

# Backward Stochastic Differential Equations Associated to a Symmetric Markov Process

Vlad Bally, Etienne Pardoux, L. Stoica

► **To cite this version:**

Vlad Bally, Etienne Pardoux, L. Stoica. Backward Stochastic Differential Equations Associated to a Symmetric Markov Process. [Research Report] RR-4425, INRIA. 2002. inria-00072163

**HAL Id: inria-00072163**

**<https://hal.inria.fr/inria-00072163>**

Submitted on 23 May 2006

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

***Backward Stochastic Differential Equations  
Associated to a Symmetric Markov Process***

V. Bally — E. Pardoux — L. Stoica

**N° 4425**

Avril 2002

THÈME 4



*R*apport  
de recherche



## Backward Stochastic Differential Equations Associated to a Symmetric Markov Process

V. Bally , E. Pardoux , L. Stoica

Thème 4 — Simulation et optimisation  
de systèmes complexes  
Projet Mathfi

Rapport de recherche n° 4425 — Avril 2002 — 51 pages

**Abstract:** We consider a second order semi-elliptic differential operator  $L$  with measurable coefficients, in divergence form, and the semilinear parabolic PDE

$$\begin{aligned}(\partial_t + L)u(t, x) + f(t, x, u, \nabla u \sigma) &= 0, \forall 0 \leq t \leq T \\ u(T, x) &= \Phi(x)\end{aligned}$$

and employ the symmetric Markov process of infinitesimal operator  $L$  in order to give a probabilistic interpretation for the solution  $u$ , i.e. we solve the corresponding BSDE. We obtain also a representation theorem for martingales which represents a generalization of the representation theorem given by Fukushima for additive functional martingales. This permits us to solve general (non-Markov) BSDE's with semi-linear terms. The nonlinear term  $f$  satisfies a monotonicity condition with respect to  $u$  and a Lipschitz condition with respect to  $\nabla u$ . Finally we prove a comparison theorem and use it in order to solve a stochastic control problem.

**Key-words:** Symetric Markov processes, Semi-linear PDE's, Backward stochastic differential equations, Fukushima stochastic calculus

# Equations différentielles Stochastiques Rétrogrades Associées à un Processus de Markov Symétrique

**Résumé :** On considère un opérateur différentiel de second ordre semi-elliptique  $L$  avec des coefficients mesurables, en forme divergente, et l'EDP parabolique

$$\begin{aligned}(\partial_t + L)u(t, x) + f(t, x, u, \nabla u\sigma) &= 0, \forall 0 \leq t \leq T \\ u(T, x) &= \Phi(x)\end{aligned}$$

et on emploie le processus de Markov symétrique d'opérateur infinitesimal  $L$  pour donner une interprétation probabiliste pour la solution  $u$  ce qui revient à résoudre l'EDSR correspondante. On obtient aussi un théorème de représentation des martingales qui représente une généralisation du théorème de représentation donné par Fukushima pour des martingales fonctionnelles additives. Ceci permet de résoudre l'EDSR général (dans le contexte non Markovien). Le terme non linéaire  $f$  satisfait une condition de monotonie par rapport à  $y = u$  et de Lipschizianité par rapport à  $z = \nabla u\sigma$ . Nous démontrons aussi un théorème de comparaison qu'on emploie pour résoudre un problème de contrôle stochastique.

**Mots-clés :** Processus de Markov symétriques, EDP semi-linéaires, Equations différentielles stochastiques rétrogrades, Calcul stochastique de Fukushima

## 1 Introduction

Backward Stochastic Differential Equations (in short BSDE's) have been introduced (in their actual form) by E. Pardoux and S. Peng [10] and found applications in stochastic control and mathematical finance. One motivation is also to give a probabilistic interpretation to the solutions of a system of parabolic semilinear PDE's (see [11]) of the following form

$$(\partial_t + L)u + f(t, x, u, \nabla u \sigma) = 0 \text{ on } [0, T] \times R^N, \quad u_T = \Phi.$$

If  $f$  and the coefficients of the second order differential operator  $L$  are sufficiently smooth then the PDE has a classical solution and one may construct the pair of processes  $Y_s^{t,x} =: u(s, X_s^{t,x}), Z_s^{t,x} =: \nabla u \sigma(s, X_s^{t,x})$  where  $X_s^{t,x}, t \leq s \leq T$ , is the diffusion process with infinitesimal operator  $L$  which starts from  $x$  at moment  $t$ . Then, using Itô's formula one checks that  $(Y_s^{t,x}, Z_s^{t,x})_{t \leq s \leq T}$  solves the BSDE

$$Y_s^{t,x} = \Phi(X_T^{t,x}) + \int_s^T f(r, X_r^{t,x}, Y_r^{t,x}, Z_r^{t,x}) dr - \int_s^T Z_s^{t,x} dB_r.$$

If  $f$  is linear then the solution of above BSDE has an explicit formula and the equality  $u(t, x) = Y_t^{t,x} = E(Y_t^{t,x})$  yields the classical Feynman Kac formula.

On the other hand, if  $f$  and the coefficients of  $L$  are just Lipschitz continuous then the BSDE still has a unique solution but the PDE has no more a classical solution, so that one has to consider generalized solutions. In a series of papers initiated by [11] one considers viscosity solutions and more recently in [2] and [1] one considers solutions in Sobolev spaces for the PDE and proves that the probabilistic interpretation given above remains true. Anyway, in both these approaches, since the coefficients are Lipschitz continuous, the Markov process  $X$  with infinitesimal operator  $L$  is a diffusion process which satisfies a SDE and so one may use the flow  $X_s^{t,x}$  and its associated stochastic calculus.

In this paper we consider a semi-elliptic symmetric second order differential operator  $L$  (which is written in divergence form; no nondegeneracy condition is assumed) with measurable coefficients. Then  $X$  is a symmetric Markov process which does no more satisfy a SDE. There exists no more a flow representing  $X$  and the usual stochastic calculus corresponding to SDE's has to be replaced by the calculus associated to such a process by Fukushima (see [4]). We prove that the above PDE has a unique solution  $u$  in some Dirichlet space and the stochastic interpretation is given as  $u(t, x) = Y_0^t, P^x - a.s$ , where  $P^x$  is the law of the process  $X$  starting from  $x$  and  $Y$  is the solution of a BSDE (analogous to that considered above) on  $[0, T - t]$ . We also solve a general BSDE, give a comparison theorem and use them in order to solve a stochastic control problem associated to the Markov process  $X$  and to establish the link between this control problem and the corresponding Hamilton Jacobi Bellman equation.

The paper is organized as follows. In Sections 2 and 3 we use analytical methods to solve the above PDE written in variational form

$$\int_0^T ((u_t, \partial_t \varphi) + e(u_t, \varphi_t)) dt = \int_0^T (f(t, \cdot, u_t, \nabla u_t \sigma), \varphi_t) dt + (\Phi, \varphi_T),$$

where  $\varphi$  is a test function. Here  $e$  denotes the energy associated to the operator  $L$ . Note that  $\nabla u_t$  does not exist in our situation, but we are able to give a weak definition of  $\nabla u_t \sigma$ . Moreover we prove that the above weak equation is equivalent with the mild equation in  $L^2$  sense

$$u(t, \cdot) = P_{T-t} \Phi(\cdot) + \int_t^T P_{s-t} f(s, \cdot, u_s, \nabla u_s \sigma)(\cdot) ds. \quad (1)$$

We assume that  $f$  is Lipschitz continuous with respect to its last variable, and satisfies a monotonicity condition with respect to  $u$ .

In Sections 4 and 5 we extend to time- space functions the representation theorem given in [4]. More precisely we prove that if  $u$  is the solution of the above equation then

$$\begin{aligned} u(s+t, X_t) - u(s, X_0) &= \sum_{i=1}^N \int_0^t \phi_i(s+r, X_r) dM^i(r) - \\ &\quad - \int_0^t f(s+r, X_r, u(s+r, X_r), D_\sigma u(s+r, X_r)) dr. \end{aligned} \quad (2)$$

The martingales  $M^i$  which appear in the above formula are the "coordinate martingales" introduced by Fukushima,  $\phi^i$  is a generalization of  $\frac{\partial u}{\partial x^i}$  and  $D_\sigma u$  is a generalization of  $\nabla u \sigma$ . Note also that in the classical situation (Lipschitz continuous coefficients)  $dM = \sigma dB$  where  $B$  is a Brownian motion. So the above formula is the analogue of the standard BSDE.

In Section 4 we prove that (2) holds under  $P^m$  where  $m$  is the Lebesgue measure (which represents the invariant measure for the process  $X$ ) and in Section 5 we strengthen this result in the sense that the equality holds under each  $P^x$  with  $x \in \mathcal{N}^c$ , where  $\mathcal{N}$  is a set of null capacity.

In Section 6 we employ the above results and a standard fixed point argument to prove existence and uniqueness for the solutions of the general BSDE

$$Y_t = \xi + \int_t^T f(s, \omega, Y_s, Z_s \sigma(X_s)) ds - \int_t^T Z_s \cdot dM_s,$$

where again the function  $f$  need not be Lipschitz continuous with respect to  $y$ ; only monotonicity suffices. (The data  $\xi$  and  $f(s, \omega, 0, 0)$  are assumed to belong to  $L^p(\Omega)$ , resp.  $L^p(\Omega; L^1(0, T))$ ,  $p > 1$ .) The equality holds for all  $t \in [0, T]$ , and  $P^x - a.s$  for every  $x \in \mathcal{N}^c$ . Note that the exceptional set  $\mathcal{N}$  may be chosen independently of  $\xi$ . The interest of such a result is the following: one may choose the exceptional set  $\mathcal{N}$  such that, if the process  $X$  starts from a point of  $\mathcal{N}^c$ , it remains always in this set. So  $\mathcal{N}^c$  represents a natural state space for the Markov process  $X$ . In some sense this would replace the lack of the flow. As a consequence one finds a version of the solution which satisfies the mild equation (1) pointwise.

Finally, in Section 7 we prove a comparison result and use it in order to solve a stochastic control problem.

## 2 Preliminaries

Given a bounded measurable function  $\sigma : R^N \rightarrow R^N \otimes R^n, \sigma = (\sigma_k^i)$  ( with upper index  $i = 1, \dots, N$ , and lower index  $k = 1, \dots, n$ ) we define  $a := \sigma\sigma^* = (a^{i,j})$ , where  $a^{i,j} = \sum_k \sigma_k^i \sigma_k^j$  and note (see the Appendix) that there exists a measurable function  $\tau : R^N \rightarrow R^n \otimes R^N$  such that

$$\sigma^* \sigma \tau = \sigma^*, \quad \sigma \tau \sigma = \sigma, \quad \|\sigma \tau\| = \|\tau \sigma\| = 1.$$

The matrix field  $\tau$  is fixed through the paper. Further introduce the bilinear form

$$e(u, v) = \int_{R^N} a^{i,j}(x) \frac{\partial u}{\partial x^i}(x) \frac{\partial v}{\partial x^j}(x) dx, \quad u, v \in C_c^\infty(R^N),$$

where  $C_c^\infty(R^N)$  is the space of infinitely differentiable functions with compact support and the convention of summation on repeated indices is in force. We also define

$$e_1(u, v) = (u, v) + e(u, v),$$

where  $(\cdot, \cdot)$  is the usual scalar product in  $L^2$ . We use  $e(u)$  for  $e(u, u)$  and  $e_1(u)$  for  $e_1(u, u)$ . We *assume* that  $(C_c^\infty(R^N), e_1)$  is closable (see [4] for detailed definitions and notation). Some general criteria imposing conditions on  $a$  in order that  $e_1$  be closable are given in Section 3.1 of [4]: essentially certain partial derivatives of  $a$  have to be locally in  $L^2$  or  $a$  has to verify a local ellipticity assumption. More precisely, a sufficient condition is that for each  $x \in R^N$  there is some  $r > 0$  such that one of the following two conditions holds:

- i)  $\frac{\partial a^{i,j}}{\partial x^i} \in L^2(B_r(x)), 1 \leq i, j \leq N$ ,
- ii) There exists some  $c > 0$  such that  $a(y) \geq cI$  for every  $y \in B_r(x)$ .

The domain of the form  $(C_c^\infty(R^N), e_1)$  ( i.e. the closure of  $C_c^\infty(R^N)$  with respect to  $e_1$ ) is denoted by  $F$ , the associated semigroup is  $(P_t)_{t \geq 0}$ , - which is a symmetric Markovian semigroup that admits the Lebesgue measure  $m$  as an invariant measure. The infinitesimal operator of this semigroup is denoted by  $(D(L), L)$ . We recall that  $\forall u \in L^2, \forall t > 0, P_t u \in D(L) \subset F$  and, if  $u \in D(L), v \in F$  then  $e(u, v) = -(Lu, v)$ . A very useful inequality is (see Lemma 1.3.3 in [4])

$$e(P_t u) \leq \frac{1}{2t} \|u\|_2^2, \quad \forall t > 0, u \in L^2. \quad (3)$$

We introduce the space of functions

$$C_T = \left\{ \varphi : [0, T] \times R^N \rightarrow R / \varphi_t \in F \text{ for almost each } t, \int_0^T e(\varphi_t, \varphi_t) dt < \infty, \right. \\ \left. t \rightarrow \varphi_t \text{ is differentiable in } L^2 \text{ and } t \rightarrow \partial_t \varphi_t \text{ is } L^2 - \text{continuous on } [0, T] \right\},$$

which turns out to be the appropriate space of test functions. Denote by  $C_{T,0}$  the space of functions  $\varphi \in C_T$  such that  $\varphi_0 = 0$ . For  $\varphi \in C_T$  we define

$$\|\varphi\|_T := \left( \sup_{t \leq T} \|\varphi_t\|_2^2 + \int_0^T e(\varphi_t) dt \right)^{\frac{1}{2}}$$



and we denote by  $\widehat{F}$  the completion of  $C_T$  with respect to  $\|\cdot\|_T$ .

