{R}^d) \times \mathbb{H}_T^2(\mathbb{R}^{n\times d})$ which solves (2.1).*

We often refer to such a solution as a square-integrable solution. A proof can be found in Pardoux and Peng (1990). We give here a shorter direct proof using useful a priori estimates.

A Priori Estimates.

PROPOSITION 2.1. *Let $((f^i, \xi^i);$ i=1,2) be two standard parameters of the BSDE and $((Y^i, Z^i);$ i=1,2) be two square-integrable solutions. Let C be a Lipschitz constant for f^1, and put $\delta Y_t = Y_t^1 - Y_t^2$ and $\delta_2 f_t = f^1(t, Y_t^2, Z_t^2) - f^2(t, Y_t^2, Z_t^2)$. For any (λ, μ, β) such that $\mu > 0$, $\lambda^2 > C$, and $\beta \ge C(2 + \lambda^2) + \mu^2$, it follows that*

$$||\delta Y||_\beta^2 \;\le\; T\left[e^{\beta T}\mathbb{E}(|\delta Y_T|^2) + \frac{1}{\mu^2}||\delta_2 f||_\beta^2\right], \tag{2.3}$$

$$||\delta Z||_\beta^2 \;\le\; \frac{\lambda^2}{\lambda^2 - C}\left[e^{\beta T}\mathbb{E}(|\delta Y_T|^2) + \frac{1}{\mu^2}||\delta_2 f||_\beta^2\right]. \tag{2.4}$$

Proof. Let $(Y, Z) \in \mathbb{H}_T^2(\mathbb{R}^d) \times \mathbb{H}_T^2(\mathbb{R}^{n\times d})$ be a solution of (2.1). Using (2.2), we derive that

$$|Y_t| \le |\xi| + \int_0^T |f(s, Y_s, Z_s)|\, ds \;+\; \sup_t \left|\int_t^T Z_s^*\, dW_s\right|.$$

It follows from Burkholder–Davis–Gundy inequalities (Karatzas and Shreve 1987, Theorem 3.28) that

$$\mathbb{E}\left[\sup_t |\int_t^T Z_s^* \, dW_s|^2\right] \le 2\mathbb{E}\left[|\int_0^T Z_s^* \, dW_s|^2\right] + 2\mathbb{E}\left[\sup_t |\int_0^t Z_s^* \, dW_s|^2\right] \le 4\mathbb{E}\left[\int_0^T |Z_s|^2 \, ds\right].$$

Now since (f, ξ) are standard parameters, $|\xi| + \int_0^T |f(s, Y_s, Z_s)| \, ds$ belongs to $\mathbb{L}_T^{2,1}$ and $\sup_{s \le T} |Y_s| \in \mathbb{L}_T^{2,1}$.

Now consider (Y^1, Z^1) and (Y^2, Z^2), the two solutions associated with (f^1, ξ^1) and (f^2, ξ^2), respectively. From Itô's formula applied from $s = t$ to $s = T$ to the semimartingale $e^{\beta s}|\delta Y_s|^2$, it follows that

$$\begin{aligned} &e^{\beta t}|\delta Y_t|^2 + \beta \int_t^T e^{\beta s}|\delta Y_s|^2 \, ds + \int_t^T e^{\beta s}|\delta Z_s|^2 \, ds \\ &\quad = e^{\beta T}|\delta Y_T|^2 + 2\int_t^T e^{\beta s}\langle \delta Y_s, \ f^1(s, Y_s^1, Z_s^1) - f^2(s, Y_s^2, Z_s^2)\rangle \, ds \\ &\qquad - 2\int_t^T e^{\beta s}\langle \delta Y_s, \delta Z_s^* \, dW_s\rangle. \end{aligned}$$

Since $\sup_{s \le T}|\delta Y_s|$ belongs to $\mathbb{L}_T^{2,1}$, $e^{\beta s}\delta Z_s \, \delta Y_s$ belongs to $\mathbb{H}_T^{1,n}$ and the stochastic integral $\int_t^T e^{\beta s}\langle \delta Y_s, \delta Z_s^* \, dW_s\rangle$ is $\mathbb{P}$-integrable, with zero expectation. Moreover,

$$\begin{aligned} |f^1(s, Y_s^1, Z_s^1) - f^2(s, Y_s^2, Z_s^2)| &\le |f^1(s, Y_s^1, Z_s^1) - f^1(s, Y_s^2, Z_s^2)| + |\delta_2 f_s| \\ &\le C[|\delta Y_s| + |\delta Z_s|] + |\delta_2 f_s| \,. \end{aligned}$$

The inequality $2y(Cz + t) \le Cz^2/\lambda^2 + t^2/\mu^2 + y^2(\mu^2 + C\lambda^2)$ $(\lambda, \mu > 0)$ implies

$$\begin{aligned} &\mathbb{E}[e^{\beta t}|\delta Y_t|^2] + \beta\mathbb{E}\left[\int_t^T e^{\beta s}|\delta Y_s|^2 \, ds + \int_t^T e^{\beta s}|\delta Z_s|^2 \, ds\right] \\ &\quad \le \mathbb{E}[e^{\beta T}|\delta Y_T|^2] + \mathbb{E}\int_t^T e^{\beta s}[C|\delta Y_s|^2(2 + \lambda^2) \\ &\qquad + C\frac{|\delta Z_s|^2}{\lambda^2} + \frac{|\delta_2 f_s|^2}{\mu^2} + \mu^2|\delta Y_s|^2] \, ds \\ &\quad \le \mathbb{E}[e^{\beta T}|\delta Y_T|^2] + [C(2 + \lambda^2) + \mu^2]\mathbb{E}\int_t^T e^{\beta s}|\delta Y_s|^2 \, ds \\ &\qquad + \frac{C}{\lambda^2}\mathbb{E}\int_t^T e^{\beta s}|\delta Z_s|^2 \, ds + \frac{1}{\mu^2}\mathbb{E}\int_t^T e^{\beta s}|\delta_2 f_s|^2 \, ds. \end{aligned} \tag{2.5}$$

Choosing $\beta \geq C(2+\lambda^2)+\mu^2$ and $C<\lambda^2$, these inequalities give

$$\mathbb{E}[e^{\beta t}|\delta Y_t|^2] \leq \mathbb{E}[e^{\beta T}|\delta Y_T|^2] + \mathbb{E}\int_t^T e^{\beta s}|\delta_2 f_s|^2 \frac{1}{\mu^2}\, ds.$$

We obtain the control of the norm of the process δY by integration. Then the control of the norm of the process δZ follows by inequality (2.5). □

REMARKS. (a) In the control of the norm of δY, we can replace T by $\inf(T, [\beta - (C(2+\lambda^2)+\mu^2)]^{-1})$.

(b) By classical results on the norms of semimartingales, we prove similarly that

$$\mathbb{E}[\sup_{t\leq T}|\delta Y_t|^2] \leq K\mathbb{E}\left[|\delta Y_T|^2 + \int_0^T |\delta_2 f_t|^2\, dt\right],$$

where K is a positive constant only depending on T.

Proof of Theorem 2.1. We use a fixed-point theorem for the mapping from $\mathbb{H}^2_{T,\beta}(\mathbb{R}^d)\times\mathbb{H}^2_{T,\beta}(\mathbb{R}^{n\times d})$ into $\mathbb{H}^2_{T,\beta}(\mathbb{R}^d)\times\mathbb{H}^2_{T,\beta}(\mathbb{R}^{n\times d})$, which maps (y,z) onto the solution (Y,Z) of the BSDE with generator $f(t,y_t,z_t)$; i.e.,

$$Y_t = \xi + \int_t^T f(s,y_s,z_s)\, ds - \int_t^T Z_s^*\, dW_s\,.$$

Let us remark that the assumption that (f,ξ) are standard parameters implies that $(f(t,y_t,z_t); t\in[0,T])$ belongs to $\mathbb{H}^2_T(\mathbb{R}^d)$. The solution (Y,Z) is defined by considering the continuous version M of the square-integrable martingale $\mathbb{E}[\int_0^T f(s,y_s,z_s)\, ds + \xi|\mathcal{F}_t]$. By the martingale representation theorem for the Brownian motion (Karatzas and Shreve 1987, Theorem 4.15) there exists a unique integrable process $Z\in\mathbb{H}^{2,n\times d}_T$ such that $M_t = M_0 + \int_0^t Z_s^*\, dW_s$. Define the adapted and continuous process Y by $Y_t = M_t - \int_0^t f(s,y_s,z_s)\, ds$. Notice that Y is also given by

$$Y_t = \mathbb{E}\left[\int_t^T f(s,y_s,z_s)\, ds + \xi|\mathcal{F}_t\right].$$

The square integrability of Y follows from the above assumptions.

Let (y^1,z^1), (y^2,z^2) be two elements of $\mathbb{H}^{2,d}_{T,\beta}\times\mathbb{H}^{2,n\times d}_{T,\beta}$, and let (Y^1,Z^1) and (Y^2,Z^2) be the associated solutions. By Proposition 2.1 applied with $C=0$ and $\beta=\mu^2$, we obtain

$$||\delta Y||^2_\beta \leq \frac{T}{\beta}\mathbb{E}\int_0^T e^{\beta s}|f(s,y^1_s,z^1_s) - f(s,y^2_s,z^2_s)|^2\, ds$$

and

$$||\delta Z||_\beta^2 \leq \frac{1}{\beta}\mathbb{E}\int_0^T e^{\beta s}|f(s, y_s^1, z_s^1) - f(s, y_s^2, z_s^2)|^2\, ds\,.$$

Now since f is Lipschitz with constant C, we have

$$||\delta Y||_\beta^2 + ||\delta Z||_\beta^2 \leq \frac{2(1+T)C}{\beta}[||\delta y||_\beta^2 + ||\delta z||_\beta^2]. \tag{2.6}$$

Choosing $\beta > 2(1+T)C$, we see that this mapping Φ is a contraction from $\mathbb{H}_{T,\beta}^{2,d} \times \mathbb{H}_{T,\beta}^{2,n\times d}$ onto itself and that there exists a fixed point,[5] which is the unique continuous solution of the BSDE. □

From the proof of Proposition 2.1 (and more precisely from estimate (2.6)), we derive that the Picard iterative sequence converges almost surely to the solution of the BSDE.

COROLLARY 2.1. *Let β be such that $2(1+T)C < \beta$. Let (Y^k, Z^k) be the sequence defined recursively by* $(Y_0 = 0;\ Z_0 = 0)$ *and*

$$-\,dY_t^{k+1} = f(t, Y_t^k, Z_t^k)\,dt - (Z_t^{k+1})^*\,dW_t\,, \qquad Y_T^{k+1} = \xi\,. \tag{2.7}$$

Then the sequence (Y^k, Z^k) converges to (Y, Z), $d\mathbb{P}\otimes dt$ a.s. (and in $\mathbb{H}_{T,\beta}^2(\mathbb{R}^d) \times \mathbb{H}_{T,\beta}^2(\mathbb{R}^{n\times d})$) as k goes to $+\infty$.

Proof. Let (Y^k, Z^k) be the sequence defined recursively by (2.7). Then by (2.6),

$$||Y^{k+1} - Y^k||_\beta^2 + ||Z^{k+1} - Z^k||_\beta^2 \leq \epsilon^k K,$$

where $K = ||Y^1 - Y^0||_\beta^2 + ||Z^1 - Z^0||_\beta^2$ and $\epsilon = 2(1+T)C/\beta < 1$. Hence

$$\sum_k ||Y^{k+1} - Y^k||_\beta^2 + \sum_k ||Z^{k+1} - Z^k||_\beta^2 < +\infty,$$

and the result follows. □

REMARK. Again for Y it is possible to consider the norm $||\sup_{s\in[0,T]}|Y_s^k - Y_s|\,||_2$ instead of $||Y||_\beta$; consequently we also have that $\sup_{s\in[0,T]}|Y_s^k - Y_s|$ converges $\mathbb{P}$ a.s. to 0.

[5] Let $(\overline{Y}, \overline{Z})$ be a representation of the fixed point of the mapping Φ in the class $\mathbb{H}_{T,\beta}^2(\mathbb{R}^d)\times\mathbb{H}_{T,\beta}^2(\mathbb{R}^{n\times d})$, and choose the continuous version Y defined by $Y_t = \mathbb{E}[\int_t^T f(s, \overline{Y}_s, \overline{Z}_s)\,ds + \xi|\mathcal{F}_t] = \mathbb{E}[\int_t^T f(s, Y_s, \overline{Z}_s)\,ds + \xi|\mathcal{F}_t]$. Hence,$(Y, \overline{Z})$ is a continuous solution of the BSDE.

Linear BSDE. Theorem 2.1 applied to linear BSDEs specifies the integrability properties of the solution of the standard pricing problem (Theorem 1.1).

PROPOSITION 2.2. *Let (β, γ) be a bounded $(\mathbb{R}, \mathbb{R}^n)$-valued predictable process, φ an element of $\mathbb{H}_T^2(\mathbb{R})$, and ξ an element of $\mathbb{L}_T^2(\mathbb{R})$. Then the LBSDE*

$$-\,dY_t = [\varphi_t \,+ Y_t\,\beta_t + Z_t^*\gamma_t]\,dt \,-\, Z_t^*\,dW_t\,, \qquad Y_T = \xi, \tag{2.8}$$

has a unique solution (Y, Z) in $\mathbb{H}^2_{T,\beta}(\mathbb{R}) \times \mathbb{H}^2_{T,\beta}(\mathbb{R}^n)$ and Y_t is given by the closed formula

$$Y_t \,=\, \mathbb{E}\left[\xi\Gamma_T^t + \int_t^T \Gamma_s^t\varphi_s\,ds\,|\mathcal{F}_t\right] \mathbb{P}\text{ a.s.}, \tag{2.9}$$

where Γ_s^t is the adjoint process defined for $s \geq t$ by the forward LSDE

$$d\Gamma_s^t = \Gamma_s^t[\beta_s\,dt\,s + \gamma_s^*\,dW_s]\,, \qquad \Gamma_t^t = 1. \tag{2.10}$$

In particular, if ξ and φ are nonnegative, the process Y is nonnegative. If, in addition, $Y_0 = 0$, then, for any t, $Y_t = 0$ a.s., $\xi = 0$ a.s., and $\varphi_t = 0\; d\mathbb{P} \otimes dt$ a.s.

Proof. Since β and γ are bounded processes, the linear generator $f(t, y, z) = \varphi_t + \beta_t y + \gamma_t^* z$ is uniformly Lipschitz and the pair (f, ξ) are standard parameters. By Theorem 2.1 there exists a unique square-integrable solution (Y, Z) of the linear BSDE associated with (f, ξ). By standard calculations similar to those of Section 1.2, it follows that $\Gamma_t Y_t \,+ \int_0^t \Gamma_s\varphi_s\,ds$ is a local martingale. Now $\sup_{s\leq T}|Y_s|$ and $\sup_{s\leq T}|\Gamma_s|$ belong to $\mathbb{L}_T^{2,1}$ and $\sup_{s\leq T}|Y_s| \times \sup_{s\leq T}|\Gamma_s|$ belongs to $\mathbb{L}_T^{1,1}$. Therefore the local martingale $\Gamma_t Y_t \,+ \int_0^t \Gamma_s\varphi_s\,ds$ is uniformly integrable and equal to the conditional expectation of its terminal value. In particular, if ξ and φ are nonnegative, Y_t is also nonnegative. If, in addition, $Y_0 = 0$, then the expectation of the nonnegative variable $\xi\Gamma_T + \int_0^T \Gamma_s\varphi_s ds$ is equal to 0. So $\xi = 0$, $\mathbb{P}$ a.s., $\varphi_t = 0$, $d\mathbb{P} \otimes dt$ a.s., and $Y = 0$ a.s. $\square$

REMARK. Recall that in Section 1.2 we constructed various solutions for the LBSDE which are not square integrable. Nevertheless, all solutions bounded below still satisfy the positivity property. More precisely, let (X, Π) be a solution of (2.8) (not necessarily square integrable) with $X_t \geq -B$ $(B \geq 0)$ for any time t, where B is a square-integrable $\mathcal{F}_T$-measurable variable, and suppose that ξ and φ are nonnegative. Since $(M_t \,=\, \Gamma_t X_t \,+ \int_0^t \Gamma_s\varphi_s\,ds;\, t \,\in\, [0, T])$ is a local martingale, bounded below by the integrable variable $-\sup_{t\leq T}(\Gamma_t)B$, Fatou's lemma implies that M is a supermartingale which is minorized by $\mathbb{E}[M_T|\mathcal{F}_t]$. It follows that $X_t \geq 0$. (Notice that the square integrability of φ is not needed for this property.) Furthermore, the square-integrable solution of (2.8) (Y, Z) is the smallest of the solutions (X, Π) which are bounded by below by a square-integrable variable. Indeed, the difference $X - Y$ is a bounded below by a square-integrable variable solution of the LBSDE with terminal condition 0 and $\varphi = 0$ a.s., so $X - Y$ is nonnegative. This property is a mild extension of Theorem 1.1, which will be generalized to the case of nonlinear BSDEs in Section 2.3.

2.2. Comparison Theorem

As an immediate consequence of Proposition 2.2, we provide (in the one-dimensional case) a comparison theorem first obtained by Peng (1992a). Recall that such a property can be obtained for forward SDEs only under strong assumptions on the coefficients; in particular, the functions which appear in the coefficients of diffusion must be the same for the two equations (Karatzas and Shreve 1987).

THEOREM 2.2 (Comparison Theorem). *Let* (f^1, ξ^1) *and* (f^2, ξ^2) *be two standard parameters of BSDEs, and let* (Y^1, Z^1) *and* (Y^2, Z^2) *be the associated square-integrable solutions. We suppose that*

- $\xi^1 \geq \xi^2$ $\mathbb{P}$ a.s.
- $\delta_2 f_t = f^1(t, Y_t^2, Z_t^2) - f^2(t, Y_t^2, Z_t^2) \geq 0$, $d\mathbb{P} \otimes dt$ *a.s.*

Then we have that almost surely for any time t, $Y_t^1 \geq Y_t^2$.

Moreover the comparison is strict; that is, if, in addition, $Y_0^1 = Y_0^2$, *then* $\xi^1 = \xi^2$, $f^1(t, Y_t^2, Z_t^2) = f^2(t, Y_t^2, Z_t^2)$, $d\mathbb{P}\otimes dt$ *a.s., and* $Y^1 = Y^2$ *a.s. More generally if* $Y_t^1 = Y_t^2$ *on a set* $A \in \mathcal{F}_t$, *then* $Y_s^1 = Y_s^2$ *almost surely on* $[t, T] \times A$, $\xi_1 = \xi^2$ *a.s. on* A, *and* $f^1(s, Y_s^2, Z_s^2) = f^2(s, Y_s^2, Z_s^2)$ *on* $A \times [t, T]$ $d\mathbb{P} \otimes ds$ *a.s.*

Before proving this theorem, we deduce a sufficient condition for the nonnegativity of the BSDE solution.

COROLLARY 2.2. *If* $\xi \geq 0$ *a.s. and* $f(t, 0, 0) \geq 0$ $d\mathbb{P} \otimes dt$ *a.s., then* $Y \geq 0$ $\mathbb{P}$ *a.s. In addition, if* $Y_t = 0$ *on a set* $A \in \mathcal{F}_t$, *then* $Y_s = 0$, $f(s, 0, 0) = 0$ *on* $[t, T] \times A$, $d\mathbb{P} \otimes ds$ *a.s., and* $\xi = 0$ *almost surely on* A.

Proof of Theorem 2.5. We use the notation of Proposition 2.1. The pair $(\delta Y, \delta Z)$ is the solution of the following LBSDE:

$$
\begin{aligned}
-d\delta Y_t &= \Delta_y f^1(t)\delta Y_t + \Delta_z f^1(t)^* \delta Z_t + \delta_2 f_t\, dt - \delta Z_t^*\, dW_t, \\
\delta Y_T &= \xi^1 - \xi^2,
\end{aligned}
\tag{2.11}
$$

where $\Delta_y f^1(t) = (f^1(t, Y_t^1, Z_t^1) - f^1(t, Y_t^2, Z_t^1))/(Y_t^1 - Y_t^2)$ if $Y_t^1 - Y_t^2$ is not equal to 0, whereas $\Delta_y f^1(t) = 0$, otherwise. Also, $\Delta_z f^{1,i}(t) = (f^1(t, Y_t^2, \tilde{Z}_t^{i-1}) - f^1(t, Y_t^2, \tilde{Z}^i))/(Z_t^{1,i} - Z_t^{2,i})$ if $Z_t^{1,i} - Z_t^{2,i}$ is not equal to 0, whereas $\Delta_z f^{1,i}(t) = 0$, otherwise. Here $\tilde{Z}^i$ is the vector whose first i components are equal to those of Z^2 and whose $n-i$ others are equal to those of Z^1; that is, $\tilde{Z}_t^i = (Z_t^{2,1}, \ldots, Z_t^{2,i}, Z_t^{1,i+1}, \ldots, Z_t^{1,n})$.

Now since by assumption the generator f^1 is uniformly Lipschitz with respect to (y, z), it follows that $\Delta_y f^1$ and $\Delta_z f^{1,i}$ are bounded processes. Also, $\delta_2 f_t$ and δY_T are nonnegative. It follows from Proposition 2.2 that the unique square-integrable solution $(\delta Y, \delta Z)$ of the LBSDE (2.11) is nonnegative and satisfies

$$
\Gamma_t \delta Y_t = \mathbb{E}[(\xi^1 - \xi^2)\Gamma_T + \int_t^T \Gamma_s \delta_2 f_s\, ds | \mathcal{F}_t], \tag{2.12}
$$

where Γ is the adjoint (positive) process of the above LBSDE. Also, if $Y_t^1 = Y_t^2$ on a set $A \in \mathcal{F}_t$, then $\xi^1 = \xi^2$, $\delta_2 f_s = 0$, $d\mathbb{P} \otimes ds$ on $A \times [t, T]$, and $Y_s^1 = Y_s^2$ a.s. on $A \times [t, T]$. Thus we obtain the last point of Theorem 2.2. □

REMARK. Relax the assumptions of square integrability for the solutions of the BSDE in the comparison theorem and suppose only that there exists a square-integrable variable $B \geq 0$ such that $Y_t^1 - Y_t^2 \geq -B, t \in [0, T]$. From the remark which follows Proposition 2.2, it is easy to prove that the inequality $Y_.^1 \geq Y_.^2$ still holds and that the other properties of the comparison theorem hold, too.

2.3. Supersolution

Earlier (Definition 1.2) we introduced the notion of a superhedging strategy, which can be considered as a supersolution of a one-dimensional LBSDE, defined as follows.

DEFINITION 2.1. Suppose that $d = 1$. A *supersolution* of a BSDE associated with standard parameters (f, ξ) is a vector process (Y, Z, C) satisfying

$$-dY_t = f(t, Y_t, Z_t)\,dt - Z_t^* \,dW_t + dC_t\,, \qquad Y_T = \xi, \tag{2.13}$$

or, equivalently,

$$Y_t = \xi + \int_t^T f(s, Y_s, Z_s)\,ds - \int_t^T Z_s^* \,dW_s + \int_t^T dC_s\,, \tag{2.14}$$

where

- ξ is an $\mathbb{R}$-valued, $\mathcal{F}_T$-measurable random variable.
- $(Y_t, t \in [0, T])$ is a right-continuous, left-limited adapted real process. When Y is continuous, the solution is said to be continuous.
- Z is a predictable process which takes values in $\mathbb{R}^n$ with $\int_0^T |Z_s|^2\,ds < +\infty$ $\mathbb{P}$ a.s.
- $(C_t; t \in [0, T])$ is an increasing, adapted, right-continuous process such that $C_0 = 0$.
- Y is bounded below; that is, there exists a square-integrable $\mathcal{F}_T$-measurable variable $B > 0$ such that $Y_t \geq -B$, $t \in [0, T]$, $\mathbb{P}$ a.s.

REMARK a. Suppose that f is a linear generator with bounded coefficients β, γ and associated with $\varphi \in \mathbb{H}_T^2(\mathbb{R})$. The adjoint process is denoted by Γ. Let (Y, Z, C) be a supersolution associated with (f, ξ). Then $\Gamma_t Y_t + \int_0^t \Gamma_s \varphi_s\,ds$ is a local supermartingale. Notice that this property corresponds in finance to the fact that the discounted wealth associated with a superhedging strategy is a risk-neutral supermartingale (in this case $\varphi = 0$).

REMARK b. By the extension of the comparison theorem (Theorem 2.2), it is clear that if (f, ξ) are standard parameters the continuous supersolutions dominate the classical square-integrable solution of the BSDE. This property applied to European option pricing in the constrained case (Sections 1.3 and 3.3) shows that the upper price corresponds to the square-integrable strategy (that is, the first statement of Theorem 1.1 still holds).

REMARK c. Let (Y^1, Z^1) and (Y^2, Z^2) be two square-integrable solutions of BSDEs with standard parameters satisfying the assumptions of the comparison theorem 2.2. Then there exists an increasing process C^2 such that (Y^2, Z^2, C^2) is a supersolution for the BSDE with parameters (f^1, ξ^1).

An important property is that the infimum of two continuous supersolutions is still a supersolution. More precisely,

PROPOSITION 2.3. *Let (Y^1, Z^1, C^1) and (Y^2, Z^2, C^2) be two continuous supersolutions of the BSDEs with parameters (f^1, ξ^1) and (f^2, ξ^2). Then there exists (Z^*, C^*) such that $(Y^* = Y^1 \wedge Y^2, Z^*, C^*)$ is a supersolution of the backward equation with terminal condition $\xi^* = \xi^1 \wedge \xi^2$ and generator $f^*(t, y, z) = \mathbf{1}_{Y_t^1 \leq Y_t^2} f^1(t, y, z) + \mathbf{1}_{Y_t^2 < Y_t^1} f^2(t, y, z)$. In particular, if $f^1 = f^2$, then $(Y^* = Y^1 \wedge Y^2, Z^*, C^*)$ is a supersolution of the BSDE with parameters (f^*, ξ^*).*

Proof. Recall the Tanaka formula (Karatzas and Shreve 1987) for the minimum of two continuous semimartingales Y^1 and Y^2,

$$dY_t^1 \wedge Y_t^2 = 1_{Y_t^1 \leq Y_t^2} dY_t^1 + 1_{Y_t^2 < Y_t^1} dY_t^2 - dL_t ,$$

where L is a local time—that is, a continuous increasing process with support included in $\{t \in [0, T], Y_t^1 = Y_t^2\}$. Then $Y^* = Y^1 \wedge Y^2$ satisfies

$$\begin{aligned} -dY_t^* \;=\; & 1_{Y_t^1 \leq Y_t^2}[f^1(t, Y_t^1, Z_t^1)\, dt \;-\; Z_t^1\, dW_t \;+ dC_t^1] \\ & + 1_{Y_t^2 < Y_t^1}[f^2(t, Y_t^2, Z_t^2)\, dt \;-\; Z_t^2\, dW_t \;+ dC_t^2] + dL_t . \end{aligned}$$

Put $Z_t^* = 1_{Y_t^1 \leq Y_t^2} Z_t^1 + 1_{Y_t^2 < Y_t^1} Z_t^2$ and $dC_t^* = 1_{Y_t^1 \leq Y_t^2} dC_t^1 + 1_{Y_t^2 < Y_t^1} dC_t^2 + dL_t$. Then $(Y^* = Y^1 \wedge Y^2, Z^*, C^*)$ is a supersolution with parameters (f^*, ξ^*) since $Y^1 \wedge Y^2$ is bounded below. □

COROLLARY 2.3. *Let (Y^1, Z^1, C^1) be a continuous positive supersolution of a BSDE with parameters (f^1, ξ^1). Then the increasing process $1_{\{Y_t^1 = 0\}} dC_t^1$ is absolutely continuous with respect to the positive measure $f^1(t, 0, 0)^- dt$.*

Proof. The above calculation applied with $f^2 = 0$, $\xi^2 = 0$, $C^2 = 0$, $Y^2 = 0$, and $Z^2 = 0$ yields to

$$0 = 1_{Y_t^1 = 0}[f^1(t, Y_t^1, Z_t^1)\, dt \;-\; Z_t^1\, dW_t \;+ dC_t^1] + dL_t .$$

Hence $1_{Y_t^1 = 0} Z_t^1\, dW_t = 0$ and $1_{Y_t^1 = 0}[f^1(t, 0, 0)\, dt \;+ dC_t^1] + dL_t = 0$. It follows that on

$\{Y_t^1 = 0\}$, $f^1(t, 0, 0)$ is negative and

$$dL_t + 1_{Y_t^1=0} dC_t^1 = 1_{Y_t^1=0} f^1(t, 0, 0)^- \, dt,$$

and so the result follows. □

2.4. Flow and BSDE Dependence upon Parameters

In this section we study the properties of continuity and differentiability of the solutions of BSDEs depending on parameters and the flow properties of a BSDE. These results follow essentially from the a priori estimates.

Continuity and Differentiability. Let $(f(\alpha, \cdot), \xi(\alpha, \cdot), \alpha \in \mathbb{R})$ be a family of standard parameters of a BSDE whose solutions are denoted by (Y^α, Z^α). For notational convenience, we often write (Y^0, Z^0) for $(Y^{\alpha_0}, Z^{\alpha_0})$. Let us make the following hypotheses:

1. The family $f(\alpha, \cdot), \alpha \in \mathbb{R}$, is equi-Lipschitz; i.e., there exists $C > 0$ such that, $d\mathbb{P} \otimes dt$ a.s.,

$$\forall \alpha \in \mathbb{R}, \quad |f(\alpha, \omega, t, y_1, z_1) - f(\alpha, \omega, t, y_2, z_2)| \le C(|y_1 - y_2| + |z_1 - z_2|).$$

2. The function $\alpha \mapsto (f(\alpha, \cdot), \xi(\alpha))$ is "continuous"; i.e., for each α_0, $f(\alpha, t, Y_t^0, Z_t^0) - f(\alpha^0, t, Y_t^0, Z_t^0)$ converges to 0 in $\mathbb{H}^2_{T,\beta}(\mathbb{R}^d)$ and $\xi(\alpha) - \xi(\alpha_0)$ converges to 0 in $\mathbb{L}^2_T(\mathbb{R}^d)$ as $\alpha \to \alpha_0$.
3. $f(\cdot, \omega, t, y, z)$ and $\xi(\cdot, \omega)$ are equi-Lipschitz with respect to α.
4. $\forall \alpha \in \mathbb{R}$, $f(\alpha, \cdot)$ is differentiable with respect to (y, z) with uniformly bounded derivatives denoted by $\partial_y f(\alpha, y, z)$ and $\partial_z f(\alpha, y, z)$ which are uniformly continuous; that is, $\forall \varepsilon > 0 \, \exists \eta$ such that $d\mathbb{P} \otimes dt$ a.s.,

$$|h| < \eta \Rightarrow \forall (\alpha, y, z), |\partial_y f(\alpha, \omega, t, y + h, z) - \partial_y f(\alpha, \omega, t, y, z)| < \varepsilon$$

   (and the same holds for $\partial_z f(\alpha, y, z)$). Such assumptions hold if, for example, f is twice differentiable with bounded second derivatives .
5. The function $\alpha \mapsto (f(\alpha, \cdot), \xi(\alpha, \cdot))$ is differentiable; i.e., for each α_0 the functions $\alpha \mapsto f(\alpha, \cdot, Y^0_\cdot, Z^0_\cdot), \mathbb{R} \to \mathbb{H}^2_{T,\beta}(\mathbb{R}^d)$ and $\alpha \mapsto \xi(\alpha, \cdot), \mathbb{R} \to \mathbb{L}^2_T(\mathbb{R}^d)$ are differentiable at α_0 with derivative $\partial_\alpha f(\alpha_0, \cdot, Y^0_\cdot, Z^0_\cdot)$.

PROPOSITION 2.4. *Let $(f(\alpha, \cdot), \xi(\alpha, \cdot), \alpha \in \mathbb{R})$ be a family of standard parameters of a BSDE with solutions denoted by (Y^α, Z^α).*

1. *Suppose these parameters satisfy hypotheses 1 and 2. Then the function $\alpha \mapsto (Y^\alpha, Z^\alpha), \mathbb{R} \to \mathbb{H}^2_{T,\beta}(\mathbb{R}^d) \times \mathbb{H}^2_{T,\beta}(\mathbb{R}^{n\times d})$, is continuous. Moreover if hypothesis 3 holds,*[6] *there exists a bicontinuous version of $(\alpha, t) \mapsto (Y_t^\alpha)$.*

[6] It is sufficient to suppose that for any $\alpha, \gamma \in \mathbb{R}$ the spread between the corresponding solutions (Y^α), (Y^γ) satisfies the inequality

$$\mathbb{E}[\sup_{t \le T} |Y_t^\alpha - Y_t^\gamma|^2] \le M(1 + |\alpha|^2)|\alpha - \gamma|^2$$

for a constant $M > 0$.

2. *Suppose these parameters satisfy hypotheses 4 and 5. Then the function* $\alpha \mapsto (Y^\alpha, Z^\alpha)$; $\mathbb{R} \to \mathbb{H}^2_{T,\beta}(\mathbb{R}^d) \times \mathbb{H}^2_{T,\beta}(\mathbb{R}^{n\times d})$, *is differentiable with derivatives given by* $(\partial_\alpha Y^\alpha, \partial_\alpha Z^\alpha)$, *the solution of the following BSDE:*

$$
\begin{aligned}
(2.15)\quad -d(\partial_\alpha Y^\alpha_t) &= [\partial_y f(\alpha, t, Y^\alpha_t, Z^\alpha_t)\partial_\alpha Y^\alpha_t + \langle \partial_z f(\alpha, t, Y^\alpha_t, Z^\alpha_t), \partial_\alpha Z^\alpha_t\rangle]\, dt \\
&\quad + \partial_\alpha f(\alpha, t, Y^\alpha_t, Z^\alpha_t)\, dt - \left(\partial_\alpha Z^\alpha_t\right)^*\, dW_t, \\
\partial_\alpha Y^\alpha_T &= \partial_\alpha \xi^\alpha ,
\end{aligned}
$$

where[7] $\langle \partial_z f, \partial_\alpha Z^\alpha\rangle = (\langle \partial_z f^i, \partial_\alpha Z^\alpha\rangle)_{1\le i\le d}$.

Proof.[8] Property 1 is an immediate consequence of the a priori estimates. Let us prove the second one. By hypothesis 3 and the a priori estimates (Proposition 2.1), it follows that for a constant $M > 0$

$$\mathbb{E}[\sup_{t\le T} |Y^\alpha_t - Y^\gamma_t|^2] \le M|\alpha - \gamma|^2 .$$

The existence of a bicontinuous version follows from Kolmogorov's criteria (Karatzas and Shreve 1987, p. 53; Revuz and Yor 1991, Chapter VI, Proposition 1.3).

Let us show that if hypotheses 4 and 5 hold, then for each $\alpha_0 \in \mathbb{R}$ the function $\alpha \mapsto (Y^\alpha, Z^\alpha)$, $\mathbb{R} \mapsto \mathbb{H}^{2,d}_{T,\beta} \times \mathbb{H}^{2,n\times d}_{T,\beta}$ is differentiable at α_0. For notational convenience, we can assume that $\alpha_0 = 0$ and that the dimensions n and d are equal to one. Put $\Delta_\alpha Y_t = \alpha^{-1}(Y^\alpha_t - Y^0_t)$ and $\Delta_\alpha Z_t = \alpha^{-1}(Z^\alpha_t - Z^0_t)$. Then

$$
\begin{aligned}
-d\Delta_\alpha Y_t &= \alpha^{-1}[f(\alpha, t, Y^\alpha_t, Z^\alpha_t) - f(0, t, Y^0_t, Z^0_t)]\, dt - \Delta_\alpha Z^*_t\, dW_t, \\
\Delta_\alpha Y_T &= \alpha^{-1}[\xi(\alpha) - \xi(0)] .
\end{aligned}
$$

Hence, as in the proof of Theorem 2.5, we treat this equation as a linear one:

$$-d\Delta_\alpha Y_t = \psi(\alpha, t, \Delta_\alpha Y^\alpha_t, \Delta_\alpha Z^\alpha_t)\, dt - (\Delta_\alpha Z^\alpha)^*_t\, dW_t ,$$

where ψ is defined by $\psi(\alpha, t, y, z) = A_\alpha(t)\, y + B_\alpha(t)\, z + \varphi_\alpha(t)$ and where, for $\alpha \ne 0$,

$$
(2.16)\qquad A^\alpha(t) = \begin{cases} \dfrac{f(\alpha, t, Y^\alpha_t, Z^\alpha_t) - f(\alpha, t, Y^0_t, Z^\alpha_t)}{Y^\alpha_t - Y^0_t} & \text{if } Y^\alpha_t \ne Y^0_t, \\ \partial_y f(\alpha, t, Y^0_t, Z^\alpha_t) & \text{otherwise,} \end{cases}
$$

$$
(2.17)\qquad B^\alpha(t) = \begin{cases} \dfrac{f(\alpha, t, Y^0_t, Z^\alpha_t) - f(\alpha, t, Y^0_t, Z^0_t)}{Z^\alpha_t - Z^0_t} & \text{if } Z^\alpha_t \ne Z^0_t, \\ \partial_z f(\alpha, t, Y^0_t, Z^0_t) & \text{otherwise,} \end{cases}
$$

[7]We use the notation $\langle \partial_z f^i, \partial_\alpha Z^\alpha\rangle = \sum_{1\le k\le n, 1\le l\le d} \partial_{z_{k,l}} f^i \partial_\alpha Z^\alpha_{k,l}$.
[8]We thank Martin Schweizer for his remark concerning this proof.

and

$$\varphi_\alpha(t) = \frac{1}{\alpha}\left(f(\alpha, t, Y_t^0, Z_t^0) - f(0, t, Y_t^0, Z_t^0)\right).$$

Put $\psi(0, t, y, z) = \partial_y f(0, t, Y_t^0, Z_t^0)\, y + \partial_z f(0, t, Y_t^0, Z_t^0)\, z + \partial_\alpha f(0, t, Y_t^0, Z_t^0)$.

By property 1 of this proposition, (Y^α, Z^α) converges to (Y^0, Z^0) in $\mathbb{H}^{2,1}_{T,\beta} \otimes \mathbb{H}^{2,1}_{T,\beta}$. We now have to prove that $(\Delta_\alpha Y, \Delta_\alpha Z)$ converges to $(\partial_\alpha Y^0, \partial_\alpha Z^0)$, the solution of the above BSDE, as α goes to 0. To use the same convergence argument, we must show that $\psi(\alpha, \cdot, \partial_\alpha Y^0, \partial_\alpha Z^0)$ converges to $\psi(0, \cdot, \partial_\alpha Y^0, \partial_\alpha Z^0)$ in $\mathbb{H}^{2,1}_{T,\beta} \otimes \mathbb{H}^{2,1}_{T,\beta}$ as α goes to 0. Notice that $A^\alpha(t) = \int_0^1 \partial_y f(\alpha, t, Y_t^0 + \lambda(Y_t^\alpha - Y_t^0), Z_t^\alpha)\, d\lambda$. Consequently,

$$\begin{aligned}
&\mathbb{E}\int_0^T (A^\alpha(t) - \partial_y f(\alpha, t, Y_t^0, Z_t^\alpha))^2 (\partial_\alpha Y_t^0)^2\, dt \\
&\qquad \le \mathbb{E}\int_0^T \int_0^1 (\partial_y f(\alpha, t, Y_t^0 + \lambda(Y_t^\alpha - Y_t^0), Z_t^\alpha) \\
&\qquad\qquad\qquad - \partial_y f(\alpha, t, Y_t^0, Z_t^\alpha))^2 (\partial_\alpha Y_t^0)^2\, d\lambda\, dt.
\end{aligned}$$

Splitting this integral into two terms on the sets $\{|Y_t^\alpha - Y_t^0| \le \eta\}$ and $\{|Y_t^\alpha - Y_t^0| > \eta\}$ and using that, by hypothesis 4, $\partial_y f(\alpha, t, y, z)$ is uniformly continuous and bounded (by a constant K), it follows that for each $\epsilon > 0$ there exists $\eta > 0$ such that

$$||(A_\alpha(t) - \partial_y f(\alpha, \cdot, Y^0, Z^\alpha))\partial_\alpha Y^0||_2^2 \le \epsilon^2 ||\partial_\alpha Y^0||_2^2 + K^2 \mathbb{E}\int_0^T \mathbf{1}_{\{|Y_t^\alpha - Y_t^0| > \eta\}} |\partial_\alpha Y_t^0|^2\, dt.$$

Now split the last term into two parts corresponding to the set $\{|\partial_\alpha Y_t^0| \le M\}$ and its complement. Then by applying the Markov inequality to $Y_t^\alpha - Y_t^0$, we have

$$\mathbb{E}\int_0^T \mathbf{1}_{\{|Y_t^\alpha - Y_t^0| > \eta\}} |\partial_\alpha Y_t^0|^2\, dt \le \frac{M^2}{\eta^2} ||Y^\alpha - Y^0||_2^2 + \mathbb{E}\int_0^T \mathbf{1}_{\{|\partial_\alpha Y^0| > M\}} |\partial_\alpha Y_t^0|^2\, dt.$$

By the Lebesgue theorem, since $\partial_\alpha Y^0$ is square integrable, $\mathbb{E}\int_0^T \mathbf{1}_{\{|\partial_\alpha Y^0| > M\}} |\partial_\alpha Y_t^0|^2\, dt$ converges to 0 as $M \to \infty$. Choosing M sufficiently large and using the convergence of Y^α to Y^0 in $\mathbb{H}^{2,1}_T$, it follows easily that

$$\lim_{\alpha \to 0} ||(A_\alpha(t) - \partial_y f(\alpha, \cdot, Y^0, Z^\alpha))\partial_\alpha Y^0||_2 = 0.$$

By the same method we easily see that $\lim_{\alpha \to 0} ||(\partial_y f(\alpha, \cdot, Y^0, Z^\alpha) - \partial_y f(0, \cdot, Y^0, Z^0))\, \partial_\alpha Y^0||_2 = 0$. Hence, it follows that $\lim_{\alpha \to 0} ||(A_\alpha(t) - \partial_y f(0, \cdot, Y^0, Z^0))\partial_\alpha Y^0||_2 = 0$. Similar arguments give that $\lim_{\alpha \to 0} ||(B_\alpha(t) - \partial_z f(0, t, Y^0, Z^0))\partial_\alpha Z^0||_2 = 0$. Consequently, using hypothesis 5, as α goes to 0, $\psi(\alpha, t, \partial_\alpha Y^0, \partial_\alpha Z^0)$ converges to $\psi(0, \cdot, \partial_\alpha Y^0, \partial_\alpha Z^0)$ in $\mathbb{H}^{2,1}_{T,\beta} \times \mathbb{H}^{2,1}_{T,\beta}$. By the first part of the proposition, the solution $(\Delta_\alpha Y, \Delta_\alpha Z)$ converges to $\left(\partial_\alpha Y^0, \partial_\alpha Z^0\right)$ in $\mathbb{H}^{2,1}_{T,\beta} \times \mathbb{H}^{2,1}_{T,\beta}$. $\square$

REMARK. Notice that, as for the a priori estimates, one can take the norm $\mathbb{E}[\sup_{t\leq T}|Y_t|^2]$ instead of $||Y||_\beta$ for Y. Consequently, if the parameters are differentiable, then the function $\alpha \mapsto Y^\alpha$ is differentiable for this norm.

Flow of a BSDE. Recall the dependence of the solutions of BSDE with respect to terminal condition by the notation $(Y(T,\xi), Z(T,\xi))$. We provide a flow property and some regularity results similar to the case of forward SDEs.

PROPOSITION 2.5. *Let (Y, Z) be the solution of a BSDE with standard parameters $(T,\ f,\xi)$.*

- *For any stopping time $S \leq T$,*

$$Y_t(T,\xi) = Y_t(S, Y_S(T,\xi)), \quad Z_t(T,\xi) = Z_t(S, Y_S(T,\xi)), \quad t \in [0,S]\, d\mathbb{P}\otimes dt\; a.s.$$

- *Suppose that the sequence of stopping times S_n converges a.s. to S and that the sequence of the terminal variables $\xi_n \in \mathcal{F}_{S_n}$ converges in $\mathbb{L}^2_T(\mathbb{R}^d)$ to $\xi \in \mathcal{F}_S$. Then the pair of processes $(Y(S_n,\xi_n), Z(S_n,\xi_n))$ converges in $\mathbb{H}^2_{T,\beta}(\mathbb{R}^d) \times \mathbb{H}^2_{T,\beta}(\mathbb{R}^{n\times d})$ to $(Y(S,\xi), Z(S,\xi))$.*

Proof. By conventional notation we define the solution of the BSDE with terminal condition (T,ξ) for $t \geq T$ by $(Y_t = \xi, Z_t = 0)$. So if $T' \geq T$, then $(Y_t, Z_t\ ; t \leq T')$ is the unique solution of the BSDE with standard parameters $(T',\ f(t,y,z)\mathbf{1}_{\{t\leq T\}}, \xi)$.

Now let $S \leq T$ be a stopping time, and denote by $Y_t(S,\xi_S)$ the solution of the BSDE with standard parameters $(T,\ f(t,y,z)\mathbf{1}_{\{t\leq S\}}, \xi_S)$. The processes $(Y_t(S,Y_S), Z_t(S,Y_S); t \in [0,T])$ and $(Y_{t\wedge S}(T,\xi),\ Z_t(T,\xi)1_{\{t\leq S\}}; t \in [0,T])$ are solutions of the BSDE with parameters $(T,\ f(t,y,z)1_{\{t\leq S\}}, Y_S)$. By uniqueness these processes are the same $d\mathbb{P}\otimes dt$ a.s. The convergence property results immediately from Proposition 2.4, since the parameters $(T,\ f(t,y,z)1_{\{t\leq S_n\}}, \xi_n)$ satisfy hypotheses 1 and 2 of that proposition. □

3. CONCAVE BSDES AND CONTROL PROBLEMS

In this section we are concerned with the solution of a BSDE with respect to standard parameters that are infima of standard parameters. This property can be translated to the solutions under some mild conditions. In other words,

$$Y_t(\inf f^\alpha, \inf \xi^\alpha) = \operatorname{ess\,inf} Y_t(f^\alpha, \xi^\alpha).$$

This property can be applied to some classical control problems. Under some mild conditions, the value function can be characterized as the solution of a (concave) BSDE. From this point of view, classical properties of the value function can be derived. Then we show that, conversely, the solution (Y, Z) of a BSDE (with $d = 1$) with concave generator can be considered as the value function of a control problem.

3.1. BSDE and Optimization

Solution of BSDE as Minimum or Minimax. In this section we are concerned with standard generators f (respectively terminal conditions ξ) which can be obtained as infima of standard generators f^α (respectively of terminal conditions ξ^α). From the comparison theorem the solution of the BSDE associated with (f, ξ) is less than the infimum of the solutions associated with (f^α, ξ^α). The problem is to know when the equality holds.

PROPOSITION 3.1. *Let (f, ξ) and (f^α, ξ^α) be a family of standard parameters, and let (Y, Z) and (Y^α, Z^α) be the solution of associated BSDEs. Suppose that there exists a parameter $\overline{\alpha}$ such that*

$$\begin{aligned} f(t, Y_t, Z_t) &= \operatorname*{ess\,inf}_\alpha f^\alpha(t, Y_t, Z_t) = f^{\overline{\alpha}}(t, Y_t, Z_t)\ d\mathbb{P} \otimes dt\ a.s., \\ \xi &= \operatorname*{ess\,inf}_\alpha \xi^\alpha = \xi^{\overline{\alpha}}\ \mathbb{P}\,a.s. \end{aligned} \tag{3.1}$$

Then the[9] *processes Y and Y^α satisfy*

$$Y_t = \operatorname*{ess\,inf}_\alpha Y_t^\alpha = Y_t^{\overline{\alpha}}, \quad \forall t \in [0, T],\ \mathbb{P}\,a.s. \tag{3.2}$$

Proof. (Y, Z) and (Y^α, Z^α) are solutions of two BDSEs whose generators and terminal conditions satisfy the assumptions of the comparison theorem (2.5). Hence, for any α, $Y_t \le Y_t^\alpha$ and, consequently, $Y_t \le \operatorname{ess\,inf} Y_t^\alpha$ for any time t $\mathbb{P}$ a.s.