**Remark 2.1** *i) If  $u \in \widehat{F}$  then, for almost every  $t$ ,  $u_t \in F$ ,  $\int_0^T e(u_t)dt < \infty$ , and*

$$\|u\|_T = \left( \sup_{t \leq T} \|u_t\|_2^2 + \int_0^T e(u_t)dt \right)^{1/2} < \infty.$$

*ii) If  $u \in \widehat{F}$  and  $\varphi \in F$  then  $\int_0^T u_t dt \in F$  and*

$$e \left( \int_0^T u_t dt, \varphi \right) = \int_0^T e(u_t, \varphi) dt.$$

*iii) The space  $\widehat{F}$  can be described as the space of functions  $u \in L^2([0, T] \times R^N)$  such that  $u_t \in F$  for almost every  $t \in [0, T]$ ,  $\int_0^T e(u_t)dt < \infty$ , and such that  $u_t \in L^2$  for any  $t \in [0, T]$  and the map  $t \rightarrow u_t$  is  $L^2$ -continuous on  $[0, T]$ . This follows by approximating any such function with a sequence  $(u_n)$  defined by*

$$u_{n,t} = n \int_0^{1/n} \tilde{u}_{t+s} ds, \quad t \in [0, T]$$

*the function under the integral being a suitable extension of  $u$  on a larger interval  $[0, T + \varepsilon]$  such that  $u_t \in F$  for almost every  $t \in [T, T + \varepsilon]$  (for example  $\tilde{u}_t = u_{2T-t}$ , for  $t > T$ ).*

Note that  $C_c^\infty([0, T] \times R^N)$  is not necessarily dense in  $\widehat{F}$  with respect to  $\|\cdot\|_T$ , but in a weaker sense given in the following lemma.

**Lemma 2.2** *For every  $u \in \widehat{F}$ , there exists a sequence  $u^k \in C_c^\infty([0, T] \times R^N)$ ,  $k \in N$ , such that  $\int_0^T e_1(u_t - u_t^k)dt \rightarrow 0$ .*

**Proof.** Consider the scalar product  $\bar{e}_1(u, v) = \int_0^T e_1(u_t, v_t)dt$  and let  $Q$  be the closure of  $C_c^\infty([0, T] \times R^N)$  with respect to the norm associated to  $\bar{e}_1$ . Since  $e$  is closable,  $\bar{e}_1$  is closable also (i.e. if  $u^k \rightarrow 0$  in  $L^2([0, 1] \times R^N)$  and  $(u^k)_k$  is Cauchy with respect to the norm induced by  $\bar{e}_1$ , then  $u^k \rightarrow 0$  with respect to that norm) so that  $Q \subseteq H =: \{u \in L^2([0, 1] \times R^N) : u_t \in F, \text{ for almost every } t \text{ and } \int_0^T e(u_t)dt < \infty\}$ .

Let us prove that  $Q = H$ . It suffices to check that if  $w \in H$  and  $w \perp Q$  (in the sense of  $\bar{e}_1$ ) then  $w = 0$ . Assume that  $\int_0^T e_1(w_t, \phi_t)dt = 0, \forall \phi \in C_c^\infty([0, T] \times R^N)$ . Replacing  $\phi_t$  by  $\alpha_t \phi_t$  with  $\alpha \in C_c^\infty[0, T]$  we conclude that for almost every  $t$ ,  $e_1(w_t, \psi) = 0, \forall \psi \in C_c^\infty(R^N)$ . Since  $w_t \in F$  this yields  $w_t = 0$ .

Finally, since  $\widehat{F} \subseteq H$  the proof is complete.  $\square$

## 2.1 Linear Equations

For  $f \in L^2([0, T] \times R^N)$  and  $\Phi \in L^2(R^N)$  we consider the following parabolic equation on  $[0, T] \times R^N$  with terminal condition at time  $T$  :

$$(\partial_t + L)u + f = 0, \quad u_T = \Phi. \quad (4)$$

In the case  $a \geq cI$ , with  $c > 0$ , the treatment of this equation is classical (see [7]). In what follows we slightly modify the classical arguments so as to cover the degenerate case. The semigroup  $(P_t)_{t \geq 0}$ , whose existence is assumed, will be used in an essential way.

**Definition 2.3 :** A function  $u \in \widehat{F}$  is called a strong solution of equation (4) with data  $(\Phi, f)$  if  $t \rightarrow u_t = u(t, \cdot)$  is  $L^2$ -differentiable on  $[0, T]$ ,  $\partial_t u_t \in L^2([0, T] \times R^N)$ ,  $u_t \in D(L)$  for almost every  $t \in [0, T]$ , and the equalities in (4) hold almost everywhere.

A function  $u \in \widehat{F}$  is called a weak solution of equation (4) if the following relation holds

$$\int_0^T ((u_t, \partial_t \varphi) + e(u_t, \varphi_t)) dt = \int_0^T (f_t, \varphi_t) dt + (\Phi, \varphi_T) \quad (5)$$

for every  $\varphi \in C_{T,0}$ .

One can see by direct verification that any strong solution is a weak solution. We note also that the relation defining a weak solution is equivalent to

$$\int_{t_0}^T ((u_t, \partial_t \varphi) + e(u_t, \varphi_t)) dt = \int_{t_0}^T (f_t, \varphi_t) dt + (\Phi, \varphi_T) - (u_{t_0}, \varphi_{t_0}), \quad (6)$$

for every  $t_0 \in [0, T]$  and every  $\varphi \in C_T$ , as one can directly check. Below we give sufficient conditions for the existence and uniqueness of a strong solution.

**Proposition 2.4** i) If  $\Phi \in L^2(R^N)$  then  $t \rightarrow P_{T-t}\Phi$  is  $L^2$ -continuous on  $[0, T]$ ,  $L^2$ -differentiable on  $[0, T]$  and  $\partial_t P_{T-t}\Phi = -LP_{T-t}\Phi$ .

ii) Let  $f : [0, T] \times R^N \rightarrow R$  be a function such that  $t \rightarrow f_t$  is differentiable in  $L^2$  and  $t \rightarrow \partial_t f_t$  is  $L^2$ -continuous on  $[0, T]$ . Then the function  $w(t, x) =: \int_t^T P_{s-t} f_s(x) ds$  is  $L^2$ -differentiable on  $[0, T]$  and

$$\partial_t w(t, x) = -P_{T-t} f_T(x) + \int_t^T P_{s-t} \partial_s f_s(x) ds.$$

Moreover,  $t \rightarrow \partial_t w(t, x)$  is  $L^2$ -continuous on  $[0, T]$ .

iii) Assume that  $\Phi \in L^2(R^N)$ ,  $f$  satisfies the hypothesis in ii) and define

$$u(t, x) = P_{T-t}\Phi(x) + \int_t^T P_{s-t} f_s(x) ds.$$

Then  $u$  is a strong solution of (4).

**Proof.** i) The assertion is a consequence, among other things, of the fact that  $P_t \Phi \in \mathcal{D}(L)$  for each  $t > 0$ .

ii) In order to compute the derivative  $\partial_t w$  one writes

$$\begin{aligned} w_{t+r} - w_t &= \int_0^{T-t-r} P_l f_{t+r+l} dl - \int_0^{T-t} P_l f_{t+l} dl \\ &= \int_0^{T-t} P_l (f_{t+r+l} - f_{t+l}) - \int_{T-t-r}^{T-t} P_l f_{t+r+l} dl. \end{aligned}$$

Then one deduces that

$$\frac{1}{r} (w_{t+r} - w_t) \rightarrow \int_t^T P_{s-t} \partial_s f_s ds - P_{T-t} f_T.$$

Let us now prove iii). By i)  $LP_{s-t} f_s = \partial_s (P_{s-t} f_s) - P_{s-t} \partial_s f_s$  and further, by i) and ii),

$$\begin{aligned} Lu_t &= LP_{T-t} \Phi + \int_t^T LP_{s-t} f_s ds \\ &= -\partial_t P_{T-t} \Phi - f_t + P_{T-t} f_T - \int_t^T P_{s-t} \partial_s f_s ds = -\partial_t u_t - f_t. \end{aligned}$$

□

**Proposition 2.5** Assume that  $f \in L^2([0, T] \times \mathbb{R}^N)$  and  $\Phi \in L^2(\mathbb{R}^N)$ . Then the equation (4) has a unique weak solution  $u \in \hat{F}$  and it is given by

$$u(t, x) = P_{T-t} \Phi(x) + \int_t^T P_{s-t} f_s(x) ds. \quad (7)$$

Moreover, the solution satisfies the following relations

$$\|u_t\|_2^2 + 2 \int_t^T e(u_s) ds = 2 \int_t^T (f_s, u_s) ds + \|\Phi\|_2^2, \quad 0 \leq t \leq T \quad (8)$$

$$\|u\|_T^2 \leq e^T \left( \|\Phi\|_2^2 + \|f\|_{L^2([0, T] \times \mathbb{R}^N)}^2 \right). \quad (9)$$

**Proof** Uniqueness. Let  $u = v - w$  where  $v, w \in \hat{F}$  are two weak solutions of (5). Then  $u$  satisfies

$$\int_{t_0}^T ((u_t, \partial_t \varphi) + e(u_t, \varphi_t)) dt = -(u_{t_0}, \varphi_{t_0}), \quad \forall t_0 \geq 0, \varphi \in C_T. \quad (*)$$

Define

$$u^\varepsilon(t, x) = \frac{1}{\varepsilon} \int_0^\varepsilon u_{t+s}(x) ds$$

with the convention  $u(t, x) = 0$  for  $T \leq t \leq T + \varepsilon$ . We check that  $u^\varepsilon$  fulfills (\*) also. We denote  $\varphi_r^s =: \varphi_{r-s}$  and employ (\*) in order to get

$$\begin{aligned}
 & \int_{t_0}^T (u_t^\varepsilon, \partial_t \varphi_t) dt = \\
 &= \frac{1}{\varepsilon} \int_0^\varepsilon \int_{t_0}^T (u_{t+s}, \partial_t \varphi_t) dt ds = \frac{1}{\varepsilon} \int_0^\varepsilon \int_{t_0+s}^T (u_t, \partial_t \varphi_t^s) dt ds \\
 &= -\frac{1}{\varepsilon} \int_0^\varepsilon \left[ \int_{t_0+s}^T e(u_t, \varphi_t^s) dt + (u_{t_0+s}, \varphi_{t_0+s}^s) \right] ds = \\
 &= -\frac{1}{\varepsilon} \int_0^\varepsilon \left[ \int_{t_0}^T e(u_{t+s}, \varphi_t) dt + (u_{t_0+s}, \varphi_{t_0}) \right] ds = \\
 &= -\int_{t_0}^T e\left(\frac{1}{\varepsilon} \int_0^\varepsilon u_{t+s} ds, \varphi_t\right) dt - \left(\frac{1}{\varepsilon} \int_0^\varepsilon u_{t_0+s} ds, \varphi_{t_0}\right) = \\
 &= -\int_{t_0}^T e(u_t^\varepsilon, \varphi_t) dt - (u_{t_0}^\varepsilon, \varphi_{t_0}).
 \end{aligned}$$

Since  $t \rightarrow u_t$  is continuous in  $L^2$ , it follows that  $t \rightarrow u_t^\varepsilon$  is differentiable in  $L^2$  and  $\partial_t u_t^\varepsilon$  is continuous. So, the function  $u^\varepsilon$  is in  $C_T$  and we may write

$$\int_{t_0}^T (u_t^\varepsilon, \partial_t u_t^\varepsilon) + e(u_t^\varepsilon, u_t^\varepsilon) dt = - (u_{t_0}^\varepsilon, u_{t_0}^\varepsilon).$$

Since  $\partial_t (u_t^\varepsilon, u_t^\varepsilon) = 2(u_t^\varepsilon, \partial_t u_t^\varepsilon)$ , it follows that

$$\int_{t_0}^T (u_t^\varepsilon, \partial_t u_t^\varepsilon) dt = \frac{1}{2} \int_{t_0}^T \partial_t (u_t^\varepsilon, u_t^\varepsilon) dt = -\frac{1}{2} (u_{t_0}^\varepsilon, u_{t_0}^\varepsilon)$$

and we get  $\frac{1}{2}(u_{t_0}^\varepsilon, u_{t_0}^\varepsilon) + \int_{t_0}^T e(u_t^\varepsilon, u_t^\varepsilon) dt = 0$ , that is  $u_{t_0}^\varepsilon = 0$ . Since  $u_{t_0} = \lim_{\varepsilon} u_{t_0}^\varepsilon$  in  $L^2$  the uniqueness follows.

Existence. Let us first consider the case when  $f$  satisfies the conditions ii) of the preceding proposition and  $\Phi \in \mathcal{D}(L)$ . Then we know that the solution is given by the relation (7). Moreover, since  $\Phi \in \mathcal{D}(L)$ , one has that  $u \in C_T$ , and hence this function may be introduced in the relation(6) as a test function. This establishes the relation (8). In order to deduce the last estimate in the statement, one uses (8) and the following consequence of Schwartz's inequality

$$\int_t^T (f_s, u_s) ds \leq \frac{1}{2} \int_t^T \|f_s\|_2^2 ds + \frac{1}{2} \int_t^T \|u_s\|_2^2 ds.$$

Then (9) follows from Gronwall's lemma.

In order to obtain the result for general data  $f, \Phi$ , one takes a sequence of functions  $(f_k)_k \subset C_c^\infty([0, T] \times R^N)$  such that  $f_k \rightarrow f$  in  $L^2([0, T] \times R^N)$  and a sequence of functions

$(\Phi_k)_k \subset D(L)$  such that  $\Phi_k \rightarrow \Phi$  in  $L^2(R^N)$ . Let  $u_k$  denote the solution corresponding to the data  $(f_k, \Phi_k)$ . Then the relations (8) and (9) can be used to deduce that  $(u_k)$  is a Cauchy sequence in  $\widehat{F}$ . The limit  $u = \lim_k u_k$  is the solution corresponding to the data  $(f, \Phi)$  and it satisfies the relations (8) and (9).  $\square$

### 3 The Nonlinear Equation

In this section we deal with the nonlinear version of equation (4), which takes the form

$$(\partial_t + L)u(t, x) + f(t, x, u(t, x), \nabla u(t, x)\sigma) = 0, \quad u(T, x) = \Phi(x). \quad (10)$$

We look for a weak solution of this equation in the space  $\widehat{F}$ . Note that if  $a \geq cI$ , then  $F = H^1$  and, since  $u_t \in F$  for almost any  $t$ , the gradient  $\nabla u(t, x)$  is well defined in the above equation. But in our general situation this is not necessarily true, so that we first have to give a sense to  $\nabla u\sigma$ . Set  $D_\sigma\varphi =: \nabla\varphi\sigma$  for any  $\varphi \in C_c^\infty(R^N)$ , define  $V_0 = \{D_\sigma\varphi : \varphi \in C_c^\infty(R^N)\}$ , and let  $V$  be the closure of  $V_0$  in  $L^2(R^N; R^n)$ .