We now prove the equality using the uniqueness theorem for BSDEs and the existence of a parameter $\overline{\alpha}$ such that $f(t, Y_t, Z_t) = f^{\overline{\alpha}}(t, Y_t, Z_t)$ and $\xi = \xi^{\overline{\alpha}}$ a.s. Hence, (Y, Z) and $(Y^{\overline{\alpha}}, Z^{\overline{\alpha}})$ are both solutions of the same BSDE with parameters $(f^{\overline{\alpha}}, \xi^{\overline{\alpha}})$; therefore, they are the same. So,

$$\operatorname{ess\,inf} Y_t^\alpha \ge Y_t = Y_t^{\overline{\alpha}} \ge \operatorname{ess\,inf} Y_t^\alpha \quad \forall t \in [0, T],\ \mathbb{P}\text{ a.s.} \qquad \square$$

COROLLARY 3.1. *The same result holds if the generators only satisfy the following:*

- *The generators f^α are equi-Lipschitz with the same constant C.*
- *For each $\epsilon > 0$, there exists a control α^ϵ such that*

$$\begin{aligned} f(t, Y_t, Z_t) &= \operatorname{ess\,inf} f^\alpha(t, Y_t, Z_t) \ge f^{\alpha^\varepsilon}(t, Y_t, Z_t) - \varepsilon, d\mathbb{P} \otimes dt\ a.s., \\ \xi &= \operatorname*{ess\,inf}_\alpha \xi^\alpha \ge \xi^{\alpha^\varepsilon} - \varepsilon, \mathbb{P}\,a.s. \end{aligned} \tag{3.3}$$

Proof. We suppose that the generators satisfy (3.3). Put $\delta Y_t = Y_t - Y_t^{\alpha^\varepsilon}$ and $\delta Z_t = Z_t - Z_t^{\alpha^\varepsilon}$. Using the same arguments as in the proof of the comparison theorem, we derive

[9] Dellacherie (1977) introduced the notion of ess inf of processes in the following manner:

- a process U is said to minorize the process U^α if $\{\omega; \exists t \in [0, T] U_t(\omega) > U_t^\alpha(\omega)\}$ is a $\mathbb{P}$-null set.
- a process U is said to be ess inf U^α if, for any α, U minorizes U^α, and if a process V which minorizes U^α for each α minorizes U. Moreover, for right-continuous left-limited processes U^α, ess inf U^α exists and there exists a denumberable family (α_n) such that $U = \inf U^{\alpha_n}$.

that $(\delta Y, \delta Z)$ is the solution of the following LBSDE :

$$
\begin{aligned}
(3.4) \qquad -d\delta Y_t &= \Delta_y f(t)\delta Y_t + \langle \Delta_z f(t), \delta Z_t\rangle + \delta f_t^{\varepsilon}\, dt - \delta Z_t^*\, dW_t, \\
\delta Y_T &= \xi - \xi^{\alpha^{\varepsilon}},
\end{aligned}
$$

where $\Delta_y f(t)$ and $\Delta_z f(t)$ are predictable processes bounded by the Lipschitz constant C of f and

$$\delta f_t^{\varepsilon} = f(t, Y_t, Z_t) - f^{\alpha^{\varepsilon}}(t, Y_t, Z_t).$$

It follows that

$$(3.5) \qquad \delta Y_t = \mathbb{E}\left[\int_t^T \Gamma_{t,s}\delta f_s^{\varepsilon}\, ds + \Gamma_{t,T}\delta Y_T | \mathcal{F}_t\right],$$

where $\Gamma_{t,.}$ is the adjoint process (positive) of the above LBSDE; that is,

$$(3.6) \qquad d\Gamma_s = \Gamma_s[\Delta_y f(s)\, ds + \Delta_z f(s)^*\, dW_s], \qquad \Gamma_t = 1\,.$$

By (3.3) we have

$$\delta Y_t \geq -\varepsilon\mathbb{E}\left[\int_t^T \Gamma_{t,s}\, ds + \Gamma_{t,T} | \mathcal{F}_t\right] \geq -\varepsilon(T+1)e^{CT},$$

where C is the Lipschitz constant for the f^{α}'s and the result follows. □

Similar results were extended to minimax problems in Hamadene and Lepeltier (1994) in connection with stochastic differential games. These techniques are also useful for solving optimization problems associated with recursive utilities (Quenez 1993; El Karoui, Peng, and Quenez 1994).

COROLLARY 3.2. *Let (f, ξ) and $(f^{\alpha,\beta}, \xi^{\alpha,\beta})$ be a family of standard parameters, and let (Y, Z) and $(Y^{\alpha,\beta}, Z^{\alpha,\beta})$ be the associated solutions. Suppose that f and $f^{\alpha,\beta}$ (resp. ξ and $\xi^{\alpha,\beta}$) are linked by a minimax relation and that there exists a pair of parameters $(\overline{\alpha}, \overline{\beta})$ such that the following formulation of the Isaac condition holds:*

$$
\begin{aligned}
(3.7) \qquad f(t, Y_t, Z_t) &= \operatorname{ess\,inf}_{\alpha} \sup_{\beta} f^{\alpha,\beta}(t, Y_t, Z_t) = f^{\overline{\alpha},\overline{\beta}}(t, Y_t, Z_t), d\mathbb{P}\otimes dt\ a.s., \\
\xi &= \operatorname{ess\,inf}_{\alpha} \sup_{\beta} \xi^{\alpha,\beta} = \xi^{\overline{\alpha},\overline{\beta}}, \mathbb{P}\, a.s.
\end{aligned}
$$

Then the solutions Y_t and $Y_t^{\alpha,\beta}$ are also linked by a minimax relation with saddle point $(\overline{\alpha}, \overline{\beta})$; that is, the Isaac condition is satisfied:

$$(3.8) \qquad Y_t = \operatorname{ess\,inf}_{\alpha} \sup_{\beta} Y_t^{\alpha,\beta} = Y_t^{\overline{\alpha},\overline{\beta}} = \operatorname{ess\,sup}_{\beta} \inf_{\alpha} Y_t^{\alpha,\beta} \quad \forall t \in [0, T],\ \mathbb{P}\, a.s.$$

Proof. Use the fact that $(\overline{\alpha}, \overline{\beta})$ is a saddle point; i.e.,

$$\operatorname{ess\,sup}_{\beta} f^{\overline{\alpha},\beta}(Y_t, Z_t) \geq f(t, Y_t, Z_t) = f^{\overline{\alpha},\overline{\beta}}(t, Y_t, Z_t) \geq \operatorname{ess\,inf}_{\alpha} f^{\alpha,\overline{\beta}}(t, Y_t, Z_t)\,.$$

The same inequalities hold for the terminal conditions. As a consequence of the previous proposition, the same inequalities hold for the solutions:

$$\operatorname{ess\,sup}_{\beta} Y_t^{\overline{\alpha},\beta} \geq Y_t = Y_t^{\overline{\alpha},\overline{\beta}} \geq \operatorname{ess\,inf}_{\alpha} Y_t^{\alpha,\overline{\beta}} \quad \forall t \in [0, T],\ \mathbb{P} \text{ a.s.}$$

These inequalities imply that the Isaac condition is satisfied for these processes. □

Stochastic Control Problems. A number of stochastic control problems (Krylov 1980; El Karoui 1981; Elliott 1982; Davis 1973; El Karoui and Jeanblanc-Picqué 1988) are specified in the following manner: the laws of the controlled process belong to a family of equivalent measures whose densities are

$$dH_t^u = H_t^u[d(t, u_t)\,dt + n(t, u_t)^*\,dW_t], \qquad H_0^u = 1\,, \tag{3.9}$$

where $d(t, u)$ and $n(t, u)$ are predictable processes uniformly bounded by δ_t and ν_t respectively. A feasible control $(u_t, t \in [0, T])$ is a predictable process valued in a (Polish) space U. The set of feasible controls is denoted by $\mathcal{U}$. The problem is to minimize over all feasible control processes u the objective function

$$J(u) = \mathbb{E}\left[\int_0^T H_t^u\, k(t, u_t)\,dt + H_T^u K(u_T)\right], \tag{3.10}$$

where $K(\cdot, u_T)$ is the terminal condition and $k(\cdot, t, u_t)$ is the running cost associated with the control process u. The processes $(k(\omega, t, u), t \in [0, T]))$ (respectively the terminal conditions $K(\omega, u)$) are assumed to be measurable with respect to $\mathcal{P} \otimes \mathcal{B}(U)$ (respectively $\mathcal{F}_T \otimes \mathcal{B}(U)$), where $\mathcal{B}(U)$ is the Borelian σ-algebra on U; furthermore, they are assumed to be uniformly bounded by a square-integrable process $(k_t; t \in [0, T])$ (respectively by a square-integrable variable χ). We also suppose that δ, ν, k, and χ are bounded.

The controller acts on the law of processes by change of equivalent probability measures with Radon-Nikodym derivatives given by

$$dL_s^u = L_s^u\, n(s, u_s)^*\, dW_s$$

and by a controlled discount factor with bounded rate $d(s, u_s)$; that is,

$$dD_s^u = D_s^u\, d(s, u_s)\, ds, \qquad \text{with } H_t^u = D_t^u L_t^u.$$

Let us denote by $\mathbb{Q}^u$ the probability measure with density L_T^u on $\mathcal{F}_T$. Then the objective function can be written

$$J(u) = \mathbb{E}_{\mathbb{Q}^u}\left[\int_0^T D_t^u\, k(t, u_t)\, dt + D_T^u K(u_T)\right]. \tag{3.11}$$

Notice that, by Proposition 2.2, $J(u) = Y_0^u$, where (Y^u, Z^u) is the solution of the linear BSDE associated with standard parameters (f^u, ξ^u), where

$$f^u(t, y, z) = k(t, u_t) + d(t, u_t) \cdot y + n(t, u_t)^* z, \qquad \xi^u = K(u_T).$$

The process H^u corresponds to the adjoint process associated with (Y^u, Z^u) and

$$Y_t^u = \mathbb{E}\left[\int_t^T H_{t,s}^u\, k(s, u_s)\, dt + H_{t,T}^u \xi^u \mid \mathcal{F}_t\right].$$

The previous results yield the verification theorem that is a sufficient condition for a process to be the value function.

PROPOSITION 3.2 (Verification Theorem). *The parameters (f, ξ) defined by*

$$f(t, y, z) = \text{ess inf}\{f^u(t, y, z) \mid u \in \mathcal{U}\}, \qquad \xi = \text{ess inf}\{\xi^u \mid u \in \mathcal{U}\},$$

are standard parameters. Let (Y, Z) be the solution of the BSDE associated with terminal condition ξ. Then Y is the value function Y^* of the control problem; that is, for each $t \in [0, T]$,

$$Y_t = Y_t^* = \text{ess inf}\{Y_t^u \mid u \in \mathcal{U}\}.$$

Proof. To show that f is a standard generator, we have to overcome measurability questions: for given (ω, t), $f(\omega, t, y, z) = \inf\{k(\omega, t, u) + d(\omega, t, u)y + n(\omega, t, u)^* z \mid u \in \mathcal{U}\}$ is a concave function with respect to (y, z), with bounded derivatives. By taking the minimum only over a denumerable dense family $\{(y_n, z_n)\}$, we define, for each n, a measurable process $f(t, y_n, z_n)$ and a $d\mathbb{P} \otimes dt$-null set N such that, for $(\omega, t) \in N^c$, $f(\omega, t, y_n, z_n) = \inf\{k(\omega, t, u) + d(\omega, t, u)y_n + n(\omega, t, u)^* z_n, \mid u \in \mathcal{U}\}$. For $(\omega, t) \in N^c$, $f(\omega, t, y, z)$ is defined as the limit of the Cauchy sequence $f(\omega, t, y_n, z_n)$ as (y_n, z_n) goes to (y, z). So, the infimum of the linear generators defines a standard generator f. $\square$

To apply the previous results on the infimum of standard generators, we will use the following lemma.

LEMMA 3.1. *For each $\epsilon > 0$ there exists a feasible control u^ϵ such that*

$$\begin{aligned} f(t, Y_t, Z_t) &= \text{ess inf}\, f^u(t, Y_t, Z_t) \geq f^{u^\varepsilon}(t, Y_t, Z_t) - \varepsilon,\ d\mathbb{P} \otimes dt \quad \text{a.s.}, \\ \xi &= \text{ess inf}\, \xi^u \geq \xi^{u^\varepsilon} - \varepsilon, a.s. \end{aligned} \tag{3.12}$$

Proof. For each $(\omega, t) \in \Omega \times [0, T[$, the sets given by

$$\{u \in U \mid f(t, Y_t(\omega), Z_t(\omega)) \geq k(t, \omega, u) + d(t, \omega, u) \cdot Y_t(\omega) + n(t, \omega, u)^* Z_t(\omega) - \varepsilon\}$$

and $\{u \in U \mid \xi(\omega) \geq K(\omega, u) - \varepsilon\}$ are nonempty. Hence, by a measurable selection theorem (see, for example, Dellacherie 1972 or Benes 1970, 1971) and since Y and Z are predictable processes and k, d, n, and K are measurable, there exists a U-valued predictable processes u^ε such that

$$\begin{aligned} f(t, Y_t, Z_t) &= \operatorname{ess\,inf} f^u(t, Y_t, Z_t) \geq f^{u^\varepsilon}(t, Y_t, Z_t) - \varepsilon,\ d\mathbb{P} \otimes dt \text{ a.s.}, \\ \xi &= \operatorname{ess\,inf} \xi^u \geq \xi^{u^\varepsilon} - \varepsilon, \mathbb{P} \text{ a.s.} \qquad \square \end{aligned} \tag{3.13}$$

Proof of the Verification Theorem. Corollary 3.1 and Lemma 3.1 give the desired result directly. □

Recall that the main tool in stochastic control is the *principle of dynamic progamming* (see Fleming and Rishel 1975 or El Karoui 1981). However, in the context of BSDEs it is nothing else than the flow property (2.11). Using the same notation as in (2.11), for a stopping time $S \leq T$ and an $\mathcal{F}_S$-measurable variable ξ_S, we denote by $Y_t^u(S, \xi_S)$ the solution of the BSDE with standard parameters $(T,\ f^u(t, y, z)1_{\{t \leq S\}}, \xi_S)$.

PROPOSITION 3.3. *The value function (Y_t) satisfies the dynamic programming principle: for any time t and any stopping time S with $t \leq S \leq T$,*

$$Y_t(T, \xi) = \operatorname*{ess\,inf}_{u \in \mathcal{U}} Y_t^u(S, Y_S(T, \xi))\ \mathbb{P}\ a.s.,$$

which can also be written

$$Y_t(T, \xi) = \operatorname*{ess\,inf}_{u \in \mathcal{U}} \mathbb{E}\left[\int_t^S H_{t,s}^u\, k(s, u_s)\, dt + H_{t,S}^u Y_S(T, \xi) \mid \mathcal{F}_t\right] \mathbb{P}\ a.s.$$

Actually, the optimization problem is to find a 0-optimal control u^0 which achieves the minimum for the problem $\inf\{Y_0^u \mid u \in \mathcal{U}\}$; that is, $Y_0^* = Y_0^{u^0}$. The comparison theorem gives a *criterion* for finding 0-optimal controls.

COROLLARY 3.3 (Optimality Criterion). *A control $(u_s^0,\ 0 \leq s \leq T)$ is 0-optimal if and only if*

$$\begin{aligned} f(s, Y_s, Z_s) &= f^{u^0}(s, Y_s, Z_s) \quad d\mathbb{P} \otimes ds\ a.s., \\ \xi &= \xi^{u^0}\ \mathbb{P}\ a.s. \end{aligned} \tag{3.14}$$

In this case, u^0 is also optimal for the problem starting at time t; that is, $Y_t^* = Y_t^{u^0}$.

Proof. It is an easy consequence of the second part of the comparison theorem 2.2. □

From the verification theorem and Remark b, Section 2.3, (Y, Z) is a subsolution of the BSDEs with parameters (f^u, ξ^u). For each feasible control u, $H_t^u Y_t + \int_0^t H_s^u\, k(s, u_s)\, ds$ is a uniformly integrable submartingale with increasing process given by $\int_0^t H_s^u K_s^u\, ds$, where

$$K_t^u = -f(t, Y_t, Z_t) + f^u(t, Y_t, Z_t).$$

Furthermore, the optimality criterion yields that u^0 is 0-optimal if and only if $\xi = \xi^{u^0}$ and $K_t^{u^0} = 0$—in other words, if and only if $Y_T = K(u_T^0)$ and $H_t^{u^0} Y_t + \int_0^t H_s^{u^0}\, k(s, u_s^0)\, ds$ is a martingale. Consequently, the previous results correspond to the classical properties of the value function (El Karoui 1981, Theorem 3.2).

In control theory the processes δ and ν are not necessarily bounded but are integrable enough to guarantee that the family (H_T^u) is uniformly bounded in $\mathbb{L}_T^{2,1}$ (El Karoui 1981, p. 297). If f is not a standard generator, it is not possible to use the BSDE equations. The direct study of the value function Y^* gives that Y^* is the greatest process equal to ξ at time T such that $H_t^u Y_t^* + \int_0^t H_s^u\, k(s, u_s)\, ds$ is a uniformly integrable submartingale for any feasible control u. Using arguments of weak convergence about a minimizing control sequence, it is proved in El Karoui (1981) that Y_t^* is a solution (not necessarily square integrable) of the BSDE associated with terminal condition ξ and with generator f and that it is the maximal solution.

Notice that in this example the generator f is concave. We will see in the next section that, conversely, a concave BSDE is always associated with a control problem.

Concave BSDE as Infimum. Here we fix some notation and recall a few properties of convex analysis (whose proofs are, for example, in Ekeland and Teman 1976 and Ekeland and Turnbull 1979) in order to show that a concave generator is an infimum of linear generators. Let $f(t, y, z)$ be a standard generator of a BSDE, concave with respect to y, z, and let $F(t, \beta, \gamma)$ be the polar process associated with f:

$$F(\omega, t, \beta, \gamma) = \sup_{(y,z)\in\mathbb{R}\times\mathbb{R}^n} [f(\omega, t, y, z) - \beta\, y - \gamma^* z]. \tag{3.15}$$

The *effective domain of* F is, by definition,

$$\mathcal{D}_F = \{(\omega, t, \beta, \gamma) \in \Omega \times [0, T] \times \mathbb{R} \times \mathbb{R}^n \mid F(\omega, t, \beta, \gamma) < +\infty\}.$$

Notice[10] that, since by assumption f is uniformly Lipschitz with Lipschitz constant C, the (ω, t)-section of $\mathcal{D}_F$, denoted by $\mathcal{D}_F^{(\omega,t)}$, is included in the bounded domain $K = [-C, C]^{n+1}$

[10] Indeed, if, for example, β satisfies $|\beta| > C$, then

$$f(\omega, t, y, z) - \beta\, y - \gamma^* z \ge -C|y| + f(\omega, t, 0, z) - \beta y - \gamma^* z.$$

Now, $\sup_{y\in\mathbb{R}}[-C|y| - \beta y] = +\infty$. Hence, $(\beta, \gamma) \notin \mathcal{D}_F^{(\omega,t)}$.

of $\mathbb{R} \times \mathbb{R}^n$. Since f is concave, f is continuous with respect to (y, z), and (f, F) satisfies the conjugacy relation

$$f(\omega, t, y, z) = \inf\{F(\omega, t, \beta, \gamma) + \beta y + \gamma^* z \mid (\beta, \gamma) \in \mathcal{D}_F^{(\omega,t)}\}.$$

For every (ω, t, y, z) the infimum is achieved in this relation by a pair (β, γ) which depends on (ω, t).[11]

We want to associate with the polar process F a wide enough family of linear *standard* generators $f^{\beta,\gamma}$ such that the assumptions of Proposition 3.1 hold. Let

$$f^{\beta,\gamma}(t, y, z) = F(t, \beta_t, \gamma_t) + \beta_t y + \gamma_t^* z,$$

where (β, γ) are predictable processes, called *control parameters*. Recall that by the conjugacy relation f is also the infimum of $f^{\beta,\gamma}$. To ensure that $f^{\beta,\gamma}$ is a standard generator, it is sufficient to suppose that (β, γ) belongs to $\mathcal{A}$, defined by

$$\mathcal{A} = \left\{(\beta, \gamma) \in \mathcal{P}, \text{K-valued } \mid \mathbb{E}\int_0^T F(t, \beta_t, \gamma_t)^2 \, dt < +\infty\right\}.$$

$\mathcal{A}$ is said to be the set of *admissible control parameters*. Let (Y, Z) be the unique solution of the BSDE with concave standard generator f and terminal value ξ. To apply Proposition 3.1, we must show the following lemma (which is similar to Lemma 3.1).

LEMMA. There exists an optimal control $(\overline{\beta}, \overline{\gamma}) \in \mathcal{A}$ such that

$$f(t, Y_t, Z_t) = f^{\overline{\beta},\overline{\gamma}}(t, Y_t, Z_t) \, d\mathbb{P} \otimes dt \ a.s.$$

Proof. Recall that for each (t, ω, y, z) the infimum in the conjugacy relation is achieved, since f is concave uniformly Lipschitz. Also, by a measurable selection theorem and since $f(\cdot, Y, Z)$, Y, and Z are predictable processes, there exists a pair of predictable (bounded) processes $(\overline{\beta}, \overline{\gamma})$ such that

$$f(t, Y_t, Z_t) = f^{\overline{\beta},\overline{\gamma}}(t, Y_t, Z_t), d\mathbb{P} \otimes dt \ \text{a.s.}$$

[11] Indeed, for fixed (ω, t) there exists a sequence $(\beta^k, \gamma^k)_{k\in\mathbb{N}} \in \mathcal{D}_F^{(\omega,t)}$ such that

$$f(\omega, t, y, z) = \lim_{k\to+\infty}\{F(\omega, t, \beta^k, \gamma^k) + \beta^k y + (\gamma^k)^* z\}.$$

Now since the sequence (β^k, γ^k) is bounded, there exists a subsequence still denoted by (β^k, γ^k) which converges in K to (β, γ). Also, (β, γ) achieves the infimum since F is continuous and

$$F(\omega, t, \beta, \gamma) + \beta y + \gamma^* z = \lim_{k\to+\infty}[F(\omega, t, \beta^k, \gamma^k) + \beta^k y + (\gamma^k)^* z] = f(\omega, t, y, z).$$

Since by assumption $f(\cdot, Y, Z)$, Z, and Y are square integrable and $\overline{\beta}, \overline{\gamma}$ are bounded, $F(\cdot, \overline{\beta}, \overline{\gamma})$ belongs also to $\mathbb{H}_T^{2,1}$. Hence, the pair $(\overline{\beta}, \overline{\gamma})$, which achieves the infimum in the conjugacy relation, belongs to $\mathcal{A}$. □

For each control process $(\beta, \gamma) \in \mathcal{A}$, introduce the dual "controlled objective" processes $(Y^{\beta,\gamma}, Z^{\beta,\gamma})$ as the unique solution of the LBSDE with data $(f^{\beta,\gamma}, \xi)$. Thus Proposition 3.1 gives directly the following result.

PROPOSITION 3.4. *Let f be a concave standard generator and $f^{\beta,\gamma}$ the associated linear standard generators satisfying*

$$f = \operatorname{ess\,inf}\{f^{\beta,\gamma} \mid (\beta, \gamma) \in \mathcal{A}\} \; d\mathbb{P} \otimes dt \;\; a.s.$$

Then $\mathbb{P}$ a.s. for any time t,

$$Y_t = \operatorname{ess\,inf}\{Y_t^{\beta,\gamma} \mid (\beta, \gamma) \in \mathcal{A}\}\,.$$

Let us interpret the above result as associated with a control problem, with control set $\mathcal{A}$. From Proposition 2.2, the LBSDEs solution $Y_t^{\beta,\gamma}$ can be written using the adjoint process $(\Gamma_{t,s}^{\beta,\gamma}, t \le s \le T)$, which is the unique solution of the forward linear SDE

$$d\Gamma_s = \Gamma_s[\beta_s\, ds + \gamma_s\, dW_s] \qquad \Gamma_t = 1\,, \tag{3.16}$$

in the following manner:

$$Y_t^{\beta,\gamma} = \mathbb{E}\left[\int_t^T \Gamma_{t,s}^{\beta,\gamma}\; F(s, \beta_s,\, \gamma_s)\, ds + \Gamma_{t,T}^{\beta,\gamma}\; \xi | \mathcal{F}_t\right].$$

Here $Y^{\beta,\gamma}$ is called the *controlled objective function* of a control problem, where the running cost function is the function $F(t, \beta, \gamma)$ and the terminal cost is the random variable ξ.

3.2. Application to Recursive Utility

We come back to the example of recursive utility presented in Section 1. In an economic or financial context the generator $f(t, c_t, Y_t, Z_t)$ represents the instantaneous utility (at time t) of consumption rate $(c_t \ge 0)$. In general, we suppose that the consumption process c belongs to $\mathbb{H}_T^2(\mathbb{R}^+)$ and is such that $(f(\cdot, c, \cdot, \cdot), \xi)$ are standard parameters of BSDE (in particular,[12] $f(\cdot, c, 0, 0) \in \mathbb{H}_T^2(\mathbb{R})$).

Classical Properties. In this section we show that under natural conditions the classical properties of utilities (Section 1.4) are satisfied by recursive ones; actually, it is a direct consequence of the comparison theorem.

[12]For example, it suffices that $|f(t, c, 0, 0)| \le k_1 + k_2|c|$ $\mathbb{P}$ a.s.