**Proposition 3.1** *i) For every  $u \in F$  there is a unique element of  $V$ , which we denote by  $D_\sigma u$  such that*

$$e(u, \varphi) = \int_{R^N} \langle D_\sigma u(x), D_\sigma \varphi(x) \rangle dx, \quad \forall \varphi \in C_c^\infty(R^N).$$

*One has  $D_\sigma u\tau\sigma = D_\sigma u$ . Moreover for  $u, v \in F$*

$$e(u, v) = \int_{R^N} \langle D_\sigma u(x), D_\sigma v(x) \rangle dx. \quad (**)$$

*ii) Furthermore, if  $u \in \widehat{F}$ , there exists a measurable function  $\phi : [0, T] \times R^N \rightarrow R^N$  such that  $|\phi\sigma| \in L^2([0, T] \times R^N)$  and  $D_\sigma u_t = \phi_t\sigma$  for almost all  $t \in [0, T]$ .*

*iii) Let  $u^k, u \in \widehat{F}$ ,  $k \in N$ , such that  $u^k \rightarrow u$  in  $L^2([0, T] \times R^N)$  and  $(D_\sigma u^k)_k$  is Cauchy in  $L^2([0, T] \times R^N)$ . Then  $D_\sigma u^k \rightarrow D_\sigma u$  in  $L^2([0, T] \times R^N)$ , i.e.  $D_\sigma$  is closable.*

**Proof.** Uniqueness. Let  $v, w \in V$  which verify the relation in the statement, i.e. they may replace  $D_\sigma u$  in the right hand side of the first relation in the statement. Then  $\int \langle v - w, D_\sigma \varphi \rangle dx = 0$  for every  $\varphi \in C_c^\infty$  and, since  $v - w \in \overline{V}_0$ , this yields  $v = w$ .

Existence. We treat directly the time dependent case: for a given  $u \in \widehat{F}$  we construct a measurable function  $\phi : [0, T] \times R^N \rightarrow R^N$  such that  $|\phi\sigma| \in L^2([0, T] \times R^N)$  and, for almost every  $t$ ,  $\phi_t\sigma = D_\sigma u_t$ .

Let  $u^k \in C_c^\infty([0, T] \times R^N)$ ,  $k \in \mathbf{N}$ , be such that  $\bar{\epsilon}_1(u^k - u) \rightarrow 0$  (see Lemma2.2) and let  $\phi^k = \nabla u^k$ . We have

$$e(u_t, \varphi) = \int_{R^N} \langle D_\sigma u_t(x), D_\sigma \varphi(x) \rangle dx, \quad \forall \varphi \in C_c^\infty(R^N). \quad (*)$$

so we may define  $\psi = \lim_k \phi^k \sigma$  and  $\phi = \psi \tau$  (the matrix  $\tau$  is associated to  $\sigma$  and has been introduced at the beginning of the paper). Since  $\sigma \tau \sigma = \sigma$  and  $\|\tau \sigma\| \leq 1$ , we have  $D_\sigma u^k = \phi^k \sigma = \phi^k \sigma \tau \sigma \rightarrow \psi \tau \sigma = \phi \sigma$  in  $L^2([0, T] \times R^N; R^n)$ . Passing to a subsequence we may find a set  $\Lambda \subset [0, T]$  such that  $[0, T] \setminus \Lambda$  is negligible and for every  $t \in \Lambda$ ,  $e(u_t^k - u_t) \rightarrow 0$  and  $\|(\phi_t^k - \phi_t) \sigma\|_2 \rightarrow 0$ . We fix  $t \in \Lambda$  and write  $(*)$  for  $u_t^k$  and  $\phi_t^k$ ,

$$e(u_t^k, \varphi) = \int_{R^N} \langle \phi_t^k \sigma(x), D_\sigma \varphi(x) \rangle dx, \quad \forall \varphi \in C_c^\infty(R^N).$$

Passing to the limit we get the relation for  $u_t$  and  $\phi_t$ .

Let us now prove iii). Let  $v = \lim_k D_\sigma u^k$ . Passing to a subsequence we may assume that for almost every  $t \in [0, T]$ ,  $\|v_t - D_\sigma u_t^k\|_2 \rightarrow 0$ . We fix such a  $t$ . We take  $\varphi \in \mathcal{D}(L) \subset F$  and we write

$$\begin{aligned} (v_t, D_\sigma \varphi) &= \lim_k (D_\sigma u_t^k, D_\sigma \varphi) = \lim_k e(u_t^k, \varphi) = -\lim_k (u_t^k, L\varphi) = \\ &= -(u_t, L\varphi) = e(u_t, \varphi) = (D_\sigma u_t, D_\sigma \varphi). \end{aligned}$$

Since  $\mathcal{D}(L)$  is dense in  $F$  with respect to  $e$ ,  $\{D_\sigma \varphi : \varphi \in \mathcal{D}(L)\}$  is dense in  $V$ . It follows that  $v_t = D_\sigma u_t$ .  $\square$

**NOTATION:** We will preserve in the rest of the paper the notation  $D_\sigma u$ , introduced in this proposition for  $u \in F$ . Moreover, we denote by  $\tilde{\nabla} u$  the set of all measurable functions  $\phi : R^N \rightarrow R^N$ , such that  $\phi \sigma = D_\sigma u$  as elements of  $L^2(R^N, R^n)$ . If  $u \in \hat{F}$ , then the same notation  $\tilde{\nabla} u$  will denote the set of all measurable functions  $\phi : [0, T] \times R^N \rightarrow R^N$  such that  $|\phi \sigma| \in L^2([0, T] \times R^N)$  and  $D_\sigma u_t = \phi_t \sigma$  for almost all  $t \in [0, T]$ .

**Remark 3.2** *If  $a \geq cI$  on some open set  $D \subset R^N$  and  $u \in F$ , then  $\nabla u$  exists on  $D$  (in the  $L^2$  sense) and  $\nabla u \sigma = D_\sigma u$  on  $D$ . In the degenerate case the set  $\tilde{\nabla} u$  may contain more than one element. Obviously  $D_\sigma u \tau \in \tilde{\nabla} u$ .*

### 3.1 The case of Lipschitz conditions

Consider now a measurable function  $f : [0, T] \times R^N \times R \times R^n \rightarrow R$  such that

$$(t, x) \rightarrow f(t, x, 0, 0) \text{ is in } L^2([0, T] \times R^N), \quad (11)$$

$$|f(t, x, y, z) - f(t, x, y', z')| \leq C(|y - y'| + |z - z'|) \quad (12)$$

and a measurable function  $\Phi \in L^2(R^N)$ . With the above notation, equation (10) may be more appropriately written as

$$(\partial_t + L)u + f(\cdot, \cdot, u, D_\sigma u) = 0, \quad u_T = \Phi. \quad (13)$$

In fact, we are going to treat the weak form of this equation, which can be written as

$$\int_0^T [(u_t, \partial_t u_t) + e(u_t, \varphi_t)] dt = \int_0^T (f(t, \cdot, u_t, D_\sigma u_t), \varphi_t) dt + (\Phi, \varphi_T), \quad \forall \varphi \in C_{T,0}. \quad (14)$$

We also consider the *mild* equation (which corresponds to the explicit formula (7) in the case of linear equations):

$$u(t, x) = P_{T-t}\Phi(x) + \int_t^T P_{s-t}f(s, \cdot, u_s, D_\sigma u_s)(x)ds, \quad dx - a.s. \quad (15)$$

The meaning of the above equations is given as follows. If  $u \in \widehat{F}$ , then  $u, |D_\sigma u| \in L^2([0, T] \times R^N)$  and, on account of assumptions (11) and (12), one deduces that the function

$$f_u(t, x) := f(t, x, u(t, x), D_\sigma u(t, x))$$

belongs to  $L^2([0, T] \times R^N)$ . Having this in mind, we should look at equation (14) as being nothing else but equation (5) written with  $f_u$ . Similarly, equation (15) should be understood with  $f_u$  under  $P_{s-t}$ . Then, by Proposition 2.5, a function  $u \in \widehat{F}$  is a weak solution of the nonlinear equation (14) if and only if it solves the mild equation (15).

**Theorem 3.3** *Assume that  $f$  satisfies conditions (11) and (12) and that  $\Phi \in L^2(R^N)$ . Then equation (14) (respectively (15)) admits a unique weak (resp. mild) solution  $u \in \widehat{F}$ . The solution satisfies the following estimate (with  $C$  the Lipschitz constant in (12)),*

$$\|u\|_T^2 \leq e^{T(1+2C+C^2)} \left( \|\Phi\|_2^2 + \|f(\cdot, \cdot, 0, 0)\|_{L^2([0, T] \times R^N)}^2 \right).$$

**Proof.** We define the operator  $A : \widehat{F} \rightarrow \widehat{F}$  by

$$(Au)(t, x) = P_{T-t}\Phi(x) + \int_t^T P_{s-t}f(s, \cdot, u_s, D_\sigma u_s)(x)ds.$$

By Proposition 2.5 we know that  $Au \in \widehat{F}$ . We shall prove that, if  $T$  is sufficiently small, then  $A$  is a contraction with respect to  $\|\cdot\|_T$  and so the existence and uniqueness of the solution of (15) follows - a standard recurrence procedure permits then to extend the result to any  $T$ . First of all

$$\begin{aligned} [e(Au_t - Av_t)]^{1/2} &= [e(\int_t^T P_{s-t}(f_{u,s} - f_{v,s})ds)]^{1/2} \leq \int_t^T [e(P_{s-t}(f_{u,s} - f_{v,s}))]^{1/2} ds \leq \\ &\leq \int_t^T \|f_{u,s} - f_{v,s}\|_2 \frac{ds}{2\sqrt{s-t}}, \end{aligned}$$

so by using the hypothesis (12)

$$\begin{aligned}
 & \int_0^T e(Au_t - Av_t) dt \\
 \leq & \int_0^T \left( \int_t^T \|f_{u,s} - f_{v,s}\|_2 \frac{ds}{2\sqrt{s-t}} \right)^2 dt \leq \int_0^T dt \sqrt{T-t} \int_t^T \|f_{u,s} - f_{v,s}\|_2^2 \frac{ds}{2\sqrt{s-t}} = \\
 = & \int_0^T ds \|f_{u,s} - f_{v,s}\|_2^2 \int_0^s \frac{\sqrt{T-t}}{2\sqrt{s-t}} dt \leq KT \int_0^T ds \|f_{u,s} - f_{v,s}\|_2^2 \leq \\
 \leq & KT \int_0^T \|u_s - v_s\|_2^2 + \|D_\sigma u_s - D_\sigma v_s\|_2^2 ds = KT \int_0^T \left( \|u_s - v_s\|_2^2 + e(u_s - v_s) \right) ds \leq \\
 \leq & KT \|u - v\|_T^2.
 \end{aligned}$$

A similar but simpler calculation shows that  $\|Au_t - Av_t\|_2^2 \leq KT \|u - v\|_T^2$  and so  $\|Au - Av\|_T^2 \leq KT \|u - v\|_T^2$ . Now we take  $T$  sufficiently small in order that  $KT < 1$  and the proof of existence and uniqueness is finished. It remains to prove the estimate in the statement. We write

$$\begin{aligned}
 & \left| \int_t^T (f_{u,s}, u_s) ds \right| \leq \\
 \leq & \int_t^T |(f_s(\cdot, 0, 0), u_s)| ds + C \int_t^T \|u_s\|_2^2 ds + C \int_t^T \|D_\sigma u_s\|_2 \|u_s\|_2 ds \leq \\
 \leq & \frac{1}{2} \int_t^T \|f_s(\cdot, 0, 0)\|_2^2 ds + \left( \frac{1}{2} + C + \frac{1}{2}C^2 \right) \int_t^T \|u_s\|_2^2 ds + \frac{1}{2} \int_t^T e(u_s) ds.
 \end{aligned}$$

Then relation (8) gives

$$\|u_t\|_2^2 + \int_t^T e(u_s) ds \leq \|\Phi\|_2^2 + \int_t^T \|f_s(\cdot, 0, 0)\|_2^2 ds + (1 + 2C + C^2) \int_t^T \|u_s\|_2^2 ds,$$

and Gronwall's lemma yields the inequality of the statement.  $\square$

**Remark 3.4** *The spaces  $C_T$  and  $\hat{F}$  introduced above may be defined in a more general set-up, for arbitrary Dirichlet spaces. Moreover the content of these first sections remain valid for other Dirichlet spaces (e.g. for the Dirichlet space corresponding to the infinite dimensional Ornstein - Uhlenbeck process).*

### 3.2 Monotonicity conditions

The nonlinear equation we are going to treat now is determined by a measurable function  $f : [0, T] \times R^N \times R \times R^n \rightarrow R$ , which satisfies the Lipschitz condition with respect to  $z$ ,

$$|f(t, x, y, z) - f(t, x, y, z')| \leq C |z - z'|, \tag{16}$$



with  $t, x, y, z, z'$  arbitrary and  $C > 0$  a constant, but instead of the Lipschitz condition with respect to  $y$ , it is assumed to satisfy some weaker conditions. First one assumes that for any fixed triple  $t, x, z$  the map  $y \rightarrow f(t, x, y, z)$  is continuous. The main substitute for Lipschitz continuity is the following monotonicity condition

$$(y - y')(f(t, x, y, z) - f(t, x, y', z)) \leq \mu (y - y')^2, \quad (17)$$

with arbitrary  $t, x, y, y', z$  and a fixed constant  $\mu \geq 0$ . We will prove existence and uniqueness for the nonlinear equation with functions  $f$  satisfying the above and other supplementary integrability and boundedness conditions. In order to treat this case, it turns out that another functional analytic frame is needed. In fact we need a slight generalization of the preceding, which is next introduced.