PROPOSITION 3.5. *Let ξ^1 and ξ^2 be two terminal rewards which belong to $\mathbb{L}^2_T(\mathbb{R})$. Let c^1 and c^2 be two consumption processes which belong to $\mathbb{H}^2_T(\mathbb{R}^+)$. Let Y^{c^1,ξ^1} and Y^{c^2,ξ^2} be the recursive utilities associated with $(f(t,c^1,\cdot),\xi^1)$ and $(f(t,c^2,\cdot),\xi^2)$. Then the following are true:*

- *(Time consistency) Suppose $\xi^1=\xi^2=\xi$. If $Y_t^{c^1,\xi}\ge Y_t^{c^2,\xi}$ and $c_s^1=c_s^2$, $0\le s\le t$, $d\mathbb{P}\otimes ds$ a.s., then $Y_s^{c^1,\xi}\ge Y_s^{c^2,\xi}$, $0\le s\le t$, $\mathbb{P}$ a.s.*
- *(Monotonicity with respect to the terminal value) Suppose $c^1=c^2=c$. If $\xi^1\ge\xi^2$ $\mathbb{P}$ a.s., then $Y^{c,\xi^1}\ge Y^{c,\xi^2}$ $\mathbb{P}$ a.s.*
- *(Monotonicity with respect to the consumption) Suppose $\xi^1=\xi^2=\xi$. If the generator f is nondecreasing with respect to c, and if $c_t^1\ge c_t^2$, $d\mathbb{P}\otimes dt$ a.s., then $Y^{c^1,\xi}\ge Y^{c^2,\xi}$, $\mathbb{P}$ a.s.*
- *(Concavity) If the generator f is concave with respect to c, y, and z, then for each $\lambda\in[0,1]$, $\lambda Y^{c^1,\xi^1}+(1-\lambda)Y^{c^2,\xi^2}\le Y^{c,\xi}$, $\mathbb{P}$ a.s., where $c=\lambda c^1+(1-\lambda)c^2$ and $\xi=\lambda\xi^1+(1-\lambda)\xi^2$.*

Proof. We show how to prove the first point. Since $Y_t^{c^1,\xi}\ge Y_t^{c^2,\xi}$ $\mathbb{P}$ a.s. and since $Y^{c^1,\xi}$ and $Y^{c^2,\xi}$ have the same generator on $[0,t]$, the result follows from the comparison theorem applied between time 0 and time t. The other properties are also direct consequences of the comparison theorem. $\square$

Variational Formulation of the Recursive Utility. A natural assumption for a recursive utility is the concavity of the generator f with respect to (c,y,z). Consequently, by the results on concave BSDEs, the recursive utility can be written as the value function of a control problem.

Fix ξ, a terminal reward which belongs to $\mathbb{L}^2_T(\mathbb{R})$, and c, a consumption process in $\mathbb{H}^2_T(\mathbb{R}^+)$. Let $Y^{c,\xi}$ be the associated recursive utility—that is, the solution of the BSDE associated with generator $f(t,c_t(\cdot),\cdot,\cdot)$ and terminal value ξ. For a consumption rate c_t, let $F(t,\ c_t,\cdot,\cdot)$ be the polar function of $f(t,c_t,\cdot,\cdot)$; i.e.,

$$F(t,\ c_t,\beta,\gamma)=\sup_{(y,z)\in\mathbb{R}\times\mathbb{R}^n}[f(t,c_t,\ y,\ z)-\beta\cdot y-\gamma\cdot z].$$

Let $\mathcal{A}(c)$ be the set of admissible processes (β,γ) such that $\mathbb{E}\int_0^T F(t,c_t,\beta_t,\gamma_t)^2\,dt<+\infty$. Then by Section 3.1 the recursive utility can be written as

$$Y_t^{c,\xi}=\operatorname*{ess\,inf}_{(\beta,\gamma)\in\mathcal{A}}\mathbb{E}\left[\int_t^T\Gamma_{t,s}^{\beta,\gamma}\,F(s,c_s,\beta_s,\gamma_s)\,ds+\ \Gamma_{t,T}^{\beta,\gamma}\xi\ |\,\mathcal{F}_t\right].$$

Hence the recursive utility $Y^{c,\xi}$ can be defined through a felicity function F first introduced by Geoffard (1995) in the deterministic case. The felicity function $F(t,c,\beta,\gamma)$ at some current time t, expressed in terms of current time, is a function of current consumption c, current rate $-\beta$, and risk premium $-\gamma$. This function can be thought as an ex post felicity when the agent knows the current rate and the risk premium.

Notice that the adjoint processes $\Gamma^{\beta,\gamma}$ can be interpreted as a deflator (Duffie 1992 or Duffie, Geoffard, and Skiadias 1992). Also, the process $Y_t^{\beta,\gamma}$ can be interpreted as an ex post utility, when the deflator is given by $\Gamma_t^{\beta,\gamma}$. Hence, the utility is equal to the minimum of ex post utilities over all price deflators. Ex ante the optimal deflator is the one that minimizes the agent's ex post utility.

Concerning the wealth process associated with some portfolio consumption strategy, we have the same kind of interpretation, as we shall to see in the next section.

3.3. Application to European Option Pricing in the Constrained Case

A general setting of the wealth equation (which extends the examples of Section 1.3) is

$$-\,dX_t = b(t, X_t, \sigma_t^* \pi_t)\, dt - \pi_t^* \sigma_t \, dW_t \,. \tag{3.17}$$

Here b is a real process defined on $\Omega \times [0, T] \times \mathbb{R} \times \mathbb{R}^n$ satisfying the standard hypotheses of a generator. The classical case (Section 1.2) corresponds to a linear functional

$$b(t, x, z) = -r_t x - z^* \theta_t \,,$$

where θ is the bounded risk premium vector and r is the bounded spot-rate process. Notice that, since b is Lipschitz, given an initial investment x and a risky portfolio π, there exists a unique wealth process solution of the forward equation (3.17) with initial value x.

A price system is a mapping Ψ which maps a contingent claim $\xi \geq 0$ onto its (predictable) price process $(\Psi_t(\xi),\ 0 \leq t \leq T)$ such that

- At any time t, the price $\Psi_t(\xi)$ for a positive contingent claim ξ is positive.
- At any time t, the price $\Psi_t(\xi)$ is an increasing function with respect to ξ.
- No-arbitrage holds for these nonlinear strategies; i.e., if $\xi^1 \geq \xi^2$, and if the prices X_t^1 and X_t^2 coincide on an event $A \in \mathcal{F}_t$, then on A, $\xi^1 = \xi^2$, a.s.

Furthermore, a price system Ψ is admissible for the sellers if at any time t the price $\Psi_t(\xi)$ is a convex function with respect to ξ.

Let $\xi \geq 0$ be a square-integrable contingent claim. As in the classical case the price for the contingent claim ξ is the wealth process X associated with an admissible strategy which finances ξ; i.e., $(X_. \sigma^*_. \pi_.)$ is the square-integrable solution of BSDE (3.17) with standard parameters (b, ξ).

Let us show that, under some conditions, this price rule $\Psi_.(\xi) = X_.$ defines a price system admissible for the sellers. Actually, the comparison theorem gives sufficient conditions so that these different properties hold:

- The price is increasing with respect to the contingent claim, and the property of no-arbitrage corresponds exactly to the strict comparison theorem.
- Suppose $b(t, 0, 0) \geq 0$, $d\mathbb{P} \otimes dt$ a.s. Then the price is positive and it is smaller than any supersolution of BSDE (3.17).
- Suppose b is convex with respect to (x, z). Then the price system is convex with respect to the terminal value.

This convexity property holds in all examples of Section 1.3, where the prices are viewed as superprices. In the next section we give a variational formulation of this price system when b is assumed to be convex.

Variational Formulation of the Price System. Suppose that the generator b is convex with respect to (x, z). Let $B(t, \beta, \gamma)$ be the polar process associated with b:

$$B(t, \beta, \gamma) = \inf_{(x,z)\in\mathbb{R}\times\mathbb{R}^n} [b(t, x, z) + \beta.x + \gamma^* z]. \tag{3.18}$$

The *effective domain of B* is, by definition, $\mathcal{D}_B^t = \{(\beta, \gamma) \mid B(t, \beta, \gamma) > -\infty\}$. Since b is convex, (b, B) satisfies the conjugacy relation

$$b(t, x, z) = \sup\{B(t, \beta, \gamma) - \beta x - \gamma.z \mid (\beta, \gamma) \in \mathcal{D}_B^t\}.$$

The results on concave BSDEs give

PROPOSITION 3.6. *Let (X, π) be the hedging strategy for ξ such that $(X, \sigma^*\pi)$ is the unique solution of BSDE (3.17), with convex standard parameters (b, ξ). Then X_t can be written as the maximum of ex post prices over all feasible deflators; that is,*

$$X_t = \operatorname{ess\,sup}\{X_t^{\beta,\gamma} \mid (\beta, \gamma) \in \mathcal{A}\},$$

where $\mathcal{A}$ is the set of (β, γ) feasible control parameters, defined by

$$\mathcal{A} = \left\{(\beta, \gamma) \in \mathcal{P} \mid \mathbb{E}\int_0^T B(t, \beta_t, \gamma_t)^2\, dt < +\infty\right\}$$

and where, for each pair of control parameters $(\beta, \gamma) \in \mathcal{A}$, the ex post strategy $(X^{\beta,\gamma}, \pi^{\beta,\gamma})$ corresponds to the unique solution of the LBSDE

$$\begin{aligned} -dX_t^{\beta,\gamma} &= (B(t, \beta_t, \gamma_t) - \beta_t X_t^{\beta,\gamma} - (\gamma_t)^*\sigma_t^*\pi_t^{\beta,\gamma})\, dt - (\pi^{\beta,\gamma})_t^*\sigma_t\, dW_t, \\ X_T^{\beta,\gamma} &= \xi. \end{aligned} \tag{3.19}$$

The ex post strategy $(X^{\beta,\gamma}, \pi^{\beta,\gamma})$ is a classical hedging strategy against the claim ξ in a fictitious market, with bounded interest rate process β, bounded risk premium process γ, and cost function B. Furthermore, the price of the contingent claim ξ is the standard price in an optimal fictitious market associated with $(\overline{\beta}, \overline{\gamma})$ which achieves the supremum in the conjugacy relation

$$b(t, X_t, \sigma_t^*\pi_t) = B(t, \overline{\beta}_t, \overline{\gamma}_t) - \overline{\beta}_t X_t - (\overline{\gamma})_t^*\sigma_t^*\pi_t\, d\mathbb{P}\otimes dt\ a.s. \tag{3.20}$$

We remark that the only difference in the nonconstrained case is the fact that the optimal fictitious market depends on the claim to be priced (and also the introduction of a cost function).

EXAMPLE. We come back to Example 1.1 (Section 1.3, hedging claims with higher interest rate for borrowing) and solved by Cvitanic and Karatzas (1993) under slightly different assumptions. Here we suppose the matrix $(\sigma^*)^{-1}$ to be a bounded process. The hedging strategy (wealth, portfolio) (X, π) satisfies

$$(3.21)\quad dX_t = r_t X_t\, dt + \pi_t^* \sigma_t \theta_t\, dt + \pi_t^* \sigma_t\, dW_t - (R_t - r_t)\left(X_t - \sum_{i=1}^n \pi_t^i\right)^- dt,$$
$$X_T = \xi\,.$$

Like the other coefficients, the process R $(R_t \geq r_t)$ is supposed to be bounded. The generator b of this LBSDE is given by the convex process

$$b(t, x, \sigma_t^* \pi) = -r_t\, x - \pi^* \sigma_t \theta_t + (R_t - r_t)(x - \pi_t^* \mathbf{1})^-\,.$$

The polar process $B(t, \beta, \gamma)$ associated with b is given by

$$(3.22)\qquad B(t, \beta, \gamma) = \begin{cases} 0 & \text{if } \gamma = \theta_t + \sigma_t^{-1}(r_t - \beta)\mathbf{1} \text{ and } r_t \leq \beta \leq R_t, \\ -\infty & \text{otherwise.} \end{cases}$$

By Proposition 3.6, it follows that the unique solution $(X, \sigma^* \pi)$ of the BSDE (5.14) satisfies

$$X_t = \operatorname{ess\,sup}\{X_t^\beta \mid r_t \leq \beta_t \leq R_t\},$$

where

$$-dX_t^\beta = -\beta_t X_t^\beta - [\sigma_t \theta_t + (r_t - \beta_t)\mathbf{1}]^* \pi_t^\beta\, dt - (\pi_t^\beta)^* \sigma_t\, dW_t, \qquad X_T = \xi.$$

REMARK. The problems with constraints on the wealth (El Karoui et al. 1995) or on the portfolio (El Karoui and Quenez 1995; Cvitanic and Karatzas 1992, 1993) can be formulated formally in the same way but with a generator which can be infinite, with nonbounded effective domain. For example, the case of an incomplete market (Section 1.3) and, more generally, the case of the portfolio process π_t being constrained to take values in a convex set K (Cvitanic and Karatzas 1992) corresponds formally to an upper price $(X_t, 0 \leq t \leq T)$ solution of BSDE (3.17) with generator

$$b(t, x, \sigma_t^* \pi) = -r_t x - \pi^* \sigma_t \theta_t + \mathbf{1}_K(\pi)\,,$$

where $\mathbf{1}_K(\pi)$ is the indicator function of K in the sense of convex analysis, namely equal to 0 if $\pi \in K$ and equal to ∞ otherwise. Notice that the example of an incomplete market corresponds to $K = \{\pi \in \mathbb{R}^n \mid \pi_k = 0,\ j \leq k \leq n.\}$

The variational formulation of the price remains almost the same as the one described in Section 3.3. However, as in the example of an incomplete market, the effective domain is not bounded; moreover, the supremum is not attained and (X, π) is not the solution of a classical BSDE.

4. MARKOVIAN CASE

4.1. Forward-Backward Stochastic Differential Equations

In this section we consider the solution of certain BSDEs associated with some forward classical stochastic differential equations. For example, the forward equation can be the dynamics of some basic securities as in Section 1.1. Now suppose that the randomness of the parameters (f, ξ) of the BSDE comes from the state of the forward equation. When the initial conditions (t, x) for the forward equation are taken into account, the backward solution (Y, Z) can be viewed as a parametrized BSDE where the parameters are the data (t, x); consequently, some regularity properties of the solutions follow from regularity properties of the coefficients of the forward and backward equations.

However, the main property of forward-backward SDEs (FBSDEs) is that the solution (Y, Z) of the BSDE can be written as functions of time and the state process. The solution is said to be Markovian. When the generator f depends not on y and z but only on time and the state process, the Markov property of the forward diffusion allows one to express the BSDEs solution by means of the diffusion semigroup or as a viscosity solution of the second-order associated PDE. For Markovian standard parameters the same property holds and gives a generalization of the Feynman-Kac formula for nonlinear PDEs as stated by Peng (1991), Peng (1992b), and Pardoux and Peng (1992).

The Model. For any given $(t, x) \in [0, T] \times \mathbb{R}^p$, consider the following classical Itô stochastic differential equation defined on $[0, T]$:

$$
\begin{aligned}
dP_s &= b(s, P_s)\, ds + \sigma(s, P_s)\, dW_s, \qquad t \leq s \leq T, \\
P_s &= x, \qquad 0 \leq s \leq t\,.
\end{aligned}
\tag{4.1}
$$

The solution of (4.1) will be denoted $(P_s^{t,x},\ 0 \leq s \leq T)$. We then consider the associated BSDE

$$
\begin{aligned}
-\,dY_s &= f(s, P_s^{t,x}, Y_s, Z_s)\, ds - Z_s^*\, dW_s, \\
Y_T &= \Psi(P_T^{t,x})\,.
\end{aligned}
\tag{4.2}
$$

The solution of (4.2) will be denoted $\{(Y_s^{t,x}, Z_s^{t,x}),\ 0 \leq s \leq T\}$. The coupled system (4.1) and (4.2) is said to be an FBSDE and the solution is denoted by $\{(P_s^{t,x}, Y_s^{t,x}, Z_s^{t,x}),\ 0 \leq s \leq T\}$.

Here f (resp. Ψ) is an $\mathbb{R}^d$-valued Borel function defined on $[0, T] \times \mathbb{R}^p \times \mathbb{R}^d \times \mathbb{R}^{n \times d}$ (resp. on $\mathbb{R}^p$), and b (resp. σ) is an $\mathbb{R}^p$-valued (resp. $\mathbb{R}^{p \times n}$-valued) function defined on $[0, T] \times \mathbb{R}^p$. Standard Lipschitz assumptions are required on the coefficients; that is, there exists a Lipschitz constant $C > 0$ such that

$$
\begin{aligned}
|\sigma(t, x) - \sigma(t, y)| + |b(t, x) - b(t, y)| &\leq C(1 + |x - y|), \\
|\ f(t, x, y_1, z_1) - f(t, x, y_2, z_2)\ | &\leq C[|\ y_1 - y_2\ | + |\ z_1 - z_2\ |].
\end{aligned}
$$

Finally, we will suppose that there exists a constant C such that, for each $(s, x, , y, z)$,

$$\begin{aligned} |\sigma(t,x)| + |b(t,x)| &\le C(1+|x|), \\ |f(t,x,y,z)| + |\Psi(x)| &\le C(1+|x|^p) \end{aligned}$$

for real $p \ge 1/2$.

Properties of Solutions of BSDEs Associated with Some FSDEs. These equations are in a quite complex way examples of BSDEs parametrized by the initial conditions (t, x) of the FSDE. The parametrized generator is given here by $f(s, P_s^{t,x}(\omega), y, z)$ and the terminal condition by $\xi(\omega) = \Psi(P_T^{t,x}(\omega))$. As in Proposition 2.4, regularity properties of the solutions follow from regularity properties of the parameters of the BSDE.

PROPOSITION 4.1. *1. For each $t \in [0, T]$ and $x \in \mathbb{R}^p$ there exists $C \ge 0$ such that*

$$E\left(\sup_{0\le s\le T} |Y_s^{t,x}|^2\right) + E\left(\int_0^T |Z_s^{t,x}|^2\, ds\right) \le C(1+|x|^2)\,. \tag{4.3}$$

2. Suppose that f and Ψ are globally Lipschitz with respect to x, uniformly in t concerning f. Then for each $t, t' \in [0, T]$, $t \le t'$, and $x, x' \in \mathbb{R}^p$, there exists $C \ge 0$ such that

$$\begin{aligned} &E[\sup_{0\le s\le T} |Y_s^{t,x} - Y_s^{t',x'}|^2] + E\left[\int_0^T |Z_s^{t,x} - Z_s^{t',x'}|^2\, ds\right] \\ &\quad \le C(1+|x|^2)(|x-x'|^2 + |t-t'|)\,. \end{aligned} \tag{4.4}$$

3. If b, σ, f, and Ψ are twice continuously differentiable with respect to x with uniformly bounded derivatives, then for each t the function $x \mapsto (Y_.^{t,x}, Z_.^{t,x}), \mathbb{R} \mapsto H^2_{T,\beta}(\mathbb{R}^d) \times H^2_{T,\beta}(\mathbb{R}^{n\times d})$, is differentiable. Let the matrices of first-order partial derivatives of $P^{t,x}$, $Y^{t,x}$, and $Z^{i,t,x}$ with respect to x be denoted by the $p \times p$ matrix $\partial_x P^{t,x}$, by the $d \times p$ matrix $\partial_x Y^{t,x}$, and by the $d \times p$ matrix $\partial_x Z^{i,t,x}$ respectively (where $Z^{i,t,x}$ is the i^{th} line of the matrix $Z^{t,x}$). Then, for $t \le s \le T$,

$$\begin{aligned} -d\partial_x Y_s^{t,x} &= [\partial_y f(s, P_s^{t,x}, Y_s^{t,x}, Z_s^{t,x})\partial_x Y_s^{t,x} \\ &\quad + \partial_z f(s, P_s^{t,x}, Y_s^{t,x}, Z_s^{t,x})\partial_x Z_s]\, ds \\ &\quad + \partial_x f(s, P_s, Y_s, Z_s)\partial_x P_s\, ds - \sum_{1\le i\le n} \partial_x Z_s^{i,t,x}\, dW_s^i\,, \\ \partial_x Y_T^{t,x} &= \partial_x \Psi(P_T^{t,x})\partial_x P_T^{t,x}\,. \end{aligned} \tag{4.5}$$

Proof. As in Proposition 2.4, the results follow from regularity properties of the standard parameters given by $(f(s, P_s^{t,x}, y, z), \Psi(P_T^{t,x}))$, so for this proof regularity properties with respect to (t, x) are required on $P_.^{t,x}$.

Using the classical martingale inequalities (Karatzas and Shreve 1987, Theorems 3.28 and 2.9), we derive by classical techniques that, for each $t, t' \in [0, T]$, $t \le t'$, and $x, x' \in \mathbb{R}^p$, there exists $C \geq 0$ such that

$$E\left(\sup_{0\leq s\leq T}|P_s^{t,x}|^2\right) \leq C(1+|x|^2), \tag{4.6}$$
$$E\left(\sup_{0\leq s\leq T}|P_s^{t,x} - P_s^{t',x'}|^2\right) \leq C(1+|x|^2)\left(|x-x'|^2+|t-t'|\right).$$

Using the first inequality and the a priori estimates, we easily obtain inequality (4.3). Suppose now that f, Ψ are globally Lipschitz with respect to (x, y, z). Then, from the above inequalities, the a priori estimates, and the fact that f is Lipschitz with respect to x, we obtain inequality (4.4).

It remains to show the third statement. Suppose that b, σ, f, and Ψ are continuously differentiable with respect to x, y, and z with uniformly bounded derivatives. Recall (Krylov 1980) that the process $P^{t,x}_{.}$ is differentiable with respect to x and that the matrix of the first-order derivatives $\partial_x P_s^{t,x}$ is a solution of the FSDE

$$d\partial_x P_s^{t,x} = \partial_x b(s, P_s^{t,x})\,\partial_x P_s^{t,x}\,ds + \partial_x\sigma_i(s, P_s^{t,x})\,\partial_x P_s^{t,x}\,dW_s^i, \qquad \partial_x P_0 = I,$$

where we use the convention of summation over the repeated index i, from $i = 1$ to $i = p$, and σ_i denotes the ith column of the matrix σ. Then, using exactly the same arguments as in the proof of Proposition 2.4, the result easily follows. $\square$

REMARK. If b, σ, f, and Ψ are continuously differentiable with respect to x, y, and z with uniformly bounded derivatives, then the solution $\{(Y_s^{t,x}, Z_s^{t,x}); s \in [t, T]\}$ is also differentiable in Malliavin's sense (see Section 5.2) and there exists a version of the Malliavin derivative denoted by $(D_\theta Y_s^{t,x}, D_\theta Z_s^{t,x}, 0 \leq \theta, t \leq s \leq T)$ which satisfies $D_s Y_s^{t,x} = Z_s^{t,x}$ $d\mathbb{P}\otimes ds$-almost surely. This property is very useful for computing or estimating Z (which corresponds to the hedging portfolio in the pricing theory of contingent claims).

An important property is that $Y_t^{t,x}$ is deterministic. More precisely, the measurability properties of $\{P_s^{t,x}; s \in [t, T]\}$ still hold for the solution $\{(Y_s^{t,x}, Z_s^{t,x}); s \in [t, T]\}$.

PROPOSITION 4.2. *The solution* $\{(Y_s^{t,x}, Z_s^{t,x}); s \in [t, T]\}$ *of (4.2) is adapted to the future* σ*-algebra of* W *after* t*; that is, it is* $\mathcal{F}_s^t$*-adapted where for each* $s \in [t, T]$, $\mathcal{F}_s^t = \sigma(W_u - W_t, t \leq u \leq s)$. *In particular,* $Y_t^{t,x}$ *is deterministic. Consequently,* $Y_s^{t,x} = Y_t^{t,x}$ *and* $Z_s^{t,x} = 0$ *for* $0 \leq s \leq t$.

Proof. Recall first that pathwise (and strong) uniqueness holds for (4.1) (Karatzas and Shreve 1987, pp. 285, 287, 301). Consider now the translated Brownian motion W' and its associated filtration defined by $W'_s = W_{t+s} - W_t$, $0 \leq s \leq T-t$; $\mathcal{F}'_s = \mathcal{F}_{t+s}^t$, $0 \leq s \leq T-t$. Let $(P_s'^{(0,x)}, 0 \leq s \leq T-t)$ be the $(\mathcal{F}'_s)$-adapted solution of the SDE

$$dP'_s = b(s, P'_s)\,ds + \sigma(s, P'_s)\,dW'_s, \qquad P'_0 = x. \tag{4.7}$$

By uniqueness for the FSDE, $P_s^{t,x} = P_{s-t}^{\prime(0,x)}$, $0 \leq s \leq T-t$, a.s. (consequently $P_s^{t,x}$ is $\mathcal{F}_s^t$-adapted). We then consider the associated $\mathcal{F}'$-adapted solution $(Y_s', Z_s', \ 0 \leq s \leq T-t)$ of the BSDE

$$-\,dY_s' = f(s+t, P_s', Y_s', Z_s')\,ds - (Z_s')^*\,dW_s', \qquad Y_{T-t}' = \Psi(P_{T-t}')\,. \tag{4.8}$$

Hence, $\{(Y_{s-t}', Z_{s-t}')\,,\ t \leq s \leq T)\}$ is a solution of (4.2) on $[t, T]$. By uniqueness, $(Y_{s-t}', Z_{s-t}') = (Y_s^{t,x}, Z_s^{t,x})\,,\ t \leq s \leq T$. Consequently, $\{(Y_{s-t}', Z_{s-t}')\,,\ t \leq s \leq T)\}$ is $\mathcal{F}_s^t$-adapted. □

Markov Properties of Solutions of BSDEs Associated with Some FSDEs. Another way to prove that $Y_t^{t,x}$ is deterministic, and more generally that $Y_s^{t,x}$ is a deterministic function of $P_s^{t,x}$, is to use the iterative construction of the solution of the standard BSDE by noticing that if f does not depend on y, z then the property follows from the Markov property of the forward diffusion $P^{t,x}$:

$$Y_s^{t,x} = \mathbb{E}\left[\Psi(P_T^{t,x}) + \int_s^T f(r, P_r^{t,x})\,dr|\mathcal{F}_s\right] = \Phi(s, P_s^{t,x})\,,$$

where

$$\Phi(s, y) = \mathbb{E}\left[\Psi(P_T^{s,y}) + \int_s^T f(u, P_u^{s,y})\,du\right].$$

The solution of the BSDE is said to be *Markovian*. Furthermore, the process $Z^{t,x}$ associated with $Y^{t,x}$ by the martingale representation theorem is also a deterministic function of $P^{t,x}$. This result can be deduced from Cinlar et al.'s (1980) study on the functional additive martingale of a diffusion process (see also Dellacherie and Meyer 1980, pp. 241–244). In our notation, Theorem 6.27 in Cinlar et al. (1980) can be written as follows.