Let  $\theta \in C^1(\mathbb{R}^N)$  be such that  $0 \leq \theta(x) \leq |x|$ , and  $\theta(x) = |x|$ , if  $|x| \geq 1$  and define for  $\rho \in \mathbb{R}_+$ ,  $\pi_\rho(x) = \exp[-\rho\theta(x)]$ . These functions may be taken as weights and modify the functional spaces previously introduced as follows. Instead of the Lebesgue measure we will use the measure  $m_\rho(dx) = \pi_\rho(x) dx$ . Consider  $L_\rho^2 = L^2(\mathbb{R}^N, dm_\rho)$  with the scalar product  $(u, v)_\rho = \int uv dm_\rho$  and the corresponding norm  $\|\cdot\|_\rho$ . Similarly define the weighted energy  $e_\rho(u, v) = \int \langle D_\sigma u, D_\sigma v \rangle dm_\rho$  and the space  $F_\rho$  which is the completion of  $C_c(\mathbb{R}^N)$  with respect to  $e_{\rho,1}(\cdot, \cdot) = (\cdot, \cdot)_\rho + e_\rho(\cdot, \cdot)$ . Then define the time-space norm  $\|u\|_T^2 = \sup_{t \leq T} \|u_t\|_\rho^2 + \int_0^T e_\rho(u_t) dt$  and the space  $\widehat{F}_\rho$  of time-space functions obtained as the closure of the space  $\mathcal{C}_{\rho,T}$  of all functions  $u : [0, T] \rightarrow F_\rho$ , which are  $L_\rho^2$ -differentiable with  $L_\rho^2$ -continuous derivative, and such that  $\int_0^T e_\rho(u_t) dt < \infty$  (analogous to  $\mathcal{C}_T$ ). One should mention that our assumption concerning closability of the form  $(C_c(\mathbb{R}^N), e_1)$  implies that for each  $\rho > 0$ , the form  $(C_c(\mathbb{R}^N), e_{\rho,1})$  is also closable. Also note that if  $u \in F_\rho$ , then  $u^+ \in F_\rho$  and it is not difficult to show that  $D_\sigma u^+ = 1_{\{u>0\}} D_\sigma u$  a.e.

We denote by  $(P_{\rho,t})_{t \geq 0}$  the semigroup associated to the Dirichlet form  $e_\rho$ , and  $L_\rho$  its generator. Clearly, one has  $L_\rho = L - B_\rho$ , where  $B_\rho u = \rho \langle D_\sigma \theta, D_\sigma u \rangle$  is a mapping  $B_\rho : F_\rho \rightarrow L_\rho^2$  that satisfies  $\|B_\rho u\|_\rho \leq c \rho e_\rho(u)^{1/2}$ ,  $u \in F_\rho$ , with a constant  $c$  which depends only of our function  $\theta$  and of the matrix  $\sigma$ . When  $\rho = 0$ , of course one has  $m_\rho = m$ ,  $B_\rho = 0$  and everything is as in the previous section. The linear evolution equation which is naturally associated to  $L_\rho$  takes the form

$$(\partial_t + L_\rho)u + f = 0, \quad u_T = \Phi, \quad (18)$$

with  $f \in L^2([0, T] \times \mathbb{R}^N, dt \otimes dm_\rho)$  and  $\Phi \in L_\rho^2$ . The weak solution is a function  $u \in \widehat{F}_\rho$  satisfying the relation

$$\int_{t_0}^T ((u_t, \partial_t \varphi)_\rho + e_\rho(u_t, \varphi_t)) dt = \int_{t_0}^T (f_t, \varphi_t)_\rho dt + (\Phi, \varphi_T)_\rho - (u_{t_0}, \varphi_{t_0})_\rho,$$

for all  $\varphi \in \mathcal{C}_{\rho,T}$ . The statement of Proposition 2.5 remains valid in the sense that everything should be modified with a  $\rho$ -parameter.

One technical ingredient for the treatment of the monotonicity condition is the following lemma.

**Lemma 3.5** *If  $(f, \Phi)$  satisfy the above mentioned integrability conditions,  $u$  is a weak solution of equation (18),  $u^+(t, x) = \max(u(t, x), 0)$  and  $0 \leq t_1 < t_2 \leq T$ , then one has*

$$\|u_{t_1}^+\|_\rho^2 + 2 \int_{t_1}^{t_2} e_\rho(u_s^+) ds = 2 \int_{t_1}^{t_2} (f_s, u_s^+)_\rho ds + \|u_{t_2}^+\|_\rho^2.$$

**Proof.** In order to deduce the relation of the statement one may say that roughly,  $u^+$  has to be introduced as a test function in the weak equation. Since the positive part of  $u$ , the function  $u^+$ , is not necessarily a true test function, we will use an approximation of it as follows. Let  $h$  be defined by

$$h(t) = 0, \text{ if } |t| \geq 1, \quad h(t) = t + 1, \text{ if } t \in (-1, 0), \quad h(t) = 1 - t, \text{ if } t \in (0, 1)$$

Then set  $h_k(t) = kh(kt)$ ,  $k \in \mathbb{N}$ . One obviously has

$$\int h = 1, \quad \int h_k = 1, \quad h' = 1_{(-1,0)} - 1_{(0,1)}, \quad h'_k(t) = k^2 h'(kt).$$

We define

$$u_{k,t} = \int u_s^+ h_k(t-s) ds,$$

which clearly make sense for large  $k$  if  $t \in (0, T)$ . The function  $u_k$  may be used as test function over an interval  $(t_1, t_2)$  with  $0 < t_1 < t_2 < T$ . Therefore we may write

$$\int_{t_1}^{t_2} [(u_s, \partial_s u_{k,s})_\rho + e_\rho(u_s, u_{k,s})] ds = \int_{t_1}^{t_2} (f_s, u_{k,s})_\rho ds + (u_{t_2}, u_{k,t_2})_\rho - (u_{t_1}, u_{k,t_1})_\rho$$

and pass to the limit in this relation. The only thing that remains to be proved is that

$$\lim_{k \rightarrow \infty} \int_{t_1}^{t_2} (u_t, \partial_t u_{k,t})_\rho dt = \frac{1}{2} [\|u_{t_2}^+\|_\rho - \|u_{t_1}^+\|_\rho].$$

In order to check this relation we will write  $v_{k,t} = \partial_t u_{k,t}$  and observe that

$$v_{k,t} = \int u_s^+ h'_k(t-s) ds.$$

So, the expression we have to examine may be written as follows

$$\int_{t_1}^{t_2} \int (u_t, u_s^+)_\rho h'_k(t-s) ds dt.$$

Now we assert that

$$\lim_k \int_{t_1}^{t_2} \int (u_t^-, u_s^+)_\rho h'_k(t-s) ds dt = 0. \quad (*)$$

Indeed, the simple inequality  $u_t^- u_s^+ \leq (u_t - u_s)^2$ , reduces the problem to the estimation of  $\|u_t - u_s\|_2^2$ .

At this point we make the supplementary assumption that the function  $f$  satisfies the assumptions of Proposition 2.4 (that is  $t \rightarrow f_t$  is  $L^2$ -differentiable and  $t \rightarrow \partial_t f$  is  $L^2$ -continuous). One can easily see that there is no real loss of generality in doing so, because once the lemma is proved, the function  $f$  may be approximated by such functions and the relation of the lemma passes to the limit.

Then, on account of assertions i) and ii) of Proposition 2.4, the function  $t \rightarrow u_t$  is  $L^2$ -differentiable with bounded derivative, and so one has  $\|u_t - u_s\|_2^2 \leq \text{const.} |t - s|^2$ . This estimate together with the above mentioned properties of  $h'_k$  imply relation (\*).

On the other hand, for any continuous symmetric function  $j : [0, T] \times [0, T] \rightarrow R$  one has

$$\lim_k \int_{t_1}^{t_2} \int j(t, s) h'_k(t - s) ds dt = \frac{1}{2} [j(t_2, t_2) - j(t_1, t_1)],$$

which leads to

$$\lim_k \int_{t_1}^{t_2} \int (u_t^+, u_s^+)_{\rho} h'_k(t - s) ds dt = \frac{1}{2} [\|u_{t_2}^+\|_{\rho} - \|u_{t_1}^+\|_{\rho}].$$

The proof is complete.  $\square$

The nonlinear equation with parameter  $\rho$  corresponding to the above linear equation looks as follows

$$(\partial_t + L_{\rho})u + f(\cdot, \cdot, u, D_{\sigma}u) = 0, \quad u_T = \Phi, \quad (19)$$

with the function  $\Phi$  taken in  $L_{\rho}^2$ . In the case of Lipschitz conditions with respect to  $y$  and  $z$  one treats this equation exactly as we did in the case with  $\rho = 0$ . In the case of monotonicity condition the contraction principle works similarly, except that now the main problem is to ensure that  $f(\cdot, \cdot, u, D_{\sigma}u)$  belongs to  $L^2(dt \otimes dm_{\rho})$ , in order for the notion of weak solution to make sense, and then to be able to yield by iteration a sequence to which one can apply the contraction argument. The main point is the following comparison result.

**Proposition 3.6** *Let  $f^i : [0, T] \times R^N \times R \times R^n \rightarrow R$ , and  $u_i \in \widehat{F}_{\rho}$ ,  $i = 1, 2$  be such that  $f^1(\cdot, \cdot, u^1, D_{\sigma}u^1), f^2(\cdot, \cdot, u^2, D_{\sigma}u^2) \in L^2([0, T] \times R^N, dt \otimes dm_{\rho})$ . Assume that  $f^1$  satisfies the conditions (17) and (16) and that the following inequality holds  $f^1(\cdot, \cdot, u^2, D_{\sigma}u^2) \leq f^2(\cdot, \cdot, u^2, D_{\sigma}u^2)$ . If  $u^i$  is a solution of equation (19) with function  $f^i$  and final condition  $\Phi^i, i = 1, 2$  and moreover  $\Phi^1 \leq \Phi^2$ , then one has*

$$u^1 \leq u^2.$$

**Proof.** Set  $v = u^1 - u^2$ ,  $\Psi = \Phi^1 - \Phi^2$  and  $g = f^1(\cdot, \cdot, u^1, D_{\sigma}u^1) - f^2(\cdot, \cdot, u^2, D_{\sigma}u^2)$ . Since  $v$  is a solution of equation (18) with data  $(\Psi, g)$  one can apply the preceding lemma and get

$$\|v_t^+\|_{\rho}^2 + 2 \int_t^T e_{\rho}(v_s^+) ds = 2 \int_t^T (f_s, v_s^+)_{\rho} ds + \|v_T^+\|_{\rho}^2.$$

Then we use the conditions of our statement and get

$$\begin{aligned} f v^+ &= (f^1(\cdot, \cdot, u^1, D_\sigma u^1) - f^1(\cdot, \cdot, u^2, D_\sigma u^1)) v^+ + \\ &\quad (f^1(\cdot, \cdot, u^2, D_\sigma u^1) - f^1(\cdot, \cdot, u^2, D_\sigma u^2)) v^+ + \\ &\quad (f^1(\cdot, \cdot, u^2, D_\sigma u^2) - f^2(\cdot, \cdot, u^2, D_\sigma u^2)) v^+ \\ &\leq \mu (v^+)^2 + C |D_\sigma v^+| v^+. \end{aligned}$$

Since  $v_T^+ = 0$ , the preceding equality leads to the estimate

$$\|v_t^+\|_\rho^2 + \int_t^T e_\rho(v_s^+) ds \leq (2\mu + C^2) \int_t^T \|v_s^+\|_\rho^2 ds.$$

Gronwall's lemma implies  $v^+ = 0$ .  $\square$

Now, in order to treat the nonlinear equation under the monotonicity condition, we are imposing an integrability condition. Namely we shall assume that for some  $r > 0$  and  $\rho \geq 0$  the function

$$g_r(t, x) = \sup_{|y| \leq r} |f(t, x, y, 0)|, \quad (t, x) \in [0, T] \times R^N \quad (20)$$

belongs to  $L^2(dt \times dm_\rho)$ .

**Theorem 3.7** *Let  $f : [0, T] \times R^N \times R \times R^n \rightarrow R$  be a measurable function which is continuous with respect to  $y$  and satisfies conditions (16) and (17). Assume that the functions  $f^\circ(t, x) = f(t, x, 0, 0)$  and  $\Phi$  are bounded and the function defined in (20) belongs to  $L^2(dt \times dm_\rho)$  for some  $r > 0$  and some  $\rho \geq 0$ . If  $\|\Phi\|_\infty \vee \left( e^{\mu T} \|\Phi\|_\infty + \frac{e^{\mu T} - 1}{\mu} \|f^\circ\|_\infty \right) \leq r$ , then the nonlinear equation (19) has a unique weak solution  $u \in \widehat{F}_\rho$ . This solution satisfies the following estimates,*

$$\|u\|_{\rho, T}^2 \leq K \left[ \|\Phi\|_\rho^2 + \|f^\circ\|_{L^2(dt \otimes m_\rho)}^2 \right], \quad (i)$$

with a constant  $K$ , which depends only of  $T$  and of the constants  $C$  and  $\mu$  from conditions (16) and (17) and

$$\|u\|_\infty \leq \|\Phi\|_\infty \vee \left( e^{\mu T} \|\Phi\|_\infty + \frac{e^{\mu T} - 1}{\mu} \|f^\circ\|_\infty \right). \quad (ii)$$

Before proving the theorem we will establish some lemmas. First we remark that  $\|\Phi\|_\rho \leq K \|\Phi\|_\infty$  and  $\|f^\circ\|_{L^2(dt \otimes m_\rho)} \leq \sqrt{T} K \|f^\circ\|_\infty$  with  $K = \left( \int \pi_\rho(x) dx \right)^{\frac{1}{2}}$ . Note also that the conditions imposed to  $f$  ensure that if  $u \in \widehat{F}_\rho$  and  $\|u\|_\infty \leq r$ , then  $f(\cdot, \cdot, u, D_\sigma u) \in L^2(dt \otimes m_\rho)$ . Indeed this follows from the estimate

$$|f(\cdot, \cdot, u, D_\sigma u)| \leq C |D_\sigma u| + g_r.$$

Now we will prove the estimates from the statement of the theorem in the form of "a priori estimates" in the next two lemmas.

**Lemma 3.8** *Suppose that  $f$  fulfills the conditions (16) and (17). Let  $u \in \widehat{F}_\rho$  be such that  $f(\cdot, \cdot, u, D_\sigma u) \in L^2(dt \otimes m_\rho)$  and so that it is a solution of (19). Then it satisfies the estimate (i) of the theorem.*

**Proof.** One starts with the relation (8), which gives

$$\|u_t\|_\rho^2 + 2 \int_t^T e_\rho(u_s) ds = 2 \int_t^T (f(s, \cdot, u_s, D_\sigma u_s), u_s)_\rho ds + \|\Phi\|_\rho^2$$

and use relations (17) and (16) in order to estimate the integral in the right hand side, where one writes

$$\begin{aligned} f(\cdot, \cdot, u, D_\sigma u) u &= [f(\cdot, \cdot, u, D_\sigma u) - f(\cdot, \cdot, 0, D_\sigma u)] u + \\ &+ [f(\cdot, \cdot, 0, D_\sigma u) - f^\circ] u + f^\circ u. \end{aligned}$$

This leads to

$$\begin{aligned} \|u_t\|_\rho^2 + \int_t^T e_\rho(u_s) ds &\leq (2\mu + C^2 + 1) \int_t^T \|u_s\|_\rho^2 ds + \\ &+ \|\Phi\|_\rho^2 + \|f^\circ\|_{L^2(dt \otimes m_\rho)}^2. \end{aligned}$$

The asserted estimate then follows from Gronwall's lemma.  $\square$

The second estimate of the theorem follows from the next lemma which shows a more precise inequality.

**Lemma 3.9** *Suppose that  $f$  fulfills the conditions (16) and (17). Let  $u \in \widehat{F}_\rho$  be such that  $f(\cdot, \cdot, u, D_\sigma u) \in L^2(dt \otimes m_\rho)$  and so that it is a solution of (19). Assume that  $f^\circ \leq a$  and that  $\Phi \leq b$ , where  $a, b$  are two nonnegative constants. Then one has*

$$u \leq b \vee \left( b e^{\mu T} + \frac{a}{\mu} (e^{\mu T} - 1) \right).$$

If  $|f^\circ| \leq a$  and  $|\Phi| \leq b$ , then

$$|u| \leq b \vee \left( b e^{\mu T} + \frac{a}{\mu} (e^{\mu T} - 1) \right).$$

**Proof.** We will compare the solution  $u$  with the solution of the following ordinary differential equation

$$y' + a + \mu y = 0, \quad y_T = b.$$

A solution of this equation may be viewed as a solution (constant in  $x$ ) of (19), where  $f(t, x, y, z) = a + \mu y$  and  $\Phi \equiv b$ . The solution of the above ordinary differential equation is  $y = \left( b + \frac{a}{\mu} \right) e^{\mu(T-t)} - \frac{a}{\mu}$ . The first estimate now follows from Proposition 3.6.

In order to deduce the second estimate we first note that the monotonicity condition is valid for the function  $(t, x, y, z) \rightarrow -f(t, x, -y, z)$  with the same constant  $\mu$ . Therefore the conclusions deduced from the monotonicity for a solution  $u$ , also hold for  $-u$ . This implies the second inequality.  $\square$

**Proof of the theorem.** *Uniqueness.* Let  $u_1, u_2 \in \widehat{F}_\rho$  be such that  $f(\cdot, \cdot, u_1, D_\sigma u_1), f(\cdot, \cdot, u_2, D_\sigma u_2) \in L^2(dt \times m_\rho)$  and such that equation (19) is satisfied by both  $u_1$  and  $u_2$ . Writing relation (8) for the difference  $u_1 - u_2$  we will get

$$\begin{aligned} & \|u_{1,t} - u_{2,t}\|_\rho^2 + 2 \int_t^T e_\rho(u_{1,s} - u_{2,s}) ds = \\ & = 2 \int_t^T (f(s, \cdot, u_{1,s}, D_\sigma u_{1,s}) - f(s, \cdot, u_{2,s}, D_\sigma u_{2,s}), u_{1,s} - u_{2,s})_\rho ds. \end{aligned}$$

The expression under the integral in the right side is transformed writing  $f(s, \cdot, u_{1,s}, D_\sigma u_{1,s}) - f(s, \cdot, u_{2,s}, D_\sigma u_{2,s})$  in the form

$$f(s, \cdot, u_{1,s}, D_\sigma u_{1,s}) - f(s, \cdot, u_{2,s}, D_\sigma u_{1,s}) + f(s, \cdot, u_{2,s}, D_\sigma u_{1,s}) - f(s, \cdot, u_{2,s}, D_\sigma u_{2,s})$$

and then using (17) and (16) one gets

$$\begin{aligned} & \|u_{1,t} - u_{2,t}\|_\rho^2 + 2 \int_t^T e_\rho(u_{1,s} - u_{2,s}) ds \leq \\ & \leq 2 \int_t^T \mu \|u_{1,s} - u_{2,s}\|_\rho^2 ds + 2 \int_t^T C (|D_\sigma u_{1,s} - D_\sigma u_{2,s}|, |u_{1,s} - u_{2,s}|)_\rho ds \\ & \leq (2\mu + C^2) \int_t^T \|u_{1,s} - u_{2,s}\|_\rho^2 ds + \int_t^T e_\rho(u_{1,s} - u_{2,s}) ds. \end{aligned}$$

Gronwall's lemma now implies that  $u_1 = u_2$ .

*Existence.* The proof of existence will be split into two steps.

*First step.* We first consider the case where the function  $f$  is *bounded*. Then we regularise  $f$  with respect to the variable  $y$ , by convolution

$$f_k(t, x, y, z) = k \int_{\mathbb{R}} f(t, x, y', z) \varphi(k(y - y')) dy',$$

where  $\varphi$  is a smooth nonnegative function of compact support such that  $\int \varphi = 1$ . Then  $f = \lim_k f_k$ , and for each  $k$ , the function  $\partial_y f_k$  is uniformly bounded. Thus the functions  $f_k$  satisfy the Lipschitz condition with respect to both  $y$  and  $z$ , and consequently each determine a solution  $u_k \in \widehat{F}_\rho$ . Since all functions  $f_k$  satisfy the monotonicity condition with the same constant  $\mu$  and  $f_k(\cdot, \cdot, 0, 0) = f^\circ$  are bounded by the same constant, one deduces from Lemma 3.9 that the solutions  $u_k$  are uniformly bounded by  $\|\Phi\|_\infty \vee \left( e^{\mu T} \|\Phi\|_\infty + \frac{e^{\mu T} - 1}{\mu} \|f^\circ\|_\infty \right) \leq r$ .

Now let  $a, b > 0$  and set

$$d_{k,a,b}(t, x) = \sup_{|y| \leq a, |z| \leq b} |f(t, x, y, z) - f_k(t, x, y, z)|.$$

Obviously one has  $|d_{k,a,b}| \leq 2M$ . Moreover, on account of the  $y$ -continuity and of the uniform  $z$ -continuity, one remarks that for fixed  $t, x, a, b$ , the family of functions

$$\{f(t, x, \cdot, z) / |z| \leq b\}$$

is equicontinuous, and hence compact in  $\mathcal{C}[-a, a]$ . Since the convolution operators approach the identity uniformly on compacts, one has

$$\lim_{k \rightarrow \infty} d_{k,a,b}(t, x) = 0,$$

which implies  $\lim_{k \rightarrow \infty} d_{k,a,b} = 0$  in  $L^2(dt \times m_\rho)$ .

Then we write, for any  $u \in \widehat{F}_\rho$  such that  $|u| \leq r$ ,

$$\begin{aligned} |f(\cdot, \cdot, u, D_\sigma u) - f_k(\cdot, \cdot, u, D_\sigma u)| &\leq d_{k,r,b} \mathbf{1}_{\{|D_\sigma u| \leq b\}} + 2M \mathbf{1}_{\{|D_\sigma u| > b\}} \\ &\leq d_{k,r,b} + \frac{2M}{b} |D_\sigma u| \end{aligned} \quad (*)$$

Now we write relation (8) for the difference  $u_l - u_k$

$$\begin{aligned} \|u_{l,t} - u_{k,t}\|_\rho^2 + 2 \int_t^T e_\rho(u_{l,s} - u_{k,s}) ds &= \\ = 2 \int_t^T (f_l(s, \cdot, u_{l,s}, D_\sigma u_{l,s}) - f_k(s, \cdot, u_{k,s}, D_\sigma u_{k,s}), u_{l,s} - u_{k,s})_\rho ds. \end{aligned} \quad (**)$$

and express the quantity  $f_l(s, \cdot, u_{l,s}, D_\sigma u_{l,s}) - f_k(s, \cdot, u_{k,s}, D_\sigma u_{k,s})$  in the form

$$\begin{aligned} f_l(s, \cdot, u_{l,s}, D_\sigma u_{l,s}) - f_k(s, \cdot, u_{l,s}, D_\sigma u_{l,s}) &+ f_k(s, \cdot, u_{l,s}, D_\sigma u_{l,s}) - f_k(s, \cdot, u_{k,s}, D_\sigma u_{l,s}) + \\ &+ f_k(s, \cdot, u_{k,s}, D_\sigma u_{l,s}) - f_k(s, \cdot, u_{k,s}, D_\sigma u_{k,s}). \end{aligned}$$

Using relations (\*), (17) and (16) one dominates the expression of the right hand side in (\*\*) by

$$\begin{aligned} &2 \int_t^T \left( d_{l,r,b} + d_{k,r,b} + \frac{4M}{b} |D_\sigma u_{l,s}|, u_{l,s} - u_{k,s} \right)_\rho ds + 2\mu \int_t^T \|u_{l,s} - u_{k,s}\|_\rho^2 ds + \\ &+ 2 \int_t^T C (|D_\sigma u_{l,s} - D_\sigma u_{k,s}|, |u_{l,s} - u_{k,s}|)_\rho ds \leq \\ &\leq \int_t^T \|d_{l,r,b}\|_\rho^2 ds + \int_t^T \|d_{k,r,b}\|_\rho^2 ds + \frac{1}{b^2} \int_t^T \|D_\sigma u_{l,s}\|_\rho^2 ds + K \int_t^T \|u_{l,s} - u_{k,s}\|_\rho^2 ds + \end{aligned}$$

$$+ \int_t^T \|D_\sigma u_{l,s} - D_\sigma u_{k,s}\|_\rho^2 ds.$$

By Lemma 3.8 we know that  $\int_0^T \|D_\sigma u_{l,s}\|_\rho^2 ds$  is bounded independently of  $l$ . The number  $b$  is at our disposal and we may choose it as large as we want. Then  $l, k$  may be chosen so that  $\int_t^T \|d_{l,r,b}\|_\rho^2 ds$  and  $\int_0^T \|d_{k,r,b}\|_\rho^2 ds$  are small as we want. Summing up we have an estimate of the form

$$\|u_{l,t} - u_{k,t}\|_\rho^2 + \int_t^T e_\rho(u_{l,s} - u_{k,s}) ds \leq \varepsilon + K \int_t^T \|u_{l,s} - u_{k,s}\|_\rho^2 ds,$$

where  $K$  depends only on  $C, \mu$  and  $T$ . Gronwall's lemma implies that  $(u_k)$  is a Cauchy sequence in  $\widehat{F}_\rho$ . Let  $u = \lim_n u_n$ .

Now we assert that  $f_k(\cdot, \cdot, u_k, D_\sigma u_k) \rightarrow f(\cdot, \cdot, u, D_\sigma u)$  in  $L^2(dt \times m_\rho)$  and this implies that  $u$  is a solution corresponding to  $f$ . Indeed, passing to a subsequence such that  $u_k \rightarrow u$  almost everywhere, we will have  $f(\cdot, \cdot, u_k, D_\sigma u) \rightarrow f(\cdot, \cdot, u, D_\sigma u)$  in  $L^2(dt \times m_\rho)$ . From condition (16) it follows that  $f(\cdot, \cdot, u_k, D_\sigma u) - f(\cdot, \cdot, u_k, D_\sigma u_k) \rightarrow 0$  in  $L^2(dt \times m_\rho)$ . Finally, the relation (\*) leads to  $f(\cdot, \cdot, u_k, D_\sigma u_k) - f_k(\cdot, \cdot, u_k, D_\sigma u_k) \rightarrow 0$ , and the proof of the first step is finished.

Second step. We now treat the general case. Again we construct a sequence of functions approximating our function  $f$ . For  $k \in N$  define

$$f_k(t, x, y, z) = ((f(t, x, y, z) - f^\circ(t; x)) \wedge k) \vee (-k) + f^\circ(t, x).$$

Each such function satisfies conditions (17), (16) with the same constants as  $f$  and  $f_k(\cdot, \cdot, 0, 0) = f^\circ$ . Moreover they are bounded and we may apply the preceding step. Let  $u_k$  be the solution of equation (19) with  $f_k$  and the final condition  $\Phi$ . These functions all satisfy the estimates from the statement, so that they are uniformly bounded again:  $\|u_k\|_\infty \leq r$  and  $\|u_k\|_{\rho, T} \leq \text{const}$ . We will check that they are Cauchy in  $\widehat{F}_\rho$ . Again one starts with the relation (8), written for the difference  $u_k - u_{k+1}$ ,

$$\begin{aligned} & \|u_{k,t} - u_{k+1,t}\|_\rho^2 + 2 \int_t^T e_\rho(u_{k,s} - u_{k+1,s}) ds = \quad (**) \\ & = 2 \int_t^T (f_k(s, \cdot, u_{k,s}, D_\sigma u_{k,s}) - f_{k+1}(s, \cdot, u_{k+1,s}, D_\sigma u_{k+1,s}), u_{k,s} - u_{k+1,s})_\rho ds. \end{aligned}$$

In order to estimate the right hand side we will use the hypothesis that  $g_r \in L^2$ . First remark that the relations (16) and (20) imply

$$f_k(\cdot, \cdot, y, z) \mathbf{1}_{\{|y| \leq r, g_r \leq \frac{k}{2}, |z| \leq \frac{k}{2C}\}} = f(\cdot, \cdot, y, z) \mathbf{1}_{\{|y| \leq r, g_r \leq \frac{k}{2}, |z| \leq \frac{k}{2C}\}}.$$

This in turn leads to

$$|f_k(\cdot, \cdot, u_k, D_\sigma u_k) - f_{k+1}(\cdot, \cdot, u_k, D_\sigma u_k)| \leq \mathbf{1}_{\{g_r > \frac{k}{2}\} \cup \{|D_\sigma u_k| > \frac{k}{2C}\}}$$



Since the  $L^2$  bounds of  $D_\sigma u_k$  are uniform with respect to  $k$ , one has

$$\int \int 1_{\{g_r > \frac{k}{2}\} \cup \{|D_\sigma u_k| > \frac{k}{2C}\}} dt dm_\rho \leq K \frac{1}{k^2},$$

and hence

$$\int \int (f_k(\cdot, \cdot, u_k, D_\sigma u_k) - f_{k+1}(\cdot, \cdot, u_k, D_\sigma u_k)) (u_k - u_{k+1}) dt dm_\rho \leq K \frac{1}{k^2}.$$

The right hand side of relation (\*\*) is therefore dominated by

$$\int_t^T \left( \mu \|u_{k,s} - u_{k+1,s}\|_\rho^2 + C \|u_{k,s} - u_{k+1,s}\|_\rho \|D_\sigma u_{k,s} - D_\sigma u_{k+1,s}\|_\rho \right) ds + K \frac{1}{k^2}.$$

Gronwall's lemma shows that our sequence is Cauchy and has a limit  $u$  in  $\widehat{F}_\rho$ . Choosing a subsequence  $k_l$  so that  $u_{k_l}$  and  $D_\sigma u_{k_l}$  converge  $dt \otimes m_\rho$ -almost everywhere, one immediately sees that

$$f_{k_l}(\cdot, \cdot, u_{k_l}, D_\sigma u_{k_l}) \rightarrow f(\cdot, \cdot, u, D_\sigma u),$$

almost everywhere. But these functions are uniformly bounded in  $L^2(dt \otimes m_\rho)$ , which implies weak convergence. Then one may pass to the limit in the relation of weak solutions and deduce that  $u$  is a weak solution of (19) with data  $(f, \Phi)$ . This concludes the proof.  $\square$

Now let us return to equation (13) with our initial operator  $L$ . That equation may be written in the following equivalent form

$$(\partial_t + L_\rho) u + f'(\cdot, \cdot, u, D_\sigma u) = 0, \quad u_T = \Phi,$$

with an arbitrary  $\rho \in (0, 1)$  and with  $f'(t, x, y, z) = f(t, x, y, z) + \rho \sum_{i,k} \sigma_k^i(x) \partial_i \theta(x) z_k$ . The function  $f'$  satisfies the conditions (17) and (16), provided  $f$  does (the constant in the condition (16) being changed). Note also that  $f'(t, x, y, 0) = f(t, x, y, 0)$ . The conditions to apply the theorem are fulfilled so that we get existence and uniqueness for solutions of equation (13). The formal result is as follows.