LEMMA 4.1. *Let $\mathcal{B}_e$ be the filtration on $\mathbb{R}^d$ generated by the functions*

$$\mathbb{E}\int_t^T \phi(s, P_s^{t,x})\,ds,$$

where ϕ is a continuous bounded $\mathbb{R}^d$-valued function. Then for any $\mathcal{B}_e$-measurable f and Ψ such that

$$\mathbb{E}\int_0^T |f(s, P_s^{t,x})|^2\,ds < +\infty, \qquad \mathbb{E}[|\Psi(P_T^{t,x})|^2] < +\infty\,,$$

the process $Y_s^{t,x} = \mathbb{E}[\Psi(P_T^{t,x}) + \int_s^T f(r, P_r^{t,x})\,dr|\mathcal{F}_s]$ admits a continuous version given by $Y_s^{t,x} = u(s, P_s^{t,x})$, where $u(t,x) = \mathbb{E}[\Psi(P_T^{t,x}) + \int_t^T f(r, P_r^{t,x})\,dr]$ is $\mathcal{B}_e$-measurable. Moreover, $\int_t^s f(r, P_r^{t,x})\,dr + Y_s^{t,x}$ is an additive square-integrable martingale which admits *the representation*

$$\int_t^s f(r, P_r^{t,x})\,dr + Y_s^{t,x} = \int_t^s d(r, P_r^{t,x})^* \sigma(r, P_r^{t,x})\,dW_r, \qquad t \le s \le T,\ \mathbb{P}\,a.s.,$$

where $d(t,x)$ *is* $\mathcal{B}([0,T]) \otimes \mathcal{B}_e(\mathbb{R}^{p\times d})$*-measurable.*

We now consider a BSDE associated with an FSDE whose data satisfy the general assumption at the beginning of the section and prove that the solution at time s, $(Y_s^{t,x}, Z_s^{t,x})$, is Markovian in the sense that both of these processes only depend on s and $P_s^{t,x}$.

THEOREM 4.1. *There exist two* $\mathcal{B}([0,T])\otimes\mathcal{B}_e(\mathbb{R}^d)$*- and* $\mathcal{B}([0,T])\otimes\mathcal{B}_e(\mathbb{R}^{p\times d})$*-measurable deterministic functions* $u(t,x)$ *and* $d(t,x)$*, respectively, such that the solution* $(Y^{t,x}, Z^{t,x})$ *of BSDE (4.2) is*

$$Y_s^{t,x} = u(s, P_s^{t,x}), \quad Z_s^{t,x} = \sigma^*(s, P_s^{t,x}) d(s, P_s^{t,x}), \qquad t \le s \le T,\ d\mathbb{P}\otimes ds\, a.s.$$

Furthermore, for any $\mathcal{F}_t$*-measurable random variable* $\chi \in \mathbb{L}^2(\mathbb{R}^p)$*, the solution* $(Y_s^{t,\chi}, Z_s^{t,\chi})$ *is given by* $(u(s, P_s^{t,\chi}), \sigma^*(s, P_s^{t,\chi}) d(s, P_s^{t,\chi}))$*, for* $s \ge t$, $d\mathbb{P}\otimes ds\, a.s.$

Proof. This result can be established by the iterative procedure used in the proof of the existence of the solution of a backward equation (Theorem 2.1 and Corollary 2.1). This procedure gives a recursive construction of the solution $(Y^{t,x}, Z^{t,x})$ from the sequence $(Y^{(t,x),k}, Z^{(t,x),k})$ defined by $Y^{(t,x),0} = 0$, $Z^{(t,x),0} = 0$, and

$$-dY_s^{k+1} = f(s, P_s^{t,x}, Y_s^k, Z_s^k)\,ds - (Z_s^{k+1})^*\,dW_s, \qquad Y_T^{k+1} = \Psi(P_T^{t,x}).$$

We know from Corollary 2.1 that the sequence $(Y^{(t,x),k}, Z^{(t,x),k})$ converges $d\mathbb{P}\otimes ds$ a.s. to $(Y^{t,x}, Z^{t,x})$ the unique square-integrable solution of the BSDE; for $Y^{t,x}$ we also have that $\sup_{s\in[t,T]} |Y_s^{(t,x),k} - Y_s^{(t,x)}|$ converges $\mathbb{P}$ a.s. to 0. By Lemma 4.1, the theorem holds for parameters (f, ξ) depending only on $(s, P_s^{t,x})$. By applying Lemma 4.1, we conclude by recursion that there exists some $\mathcal{B}_e$-measurable functions u_k, d_k such that

$$Y_s^{(t,x),k} = u_k(s, P_s^{t,x}), \qquad Z_s^{(t,x),k} = \sigma(s, P_s^{t,x})^* d_k(s, P_s^{t,x}).$$

Put

$$u^i(s,x) = \limsup_{k\to+\infty} u_k^i(s,x), \qquad d^{i,j}(s,x) = \limsup_{k\to+\infty} d_k^{i,j}(s,x),$$

where $u = (u^i)_{1\le i\le d}$ and $d = (d^{i,j})_{1\le i\le p, 1\le j\le d}$. Notice that from the a.s. convergence of the sequence $(Y^{(t,x),k}, Z^{(t,x),k})$ to $(Y^{t,x}, Z^{t,x})$, it follows that $\mathbb{P}$ a.s., $\forall s \in [t,T]$,

$$u^i(s, P_s^{t,x}) = (\limsup_{k\to\infty} u_k^i)(s, P_s^{t,x}) = \limsup_{k\to\infty} (u_k^i(s, P_s^{t,x})) = \lim_{k\to\infty} Y_s^{i,(t,x),k} = Y_s^{i,(t,x)}.$$

The same properties hold for d, and we derive that $d(s, P_s^{t,x}) = Z_s^{(t,x)}$, $d\mathbb{P} \otimes ds$ a.s. $\square$

By Proposition 4.1, it follows that if the coefficients are supposed to be regular, then the function u satisfies some additional regularity properties. In particular, if the coefficients are differentiable, recall that the solution is differentiable in the usual sense and in Malliavin's sense and that Z_s is given (almost surely) by the Malliavin derivative $D_s Y_s$ (see Section 5.1 for a precise definition of $D_s Y_s$). In this case the function d can be written as a function of $\partial_x u$ and σ.

COROLLARY 4.1. *We suppose that b, σ, f, and Ψ are globally Lipschitz with respect to (x, y, z), uniformly in t concerning f. Then u is locally Lipschitz in x and $1/2$-Hölder continuous in t. Furthermore, if b, σ, f, and Ψ are continuously differentiable with respect to (x, y, z) with uniformly bounded derivatives, then for $0 \leq t \leq s \leq T$, $x \in \mathbb{R}^p$,*

$$Z_s^{t,x} = \sigma(s, P_s^{t,x})^* \partial_x u(s, P_s^{t,x})\, d\mathbb{P} \otimes ds \ a.s.$$

Proof. The first statement is a direct consequence of (4.3) and (4.4). Let us show the second one. Recall that $Y^{t,x}$ is differentiable with respect to x in $\mathbb{H}^{2,d}_{T,\beta}$ and in S^2; hence, u is differentiable with respect to x. Since $Y_s^{t,x} = u(s, P_s^{t,x})$, it follows by the chain rule (see Nualart 1986, p. 90; 1995, Proposition 1.2.2) that $D_s Y_s^{t,x} = D_s P_s^{t,x}\, \partial_x u(s, P_s^{t,x})$. Then, using the fact that $Z_s^{t,x} = D_s Y_s^{t,x}$ and $D_s P_s^{t,x} = \sigma(s, P_s^{t,x})^*$ almost surely, the result follows. $\square$

BSDE and Partial Differential Equations. In this section we study the relation between these forward-backward equations and partial differential equations (PDE). We first give a generalization of the Feynman-Kac formula stated by Pardoux and Peng (1992). Then we show that, conversely, under smoothness conditions the function $u(t, x) = Y_t^{t,x}$ is a solution in some sense of a PDE.

PROPOSITION 4.3 (Generalization of the Feynman-Kac Formula). *Let v be a function of class $\mathcal{C}^{1,2}$ (or smooth enough to be able to apply Itô's formula to $v(s, P_s^{t,x})$) and suppose that there exists a constant C such that, for each (s, x),*

$$|v(s, x)| + |\sigma(s, x)^* \partial_x v(s, x)| \leq C(1 + |x|).$$

Also, v is supposed to be the solution of the following quasilinear parabolic partial differential equation:

$$\begin{aligned} &\partial_t v(t, x) + \mathcal{L} v(t, x) + f(t, x, v(t, x), \sigma(t, x)^* \partial_x v(t, x)) = 0, \\ &v(T, x) = \Psi(x), \end{aligned} \tag{4.9}$$

where $\partial_x v$ is the gradiant of v and $\mathcal{L}_{(t,x)}$ denotes the second-order differential operator

$$\mathcal{L}_{(t,x)} = \sum_{i,j} a_{ij}(t, x) \partial^2_{x_i x_j} + \sum_i b_i(t, x) \partial_{x_i}, \qquad a_{ij} = \tfrac{1}{2}[\sigma\sigma^*]_{ij}.$$

Then $v(t,x) = Y_t^{(t,x)}$, where $\{(Y_s^{t,x}, Z_s^{t,x}), t \leq s \leq T)\}$ is the unique solution of BSDE (4.2). Also, $(Y_s^{t,x}, Z_s^{t,x}) = \left(v(s, P_s^{t,x}), \sigma(t, P_s^{t,x})^* \partial_x v(t, P_s^{t,x})\right), t \leq s \leq T$.

Proof. By applying Itô's formula to $v(s, P_s^{t,x})$ we have

$$d\, v(s, P_s^{t,x}) = \left(\partial_t v(s, P_s^{t,x}) + \mathcal{L} v(s, P_s^{t,x})\right) ds + \partial_x v(s, P_s^{t,x})^* \sigma(s, P_s^{t,x})\, dW_s.$$

Since v solves (4.9), it follows that

$$\begin{aligned} -d\, v(s, P_s^{t,x}) &= f(s, P_s^{t,x}, v(s, P_s^{t,x}), \sigma(s, P_s^{t,x})^* \partial_x v(s, P_s^{t,x}))\, ds \\ &\quad - \partial_x v(s, P_s^{t,x})^* \sigma(s, P_s^{t,x})\, dW_s \end{aligned}$$

with $v(T, P_T^{t,x}) = \Psi(P_T^{t,x})$. Thus, $\{v(s, P_s^{t,x}), \sigma(s, P_s^{t,x})^* \partial_x v(s, P_s^{t,x}),\ s \in [0,T]\}$ is equal to the unique solution of BSDE (4.2), and the result is obtained. □

REMARK. Ma, Protter, and Yong (1994) use this point of view to study some more general FBSDEs of the type

$$\text{(4.10)} \quad \begin{aligned} &dP_s^{t,x} = b(s, P_s^{t,x}, Y_s^{t,x}, Z_s^{t,x})\, ds + \sigma(s, P_s^{t,x}, Y_s^{t,x})\, dW_s, \qquad t \leq s \leq T, \\ &P_s^{t,x} = x, \qquad 0 \leq s \leq t, \end{aligned}$$

$$\text{(4.11)} \quad \begin{aligned} -dY_s^{t,x} &= f(s, P_s^{t,x}, Y_s^{t,x}, Z_s^{t,x})\, ds - (Z_s^{t,x})^*\, dW_s \\ Y_T^{t,x} &= \Psi(P_T^{t,x}). \end{aligned}$$

Their motivation is to prove the existence and the uniqueness of an adapted solution (X, Y, Z) of this system. Antonelli (1993) had already proved the existence of such a solution, using a fixed-point theorem in the case where b does not depend on Z and only under the assumption $CT < 1$, where C is the Lipschitz constant of f (it means that there exists a solution on a small interval only). He also gave some examples for which this condition is not satisfied and there is no solution of the FBEs. Thus, Ma et al. (1994) are concerned with showing the existence of an adapted solution (X, Y, Z) of (4.10), (4.11) without the assumption $CT < 1$. Their method is the following: they know by analysis results that, under some strong assumptions on the coefficients, there exists a classical solution of the associated PDE. Using this solution, they state, by a verification method, the existence and uniqueness of the system of forward-backward equations.

Their method was used in the mathematical finance setting in a recent preprint, "Hedging options for a large investor and forward-backward SDE's," by Cvitanic and Ma (1994) and in "Black's consol rate conjecture," by Duffie, Ma, and Yong (1994).

We now show that, conversely, in certain cases the solution of the BSDE (4.2) corresponds to the solution of the PDE (4.3). If $d = 1$, we can use the comparison theorem to show that if b, σ, f, and Ψ satisfy the assumptions at the beginning of the section and if f and Ψ are supposed to be uniformly continuous with respect to x, then $u(t,x)$ is a viscosity solution of (4.9) (Peng 1992b; Pardoux and Peng 1992).

THEOREM 4.2. *We suppose that $d = 1$ and that f and Ψ are uniformly continuous with respect to x. Then the function u defined by $u(t,x) = Y_t^{t,x}$ is a viscosity solution of PDE (4.9).*

Furthermore, if we suppose that for each $R > 0$ there exists a continuous function $m_R\colon \mathbb{R}_+ \to \mathbb{R}_+$ such that $m_R(0) = 0$ and

$$|f(t,x,y,z) - f(t,x',y,z)| \leq m_R(|x-x'|(1+|z|)), \tag{4.12}$$

for all $t \in [0,T]$, $|x|$, $|x'| \leq R$, and $|z| \leq R$ for $z \in \mathbb{R}^n$, then u is the unique viscosity solution of PDE (4.9).

Before giving the proof, recall the definition of a viscosity solution (Fleming and Soner 1993).

DEFINITION 4.1. Suppose $u \in \mathcal{C}([0,T] \times \mathbb{R}^p)$ satisfies $u(T,x) = \Psi(x)$, $x \in \mathbb{R}^p$. Then u is called a *viscosity subsolution* (resp. *supersolution*) of PDE (4.9) if, for each $(t,x) \in [0,T] \times \mathbb{R}^p$ and $\phi \in \mathcal{C}^{1,2}([0,T] \times \mathbb{R}^p)$ such that $\phi(t,x) = u(t,x)$ and (t,x) is a minimum (resp. maximum) of $\phi - u$,

$$\partial_t\phi(t,x) + \mathcal{L}\phi(t,x) + f(t,x,\phi(t,x),\sigma(t,x)^*\partial_x\phi(t,x)) \geq 0$$

(resp.

$$\partial_t\phi(t,x) + \mathcal{L}\phi(t,x) + f(t,x,\phi(t,x),\sigma(t,x)^*\partial_x\phi(t,x)) \leq 0).$$

Moreover, u is called a *viscosity solution* of PDE (4.9) if it is both a viscosity subsolution and a viscosity supersolution of PDE (4.9).

Proof of Theorem 4.2. The continuity of the function u with respect to (t,x) follows from Corollary 4.1. Now we show that u is a viscosity subsolution of (4.9) (the proof is the same if u is a supersolution). Let $(t,x) \in [0,T] \times \mathbb{R}^p$ and $\phi \in \mathcal{C}^{1,2}([0,T] \times \mathbb{R}^p)$ be such that $\phi(t,x) = u(t,x)$ and $\phi \geq u$ on $[0,T] \times \mathbb{R}^p$. We can suppose without loss of generality that ϕ is C^∞ and has bounded derivatives.[13]

For $h \geq 0$ we have $\phi(t+h, P_{t+h}^{t,x}) \geq u(t+h, P_{t+h}^{t,x}) = Y_{t+h}^{t,x}$, so one could think of letting h tend to 0 in the inequality

$$\phi(t+h, P_{t+h}^{t,x}) - \phi(t,x) - \int_t^{t+h} f(s, P_s^{t,x}, Y_s^{t,x}, Z_s^{t,x})\, ds + \int_t^{t+h} Z_s^{t,x}\, dW_s \geq 0.$$

But we do not know if the process $Z_s^{t,x}$ converges to $\sigma(t,x)^*\partial_x\phi(t,x)$.

[13]Indeed, if $\phi \in \mathcal{C}^{1,2}([0,T] \times \mathbb{R}^p)$ such that $\phi(t,x) = u(t,x)$ and (t,x) is a minimum of $\phi - u$, then it is possible to construct a sequence of functions $\phi_n \in \mathcal{C}^\infty$ with bounded derivatives such that ϕ_n (respectively its first and second derivatives) converges to ϕ (respectively its first and second derivatives) as n tends to infinity, uniformly on compacts.

Now let $\{(\overline{Y}_s, \overline{Z}_s), t \le s \le t+h\}$ be the solution of the BSDE

$$\overline{Y}_s = \phi(t+h, P^{t,x}_{t+h}) + \int_s^{t+h} f(r, P^{t,x}_r, \overline{Y}_r, \overline{Z}_r)\, dr - \int_s^{t+h} \overline{Z}_r\, dW_r\,.$$

Note that $(\overline{Y}, \overline{Z})$ has the same generator as (Y, Z), but the terminal condition is given by $\phi(t+h, P^{t,x}_{t+h})$ (which is greater than $Y_{t+h} = u(t+h, P^{t,x}_{t+h})$). By the comparison theorem and continuity of the processes, it follows that $\overline{Y}_t \ge Y^{t,x}_t = u(t,x) = \phi(t,x)$. Then we have to show that, by letting h tend to 0, $\overline{Y}_s$ tends to $\phi(t,x)$ and $\overline{Z}_s$ tends to $\partial_x\phi(t,x)^*\sigma(t,x)$. Actually, a development of $\overline{Y}$ until the first order suffices to obtain the result.

First, put $G(s,x) = \partial_s\phi(s,x) + \mathcal{L}\phi(s,x) + f(s,x,\phi(s,x), \partial_x\phi(s,x)\sigma(s,x))$ for $t \le s \le t+h$. Notice that we want to show that $G(t,x) \ge 0$. Now put $\tilde{Y}_s = \overline{Y}_s - \phi(s, P^{t,x}_s) - \int_s^{t+h} G(r,x)\, dr$, $\tilde{Z}_s = \overline{Z}_s - \partial_x\phi\sigma(s, P^{t,x}_s)$. We show that $\tilde{Y}_t = h\epsilon(h)$, where $\epsilon(h) \to 0$ as $h \to 0$. By Itô's formula, $(\tilde{Y}_s, \tilde{Z}_s)$, $t \le s \le t+h$, is the unique solution of the BSDE

$$\begin{aligned}(4.13)\quad \tilde{Y}_s = &\int_s^{t+h} f\Big(r, P^{t,x}_r,\ \phi(r, P^{t,x}_r) + \tilde{Y}_r + \int_r^{t+h} G(v,x)\, dv,\ \partial_x\phi\sigma(r, P^{t,x}_r) + \tilde{Z}_r\Big)\, dr \\ &+ \int_s^{t+h} [(\partial_r\phi + \mathcal{L}\phi)(r, P^{t,x}_r) - G(r,x)]\, dr - \int_s^{t+h} \tilde{Z}_r\, dW_r\,.\end{aligned}$$

We first show that $(\tilde{Y}, \tilde{Z})$ tends to $(0,0)$ as h goes to 0. By the a priori estimates applied to $(Y^1, Z^1) = (\tilde{Y}, \tilde{Z})$ and $(Y^2, Z^2) = (0,0)$, it follows that

$$\mathbb{E}\left[\sup_{t\le s\le t+h} |\tilde{Y}_s|^2\right] + E\left[\int_t^{t+h} |\tilde{Z}_s|^2\, ds\right] \le K\mathbb{E}\left[\int_t^{t+h} |\delta(r,h)|^2\, dr\right],$$

where

$$\begin{aligned}\delta(r,h) = &-G(r,x) + (\partial_r\phi + \mathcal{L}\phi)(r, P^{t,x}_r) \\ &+ f\left(r, P^{t,x}_r,\ \phi(r, P^{t,x}_r) + \int_r^{t+h} G(v,x)\, dv,\ \sigma(r, P^{t,x}_r)^*\partial_x\phi(r, P^{t,x}_r)\right).\end{aligned}$$

Now since $\sup_{t\le s\le t+h} E(|P^{t,x}_s - x|^2) \to 0$ as $h \to 0$, and since all the coefficients as well as ϕ and its derivatives are uniformly continuous with respect to x, it follows that

$$\lim_{h\to 0} \sup_{t\le r\le t+h} \mathbb{E}[|\delta(r,h)|^2] = 0\,.$$

Hence we obtain

$$(4.14)\quad \mathbb{E}\left[\sup_{t\le s\le t+h} |\tilde{Y}_s|^2\right] + E\left[\int_t^{t+h} |\tilde{Z}_s|^2\, ds\right] \le K\mathbb{E}\left[\int_t^{t+h} |\delta(r,h)|^2\, dr\right] \le h\epsilon(h)\,,$$

where $\epsilon(h) \to 0$ as $h \to 0$. Consequently, we only have $\mathbb{E}[\int_t^{t+h}(|\tilde{Y}_s|+|\tilde{Z}_s|)\,ds] = h\sqrt{\epsilon}(h)$ (by the Cauchy-Schwartz inequality), and this estimate is not sufficient for $\tilde{Y}$; we need to have $\tilde{Y}_t = h\epsilon(h)$.

Note that by taking the expectation in (4.13), $\tilde{Y}_t = E(\tilde{Y}_t) = \mathbb{E}[\int_t^{t+h} \delta'(r,h)\,dr]$, where

$$\begin{aligned}\delta'(r,h) \;=\; & -G(r,x) + (\partial_r\phi + \mathcal{L}\phi)(r, P_r^{t,x}) \\ & + f(r, P_r^{t,x},\, \phi(r, P_r^{t,x}) + \tilde{Y}_r + \int_r^{t+h} G(v,x)\,dv\,,\, \sigma^*(r, P_r^{t,x})\partial_x\phi(r, P_r^{t,x}) + \tilde{Z}_r).\end{aligned}$$

Since f is Lipschitz, $|\delta'(r,h) - \delta(r,h)| \le K(|\tilde{Y}_r|+|\tilde{Z}_r|)$, and by (4.14), $\tilde{Y}_t = h\epsilon(h)$, where $\epsilon(h) \to 0$ as $h \to 0$. Hence, since $\overline{Y}_t \ge \phi(t,x)$, we have $\int_t^{t+h} G(r,x)\,dr \ge -h\epsilon(h)$, so

$$\frac{1}{h}\int_t^{t+h} G(r,x)\,dr \ge -\epsilon(h)\,.$$

Then by letting h tend to 0 we obtain

$$G(t,x) = \partial_t\phi(t,x) + \mathcal{L}\phi(t,x) + f(t,x,\phi(t,x), \partial_x\phi(t,x)\sigma(t,x)) \;\ge 0\,.$$

Hence, u is a viscosity solution of (4.9).

It remains to show the second statement of Theorem 4.2. Suppose that (4.12) is satisfied. Then, by the uniqueness result of Ishii and Lions (1990), (4.9) has at most one viscosity solution. The result follows. □

When $d \ge 1$, Pardoux and Peng (1992) gave the following result.

PROPOSITION 4.4. *All the functions b, σ, f, and g are assumed to be $\mathcal{C}^3$ with bounded derivatives. Then $u(t,x) = Y_t^{t,x}$ belongs to $\mathcal{C}^{1,2}\left([0,T]\times\mathbb{R}^p, \mathbb{R}^d\right)$ and it solves PDE (4.9).*

Sketch of the proof. Refer to Pardoux and Peng (1992). By Proposition 4.1, u belongs to $\mathcal{C}^{0,1}([0,T]\times\mathbb{R}^p,\mathbb{R}^d)$. The proof that u is $\mathcal{C}^2$ with respect to x needs some estimates of $\sup_s |Z_s|^2$ which can be given by the properties of the Malliavin derivative (Pardoux and Peng 1992). Furthermore,

$$u(t+h,x) - u(t,x) = [u(t+h,x) - u(t+h, P_{t+h}^{t,x})] \,+\, [u(t+h, P_{t+h}^{t,x}) - u(t,x)].$$

The second term on the right side, $u(t+h, P_{t+h}^{t,x}) - u(t,x)$, is equal to $Y_{t+h}^{t,x} - Y_t^{t,x}$, since, by Theorem 4.1, $Y_{t+h}^{t,x} = u(t+h, P_{t+h}^{t,x})$. Then, by applying Itô's formula between $s = t$

and $s = t + h$ to $u(t+h, P_s^{t,x})$, it follows that

$$u(t+h, x) - u(t, x) = -\int_t^{t+h} \mathcal{L}u(t+h, P_s^{s,x})\, ds - \int_t^{t+h} (\partial_x u)^* \sigma(s, P_s^{t,x})\, dW_s - \int_t^{t+h} f(s, P_s^{t,x}, Y_s^{t,x}, Z_s^{t,x})\, ds + \int_t^{t+h} (Z_s^{t,x})^* dW_s\,.$$

By Corollary 4.1 we have $(Z_s^{t,x})^* = (\partial_x u)^* \sigma(s, P_s^{t,x})$. Then by dividing by h and letting h tend to 0, it follows that u is differentiable with respect to t and that u is a regular solution of (4.9). □

4.2. Application to European Option Pricing in the Constrained and Markovian Cases

In this section we give a simple application to finance which shows that Markovian BSDEs are a useful tool in pricing theory since they give a generalization of the Black-Sholes formula, in the sense where the price of a contingent claim which only depends on the prices of the basic securities has the same property. Also the hedging portfolio depends only on these prices.