**Corollary 3.10** *Assume that the function  $f$  satisfies conditions (16), (17) and is continuous with respect to  $y$ . Moreover assume that the function defined in (20) belongs to  $L^2(dt \times dm_\rho)$  for some  $r > 0$  and  $\rho \geq 0$ . If  $f^\circ = f(\cdot, \cdot, 0, 0)$  and  $\Phi$  are bounded and  $\|\Phi\|_\infty \vee \left( e^{\mu T} \|\Phi\|_\infty + \frac{e^{\mu T} - 1}{\mu} \|f^\circ\|_\infty \right) \leq r$ , then equation (13) has a unique weak solution and it satisfies the following estimates*

$$\|u\|_{\rho, T}^2 \leq K \left[ \|\Phi\|_\rho^2 + \|f^\circ\|_{L^2(dt \otimes m_\rho)}^2 \right],$$

$$\|u\|_\infty \leq e^{\mu T} \|\Phi\|_\infty + \frac{e^{\mu T} - 1}{\mu} \|f^\circ\|_\infty.$$

## 4 Stochastic integral representation under $P^m$

We recall some results from [4]. First of all one associates to the Dirichlet form  $(F, e)$  a Markov process  $(\Omega, \mathcal{F}, \mathcal{F}_t, X_t, P^x, \theta_t, t \geq 0)$ . We stress that the filtration  $(\mathcal{F}_t)_t$  is associated in the canonical way to the process  $(X_t)_t$ . For  $u \in \mathcal{D}(L)$

$$M_t^u =: u(X_t) - u(X_0) - \int_0^t Lu(X_s) ds$$

is a martingale under  $P^m$  and under  $P^x$  for quasi- every point  $x \in \mathbf{R}^N$  (we recall that  $m$  is the Lebesgue measure, which is an invariant measure for the semigroup  $(P_t)_{t \geq 0}$  and that  $P^m =: \int P^x dx$ ). This may also be expressed as

$$u(X_t) - u(X_0) = M_t^u + N_t^u$$

with  $N_t^u = \int_0^t Lu(X_s) ds$  and so one may think of this as being a generalization of the Doob-Meyer decomposition of the semimartingale  $u(X_t) - u(X_0)$ . An important result from [4] asserts that the above decomposition holds for any  $u \in F$  with the change that then  $N_t^u$  is no more a finite variation process but a process of null energy (see Ch.5 of [4] for the definition and proofs). Moreover the decomposition holds true for  $u \in F_{loc}$  but then  $M_t^u$  is only a local martingale. The coordinate functions  $u_i(x) = x^i, 1 \leq i \leq N$ , are in  $F_{loc}$  (but generally not in  $\mathcal{D}(L)$ ) and we denote  $M^i =: M^{u_i}$ . Note that for smooth coefficients  $\sigma_k^i, X$  solves the stochastic differential equation  $dX_t^i = \sum_{j=1}^n \sigma_j^i(X_t) dB_t^j + \sum_{j=1}^n \partial_j a^{ij}(X_t) dt$ , where  $B = (B^1, \dots, B^n)$  is a Brownian motion. So  $M^i(t) = \int_0^t \sum_{j=1}^n \sigma_j^i(X_s) dB_s^j$ . In the general case such a representation is not available but one proves that under  $P^m$  or  $P^x$ , for quasi every point  $x$ ,

$$\langle M^i, M^j \rangle (t) = 2 \int_0^t \sum_{k=1}^n \sigma_k^i(X_s) \sigma_k^j(X_s) ds \quad (21)$$

and, more generally, for every  $u, v \in F$

$$\langle M^u, M^v \rangle (t) = 2 \int_0^t \left( \sum_{i,j=1}^N \sum_{k=1}^m \sigma_k^i \sigma_k^j \frac{\partial u}{\partial x^i} \frac{\partial v}{\partial x^j} \right) (X_s) ds. \quad (22)$$

In particular, since  $m = mP_s$ ,

$$E^m(\langle M^u, M^v \rangle (t)) = 2t e(u, v), \quad (23)$$

for  $u, v \in C_c^\infty$ . Since any  $u \in F$  is a limit in energy of  $C_c^\infty$ -functions, this formula remains true for any  $u, v \in F$ .

Moreover for a function  $u \in C^1(\mathbf{R}^N)$ , Corollary 5.6.2 of [4] gives the following representation

$$M_t^u = \sum_{i=1}^N \int_0^t \partial_i u(X_s) dM_s^i.$$

Since we are especially interested in writing such representations for other classes of functions, we start now a detailed discussion concerning the possibility of defining expressions like the right hand side of this equality. First let us notice that if  $\phi = (\phi_1, \dots, \phi_N) : [0, T] \times \mathbf{R}^N \rightarrow \mathbf{R}^N$  is measurable and bounded, the martingale

$$\sum_{i=1}^N \int_0^t \phi_i(s, X_s) dM_s^i,$$

has the bracket

$$\left\langle \sum_{i=1}^N \int_0^t \phi_i(s, X_s) dM_s^i \right\rangle_t = 2 \int_0^t |\phi\sigma|^2(s, X_s) ds.$$

If the function  $\phi$  is no more bounded but  $|\phi\sigma| \in L^2([0, T] \times \mathbf{R}^N)$ , a martingale like the preceding can be defined by approximation and we introduce the notation

$$\int_0^t \phi(s, X_s) \cdot dM_s := \lim_{k \rightarrow \infty} \sum_{i=1}^N \int_0^t \phi_i^k(s, X_s) dM_s^i,$$

with  $\phi^k = 1_{\{|\phi| \leq k\}} \phi$ . Note that if  $\phi\sigma = 0$  almost everywhere, then this martingale vanishes. Since the matrix  $a = \sigma\sigma^*$  is not assumed to be elliptic the class of functions  $\phi \neq 0$  such that  $\phi\sigma = 0$  may be quite large. Notice also that, if  $\phi$  is unbounded, the individual terms appearing in the sum expressing the martingale do not necessarily make sense as stochastic integrals, only their sum being introduced globally by the above procedure. (For example, one may have  $E^m \int_0^T \sum_k |\sigma_k^i \phi_i(X_t)|^2 dt = \infty$  for a certain index  $i$ , while  $E^m \int_0^T \sum_k \left| \sum_j \sigma_k^j \phi_j(X_t) \right|^2 dt < \infty$ .) The point sign in the expression  $\int \phi(s, X_s) \cdot dM_s$  would remind then that the summation is in fact formal. The same notation will be used for these sums of stochastic integrals under the measures  $P^x$ , when the corresponding integrability condition  $E^x \int_0^T \sum_k \left| \sum_j \sigma_k^j \phi_j(X_t) \right|^2 dt < \infty$  is satisfied. The martingale defined this way still verifies the above bracket relation

$$\left\langle \int_0^\cdot \phi(s, X_s) \cdot dM_s \right\rangle_t = 2 \int_0^t |\phi\sigma|^2(s, X_s) ds.$$

Consider now  $u \in F$ . Then for every  $\phi \in \tilde{\nabla}u$

$$M_t^u = \int_0^t \phi(X_s) \cdot dM_s, \quad P^m - a.s.$$

In fact, let  $u^k \in C_c^1(\mathbf{R}^N)$ ,  $k \in \mathbf{N}$ , be such that  $u^k \rightarrow u$  with respect to the norm of  $e_1$ . Using iii) in Proposition 3.1 we deduce that  $\nabla u^k \sigma = D_\sigma u^k \rightarrow D_\sigma u = \phi\sigma$  in  $L^2(\mathbf{R}^N)$ .

Then one may pass to the limit in the representation theorem for  $M^{u^k}$  in order to get the representation theorem for  $M^u$ .

Coming back to the decomposition formula, on account of the preceding representation, we have the following result.

**Lemma 4.1** *If  $u \in \mathcal{D}(L)$  and  $\phi \in \tilde{\nabla}u$ . Then*

$$u(X_t) - u(X_0) = \int_0^t \phi(X_s) \cdot dM_s + \int_0^t Lu(X_s) ds \quad P^m \text{ a.s.}$$

The aim of this section is to extend this representation to time dependent functions  $u(t, x)$ . We start with a particular type of functions in the next lemma and then, in the theorem below, we prove a more general result.

**Lemma 4.2** *Let  $u : [0, T] \times R^N \rightarrow R$  be such that*

- i)  $\forall s, u_s \in \mathcal{D}(L)$  and  $s \rightarrow Lu_s$  is continuous in  $L^2$ .*
- ii)  $s \rightarrow u_s$  is differentiable in  $L^2$  and  $s \rightarrow \partial_s u_s$  is continuous in  $L^2$ .*

*Then clearly  $u \in C_T$ . Moreover, for any  $\phi \in \tilde{\nabla}u$  and any  $s, t > 0$  such that  $s + t < T$ , the following relation holds  $P^m - a.s.$ ,*

$$u(s + t, X_t) - u(s, X_0) = \int_0^t \phi(s + r, X_r) \cdot dM_r + \int_0^t (\partial_s + L)u(s + r, X_r) dr .$$

**Proof.** We prove the above relation with  $s = 0$ , the general case being similar. Let  $0 = t_0 < t_1 < \dots < t_p = t$  be a partition of the interval  $[0, t]$  and write

$$u(t, X_t) - u(0, X_0) = \sum_{k=0}^{p-1} (u(t_{k+1}, X_{t_{k+1}}) - u(t_k, X_{t_k})) .$$

Then each term of the sum is expressed, on account of the preceding lemma, as

$$\begin{aligned} u(t_{k+1}, X_{t_{k+1}}) - u(t_k, X_{t_k}) &= u(t_{k+1}, X_{t_{k+1}}) - u(t_{k+1}, X_{t_k}) + u(t_{k+1}, X_{t_k}) - u(t_k, X_{t_k}) \\ &= \int_{t_k}^{t_{k+1}} \psi^{k+1}(s, X_s) \cdot dM_s + \int_{t_k}^{t_{k+1}} Lu_{t_{k+1}}(X_s) ds + \int_{t_k}^{t_{k+1}} \partial_s u_s(X_{t_k}) ds, \end{aligned}$$

where  $\psi^{k+1} = (\psi_1^{k+1}, \dots, \psi_N^{k+1}) \in \tilde{\nabla}u_{t_{k+1}}$  and the last integral is obtained by using the Leibnitz-Newton formula for the  $L^2$ -valued function  $s \rightarrow u_s$ . Further we estimate in  $L^2$  the differences between each term in the last expression and the similar terms corresponding to the formula we have to prove. We first estimate,

$$E^m \left( \int_{t_k}^{t_{k+1}} \psi^{k+1}(s, X_s) \cdot dM_s - \int_{t_k}^{t_{k+1}} \phi(s, X_s) \cdot dM_s \right)^2 =$$

$$E^m \int_{t_k}^{t_{k+1}} |(\psi^{k+1}(X_s) - \phi(s, X_s)) \sigma|^2 ds = \int_{t_k}^{t_{k+1}} e(u_{t_{k+1}} - u_s) ds$$

Since  $s \rightarrow Lu_s$  is  $L^2$ -continuous, it follows that  $s \rightarrow u_s$  is continuous in energy, and hence the difference appearing in the last integral,  $e(u_{t_{k+1}} - u_s)$  is uniformly small, provided the partition is tight enough. From this one deduces that

$$\sum_{k=0}^{p-1} \int_{t_k}^{t_{k+1}} \psi^{k+1}(s, X_s) .dM_s \longrightarrow \int_0^t \phi(s+r, X_r) .dM_r.$$

The next difference is estimated using Minkowski's inequality

$$\left( E^m \left( \sum_{k=0}^{p-1} \int_{t_k}^{t_{k+1}} (Lu_{t_{k+1}} - Lu_s)(X_s) ds \right)^2 \right)^{\frac{1}{2}} \leq$$

$$\sum_{k=0}^{p-1} \int_{t_k}^{t_{k+1}} \left( E^m (Lu_{t_{k+1}} - Lu_s)^2(X_s) \right)^{\frac{1}{2}} ds = \sum_{k=0}^{p-1} \int_{t_k}^{t_{k+1}} \|Lu_{t_{k+1}} - Lu_s\|_2 ds,$$

so that it is similarly expressed as an integral of an uniformly small quantity.