Consider a financial market model with coefficients which only depend on time s and on the vector of stock price process P_s. Fix $(t, x) \in [0, T] \times \mathbb{R}^{n+1}$. Here, the prices of the basic securities satisfy the following equations on $[t, T]$:

$$dP_s^0 = r(s, P_s)\, P_s^0\, ds, \tag{4.15}$$

$$dP_s^i = P_s^i \left[\mu^i(s, P_s)\, dt + \sum_{j=1}^n \sigma_j^i(s, P_s)\, dW_s^j \right]. \tag{4.16}$$

Let $(P_s^{t,x}, t \leq s \leq T)$ be the vector of stock price processes: $P_s^{t,x} = (P_s^0, P_s^1, \ldots, P_s^n)$ with initial condition given by $P_t^{t,x} = x$.

In this context a general setting of the wealth equation is

$$-dX_s = b(s, P_s, X_s, \sigma(s, P_s)^* \pi_s)\, ds - \pi_s^* \sigma(s, P_s)\, dW_s. \tag{4.17}$$

Here b is an $\mathbb{R}$-valued continuous function defined on $[0, T] \times \mathbb{R}^{n+1} \times \mathbb{R} \times \mathbb{R}^n$ that is Lipschitz with respect to (x, π) uniformly in t. The classical case (Section 1.1) corresponds to

$$b(t, x, y, z) = -r(t, x)y - \theta^*(t, x)z\,,$$

where $\theta(t, x)$ is the risk premium vector: $\theta(t, x) = \sigma^{-1}(t, x)(\mu(t, x) - r(t, x)\mathbf{1})$. Consider a contingent claim $\xi = \phi(P_T^{t,x})$. Here, ϕ: $\mathbb{R}_+^{n+1} \to \mathbb{R}^+$ is Lipschitz. There exists a unique square-integrable hedging strategy $(X^{t,x}, \pi^{t,x}) \in \mathbb{H}_T^{2,d} \times \mathbb{H}_T^{2,n\times d}$ against ξ such that

$$\begin{aligned} -dX_s^{t,x} &= b(s, P_s^{t,x}, X_s^{t,x}, \sigma(s, P_s^{t,x})^* \pi_s^{t,x})\, ds - (\pi_s^{t,x})^* \sigma(s, P_s^{t,x})\, dW_s, \\ X_T^{t,x} &= \phi(P_T^{t,x}) \end{aligned} \tag{4.18}$$

and $X_s^{t,x}$ is the price of the contingent claim $\phi(P_T^{t,x})$ at time s. Then from the results of this section, the value at time s of the contingent claim ξ is

$$X_s^{t,x} = u(s, P_s^{t,x}),$$

where $u(t,x) = X_t^{t,x}$ is the unique viscosity solution of the nonlinear parabolic PDE

$$\begin{aligned}
&\frac{\partial u}{\partial t} + \sum_{i,j=1}^{n} a_{ij}(t,x) x_i x_j \frac{\partial^2 u}{\partial x_i \partial x_j}(t,x) \\
&\qquad + \sum_{i=1}^{n} \mu_i(t,x) x_i \frac{\partial u}{\partial x_i}(t,x) + r(t,x) x_0 \frac{\partial u}{\partial x_0}(t,x) \\
&\qquad = -b\left(t, x, u(t,x), \sigma^*(t,x)\left[x \tfrac{\partial u}{\partial x}\right]\right), \\
&u(T,x) \ = \phi(x),
\end{aligned} \tag{4.19}$$

where $a_{ij(t,x)} = \frac{1}{2}[\sigma\sigma^*]_{ij}(t,x)$ and $[x\frac{\partial u}{\partial x}] = (x_i \frac{\partial u}{\partial x_i}(t,x))$. Also if the function b is $\mathcal{C}^3$ with bounded derivatives, then u belongs to $\mathcal{C}^{1,2}([0,T] \times \mathbb{R}^{n+1}, \mathbb{R})$ and it is a regular solution of the PDE. Notice that the portfolio process of the hedging strategy is then

$$\pi_s^i = P_s^i \frac{\partial u}{\partial x_i}(s, P_s), \qquad t \le s \le T, \quad 1 \le i \le n.$$

5. ADDITIONAL RESULTS: GENERALIZED BSDES AND MALLIAVIN DERIVATIVES

5.1. $\mathbb{L}^p$ Solutions of BSDE and Extension of the Filtration

In this section we give some generalizations for the solutions of BSDEs. We relax the assumption that the underlying filtration is a Brownian filtration, and only suppose that $(\mathcal{F}_t)$ is a right-continuous complete filtration. Furthermore, we are interested in solving the BSDE under a p-integrability assumption of the parameters. The definition of a solution of BSDE must be extended in the following way. Consider the generalized stochastic backward differential equation (GBSDE)

$$-\,dY_t = f(t, Y_t, Z_t)\,dt \,-\, Z_t^*\,dW_t \,-\, dM_t\,, \qquad Y_T = \xi, \tag{5.1}$$

or, equivalently,

$$Y_t = \xi + \int_t^T f(s, Y_s, Z_s)\,ds \,-\, \int_t^T Z_s^*\,dW_s - \int_t^T dM_s\,, \tag{5.2}$$

where

- Y is an RCLL adapted process which takes values in $\mathbb{R}^d$.
- Z is a predictable process which takes values in $\mathbb{R}^{n\times d}$.

- M is an RCLL local martingale, $\mathbb{R}^d$-valued, orthogonal to the Brownian motion W, with $M_0 = 0$.

Suppose that ξ belongs to $\mathbb{L}^p(\mathbb{R}^d)$ (i.e., the set of $\mathcal{F}_T$-measurable random variable such that $||\xi||^p = \mathbb{E}(|\xi|^p) < +\infty$), $p > 1$, and that $f(t, 0, 0)$ belongs to $\mathbb{H}_T^p(\mathbb{R}^d)$ (i.e., $|| f(\cdot, 0, 0)||_{\mathbb{H}^p}^p = \mathbb{E}[(\int_0^T |f(s, 0, 0)|^2 \, ds|)^{p/2}] < +\infty$). If f is uniformly Lipschitz, then the parameters (f, ξ) are said to be p-standard. In what follows, we prove a result of existence and uniqueness for solutions (Y, Z) in $\mathbb{H}_T^p(\mathbb{R}^d) \times \mathbb{H}_T^p(\mathbb{R}^{n\times d})$. Buckdahn (1993) gives the most general result in this area (if $p \geq 2$), namely when the BSDE is driven by a general continuous martingale and a predictable increasing process.

Existence and uniqueness of the solution are shown by using a fixed-point theorem (as in Section 2), but instead of introducing a coefficient β the contraction is first obtained for a terminal time T sufficiently small; then for arbitrary T, the solution is obtained by subdividing the interval $[0, T]$. While the estimates in Section 2 are stated using Itô's formula and elementary algebraic calculus, the following estimates follow from martingale inequalities.

It is convenient to introduce the set $S_T^p(\mathbb{R}^d)$ of the RCLL adapted processes φ which take values in $\mathbb{R}^d$ and are such that $\|\varphi\|_{S^p}^p = \mathbb{E}[\sup_{0\leq t\leq T} |\varphi_t|^p] < +\infty$. Let us introduce $\mathcal{B}_T^p(\mathbb{R}^d, \mathbb{R}^{n\times d})$, the Banach space $S_T^p(\mathbb{R}^d) \times \mathbb{H}_T^p(\mathbb{R}^{n\times d})$ endowed with the norm $\|(Y, Z)\|_p^p = \|Y\|_{S^p}^p + \|Z\|_{\mathbb{H}_p}^p = \mathbb{E}[\sup_{0\leq t\leq T} |Y_t|^p] + \mathbb{E}[(\int_0^T |Z_t|^2 \, dt)^{p/2}]$. Note that this definition corresponds to the definition of the classical norm for semimartingales and coincides with the one of Buckdahn (1993).

THEOREM 5.1. *Fix $p > 1$ and suppose that (f, ξ) are p-standard parameters. There exist a unique pair $(Y, Z) \in \mathcal{B}_T^p$ and a unique martingale $M \in \mathbb{H}_T^p(\mathbb{R}^d)$, orthogonal to the Brownian motion, such that (Y, Z, M) solves (5.1).*

Proof. We first prove the result for T sufficiently small. Then the general case is obtained by subdividing the interval $[0, T]$ into a finite number of small intervals. As in the proof of Theorem 2.1, we use a fixed-point theorem for the mapping ϕ defined on $\mathcal{B}_T^p$ which maps (y, z) into the solution (Y, Z) of the BSDE associated with the generator $f(t, y_t, z_t)$. In other words, Y is the right-continuous version of the semimartingale $\mathbb{E}[\xi + \int_t^T f(s, y_s, z_s)\, ds|\mathcal{F}_t]$, and Z is given by the orthogonal decomposition with respect to Brownian motion for the martingale $\mathbb{E}[\xi + \int_0^T f(s, y_s, z_s)\, ds|\mathcal{F}_t]$; that is,

$$\mathbb{E}\left[\xi + \int_0^T f(s, y_s, z_s)\, ds|\mathcal{F}_t\right] = Y_0 + \int_0^t Z_s^* \, dW_s + M_t\,,$$

where M_t is an RCLL local martingale orthogonal to W.

Let us show that (Y, Z) belongs to $\mathcal{B}_T^p$ (that is, ϕ maps $\mathcal{B}_T^p$ onto itself) and that M is in $\mathbb{H}_T^p(\mathbb{R}^d)$. We have, for each t, $|Y_t| \leq \mathbb{E}[|\xi| + \int_0^T |f(s, y_s, z_s)|\, ds|\mathcal{F}_t]$, a.s. The martingale inequalities (Protter 1990, p. 174) give

$$\|Y\|_{S^p}^p \leq C_p \mathbb{E}\left[\left(|\xi| + \int_0^T |f(s, y_s, z_s)|\, ds\right)^p\right],$$

where C_p is a positive constant which depends only on the number p. Now the Cauchy-Schwartz inequality shows that

$$\int_0^T |f(s, y_s, z_s)|\, ds \leq T^{1/2} \left(\int_0^T |f(s, y_s, z_s)|^2\, ds \right)^{1/2} .$$

It follows that for another constant, still denoted by C_p,

$$\|Y\|^p_{S^p} \leq C_p \mathbb{E}\left[|\xi|^p + T^{p/2} \left(\int_0^T |f(s, y_s, z_s)|^2\, ds \right)^{p/2} \right]. \tag{5.3}$$

Since $f(\cdot, 0, 0)$ is p-integrable and f is Lipschitz with respect to y and z, it follows easily that Y belongs to S^p_T.

We now prove that Z belongs to $\mathbb{H}^p_T(\mathbb{R}^{n\times d})$ and M belongs to $\mathbb{H}^p_T(\mathbb{R}^d)$. By Burkholder-Davis-Gundy inequalities (Protter 1990, p. 174), since $p > 1$,

$$\mathbb{E}\left[\left(\int_0^T |Z_s|^2\, ds + [M]_T \right)^{p/2} \right] \leq C_p \mathbb{E}\left[\left(\left| \int_0^T Z_s\, dW_s + M_T \right| \right)^p \right],$$

where $[M]$ is the quadratic variation of the local martingale M. Since

$$M_T + \int_0^T Z_s^*\, dW_s = \xi + \int_0^T f(s, y_s, z_s)\, ds - Y_0,$$

it follows easily that

$$\mathbb{E}\left[\left(\int_0^T |Z_s|^2\, ds + [M]_T \right)^{p/2} \right] \leq C_p \mathbb{E}\left[|\xi|^p + \left(\int_0^T |f(s, y_s, z_s)|\, ds \right)^p + Y_0^p \right].$$

Using the above estimates on $\|Y\|^p_{S^p}$, we obtain

$$\|Z\|^p_{\mathbb{H}^p} \leq C_p \mathbb{E}\left[|\xi|^p + T^{p/2} \left(\int_0^T |f(s, y_s, z_s)|^2\, ds \right)^{p/2} \right], \tag{5.4}$$

$$\mathbb{E}([M]_T^{p/2}) \leq C_p \mathbb{E}\left[|\xi|^p + T^{p/2} \left(\int_0^T |f(s, y_s, z_s)|^2\, ds \right)^{p/2} \right]. \tag{5.5}$$

Hence, Z belongs to $\mathbb{H}^p_T(\mathbb{R}^{n\times d})$ and M is in $\mathbb{H}^p_T(\mathbb{R}^d)$.

It remains to show that, for a good choice of T, ϕ is a contraction. Let (y^1, z^1) and (y^2, z^2) be two elements of $\mathcal{B}^p_T$ and let (Y^1, Z^1, M^1) and (Y^2, Z^2, M^2) be the associated solutions. Since $(\delta Y, \delta Z, \delta M)$ is the solution of the BSDE associated with the generator

$f(t, y^1, z^1) - f(t, y^2, z^2)$ and with terminal condition equal to zero, the above inequalities give

$$\|\delta Y\|_{S^p}^p + \|\delta Z\|_{\mathbb{H}^p}^p \leq C_p T^{p/2} \mathbb{E}\left[\left(\int_0^T |f(s, y_s^1, z_s^1) - f(s, y_s^2, z_s^2)|^2\, ds\right)^{p/2}\right].$$

Then since f is Lipschitz with respect to (y, z) with constant C, it follows that

$$\|\delta Y\|_{S^p}^p + \|\delta Z\|_{\mathbb{H}^p}^p \leq C_p T^{p/2}\left(\mathbb{E}\left[\left(\int_0^T |\delta y_s|^2\, ds\right)^{p/2}\right] + \mathbb{E}\left[\left(\int_0^T |\delta z_s|^2\, ds\right)^{p/2}\right]\right).$$

Hence, for $T \leq 1$,

$$\|\delta Y\|_{S^p}^p + \|\delta Z\|_{\mathbb{H}^p}^p \leq C_p T^{p/2}(\|\delta y\|_{S^p}^p + \|\delta z\|_{\mathbb{H}^p}^p).$$

Choosing T so that $C_p T^{p/2} < 1$, we have that ϕ is a contraction and there exists a fixed point (Y, Z) such that (Y, Z, M) is the unique p-integrable solution of the BSDE. Here, by construction, the martingale M is given by the orthogonal decomposition with respect to the Brownian motion of the martingale

$$\mathbb{E}\left[\xi + \int_0^T f(s, Y_s, Z_s)\, ds | \mathcal{F}_t\right] = Y_0 + \int_0^t Z_s^*\, dW_s + M_t\,. \qquad \square$$

As in the case $p = 2$, we can state a priori estimates.

PROPOSITION 5.1. *Let $((f^i, \xi^i);\ i=1,2)$ be two p-standard parameters of a BSDE, and let $((Y^i, Z^i, M^i);\ i = 1, 2)$ be the associated solutions satisfying the conditions of Theorem 5.1. Let C be a Lipschitz constant for f^1. Put $\delta Y_t = Y_t^1 - Y_t^2$, $\delta Z_t = Z_t^1 - Z_t^2$, and $\delta_2 f_t = f^1(t, Y_t^2, Z_t^2) - f^2(t, Y_t^2, Z_t^2)$. Then for T small enough there exists a constant $C_{p,T}$ such that*

$$\|\delta Y\|_{S^p}^p + \|\delta Z\|_{\mathbb{H}^p}^p + \mathbb{E}([\delta M]_T^{p/2}) \leq C_{p,T}\mathbb{E}[(|\delta Y_T|^p) + (\textstyle\int_0^T |\delta_2 f_s|\, ds)^p],$$
$$\|\delta Y\|_{S^p}^p + \|\delta Z\|_p^p + \mathbb{E}([\delta M]_T^{p/2}) \leq C_{p,T}[\mathbb{E}(|\delta Y_T|^p) + T^{p/2}\|\delta_2 f\|_{\mathbb{H}^p}^p].$$

Proof. Consider (Y^1, Z^1, M^1) and (Y^2, Z^2, M^2), the two solutions associated with (f^1, ξ^1) and (f^2, ξ^2). Using the same arguments as in the proof of Theorem 5.1 concerning $(\delta Y, \delta Z, \delta M)$, it follows easily that, for $T > 0$,

$$\begin{aligned}&\|\delta Y\|_{S^p}^p + \|\delta Z\|_{\mathbb{H}^p}^p + \mathbb{E}([\delta M]_T^{p/2})\\ &\quad \leq C_p \mathbb{E}\left[(|\delta Y_T|^p) + \left(\int_0^T |f^1(s, Y_s^1, Z_s^1) - f^1(s, Y_s^2, Z_s^2)|\, ds\right)^p\right].\end{aligned}$$

Since $|f^1(s, Y_s^1, Z_s^1) - f^2(s, Y_s^2, Z_s^2)| \leq C[|\delta Y_s| + |\delta Z_s|] + |\delta_2 f_s|$, it follows that for another constant, still denoted by C_p,

$$\begin{aligned} &\|\delta Y\|_{S^p}^p + \|\delta Z\|_{\mathbb{H}^p}^p \\ &\quad \leq C_p \left(\mathbb{E}[|\delta Y_T|^p + \left(\int_0^T |\delta_2 f_s|\, ds \right)^p] + T^p \|\delta Y\|_{S^p}^p + T^{p/2} \|\delta Z\|_p^p \right). \end{aligned}$$

Choosing T so that $\max(C_p T^p, C_p T^{p/2}) < 1$, we obtain

$$\|\delta Y\|_{S^p}^p + \|\delta Z\|_p^p \leq C_{p,T} \mathbb{E}\left[(|\delta Y_T|^p) + \left(\int_0^T |\delta_2 f_s|\, ds \right)^p \right]$$

for a positive constant $C_{p,T}$. The first inequality of Proposition 5.1 follows easily (for another constant $C_{p,T}$). The second inequality follows from the Cauchy-Schwartz inequality. □

EXAMPLE. Application in Finance: Föllmer-Schweizer Decomposition in Incomplete Markets. We come back to the situation of the incomplete market as described in Section 1.3, Example 1.3, where only some primary securities may be traded, and consider additional constraints on the portfolio. Recall that σ^1 denotes the volatility matrix of the j traded securities. We assume that the matrix $(\sigma_t^1(\sigma_t^1)^*)^{-1}\sigma_t^1$ is bounded. A nonadjusted hedging strategy $(V, {}^1\pi)$ of a contingent claim $\xi \in \mathbb{L}^p(\mathbb{R}^d)$ satisfies

$$(5.6) \qquad -dV_t = b(t, V_t, (\sigma_t^1)^*({}^1\pi_t))\, dt - ({}^1\pi_t)^* \sigma_t^1\, dW_t - d\Phi_t, \qquad V_T = \xi\,,$$

where Φ is a semimartingale in $\mathbb{H}_T^p(\mathbb{R}^d)$ called the cost process. Notice that the portfolio ${}^1\pi$ is a j-dimensional process. The Föllmer-Schweizer strategy is related to the situation where Φ is a martingale orthogonal to $\int_0^{\cdot} \sigma_t^1\, dW_t$. It is easy to construct this FS-strategy as a solution of a GBSDE.

PROPOSITION 5.2. *Let $(b(t, x, z), \xi)$ be p-standard parameters and (X, Z, M) the associated p-integrable solution of*

$$(5.7) \qquad -dX_t = b(t, X_t, \Sigma_t^1 Z_t)\, dt - Z_t^*\, dW_t - dM_t, \qquad X_T = \xi\,,$$

where $\Sigma_t^1 = (\sigma_t^1)^*(\sigma_t^1(\sigma_t^1)^*)^{-1}\sigma_t^1$. Then $(X, {}^1\pi, \Phi)$, where ${}^1\pi_t = (\sigma_t^1(\sigma_t^1)^*)^{-1}\sigma_t^1 Z_t$ and $d\Phi_t = ((\mathrm{Id} - \Sigma_t^1)Z_t)^*\, dW_t + dM_t$, is the Föllmer-Schweizer strategy.

REMARK. Example 1.3 corresponds to a Brownian filtration with

$$b(t, x, z) = -r_t x - (\theta_t^1)^* z.$$

5.2. Differentiation on Wiener Space of BSDE Solutions

We study in detail the properties of differentiation on Wiener space of the solution of a BSDE in the spirit of the work of Pardoux and Peng (1992). We state in a general framework that the Malliavin derivative of the solution of BSDE is still a solution of a linear BSDE. Applying these results to finance, we show in particular that the portfolio process of a hedging strategy corresponds to the Malliavin derivative of the price process. This important property was first emphasized by Karatzas and Ocone(1992) (see also Colwell, Elliott, and Kopp 1991) in the nonconstrained case (i.e., the linear case).

Malliavin Derivative of Solution of BSDE. First, recall briefly the notion of differentiation on Wiener space (see the expository papers by Nualart 1995, Nualart and Pardoux 1988, Ikeda and Watanabe 1989, and Ocone 1988).

- $C_b^k(\mathbb{R}^k, \mathbb{R}^q)$ will denote the set of functions of class C^k from $\mathbb{R}^k$ into $\mathbb{R}^q$ whose partial derivatives of order less than or equal to k are bounded.
- Let $\mathcal{S}$ denote the set of random variables ξ of the form $\xi = \varphi(W(h^1), \ldots, W(h^k))$, where $\varphi \in C_b^\infty(\mathbb{R}^k, \mathbb{R})$, $h^1, \ldots, h^k \in \mathbb{L}^2([0, T]; \mathbb{R}^n)$, and $W(h^i) = \int_0^T \langle h_s^i, dW_s \rangle$.
- If $\xi \in \mathcal{S}$ is of the above form, we define its derivative as being the n-dimensional process

$$D_\theta \xi = \sum_{j=1}^k \frac{\partial \varphi}{\partial x_j}(W(h^1), \ldots, W(h^k))h_\theta^j, \qquad 0 \le \theta \le T .$$

For $\xi \in \mathcal{S}$, $p > 1$, we define the norm

$$\|\xi\|_{1,p} = \left[E\left\{ |\xi|^p + \left(\int_0^T |D_\theta \xi|^2 \, d\theta \right)^{p/2} \right\} \right]^{1/p} .$$

It can be shown (Nualart 1995) that the operator D has a closed extension to the space $\mathbb{D}^{1,p}$, the closure of $\mathcal{S}$ with respect to the norm $\|\cdot\|_{1,p}$. Observe that if ξ is $\mathcal{F}_t$-measurable, then $D_\theta \xi = 0$ for $\theta \in (t, T]$. We denote by $D_\theta^i \xi$, $1 \le i \le n$, the ith component of $D_\theta \xi$.

Let $\mathbb{L}_{1,p}^a(\mathbb{R}^d)$ denote the set of $\mathbb{R}^d$-valued progressively measurable processes $\{u(t,\omega), 0 \le t \le T; \omega \in \Omega\}$ such that

(i) For a.e. $t \in [0, T]$, $u(t, \cdot) \in (D_{1,p})^d$.
(ii) $(t, \omega) \to Du(t, \omega) \in (L^2([0, T]))^{n \times d}$ admits a progressively measurable version.
(iii) $\|u\|_{1,p}^a = \mathbb{E}[(\int_0^T |u(t)|^2 \, dt)^{p/2} + (\int_0^T \int_0^T |D_\theta u(t)|^2 d\theta dt)^{p/2}] < +\infty$.

Observe that for each (θ, t, ω), $D_\theta u(t, \omega)$ is an $n \times d$ matrix. Thus, $|D_\theta u(t)|^2 = \sum_{i,j} |D_\theta^i u_j(t)|^2$. Clearly $D_\theta u(t, \omega)$ is defined uniquely up to sets of $d\theta \otimes dt \otimes dP$ measure zero. Put $\|Du\|^2 = \int_0^T \int_0^T |D_\theta u(t)|^2 \, d\theta \, dt$. With this notation notice that Jensen's inequality gives

$$\mathbb{E}(\|Du\|^2)^{p/2} \le T^{p/2-1} \int_0^T \|D_\theta u\|_p^p \, d\theta, \qquad p \ge 2, \tag{5.8}$$

$$\mathbb{E}(\|Du\|^2)^{p/2} \geq T^{p/2-1}\int_0^T \|D_\theta u\|_p^p \, d\theta, \qquad p \leq 2. \tag{5.9}$$

We now show that under natural conditions the solution of a BSDE is differentiable in Malliavin's sense and that the derivative is a solution of a linear BSDE. This result generalizes the one stated by Pardoux and Peng (1992) in the Markovian case, and we give a complete proof of it.

PROPOSITION 5.3. *Suppose that $\xi \in \mathbb{D}^{1,2}$ and $f\colon \Omega \times [0,T] \times \mathbb{R}^d \times \mathbb{R}^{n\times d} \to \mathbb{R}^d$ is continuously differentiable in (y,z), with uniformly bounded and continuous derivatives and such that, for each (y,z), $f(\cdot,y,z)$ is in $\mathbb{L}^a_{1,2}(\mathbb{R}^d)$ with Malliavin derivative denoted by $D_\theta f(t,y,z)$. Let (Y,Z) be the solution of the associated BSDE. Also, suppose that*

- *$f(t,0,0) \in \mathbb{H}^4_T(\mathbb{R}^d)$ and $\xi \in \mathbb{L}^4(\mathbb{R}^d)$.*
- *$\int_0^T \mathbb{E}(|D_\theta \xi|^2)\, d\theta < +\infty$, $\int_0^T \|D_\theta f(t,Y,Z)\|_2^2 \, d\theta < +\infty$, and for any $t \in [0,T]$ and any (y^1,z^1,y^2,z^2),*

$$|D_\theta f(t,\omega,y^1,z^1) - D_\theta f(t,\omega,y^2,z^2)| \leq K_\theta(t,\omega)(|y^1-y^2| + |z^1-z^2|),$$

where for a.e. θ, $\{K_\theta(t,\cdot), 0 \leq t \leq T\}$ is an $\mathbb{R}^+$-valued adapted process satisfying $\int_0^T \|K_\theta\|_4^4 \, d\theta < +\infty$.

Then $(Y,Z) \in \mathbb{L}^2(0,T\,;(\mathbb{D}^{1,2})^d \times (\mathbb{D}^{1,2})^{n\times d})$, and, for each $1 \leq i \leq n$, a version of $\{(D^i_\theta Y_t,\ D^i_\theta Z_t)\,;\ 0 \leq \theta, t \leq T\}$ is given by

$$\begin{aligned} D^i_\theta Y_t &= 0, \quad D^i_\theta Z_t = 0, \qquad 0 \leq t < \theta \leq T\,; \\ D^i_\theta Y_t &= D^i_\theta \xi + \int_t^T \left[\partial_y f(s,Y_s,Z_s) D^i_\theta Y_s + \partial_z f(s,Y_s,Z_s) D^i_\theta Z_s + D^i_\theta f(s,Y_s,Z_s)\right] ds \\ &\quad - \int_t^T D^i_\theta Z_s \, dW_s, \qquad \theta \leq t \leq T. \end{aligned}$$

Moreover, $\{D_t Y_t\,;\ 0 \leq t \leq T\}$ defined by (ii) is a version of $\{Z_t\,;\ 0 \leq t \leq T\}$.

REMARK. If K_θ is bounded, it is sufficient to suppose that $f(t,0,0) \in \mathbb{H}^2_T(\mathbb{R}^d)$ and $\xi \in \mathbb{L}^2_T(\mathbb{R}^d)$. Furthermore, the fact that $D_t Y_t = Z_t$ reveals the relation between the wealth process and the related portfolio. This result provides an efficient tool to estimate $E\left[\sup_t |Z_t|^p\right]$ for $p \geq 2$.

Before giving the proof of this proposition, let us recall the following lemma stated by Pardoux and Peng (1992), which shows that an Itô integral is differentiable in the Malliavin sense if and only if its integrand is so. For the proof, see Pardoux and Peng (1992) or Nualart (1995, Lemma 1.3.4), but this lemma is a consequence of the commutation relation between the derivative and the Skohorod integral (Nualart 1995, Section 1.3).

LEMMA 5.1. *Let $Z \in \mathbb{H}^2_T(\mathbb{R}^n)$ be such that $\xi = \int_t^T Z_s^* \, dW_s$ satisfies $\xi \in \mathbb{D}^{1,2}$. Then $Z^i \in \mathbb{L}^2(t, T, \mathbb{D}^{1,2})$, $1 \le i \le n$, and $d\theta \otimes d\mathbb{P}$ a.s.,*

$$D^i_\theta \xi = \int_t^T D^i_\theta Z_r \, dW_r, \qquad \theta \le t,$$

$$D^i_\theta \xi = Z^i_\theta + \int_\theta^T D^i_\theta Z_r \, dW_r, \qquad \theta > t.$$

Proof of Proposition 5.3. To simplify notation we restrict ourselves to the case $d = 1$. Let (Y^k, Z^k) be the Picard iterative sequence defined recursively by $Y^0 = 0$, $Z^0 = 0$ and

$$-dY^{k+1}_t = f(t, Y^k_t, Z^k_t)\, dt - (Z^{k+1}_t)^* \, dW_t, \qquad Y^{k+1}_T = \xi\ .$$

Using the contraction mapping defined in the proof of Theorem 5.1, we know that the sequence (Y^k, Z^k) converges in $S^4_T(\mathbb{R}) \otimes \mathbb{H}^4_T(\mathbb{R}^n)$ to (Y, Z) as $k \to +\infty$, the unique solution of the BSDE.

We recursively show that $(Y^k, Z^k) \in \mathbb{L}^2(0, T\ ;\ \mathbb{D}^{1,2} \times (\mathbb{D}^{1,2})^n)$. Suppose that $(Y^k, Z^k) \in \mathbb{L}^2(0, T\ ;\ \mathbb{D}^{1,2} \times (\mathbb{D}^{1,2})^n)$ and let us show that (Y^{k+1}, Z^{k+1}) is in $\mathbb{L}^2(0, T\ ;\ \mathbb{D}^{1,2} \times (\mathbb{D}^{1,2})^n)$.

Since $\xi + \int_t^T f(s, Y^k_s, Z^k_s)\, ds \in \mathbb{D}^{1,2}$, then $Y^{k+1}_t = \mathbb{E}[\xi + \int_t^T f(s, Y^k_s, Z^k_s)\, ds | \mathcal{F}_t] \in \mathbb{D}^{1,2}$. Now, $\xi + \int_t^T f(s, Y^k_s, Z^k_s)\, ds - Y^{k+1}_t = \int_t^T (Z^{k+1}_s)^* \, dW_s$. It follows from Lemma 5.1 that $Z^{k+1} \in \mathbb{L}^2(0, T, (\mathbb{D}^{1,2})^n)$, and for $0 \le \theta \le t$, $1 \le i \le n$,

$$\begin{aligned} -dD^i_\theta Y^{k+1}_t &= [\partial_y f(t, Y^k_t, Z^k_t) D^i_\theta Y^k_t + \partial_z f(t, Y^k_t, Z^k_t) D^i_\theta Z^k_t \ +\ D_\theta f(t, Y^k_t, Z^k_t)]\, dt \\ &\quad - (D^i_\theta Z^{k+1}_t)^* \, dW_t\ , \\ D^i_\theta Y^{k+1}_T &= D^i_\theta \xi\ . \end{aligned}$$

Hereafter, to simplify notation we assume that the Brownian is one-dimensional.

We will show that $(D_\theta Y^k, D_\theta Z^k)$ converges to (Y^θ, Z^θ) in $\mathbb{L}^a_{1,2}$ $(= \mathbb{L}^2(0, T\ ;\ \mathbb{D}^{1,2} \times \mathbb{D}^{1,2}))$, where $(Y^\theta_t, Z^\theta_t, \theta \le t \le T)$ is the solution of the BSDE.

$$\begin{aligned} (5.10)\quad -dY^\theta_t &= [\partial_y f(t, Y_t, Z_t) Y^\theta_t + \partial_z f(t, Y_t, Z_t) Z^\theta_t + D_\theta f(t, Y_t, Z_t)]\, dt - Z^\theta_t \, dW_t, \\ Y^\theta_T &= D_\theta \xi. \end{aligned}$$

First, we have that $\int_0^T (\|Y^\theta\|^2_{S^2} + \|Z^\theta\|^2_2)\, d\theta < +\infty$. Indeed, the a priori estimates applied to $Y^1 = Y^\theta$ and $Y^2 = 0$ give, for a constant C,

$$\|Y^\theta\|^2_{S^2} + \|Z^\theta\|^2_2 \le C\mathbb{E}\left(|D_\theta \xi|^2 + \|D_\theta f(\cdot, Y, Z)\|^2_2\right).$$

We now turn back to our problem. Using the a priori estimates, we obtain for almost all $\theta \in [0, T]$ that

$$\|D_\theta Y^{k+1} - Y^\theta\|_{S^2}^2 + \|D_\theta Z^{k+1} - Z^\theta\|_2^2 \leq C\mathbb{E}\left[\left(\int_\theta^T |\delta_s^k|\, ds\right)^2\right],$$

where

$$\begin{aligned}\delta_s^k \;=\; & D_\theta f(s, Y_s, Z_s) - D_\theta f(s, Y_s^k, Z_s^k) + \partial_y f(s, Y_s, Z_s)Y_s^\theta \\ & -\partial_y f(s, Y_s^k, Z_s^k) D_\theta Y_s^k + \partial_z f(s, Y_s, Z_s) Z_s^\theta \; - \partial_z f(s, Y_s^k, Z_s^k) D_\theta Z_s^k\,.\end{aligned}$$

Now $\mathbb{E}(\int_\theta^T |\delta_s^k|\, ds)^2 \leq C(A_k^\theta(T) + B_k^\theta(T) + C_k^\theta(T))$, where

$$\begin{aligned}A_k^\theta(T) \;=\; & \mathbb{E}\left(\int_\theta^T |D_\theta f(s, Y_s, Z_s) - D_\theta f(s, Y_s^k, Z_s^k)|\, ds\right)^2, \\ B_k^\theta(T) \;=\; & \mathbb{E}\left(\int_\theta^T |\partial_y f(s, Y_s^k, Z_s^k)(Y_s^\theta - D_\theta Y_s^k)|\, ds\right)^2 \\ & +\mathbb{E}\left(\int_\theta^T |\partial_z f(s, Y_s^k, Z_s^k)(Z_s^\theta - D_\theta Z_s^k)|\, ds\right)^2, \\ C_k^\theta(T) \;=\; & \mathbb{E}\left(\int_\theta^T |(\partial_y f(s, Y_s, Z_s) - \partial_y f(s, Y_s^k, Z_s^k))Y_s^\theta|\, ds\right)^2 \\ & +\mathbb{E}\left(\int_\theta^T |(\partial_z f(s, Y_s, Z_s) - \partial_z f(s, Y_s^k, Z_s^k))Z_s^\theta|\, ds\right)^2.\end{aligned}$$

Moreover, $A_k^\theta(T) \leq \mathbb{E}(\int_\theta^T |K_\theta(s)|(|Y_s - Y_s^k| + |Z_s - Z_s^k|)\, ds)^2$. By the Cauchy-Schwartz inequality,

$$\begin{aligned}& \mathbb{E}\left(\int_\theta^T |K_\theta(s)| \left(|Y_s - Y_s^k|\right)\, ds\right)^2 \\ & \quad \leq \left(\mathbb{E}\left(\int_\theta^T K_\theta(s)^2\, ds\right)^2\right)^{1/2} \left(\mathbb{E}\left(\int_\theta^T |Y_s - Y_s^k|^2\, ds\right)^2\right)^{1/2}.\end{aligned}$$

Hence, $A_k^\theta(T) \leq \|K_\theta\|_4^2 \left(\|Y - Y^k\|_4^2 + \|Z - Z^k\|_4^2\right)$.

Since (Y^k, Z^k) converges to (Y, Z) in $\mathcal{B}_4$, it follows that $\lim_{k\to+\infty}(\|Y - Y^k\|_4^2 + \|Z - Z^k\|_4^2) = 0$. Therefore,

$$\lim_{k\to+\infty} \int_0^T A_k^\theta(T)\, d\theta = 0.$$

Furthermore, since $\partial_y f$ and $\partial_z f$ are bounded and continuous with respect to y and z and since $\int_0^T (\|Y^\theta\|_{S^2}^2 + \|Z^\theta\|_2^2)\,d\theta < +\infty$, it follows by the Lebesgue theorem that $\lim_{k\to+\infty} \int_0^T C_k^\theta(T)\,d\theta = 0$.

Next, since the derivatives of f are bounded,

$$B_k^\theta(T) \le CT^2 \|D_\theta Y^k - Y^\theta\|_{S^2}^2 + CT\|D_\theta Z^k - Z^\theta\|_2^2.$$

Choose T so that $\alpha = \max(CT^2, CT) < 1$. Fix a positive real $\epsilon > 0$. There exists $N > 0$ such that, for any $k \ge N$,

$$\begin{aligned}\int_0^T (\|D_\theta Y^{k+1} - Y^\theta\|_{S^2}^2 + \|D_\theta Z^{k+1} - Z^\theta\|_2^2)\,d\theta \\ \le \epsilon + \alpha \int_0^T \left(\|D_\theta Y^k - Y^\theta\|_{S^2}^2 + \|D_\theta Z^k - Z^\theta\|_2^2\right)\,d\theta\,.\end{aligned}$$

Thus, we recursively obtain, for every $k \ge N$,

$$\begin{aligned}&\int_0^T (\|D_\theta Y^k - Y^\theta\|_{S^2}^2 + \|D_\theta Z^k - Z^\theta\|_2^2)\,d\theta \\ &\quad\le \frac{\epsilon}{1-\alpha} + \alpha^k \int_0^T \left(\|D_\theta Y^0 - Y^\theta\|_{S^2}^2 + \|D_\theta Z^0 - Z^\theta\|_2^2\right)\,d\theta \\ &\quad\le \frac{\epsilon}{1-\alpha} + \alpha^k K\,,\end{aligned}$$

where K is a positive constant. Hence, since $0 \le \alpha < 1$, it follows that the sequence $(D_\theta Y^k, D_\theta Z^k)$ converges in $\mathbb{L}^2(0, T\,; (\mathbb{D}^{1,2})^2) = \mathbb{L}_{1,2}^a$ to (Y^θ, Z^θ). Consequently, since $\mathbb{L}_{1,2}^a$ is closed for the norm $\|\cdot\|_{1,2}^a$, it follows that the limit (Y, Z) belongs to $\mathbb{L}_{1,2}^a$ and that a version of $(D_\theta Y,\ D_\theta Z)$ is given by (Y^θ, Z^θ).

It remains to show that for the considered version of the Malliavin derivatives of Y and Z, $D_s Y_s = Z_s$. Notice that for $t \le s$,

$$Y_s = Y_t - \int_t^s f(r, Y_r, Z_r)\,dr + \int_t^s Z_r\,dW_r\,.$$

It follows from Lemma 5.1 that, for $t < \theta \le s$,

$$\begin{aligned}D_\theta Y_s \;=\; Z_\theta - \int_\theta^s [\partial_y f(r, Y_r, Z_r) D_\theta Y_r + \partial_z f(r, Y_r, Z_r) D_\theta Z_r + D_\theta f(r, Y_r, Z_r)]\,dr \\ + \int_\theta^s D_\theta Z_r\,dW_r\,.\end{aligned}$$

Then, by taking $\theta = s$, it follows that $D_s Y_s = Z_s$ a.s. $\square$

REMARK. This result can be easily generalized from the case $p = 2$ to the case $p \geq 2$ (but not $1 < p < 2$) by using the same arguments and inequality (5.8).

Application to the Linear Case. The notation is the same as in the section on linear BSDEs ((2.8)). Let (β, γ) be a bounded $(\mathbb{R}, \mathbb{R}^n)$-valued predictable vector process, φ an element of $\mathbb{H}_T^2(\mathbb{R})$, and ξ an element of $\mathbb{L}_T^2(\mathbb{R})$. Then we consider the solution (Y, Z) of the BSDE

$$-\,dY_t = [\varphi_t + Y_t\,\beta_t + Z_t^*\gamma_t]\,dt - Z_t^*\,dW_t, \qquad Y_T = \xi\,. \tag{5.11}$$

From Proposition 5.3, we obtain:

PROPOSITION 5.4. *Suppose*

- $\beta, \gamma \in \mathbb{L}_{1,4}^a$, $\varphi \in \mathbb{H}_T^4 \cap \mathbb{L}_{1,2}^a$, *and* $\xi \in \mathbb{L}^4 \cap \mathbb{D}_{1,2}$.
- $\int_0^T \mathbb{E}(|D_\theta\xi|^2)\,d\theta < +\infty$, $\int_0^T \|D_\theta\varphi\|_2^2\,d\theta < +\infty$, *and* $\int_0^T (\|D_\theta\beta\|_4^4 + \|D_\theta\gamma\|_4^4)\,d\theta < +\infty$.

Then $(Y, Z) \in \mathbb{L}^2(0, T\,;\,(\mathbb{D}^{1,2})^d \times (\mathbb{D}^{1,2})^{n\times d})$, *and, for each* $1 \leq i \leq n$, *a version of* $\{(D_\theta^i Y_t,\ D_\theta^i Z_t)\,;\ 0 \leq \theta, t \leq T\}$ *is given by*

(i) $D_\theta^i Y_t = 0, \quad D_\theta^i Z_t = 0, \qquad 0 \leq t < \theta \leq T\,;$

(ii)
$$D_\theta^i Y_t = D_\theta^i\xi + \int_t^T \left[\beta_s D_\theta^i Y_s + \gamma_s D_\theta^i Z_s + D_\theta^i\varphi_s + Y_s D_\theta^i\beta_s + Z_s^* D_\theta^i\gamma_s\right] ds - \int_t^T D_\theta^i Z_s\,dW_s, \qquad \theta \leq t \leq T.$$

Moreover, $\{D_t Y_t\,;\ 0 \leq t \leq T\}$ *defined by* (ii) *is a version of* $\{Z_t\,;\ 0 \leq t \leq T\}$.

REMARK. If the coefficients β and γ are bounded deterministic functions, it is sufficient to suppose that $\varphi \in \mathbb{H}_T^2 \cap \mathbb{L}_{1,2}^a$ and $\xi \in \mathbb{L}^2 \cap \mathbb{D}_{1,2}$.

Recall that from Proposition 2.2, Y can be written

$$Y_t = \mathbb{E}\left[\Gamma_{t,T}\xi + \int_t^T \Gamma_{t,s}\varphi_s\,ds\,|\mathcal{F}_t\right], \tag{5.12}$$

where $(\Gamma_{t,s}, t \leq s \leq T)$ is the adjoint process defined by the forward LSDE,

$$d\Gamma_{t,s} = \Gamma_{t,s}[\beta_s\,ds + \gamma_s^*\,dW_s], \qquad \Gamma_{t,t} = 1.$$

Our aim is now to obtain a similar expression for Z_t and, more generally, since $Z_t = D_t Y_t$, an expression for $D_\theta Y_t$. Recall that it is possible to derive in the Malliavin sense a conditional expectation (see Nualart 1986, p. 91); hence, from (5.12) it follows that, for $\theta \le t \le T$,

$$D_\theta Y_t = \mathbb{E}\left[D_\theta\left(\Gamma_{t,T}\xi + \int_t^T \Gamma_{t,s}\varphi_s\, ds\right)|\mathcal{F}_t\right].$$

Furthermore, by natural properties on the Malliavin derivative,

$$\begin{aligned} D_\theta\left(\Gamma_{t,T}\xi + \int_t^T \Gamma_{t,s}\varphi_s\, ds\right) &= D_\theta(\Gamma_{t,T})\xi + \Gamma_{t,T}\, D_\theta\xi \\ &\quad + \int_t^T D_\theta(\Gamma_{t,s})\varphi_s\, ds + \int_t^T \Gamma_{t,s}\, D_\theta\varphi_s\, ds. \end{aligned}$$

Consequently, we obtain the following (natural) property.

PROPOSITION 5.5. *For $\theta \le t \le T$,*

$$D_\theta Y_t = \mathbb{E}\left[\Gamma_{t,T} D_\theta\xi + D_\theta\Gamma_{t,T}\,\xi + \int_t^T (\Gamma_{t,s} D_\theta\varphi_s + D_\theta\Gamma_{t,s}\,\varphi_s)\, ds\, |\mathcal{F}_t\right].$$

Actually Karatzas and Ocone (1992) gave another type of expression for DY. Let us show this property in our context. Recall first that DY can also be written using the adjoint process Γ, because $(D_\theta Y, D_\theta Z)$ is a solution of a linear BSDE similar to that of (Y, Z): for $\theta \le t \le T$,

$$D^i_\theta Y_t = \mathbb{E}\left[\Gamma_{t,T} D^i_\theta\xi + \int_t^T \Gamma_{t,s}(D^i_\theta\varphi_s + Y_s D^i_\theta\beta_s + Z^*_s D^i_\theta\gamma_s)\, ds\, |\mathcal{F}_t\right]. \tag{5.13}$$

Applying this property, we obtain the following representation formula established by Karatzas and Ocone (in the case $\varphi = \beta = 0$) (Karatzas and Ocone 1992, formula 2.20, Theorem 2.5, and Corollary 2.6) under very weak integrability conditions ($\xi \in \mathbb{D}_{1,1}$, $\gamma \in \mathbb{L}^a_{1,1}, \cdots$).

PROPOSITION 5.6. *For $\theta \le t \le T$,*

$$\begin{aligned} D_\theta Y_t = \mathbb{E}\Bigg[&\Gamma_{t,T} D_\theta\xi + \int_t^T \Gamma_{t,s}(D_\theta\varphi_s + Y_s D_\theta\beta_s)\, ds \\ &+ \left(\Gamma_{t,T}\xi + \int_t^T \Gamma_{t,s}\varphi_s\, ds\right)\left(\int_t^T D_\theta\gamma^*_s\, dW_s\right)|\mathcal{F}_t\Bigg]. \end{aligned}$$

Proof. Since $\Gamma_{t,T}\xi + \int_t^T \Gamma_{t,s}\varphi_s\, ds = Y_t + \int_t^T \Gamma_{t,s} Z^*_s\, dW_s$, we have

$$\mathbb{E}\left[\left(\Gamma_{t,T}\xi + \int_t^T \Gamma_{t,s}\varphi_s\, ds\right)\left(\int_t^T D_\theta\gamma_s\, dW_s\right)|\mathcal{F}_t\right] = \mathbb{E}\left[\int_t^T \Gamma_{t,s} Z^*_s D_\theta\gamma_s\, ds|\mathcal{F}_t\right].$$

The result follows from (5.13). □

Application to Example 1.1, Section 1.3. We come back to Example 1.1 seen in Section 1.3 (hedging claims with a higher interest rate for borrowing) and studied by Cvitanic and Karatzas (1993) under slightly different assumptions. Recall that in this example the hedging strategy (X, π) (wealth, portfolio) satisfies

$$dX_t = r_t X_t\,dt + \pi_t^* \sigma_t \theta_t\,dt + \pi_t^* \sigma_t\,dW_t - (R_t - r_t)(X_t - \pi_t^* \mathbf{1})^-\,dt, \qquad X_T = \xi\,.$$

Let $(X^R, \sigma^* \pi^R)$ be the solution of the LBSDE

$$\begin{aligned} dX_t^R \;=\; & r_t X_t^R\,dt + (\pi_t^R)^* \sigma_t \theta_t\,dt \\ & + (\pi_t^R)^* \sigma_t\,dW_t - (R_t - r_t)(X_t^R - (\pi_t^R)^* \mathbf{1})\,dt\,, \qquad X_T^R = \xi. \end{aligned} \tag{5.14}$$

Notice that X^R is equal to the ex post price $X^{(\beta,\gamma)}$ (defined in the example studied in Section 3.3) for $\beta_t = R_t$ and $\gamma_t = \sigma_t \theta_t + (r_t - \beta_t)\mathbf{1}$, $d\mathbb{P} \otimes dt$ a.s.

It is interesting to find a sufficient condition which ensures that X is equal to X^R. Actually, it is easy to see that it is sufficient to have

$$(\pi_t^R)^* \mathbf{1} \geq X_t^R,\, d\mathbb{P} \otimes dt \text{ a.s.} \tag{5.15}$$

Recall now that, by Proposition 5.3, π is also a function of X given by $\pi_t = (\sigma_t^*)^{-1} D_t X_t$, $d\mathbb{P} \otimes dt$ a.s., where $(D_u X_t, 0 \leq u \leq t \leq T)$ denotes the version of the Malliavin derivative of the process X defined in Proposition 5.3. It follows that (5.15) can also be written as

$$\mathbf{1}^* (\sigma_u^*)^{-1} D_u X_t^R \geq X_u^R,\;\; d\mathbb{P} \otimes du \text{ a.s.}$$

Using the comparison theorem and the Malliavin calculus, we state the following proposition, which generalizes a property obtained by Cvitanic and Karatzas (1993).