For the last difference we write

$$\left( E^m \left( \sum_{k=0}^{p-1} \int_{t_k}^{t_{k+1}} (\partial_s u_s(X_{t_k}) - \partial_s u_s(X_s)) ds \right)^2 \right)^{\frac{1}{2}} \leq$$

$$\sum_{k=0}^{p-1} \int_{t_k}^{t_{k+1}} \left( E^m (\partial_s u_s(X_{t_k}) - \partial_s u_s(X_s))^2 \right)^{\frac{1}{2}} ds =$$

$$\sum_{k=0}^{p-1} \int_{t_k}^{t_{k+1}} \left( 2 \int_{R^N} (\partial_s u_s)^2 dm - 2 \int_{R^N} \partial_s u_s P_{s-t_k} \partial_s u_s dm \right)^{\frac{1}{2}} ds \leq$$

$$\sqrt{2} \sum_{k=0}^{p-1} \int_{t_k}^{t_{k+1}} \|\partial_s u_s\|_2^{\frac{1}{2}} \|\partial_s u_s - P_{s-t_k} \partial_s u_s\|_2^{\frac{1}{2}} ds.$$

>From the hypotheses it follows that this will tend also to zero when the partition is tight enough.

We conclude that the relation is obtained in the limit, taking a sequence of partitions of diameters tending to zero.  $\square$

**Theorem 4.3** *Let  $f \in L^2([0, T] \times R^N)$  and  $\Phi \in L^2(R^N)$  and define*

$$u(t, x) =: P_{T-t} \Phi(x) + \int_t^T P_{s-t} f(s, \cdot)(x) ds.$$

Then for each  $\phi \in \tilde{\nabla}u$  and for each  $s \in [0, T]$ , the following equation holds  $P^m - a.s.$

$$u(s+t, X_t) - u(s, X_0) = \int_0^t \phi(s+r, X_r) .dM_r - \int_0^t f(s+r, X_r)dr . \quad (24)$$

**Proof.** Assume first that  $\Phi \in D(L)$  and  $s \rightarrow f(s, \cdot)$  is differentiable in  $L^2$  and  $s \rightarrow \partial_s f(s, \cdot)$  is continuous in  $L^2$ . Then by Proposition 2.4,  $s \rightarrow u(s, \cdot)$  is differentiable in  $L^2$  and its derivative is continuous in  $L^2$ . On the other hand  $u_s \in D(L)$  and  $Lu_t = LP_{T-t}\Phi + \int_t^T LP_{s-t}f_s x ds$  is continuous in  $L^2$ . So, using Lemma 4.2, we get the relation in the statement. For the general case one proceeds by approximation.  $\square$

## 5 Stochastic integral representation under $P^x$

The additive martingale functionals are known to be represented as stochastic integrals with respect to the coordinate martingales  $M^i$  (see [4] Theorem 5.6.3). The aim of this section is to extend this representation to general martingales (which are no more additive functionals). More precisely we shall produce an exceptional set outside of which the representation theorem holds.

We start with some technical considerations.

**Lemma 5.1** *Let  $\psi : [0, T] \times R^N \rightarrow R$  be such that  $\int_0^T \|\psi(t, \cdot)\|_2 dt / \sqrt{t} < \infty$ . Then  $v(x) =: E^x(\int_0^T \psi(t, X_t) dt)$  is a quasi-continuous, q.e. finite function.*

**Proof.** Define  $v_\varepsilon(x) =: E^x(\int_\varepsilon^T \psi(t, X(t)) dt)$ . This is a quasicontinuous function - see [4] p. 144, Th 4.2.3. Assume that  $\psi \geq 0$  - if not one writes  $\psi = \psi_+ - \psi_-$ . Then  $v_\varepsilon(x) \nearrow v(x)$  for every  $x$ . On the other hand, by, for  $\eta > \varepsilon > 0$  one has; using successively Minkowski's inequality and (3),

$$e(v_\varepsilon - v_\eta)^{1/2} \leq \int_\varepsilon^\eta e(P_s \psi_s)^{1/2} ds \leq \int_\varepsilon^\eta \|\psi_s\|_2 \frac{ds}{\sqrt{s}} \rightarrow 0 \text{ as } \varepsilon, \eta \rightarrow 0.$$

Then, by Th 2.1.4 p 69 in [4] there exists a quasicontinuous, finite function  $w$  and a sequence  $\varepsilon_n \rightarrow 0$  such that  $v_{\varepsilon_n}(x) \rightarrow w(x)$  for quasi every  $x$ . We conclude that  $v(x) = w(x)$  for quasi every  $x$  and so  $x \rightarrow v(x)$  is quasicontinuous.  $\square$

A useful tool in the sequel will be the time-space Markov process  $\hat{X}_t = (t, X_t)$  with the state space  $S = [0, T] \times R^N$ , which is defined by  $\hat{X} = (\hat{\Omega}, \hat{\mathcal{F}}, \hat{\mathcal{F}}_t, \hat{X}_t, \hat{\theta}_t, \hat{P}^{t,x}, (t, x) \in S)$  with

$$\begin{aligned} \hat{\Omega} &= \{[0, T] \times \Omega\} \cup \{\omega_\Delta\}, \\ \hat{X}_t(s, \omega) &= (s+t, X_t(\omega)) \text{ if } s+t < T, \\ &= \Delta \text{ if } s+t \geq T, \\ \hat{\theta}_t(s, \omega) &= (s+t, \theta_t(\omega)) \text{ if } s+t < T, \\ &= \omega_\Delta \text{ if } s+t \geq T, \\ \hat{P}^{t,x} &= \delta_t \times P^x \end{aligned}$$

where  $\Delta$  is an isolated point and  $\omega_\Delta(t) = \Delta, \forall t \geq 0$ . Note that the semigroup of  $\widehat{X}$  is given by  $\widehat{P}_t u(s, x) = E^x(u(s+t, X_t))$ .

Now let  $\Phi : R^N \rightarrow R$  be bounded, measurable, and of compact support and set  $u(t, x) = E^x \Phi(X_{T-t})$ . Then clearly, as an element of  $L^2$ ,  $u$  may be written  $u_t = P_{T-t}$  and represents a solution of equation (4) with data  $f = 0$  and  $\Phi$ . Moreover,  $u(t, x)$  is an invariant excessive function with respect to the time-space process  $\widehat{X}$ , and consequently the process  $t \rightarrow u(t, X_t)$  is a continuous martingale under  $P^x$ , for each  $x \in R^N$ . We are going to represent this process as a stochastic integral under quasi every measure  $P^x$ . In order to do this we have to know that the stochastic integrals make sense, and so we first prove the following lemma.

**Lemma 5.2** *Let  $\Phi$  be measurable, bounded, and of compact support. Set  $u(t, x) = E^x \Phi(X_{T-t})$  and take  $\phi \in \widetilde{\nabla} u$ . There exists a polar set  $\Gamma$  such that*

$$E^x \int_0^T |\phi_t \sigma|^2(X_t) dt \leq \frac{1}{2} E^x \Phi^2(X_T), x \in \Gamma^c.$$

**Proof.** By Corollary 4.3 one has the following representation under  $P^m$  :

$$u(T, X_T) - u(0, X_0) = \int_0^T \phi(s, X_s) .dM_s.$$

By Itô's formula we get

$$\begin{aligned} \Phi^2(X_T) = u^2(T, X_T) &= |u(0, X_0)|^2 + 2 \int_0^T |\phi_t \sigma|^2(X_t) dt \\ &+ 2 \int_0^T u(s, X_s) \phi(s, X_s) .dM_s, \quad P^m - a.s. \end{aligned}$$

Conditioning with respect to  $\mathcal{F}_0$  one gets

$$E^x \Phi^2(X_T) = E^x |u(0, X_0)|^2 + 2 E^x \int_0^T |\phi_t \sigma|^2(X_t) dt, \quad P^m a.s.$$

In particular, one has

$$E^x \int_0^T n \wedge |\phi_t \sigma|^2 dt \leq \frac{1}{2} E^x \Phi^2(X_T), \quad P^m - a.s.$$

By Lemma 5.1 the function in the left hand side is quasicontinuous. The expression in the right hand side is also quasicontinuous, and hence the inequality holds for quasi- every  $x$ . Passing to the limit with  $n \rightarrow \infty$  we get the inequality as asserted in the statement.  $\square$

Now we shall fix a properly exceptional set  $\Gamma$  and a version of each coordinate martingale  $M^i$ , which is an additive functional in the strict sense on  $\Gamma^c$  (see [4], p.181) and such that the bracket formula (21) holds under each measure  $P^x, x \in \Gamma^c$ . (We recall that if  $\mathcal{N}$  is properly

exceptional, then the process started from a point  $x \in \mathcal{N}^c$  never hits  $\mathcal{N}$ . Thanks to Theorem 4.1.1 in [4] p 137, any zero capacity set is included in a properly exceptional set.) Theorem 5.4 below will throw out other parts of  $\Gamma^c$ , where these martingales are not consistently defined.

**Lemma 5.3** *Let  $\Phi$  be measurable, bounded, and of compact support. Set  $u(t, x) = E^x \Phi(X_{T-t})$  and take  $\phi \in \tilde{\nabla}u$ . There exists a polar set  $\Lambda$  containing  $\Gamma$  so that, for each  $x \in \Lambda^c$ , the following representation holds:*

$$\Phi(X_T) = u(0, X_0) + \int_0^T \phi(s, X_s) \cdot dM_s, P^x - a.s.$$

$$E^x \int_0^T |\phi_t \sigma|^2(X_t) dt < \infty.$$

**Proof.** For  $0 \leq s, t$  such that  $s + t \leq T$ , the following expression make sense under  $P^x$ , for quasi every  $x$  :

$$I_t^s = u(t + s, X_t) - u(s, X_0) - \int_0^t \phi(s + r, X_r) \cdot dM_r.$$

Our aim is to prove that for quasi every  $x$ ,  $I_T^0 = 0, P^x - a.s.$  Since  $M^i, 1 \leq i \leq N$  are additive functionals, one may write

$$I_T^0 = I_{\varepsilon_k}^0 + I_{T-\varepsilon_k}^{\varepsilon_k} \circ \theta_{\varepsilon_k}, \quad (*)$$

with a sequence  $\varepsilon_k \searrow 0$ . Define then  $h_k(x) = E^x |I_{T-\varepsilon_k}^{\varepsilon_k}|$ . The representation theorem under  $P^m$  asserts that  $\int h_k(x) dx = E^m |I_{T-\varepsilon_k}^{\varepsilon_k}| = 0$ , so that  $E^x |I_{T-\varepsilon_k}^{\varepsilon_k} \circ \theta_{\varepsilon_k}| = P_{\varepsilon_k} h_k(x) = 0$  for quasi every  $x$ . One completes the proof taking  $\Lambda$  the union of exceptional sets of these relations and letting  $\varepsilon_k \rightarrow 0$  in the equality (\*)□

**Theorem 5.4** *There exists some properly exceptional set  $\mathcal{N}$  containing  $\Gamma$  and such that the following representation result holds. For every bounded  $\mathcal{F}_\infty$ -measurable random variable  $\xi$ , there exists an  $(\mathcal{F}_t)_{t \geq 0}$ -predictable process  $\phi = (\phi_1, \dots, \phi_N) : [0, \infty) \times \Omega \rightarrow R^N$  such that for each probability measure  $\nu$ , supported by  $R^N \setminus \mathcal{N}$ , one has*

$$\xi = E^\nu (\xi / \mathcal{F}_0) + \int_0^\infty \phi_s \cdot dM_s, \quad P^\nu - a.s.,$$

$$E^\nu \int_0^\infty |\phi_s \sigma(X_s)|^2 ds \leq \frac{1}{2} E^\nu \xi^2.$$

*If another predictable process  $\phi' = (\phi'_1, \dots, \phi'_N)$  satisfies the same relations under a certain measure  $P^\nu$ , then one has  $\phi'_i \sigma(X_s) = \phi_i \sigma(X_s), dt \times dP^\nu - a.s.$*



**Proof.** Suppose that  $\mathcal{E}$  is the class of bounded random variables for which the statement holds outside some fixed exceptional set  $\mathcal{N}$ . We claim that if  $(\xi_k) \subset \mathcal{E}$  is a uniformly bounded increasing sequence and  $\xi = \lim_{k \rightarrow \infty} \xi_k$ , then  $\xi \in \mathcal{E}$ . Indeed, since  $\xi$  and  $\xi_0$  are bounded,  $E^x |\xi_k - \xi|^2 \rightarrow 0$ . Denoting by  $\phi^k$  the process which represents  $\xi_k$ , we obtain

$$E^x \int_0^\infty |(\phi_s^k - \phi_s^p) \sigma(X_s)|^2 ds \leq E^x |\xi_k - \xi_p|^2 \rightarrow 0, x \in \mathcal{N}^c.$$

In order to pass to the limit pointwise we should choose a subsequence, and this subsequence may depend of the point  $x$ . In order to obtain a sequence of representable variables that converges rapidly enough under all measures  $P^x, x \in \mathcal{N}^c$ , we are going to construct them as follows. For each  $l = 0, 1, \dots$  set

$$k_l(x) = \inf \{k / E^x (\xi - \xi_k)^2 < \frac{1}{2^l}\},$$

$$\bar{\xi}_l = \xi_{k_l(X_0)}.$$

Then one has  $\bar{\xi}_l = \xi_{k_l(x)}$  on the set  $\{X_0 = x\}$ , and consequently  $E^x (\xi - \bar{\xi}_l)^2 < \frac{1}{2^l}$  for any  $x \in \mathcal{N}^c$ . The process which represents  $\bar{\xi}_l$  is simply obtained by the formula  $\bar{\phi}^l = \phi^{k_l(X_0)}$ . With this sequence one may pass to the limit and define  $\psi_s = \limsup_{l \rightarrow \infty} \bar{\phi}_s^l \sigma(X_s)$  (where  $\limsup$  is taken on each coordinate) and  $\phi_s = \psi_s \tau(X_s)$  where  $\tau$  is the matrix that we have introduced in the beginning of the paper. (Recall that  $\sigma \tau \sigma = \sigma$ .) Then  $\phi_s \sigma(X_s) = \psi_s \tau \sigma(X_s) = \lim_l \bar{\phi}_s^l \sigma \tau \sigma(X_s) = \lim_l \bar{\phi}_s^l \sigma(X_s)$  in  $L^2([0, \infty) \times \Omega, dt \times P^x)$ , and consequently

$$E^x \left| \int_0^\infty (\bar{\phi}_s^l - \phi_s) \cdot dM_s \right|^2 \rightarrow 0.$$

This permits us to pass to the limit in the representation formula for  $\bar{\xi}_l$  and to get it for  $\xi$ .