PROPOSITION 5.7. *Suppose that the coefficients r_t, R_t, b_t, and σ_t are deterministic functions of t and suppose that $\xi \in \mathbb{D}_{1,2}$. If $\mathbf{1}^* (\sigma_u^*)^{-1} D_u \xi \geq \xi$, $d\mathbb{P} \otimes du$ a.s., then the price for ξ is $X = X^R$.*

Proof. In order to show the proposition, it is sufficient to prove that $\mathbf{1}^* (\sigma_u^*)^{-1} D_u X_t^R \geq X_u^R$, $d\mathbb{P} \otimes du$ a.s. Recall first that (X^R, π^R) is solution of the BSDE

$$\begin{aligned} & -dX_t^R = [-R_t X_t^R - (\pi_t^R)^* (\sigma_t \theta_t + (r_t - R_t)\mathbf{1})]\,dt - (\pi_t^R)^* \sigma_t\,dW_t\,, \\ & X_T^R = \xi. \end{aligned} \tag{5.16}$$

By Proposition 5.3, $(X^R, \pi^R) \in \mathbb{L}^2(0, T\,;\; (\mathbb{D}^{1,2}) \times (\mathbb{D}^{1,2})^n)$, and, for $1 \leq i \leq n$, a version of $\{(D_u^i X_t^R,\; D_u^i \pi_t^R)\,;\; 0 \leq u \leq t \leq T\}$ is

$$\begin{aligned} & -dD_u^i X_t^R = -R_t D_u^i X_t^R - (D_u^i \pi_t^R)^* (\sigma_t \theta_t + (r_t - R_t)\mathbf{1})\,dt - (D_u^i \pi_t^R)^* \sigma_t\,dW_t\,, \\ & D_u^i X_T^R = D_u^i \xi\,. \end{aligned} \tag{5.17}$$

Put $Y_t^u = \mathbf{1}^*(\sigma_u^*)^{-1} D_u X_t^R$ and $Z_t^u = (D_u \pi_t^R)(\sigma_u)^{-1}\mathbf{1}$, for $0 \le u \le t \le T$. We easily see that $(Y_t^u, Z_t^u, u \le t \le T)$ is the solution of the BSDE

$$\text{(5.18)} \qquad \begin{aligned} &-dY_t^u = -R_t Y_t^u - (Z_t^u)^*(\sigma_t\theta_t + (r_t - R_t)\mathbf{1})\,dt - (Z_t^u)^*\sigma_t\,dW_t\,,\\ &Y_T^u = \mathbf{1}^*(\sigma_u^*)^{-1} D_u \xi\,. \end{aligned}$$

Then applying the comparison theorem to (X^R, π^R) and (Y^u, Z^u), we have that $Y_u^u \ge X_u^R$, $d\mathbb{P}$, a.s., and the result easily follows. □

From this proposition we deduce the property stated by Cvitanic and Karatzas (1993, Example 9.5).

PROPOSITION 5.8. *Suppose the coefficients are deterministic functions of t and let $\xi \ge 0$ be a contingent claim of the form $\xi = \psi(P_T)$, where ψ is a given function ψ: $\mathbb{R}^n \to \mathbb{R}^+$ of class C^1 with bounded derivative and such that $\sum_{i=1}^n x^i \partial_{x_i}\psi(x) \ge \psi(x)$ (for example, if ψ is a convex function of class C^1 with $\psi(0) = 0$, this condition is satisfied). Then $X_t = X_t^R$, $0 \le t \le T$, a.s.*

Proof. By the chain rule,

$$D_t^j \xi = \sum_{i=1}^n \partial_{x_i}\psi(P_T) D_t^j (P_T^i) = \sum_{i=1}^n \partial_{x_i}\psi(P_T) P_T^i \sigma_{i,j}(t).$$

Hence, $(\sigma_t^*)^{-1} D_t \xi = \left(\partial_{x_i}\psi(P_T) P_T^i\right)$ and the result follows from Proposition 5.7. □

REMARK. The result still holds for the classical European option $\xi = (P_T^1 - K)^+$ where K is a real (positive) constant, if the law of the random variable P_T^1 is absolutely continuous with respect to the Lebesgue measure on $\mathbb{R}$ (recall that this condition is satisfied under some nondegeneracy conditions on the coefficients; see Nualart 1995, Theorem 2.3.2). Indeed, in this case, ψ is given by $\psi(x) = (x - K)^+$. ψ is of class $\mathcal{C}^1$ on $]-\infty, K[$ and $]K, +\infty[$ and $\phi'(x) \ge 0$ for any $x \ne K$. Recall that the chain rule still holds in this case (Nualart 1995, Proposition 1.23), since the law of the random variable P_T^1 is absolutely continuous with respect to Lebesgue measure. Consequently, the result follows from Proposition 5.7. Thus, the contingent claim is evaluated under a current rate R and a risk premium equal to $\theta_t + (r_t - R_t)\sigma_t^{-1}\mathbf{1}$. Notice that if $\sigma = I$, the risk premium is lower than the primitive one θ.

Application to the Markovian Case. In this section we consider the BSDE associated with a forward equation defined in Section 4.1. By the results of Section 5.2 we have that if the coefficients are differentiable, then the solution $(Y_s^{t,x}, Z_s^{t,x})$ is differentiable in Malliavin's sense.

PROPOSITION 5.9. *If the coefficients b, σ, f, and Ψ are continuously differentiable with respect to (x, y, z) with uniformly bounded derivatives, then*

- *For any $0 \le t \le s \le T$, $x \in \mathbb{R}^p$, and $(Y_s^{t,x}, Z_s^{t,x}) \in \mathbb{L}^2(0, T\,;\, (\mathbb{D}^{1,2})^d \times (\mathbb{D}^{1,2})^{n\times d})$, and for each $1 \le i \le n$ a version of $\{(D_\theta^i Y_s,\ D_\theta^i Z_s)\,;\ 0 \le \theta, t \le s \le T\}$ is given by*

$$D_\theta^i Y_s = 0, \quad D_\theta^i Z_s = 0, \qquad 0 \le \theta < t \le T \quad \text{or} \quad s < \theta \le T;$$

and for $t \le \theta \le T$, $\{(D_\theta^i Y_s,\ D_\theta^i Z_s)\,;\ \theta \le s \le T\}$ satisfies the following LBSDE:

$$\begin{aligned}(5.19)\quad -dD_\theta^i Y_s &= [\partial_y f(s, P_s, Y_s, Z_s) D_\theta^i Y_s + \partial_z f(s, P_s, Y_s, Z_s) D_\theta^i Z_s]\, ds \\ &\quad + \partial_x f(s, P_s, Y_s, Z_s) D_\theta^i P_s\, ds - D_\theta^i Z_s{}^* \, dW_s, \\ D_\theta^i Y_T &= \Psi'(P_T) D_\theta^i P_T\,.\end{aligned}$$

Moreover, $\{D_s Y_s\,;\ t \le s \le T\}$ defined by (5.19) is a version of $\{Z_s\,;\ t \le s \le T\}$.
- *For any $0 \le t \le s \le T$, $x \in \mathbb{R}^p$,*

$$Z_s^{t,x} = \partial_x Y_s^{t,x} (\partial_x P_s^{t,x})^{-1} \sigma(s, P_s^{t,x}),\ ds \otimes d\mathbb{P}\ a.s.$$

Proof. First, recall that for any $0 \le t \le s \le T$, $x \in \mathbb{R}^p$, $(P_s^{t,x}) \in \mathbb{L}^2(0, T\,;\, (\mathbb{D}^{1,2})^p)$, and for each $1 \le i \le n$, a version of $\{D_\theta^i P_s^{t,x}\,;\ 0 \le \theta, t \le s \le T\}$ is given by $D_\theta^i P_s^{t,x} = 0$, $0 \le \theta < t$, and, for $t \le \theta$, $\{D_\theta^i P_s^{t,x}; \theta \le s \le T\}$ is the unique solution of the linear SDE

$$dD_\theta^i P_s = \partial_x b(s, P_s)\, D_\theta^i P_s\, ds + \partial_x \sigma_j(s, P_s)\, D_\theta^i P_s\, dW_s^j, \qquad D_\theta^i P_\theta = \sigma_i(\theta, P_\theta)\,.$$

Moreover, from the uniqueness of the solution of the SDE satisfied by $D_\theta P$, it follows that

$$(5.20)\qquad D_\theta P_s = \partial_x P_s (\partial_x P_\theta)^{-1} \sigma(\theta, P_\theta)\,, \qquad t \le \theta \le s \le T\,, \mathbb{P}\text{ a.s.}$$

Recall now that $\sup_{t\le s\le T}(|P_s^{t,x}| + |\partial_x P_s^{t,x}|) \in \mathbb{L}^p$ for any $p \ge 1$. Then, using the assumptions made on the coefficients, we see that the hypotheses of Proposition 5.3 are satisfied, so the first statement of the proposition is proved.

It remains to show the second one. The uniqueness of the solution of BSDE (5.19) and (5.20) yield $D_\theta^i Y_s = \partial_x Y_s (\partial_x P_\theta)^{-1} \sigma_i(\theta, P_\theta)$ or, equivalently, $D_\theta Y_s = \partial_x Y_s (\partial_x P_\theta)^{-1} \sigma(\theta, P_\theta)$. Hence, by taking $\theta = s$, $D_s Y_s = \partial_x Y_s (\partial_x P_s)^{-1} \sigma(s, P_s)$. By the first statement, $Z_s = D_s Y_s$ almost surely. The second statement now easily follows. $\square$

REFERENCES

ANSEL, J. P., and C. STRICKER (1992a): "Lois de Martingale, Densités et Décomposition de Föllmer-Schweizer," *Ann. Inst. H. Poincaré,* 28, 375–392.

ANSEL, J. P., and C. STRICKER (1992b): "Unicité et Existence de la Loi Minimale," *Sém. de Probabilités XXV,* Lecture Notes in Math., vol. 1485. Berlin and New York: Springer-Verlag.

ANTONELLI, F. (1993): "Backward-Forward Stochastic Differential Equations," *Ann. Appl. Probab.* 3, 777–793.

ARKIN, V., and M. SAKSONOV (1979): "Necessary Optimality Conditions for Stochastic Differential Equations," *Soviet Math. Dokl.*, 20, 1–5.

BACK, K., and S. R. PLISKA (1987): "The Shadow Price of Information in Continuous Time Decision Problems," *Stochastics*, 22, 151–186.

BACK, K., and S. R. PLISKA (1991): "On the Fundamental Theorem of Asset Pricing with Infinite State Space," *J. Math. Econ.*, 20, 1–18.

BARDHAN, I. (1993): "Stochastic Income in Incomplete markets: hedging and optimal policies," working paper.

BENES, V. E. (1970): "Existence of Optimal Strategies Based on Specified Information for a Class of Stochastic Decision Problems," *SIAM J. Control Optim.*, 8, 179–188.

BENES, V. E. (1971): "Existence of Optimal Stochastic Control Law," *SIAM J. Control Optim.*, 9, 446–472.

BERGMAN, Y. (1991): "Option Pricing with Divergent Borrowing and Lending Rates," working paper, Brown University.

BISMUT, J. M. (1973): "Conjugate Convex Functions in Optimal Stochastic Control," *J. Math. Anal. Appl.*, 44, 384–404.

BISMUT, J. M. (1978): *Contrôle des Systèmes Linéaires Quadratiques: Applications de l'Intégrale Stochastique,* Sém. Probab. XII., Lecture Notes in Math., vol. 649. Berlin and New York: Springer-Verlag, 180–264.

BLACK, F., and M. SCHOLES (1973): "The Pricing of Options and Corporate Liabilities," *J. Political Econ.*, 3, 637–654.

BUCKDAHN, R. (1993): "Backward Stochastic Differential Equations Driven by a Martingale," preprint.

BUCKDAHN, R., and E. PARDOUX (1994): "BSDE's with Jumps and Associated Integral-Stochastic Differential Equations," preprint.

CADENILLAS, A., and I. KARATZAS (1995): "The Stochastic Maximum Principle for Linear Convex Optimal Control with Random Coefficients," *SIAM J. Control Optim.*, 33, 590–624.

CINLAR, E., J. JACOD, P. PROTTER, and M. J. SHARPE (1980): "Semimartingale and Markov Processes," *Z. Wahrsch.* 54, 161–219.

COLWELL, D., R. ELLIOTT, and E. KOPP (1991): "Martingale Representation and Hedging Policies," *Stoch. Process. Appl.*, 38, 335–345.

CVITANIC, J., and I. KARATZAS (1992): "Convex Duality in Constrained Portfolio Optimization," *Ann. Appl. Probab.*, 2, 767–818.

CVITANIC, J., and I. KARATZAS (1993): "Hedging Contingent Claims with Constrained Portfolios," *Ann. Appl. Probab.*, 3, 652–681.

CVITANIC, J., and J. MA (1994): "Hedging Options for a Large Investor and Forward-Backward SDE's," *Ann. Appl. Probab.*, forthcoming.

DAVIS, M. H. A. (1973): "On the Existence of Optimal Policies in Stochastic Control," *SIAM J. Control Optim.*, 11, 587–594.

DELBAEN, F.(1992): "Representing Martingale Measures when Asset Prices Are Continuous and Bounded," *Math. Finance*, 2, 107–130.

DELBAEN, F., P. MONAT, W. SCHACHERMAYER, M. SCHWEIZER, and C. STRICKER (1994): "Inégalités de normes avec poids et fermeture d'un espace d' intégrales stochastiques," *CRAS 1994*.

DELBAEN, F., and W. SCHACHERMAYER (1994a): "Arbitrage and Free Lunch with Bounded Risk," *Math. Finance*, 4, 343–348.

DELBAEN, F., and W. SCHACHERMAYER (1994b): "A General Version of the Fundamental Theorem of Asset Pricing," *Math. Ann.*, 300, 463–520.

DELLACHERIE, C. (1972): *Capacités et Processus Stochastiques*, Springer, 67.

DELLACHERIE, C.(1977): "Sur l'existence de certains essinf et essup de familles de processus mesurables," Sem. Probab. XII, Lecture Notes in Math., vol. 649. Berlin and New York: Springer-Verlag.

DELLACHERIE, C., and P. A. MEYER (1980): *Probabilités et Potentiel: Théorie du Potentiel, Processus de Markov*. Paris: Hermann, chaps. XII–XVI.

DUDLEY, R. M. (1977): "Wiener Functionals as Itô Integrals," *Ann. Probab.*, 5, 140–141.

DUFFIE, D. (1988): *Security Markets: Stochastic Models*. Boston: Academic Press.

DUFFIE, D. (1992): *Dynamic Asset Pricing Theory*. Princeton University Press.

DUFFIE, D., and L. EPSTEIN (1992a): "Stochastic Differential Utility," *Econometrica*, 60(2), 353–394.

DUFFIE, D., and L. EPSTEIN (1992b): "Asset Pricing with Stochastic Differential Utility," *Rev. Financial Stud.*, 5, 411–436.

DUFFIE, D., P. Y. GEOFFARD, and C. SKIADIAS (1992): "Efficient and Equilibrium Allocations with Stochastic Differential Utility," working paper, Graduate School of Business, Stanford University.

DUFFIE, D., and M. HUANG (1994): "Swap Rates and Credit Quality," working paper, Graduate School of Business, Stanford University.

DUFFIE, D., J. MA, and J. YONG (1994): "Black's Consol Rate Conjecture," *Ann. Appl. Probab.*, 5, 356–382.

DUFFIE, D., and K. SINGLETON (1994): "Econometric Modelling of Term Structures of Defaultable Bonds," working paper, Graduate School of Business, Stanford University.

DUFFIE, D., and C. SKIADIAS (1991): "Continuous-Time Security Pricing: A Utility Gradient Approach," working paper, Graduate School of Business, Stanford University.

EKELAND, I., and R. TEMAN (1976): *Convex Analysis and Variational Problems*. Amsterdam: North-Holland.

EKELAND, I., and T. TURNBULL (1979): *Infinite Dimensional Optimization and Convexity*. Chicago Lectures in Math.

ELLIOTT, R. J. (1982): *Stochastic Calculus and Applications*. New York: Springer-Verlag.

EL KAROUI, N. (1981): *Les Aspects Probabilistes du Contrôle Stochastique*. Lecture Notes in Math., vol. 816. Berlin and New York: Springer-Verlag, 73–238.

EL KAROUI, N., C. KAPOUDJIAN, E. PARDOUX, S. PENG, and M. C. QUENEZ (1995): "Reflected Solutions of Backward SDE's, and Related Obstacle Problems for PDE's," *Ann. Probab.*, forthcoming.

EL KAROUI, N., and M. JEANBLANC-PICQUÉ (1988): *Contrôle de Processus de Markov*, Sem. Probab. XXII, Lecture Notes in Math., vol. 1321. Berlin and New York: Springer-Verlag, 508–541.

EL KAROUI, N., S. PENG, and M. C. QUENEZ (1994): "Optimization of Utility Functions," working paper, Paris VI University.

EL KAROUI, N., and M. C. QUENEZ (1991): "Programmation dynamique et évaluation des actifs contingents en marché incomplet," *C. R. Acad. Sci. Paris*, Sér. 1, 313, 851–854.

EL KAROUI, N., and M. C. QUENEZ (1995): "Dynamic Programming and Pricing of Contingent Claims in Incomplete Market," *Siam J. Control Optim.*, 33, 29–66.

EPSTEIN, L., and S. ZIN (1989): "Substitution, Risk Aversion and the Temporal Behavior of Consumption and Asset Returns: A Theorical Framework," *Econometrica*, 57, 837–969.

FLEMING, W. F., and RISHEL (1975): *Deterministic and Stochastic Control*. Berlin and New York: Springer-Verlag.

FLEMING, W. F.,and M. SONER (1993): *Controlled Markov processes and Viscosity Solutions*. Berlin and New York: Springer-Verlag.

FÖLLMER, H., and Y. M. KABANOV (1995): "On the Optional Decomposition Theorem and the Lagrange Multipliers," *Finance and Stochastics*, forthcoming.

FÖLLMER, H., and M. SCHWEIZER (1990): "Hedging of Contingent Claims under Incomplete Information," in *Applied Stochastic Analysis*, eds. M. H. A. Davis and R. J. Elliot. London: Gordon and Breach.

FÖLLMER, H., and D. SONDERMANN (1986): "Hedging of Non-redundant Contingent Claims," in *Contributions to Mathematical Economics. In Honor of Gérard Debreu,* eds. W. Hildenbrand. A. Mas-Colell. Amsterdam: North-Holland.

GEOFFARD, P. Y. (1995): "Discounting and Optimizing, Utility maximization: A Lagrange Variational Formulation as a Minmax Problem," *J. Econ. Theory*.

HAMADENE, S., and J. P. LEPELTIER (1994): "Zero-Sum Stochastic Differential Games and Backward Equations," *Systems Control Lett.*, 24, 259–263.

HARRISON, M., and D. KREPS (1979): "Martingales and Arbitrage in Multiperiod Securities Markets," *J. Econ. Theory,* 20, 381–408.

HARRISON, M., and S. R. PLISKA (1981): "Martingales and Stochastic Integrals in the Theory of Continuous Trading," *Stoch. Process. Appl.*, 11, 215–260.

HARRISON, M., and S. R. PLISKA (1983): "A Stochastic Calculus Model of Continuous Trading: Complete Markets," *Stoch. Process. Appl.,* 15, 313–316.

HE, H., and N. D. PEARSON (1991): "Consumption and Portfolio Policies with Incomplete Markets and Short-Sale Constraints: The Infinite Dimensional Case," *J. Econ. Theory,* 54, 259–304.

IKEDA, N., and S. WATANABE (1989): *Stochastic Differential Equations and Diffusion Processes*, 2nd ed., Amsterdam: North-Holland, and Tokyo: Kodansha.

ISHII, H., and P. L. LIONS (1990): "Viscosity Solutions of Fully Non-linear Second Order Elliptic Partial Differential Equations," *J. Diff. Eqs.*, 83, 26–78.

JACKA, S. J. (1996): "A Martingale Representation Result and an Application to Incomplete Financial Markets," *Math. Finance*, 2, 239–250.

JOUINI, E., and H. KALLAL (1995a): "Arbitrage in Securities Markets with Short-Sales Constraints," *Math. Finance,* 5, 178–197.

JOUINI, E., and H. KALLAL (1995b): "Efficient Trading Strategies in the Presence of Market Frictions," working paper, Paris CREST.

KABANOV, Y. (1978): "On the Pontryiagin Maximum Principle for the Linear Stochastic Differential Equations," in *Probabilistic Models and Control of Economical Processes*, CEMI.

KARATZAS, I. (1989): "Optimization Problems in the Theory of Continuous Trading," *SIAM. J. Control Optim.*, 27, 1221–1259.

KARATZAS, I., J. P. LEHOCZKY, S. E. SHREVE, and G. L. XU (1989): "Martingale and Duality Methods for Utility Maximisation in an Incomplete Market," *SIAM J. Control Optim.*, 29, 702–730.

KARATZAS, I., and D. L. OCONE (1992): "A Generalized Clark Representation Formula, with Application to Optimal Portfolios," *Stochastics*, 34, 187–220.

KARATZAS, I., and S. SHREVE (1987): *Brownian Motion and Stochastic Calculus*. New York: Springer Verlag.

KORN, R. (1992): "Option Pricing in a Model with a Higher Interest Rate for Borrowing than for Lending," working paper.

KRAMKOV, D. O. (1996): "Optional Decomposition of Supermartingales and Hedging Contingent Claims in Incomplete Security Markets," *Probab. Theory Rel. Fields,* 105, 459–479.

KRYLOV, N. (1980): *Controlled Diffusion Processes*. Springer-Verlag, New York.

MA, J., P. PROTTER, and J. YONG (1994): "Solving Forward-Backward Stochastic Differential Equations Explicitly—A Four Step Scheme," *Probab. Theory Rel. Fields*, 98, 339–359.

MERTON, R. (1971): "Optimum Consumption and Portfolio Rules in a Continuous Time Model," *J. Econ. Theory*, 3, 373–413.

MERTON, R. (1973): "Theory of Rational Option Pricing," *Bell J. Econ. Manage. Sci.,* 4, 141–183.

MERTON, R. (1991): *Continuous Time Finance*. Basil Blackwell.

MONAT, P., and C. STRICKER (1995): "Fölmer-Schweizer Decomposition and Mean-Variance Hedging of General Claims," *Ann. Probab.*, 23, 605–628.

NUALART, D. (1986): *Non-causal Stochastic Integrals and Calculus,* Lecture Notes in Mathematics, vol. 1316. Springer-Verlag, 80–129.

NUALART, D. (1995): *The Malliavin Calculus and Related Topics,* Probability and Its Applications. New York and Berlin: Springer-Verlag.

NUALART, D., and E. PARDOUX (1988): "Stochastic Calculus with Anticipating Integrands," *Probab. Theory Rel. Fields*, 78, 535–581.

OCONE, D. (1988): "Malliavin's Calculus and Stochastic Integral Representation of Functionals of Diffusion Processes," *Stochastics*, 12, 161–185.

PARDOUX, E., and S. PENG (1990): "Adapted Solution of a Backward Stochastic Differential Equation," *Systems Control Lett.*, 14, 55–61.

PARDOUX, E. and S. PENG (1992): *Backward Stochastic Differential Equations and Quasilinear Parabolic Partial Differential Equations,* Lecture Notes in CIS, vol. 176. Springer-Verlag, 200–217.

PARDOUX, E. and S. PENG (1993): "Backward Doubly Stochastic Differential Equations and Systems of Quasilinear Parabolic SPDEs," preprint.

PENG, S. (1990): "A General Stochastic Maximum Principle for Optimal Control Problems," *SIAM J. Control Optim.*, 28, 966–979.

PENG, S. (1991): "Probabilistic Interpretation for Systems of Quasilinear Parabolic Partial Differential Equations," *Stochastics*, 37, 61–74.

Peng, S. (1992a): "Stochastic Hamilton-Jacobi-Bellman Equations," *SIAM J. Control Optim.*, 30, 284–304.

PENG, S. (1992b): "A Generalized Dynamic Programming Principle and Hamilton-Jacobi-Bellman Equation," *Stochastics*, 38, 119–134.

PENG, S. (1992c): "A Nonlinear Feynman–Kac Formula and Applications," in *Proceedings of Symposium of System Sciences and Control Theory,* ed. Chen and Yong. Singapore: World Scientific, 173–184.

PENG, S. (1993): "Backward Stochastic Differential Equation and Its Application in Optimal Control," *Appl. Math. Optim.*, 27, 125–144.

PROTTER, P. (1990): *Stochastic Integration and Stochastic Differential Equations: A New Approach,* Berlin: Springer-Verlag.

QUENEZ, M. C. (1993): "Méthodes de contrôle stochastique en Finance," Thèse de doctorat de l'Université Pierre et Marie Curie.

REVUZ, D., and M. YOR (1991): *Continuous Martingales and Brownian Motion*. Springer-Verlag.

SCHWEIZER, M. (1991): "Option Hedging for Semimartingales," *Stoch. Process. Appl.*, 37, 339–363.

SCHWEIZER, M. (1992): "Mean-Variance Hedging for General Claims," *The Annals of Appl. Probab.*, 2, 171–179.

SCHWEIZER, M. (1994a): "Risk-Minimizing Hedging Strategies under Restricted Information," *Math. Finance,* 4, 327–342.

SCHWEIZER, M. (1994b): "Approximating Random Variables by Stochastic Integrals," *Ann. Probab.*, 22, 1536–1575.

STRICKER, C. (1990): "Arbitrage et Lois de Martingale," *Ann. Inst. H. Poincaré,* 26, 451–460.