Consider now a countable set  $\mathcal{A} \subset C_c^\infty(R^N)$  which is dense in  $\mathcal{C}_0(R^N)$  for the topology of uniform convergence. Consider also the countable family  $\mathcal{B}$  of random variables of the form  $\xi = \prod_{i=1}^k f_i(X(t_i))$  where  $f_i \in \mathcal{A}$ ,  $0 \leq t_1 \leq \dots \leq t_k$ , with  $t_i \in Q$ , and  $k \in \mathbf{N}$ . The completion of the  $\sigma$ -algebra generated by  $\mathcal{B}$  is  $\mathcal{F}_\infty$ . Since  $\mathcal{E}$  is closed for increasing sequences, a monotone class argument reduces the proof to the case of a random variable in  $\mathcal{B}$ .

Let  $\xi \in \mathcal{B}$  be of the above form and fix  $1 \leq i \leq k$ . In order to construct the density  $\phi_t$  for  $t \in (t_{i-1}, t_i]$  we define  $g = f_i P_{t_{i+1}-t_i} f_{i+1} \dots P_{t_k-t_{k-1}} f_k$  and use the Markov property in order to get  $E^x(\xi/F_t) = \prod_{j=1}^{i-1} f_j(X(t_j)) P_{t_i-t} g(X_t)$ . We denote  $u(t, x) = P_{t_i-t} g(x)$  and using Lemma 5.3 we get a function  $\psi : [t_{i-1}, t_i] \times R^N \rightarrow R^N$  such that

$$u(t, X_t) = u(t_{i-1}, X_{t_{i-1}}) + \int_{t_{i-1}}^t \psi(s, X_s) \cdot dM_s, \quad P^x a.s. \quad (*)$$

So we define  $\phi_{k,t}(\omega) = \prod_{j=1}^{i-1} f_j(X_{t_j}(\omega)) \psi_k(t, X_t(\omega))$  for  $t \in (t_{i-1}, t_i]$ . We recall that relation (\*) holds for every  $x$  except an exceptional set of null capacity. The exceptional set  $\mathcal{N}$  in

the statement, will be the union of all these exceptional sets corresponding to the variables  $\xi \in \mathcal{B}$ .  $\square$

The exceptional set  $\mathcal{N}$  from the preceding theorem has the property that the stochastic integral representation may be extended to non bounded random variables. One may for example, represent separately the positive and the negative parts according to the following corollary.

**Corollary 5.5** *Let  $\mathcal{N}$  be the set obtained in the preceding theorem. For any  $\mathcal{F}_\infty$  -measurable nonnegative random variable  $\xi \geq 0$  there exists a predictable process  $\phi = (\phi_1, \dots, \phi_N) : [0, \infty) \times \Omega \rightarrow \mathbb{R}^N$  such that the following holds*

$$\xi = E^x(\xi/\mathcal{F}_0) + \int_0^\infty \phi_s \cdot dM_s, \quad P^x - a.s.,$$

$$E^x \int_0^\infty |\phi_s \sigma(X_s)|^2 ds \leq E^x \xi^2,$$

for each point  $x \in \mathcal{N}^c$  such that  $E^x \xi < \infty$ .

If another predictable process  $\phi' = (\phi'_1, \dots, \phi'_N)$  satisfies the same relations under a certain measure  $P^x$ , with such a point  $x$ , then one has  $\phi'_t \sigma(X_s) = \phi_t \sigma(X_s)$ ,  $dt \times dP^x - a.s.$

## 6 Backward Stochastic Differential Equations

The set  $\mathcal{N}$  of the preceding Theorem 5.4 is fixed from now on. The representation given by that theorem allows us to solve backward stochastic differential equations under all measures  $P^x$ ,  $x \in \mathcal{N}^c$ , at the same time, by using exactly the same arguments as in [10], [9]. The first ingredient in the treatment of BSDE's is the following lemma, which repeats the original idea from [10].

**Lemma 6.1** *Let  $\xi$  be an  $\mathcal{F}_T$  -measurable random variable and  $h : [0, T] \times \Omega \rightarrow \mathbb{R}$  an  $(\mathcal{F}_t)_{t \geq 0}$  -predictable process. Let  $A$  be the set of all points  $x \in \mathcal{N}^c$  for which the following integrability condition holds*

$$E^x \left( |\xi| + \int_0^T |h(s, \omega)| ds \right) < \infty.$$

Then there exists a pair  $(Y_t, Z_t)_{0 \leq t \leq T}$  of predictable processes  $Y : [0, T] \times \Omega \rightarrow \mathbb{R}$ ,  $Z : [0, T] \times \Omega \rightarrow \mathbb{R}^N$ , such that under all measures  $P^x$ ,  $x \in A$ , they have the following properties: (i)  $Y$  is continuous, (ii)  $Z$  satisfy the integrability condition

$$\int_0^T |Z_t \sigma(X_t)|^2 dt < \infty, P^x - a.s.,$$

(iii) the local martingale obtained integrating it against the coordinate martingale,  $\int_0^t Z_s \cdot dM_s$ , is an uniformly integrable martingale, (iv) and together satisfy the equation

$$Y_t = \xi + \int_t^T h(s, \omega) ds - \int_t^T Z_s \cdot dM_s, \quad P^x - a.s.$$

If another pair  $(Y'_t, Z'_t)$  of predictable processes satisfies the above conditions, under certain measure  $P^\nu$  with some measure  $\nu$  supported by  $A$ , then one has  $Y_t = Y'_t, P^\nu - a.s.$  and  $Z_t\sigma(X_t) = Z'_t\sigma(X_t), dt \times P^\nu - a.s.$

**Proof.** The representation of the positive and negative parts of the random variable

$$\xi + \int_0^T h_s ds,$$

by the preceding corollary, gives us the predictable process  $Z$  such that

$$\xi + \int_0^T h_s ds = E^{X_0} \left( \xi + \int_0^T h_s ds \right) + \int_0^T Z_s . dM_s.$$

The process  $Y$  is obtained by the formula

$$Y_t = E^{X_0} \left( \xi + \int_0^T h_s ds \right) + \int_0^t Z_s . dM_s - \int_0^t h_s ds.$$

□

This lemma allows us to introduce the following definition.

**Definition 6.2** Let  $\xi$  be an  $\mathcal{F}_T$ -measurable random variable and  $h : [0, T] \times \Omega \times R \times R^n \rightarrow R$  a measurable function such that  $(s, \omega) \rightarrow h(s, \omega, \cdot, \cdot)$  is predictable. Let  $\nu$  be a probability measure supported by  $\mathcal{N}^c$  such that  $E^\nu |\xi|^p < \infty$  for some  $p > 1$ . We say that a pair  $(Y_t, Z_t)_{0 \leq t < T}$  of predictable processes  $Y : [0, T] \times \Omega \rightarrow R, Z : [0, T] \times \Omega \rightarrow R^N$  is a solution of the backward stochastic differential equation in  $L^p(P^\nu)$  with data  $(\xi, h)$  provided that  $Y$  is continuous under  $P^\nu$  and together satisfy the integrability conditions

$$E^\nu \int_0^T |h(t, \cdot, Y_t, Z_t\sigma(X_t))| dt < \infty,$$

$$E^\nu \left( \int_0^T |Z_t\sigma(X_t)|^2 dt \right)^{\frac{p}{2}} < \infty,$$

and the following equation

$$Y_t = \xi + \int_t^T h(s, \omega, Y_s, Z_s\sigma(X_t)) ds - \int_t^T Z_s . dM_s, P^\nu - a.s. \quad (25)$$

Note that in this definition the function  $h$  may depend on the whole path  $\omega$  and need not be just a function of  $X_t(\omega)$  at time  $t$ . The relation (25) represents the backward stochastic differential equation naturally associated to the data  $(\xi, h)$  in our framework. Remark that the stochastic integral in the BSDE is not written with respect to the Brownian motion but with respect of the martingales  $M^i$ . We recall that in the case where  $L$  has Lipschitz continuous coefficients the Markov process  $X$  is usually taken as the solution of a SDE driven by a Brownian motion  $B$  and then one may take  $dM = \sigma(X)dB$ . So  $Z.dM = Z\sigma(X)dB$ . We stress that in the already standard notation concerning BSDE's  $Z$  stands for  $Z\sigma(X_t)$ .

### 6.1 Existence and uniqueness of solutions of BSDE's

In this subsection we assume that  $\xi$  is a given  $\mathcal{F}_T$ -measurable random variable and  $h : [0, T] \times \Omega \times R \times R^n \rightarrow R$  is a given measurable function such that  $(s, \omega) \rightarrow h(s, \omega, \cdot, \cdot)$  is predictable and satisfies the conditions

$$|h(s, \omega, y, z) - h(s, \omega, y, z')| \leq C |z - z'| \tag{26}$$

$$(h(s, \omega, y, z) - h(s, \omega, y', z)) (y - y') \leq \mu (y - y')^2 \tag{27}$$

with arbitrary  $s, \omega, y, y', z, z'$  and some constants  $C > 0, \mu \in R$ . (In the proofs we will often assume that  $\mu = 0$ , because one can always make the change  $Y_t^* = e^{\mu t} Y_t, Z_t^* = e^{\mu t} Z_t$  for the solutions, which correspond to the change  $\xi^* = e^{\mu T} \xi, h_t^*(y, z) = e^{\mu t} h_t(e^{-\mu t} y, e^{-\mu t} z) - e^{-\mu t} y$  for the data. The latter function  $h^*$  satisfies (27) with  $\mu = 0$ .) We will use the notation  $h^\circ(s, \omega) = h(s, \omega, 0, 0)$  and  $h'(s, \omega, y) = h(s, \omega, y, 0) - h(s, \omega, 0, 0)$ . Let  $\nu$  be a measure supported by  $\mathcal{N}^c$  and  $p > 1$  such that  $E^\nu |\xi|^p < \infty$ . One basic tool in the treatment of  $L^p$ -solutions is the following lemma.

**Lemma 6.3** *If  $(Y, Z)$  is a solution of the BSDE in  $L^p(P^\nu)$  with data  $(\xi, h)$ , then one has*

$$\begin{aligned} & e^{\mu t} |Y_t| + \int_t^T e^{\mu s} |h'_s(Y_s) - \mu Y_s| ds + \int_t^T e^{\mu s} dL_s \leq \\ & \leq e^{\mu T} |\xi| + \int_t^T e^{\mu s} |h^\circ_s| ds + C \int_t^T e^{\mu s} |Z_s \sigma(X_s)| ds - \int_t^T e^{\mu s} \text{sign}(Y_s) Z_s \cdot dM_s, \\ & e^{\mu t} |Y_t| + E^\nu \left[ \int_t^T e^{\mu s} |h'_s(Y_s) - \mu Y_s| ds + \int_t^T e^{\mu s} dL_s / \mathcal{F}_t \right] \leq \\ & E^\nu \left[ e^{\mu T} |\xi| + \int_t^T e^{\mu s} |h^\circ_s| ds + C \int_t^T e^{\mu s} |Z_s \sigma(X_s)| ds / \mathcal{F}_t \right] \end{aligned}$$

where  $L = (L_s)$  is the local time of the semimartingale  $Y$ .

**Proof.** One starts with Tanaka's formula for  $Y$ ,

$$|Y_T| - |Y_t| = - \int_t^T \text{sign}(Y_s) h_s(Y_s, Z_s \sigma(X_s)) ds + \int_t^T \text{sign}(Y_s) Z_s \cdot dM_s + L_T - L_t.$$

Then one writes Ito's formula for the product  $e^{\mu t} |Y_t|$ , obtaining

$$\begin{aligned} e^{\mu T} |\xi| - e^{\mu t} |Y_t| &= \int_t^T e^{\mu s} \text{sign}(Y_s) (\mu Y_s - h_s(Y_s, Z_s \sigma(X_s))) ds + \\ &+ \int_t^T e^{\mu s} dL_s + \int_t^T e^{\mu s} \text{sign}(Y_s) Z_s \cdot dM_s. \end{aligned}$$

The first inequality from the lemma follows from the relations (26) and (27). The second inequality is obtained from the first upon conditioning.  $\square$

The well-known quadratic relation, obtained by writing Ito's formula for  $Y^2$ , has the form

$$Y_t^2 + 2 \int_t^T |Z_s \sigma(X_s)|^2 ds = |\xi|^2 + 2 \int_t^T Y_s h_s(Y_s, Z_s \sigma(X_t)) - 2 \int_t^T Y_s Z_s \cdot dM_s. \quad (28)$$

Note that the bracket formula introduces a 2 in the second term from the right hand side. (In the classical case our framework corresponds to the infinitesimal generator  $\Delta$  not to  $\frac{1}{2}\Delta$ .) Combining this relation and the estimates of the preceding lemma, one gets the next result which justifies the name of solutions in  $L^p(P^\nu)$ . It also ensures uniqueness of solutions.

**Lemma 6.4** *There exists a constant  $K$ , which depends only on  $T, C$  and  $p$ , such that the solution  $(Y, Z)$  in  $L^p(P^\nu)$  with data  $(\xi, h)$  is estimated in terms of the data as follows*

$$\left\| \sup_{0 \leq t \leq T} e^{\mu t} |Y_t| \right\|_p + \left\| \left( \int_0^T e^{2\mu t} |Z_t \sigma(X_t)|^2 dt \right)^{\frac{1}{2}} \right\|_p \leq K \left( e^{\mu T} \|\xi\|_p + \left\| \int_0^T e^{\mu t} |h^\circ_t| dt \right\|_p \right).$$

Moreover one has the following estimate with a similar constant

$$\left\| \int_0^T e^{\mu t} |h_t(Y_t, Z_t \sigma(X_t))| dt \right\|_p \leq K' \left( e^{\mu T} \|\xi\|_p + \left\| \int_0^T e^{\mu t} |h^\circ_t| dt \right\|_p \right).$$

**Proof.** We are assuming that  $\mu = 0$ . From the second inequality of the preceding lemma and Doob's  $L^p$ -estimates for martingales, it follows that the process

$$Y^{*,t} = \sup_{t \leq s \leq T} |Y_s|$$

satisfies the estimate

$$\|Y^{*,t}\|_p \leq \frac{p}{p-1} \left( \|\xi\|_p + \left\| \int_t^T |h^\circ_s| ds \right\|_p + \left\| \int_t^T |Z_s \sigma(X_s)| ds \right\|_p \right). \quad (*)$$

On the other hand, from the above relation (28) and the conditions (26) and (27) it follows

$$\begin{aligned} Y_t^2 + 2 \int_t^T |Z_s \sigma(X_s)|^2 ds &\leq |\xi|^2 + \left( \int_t^T |h^\circ_s| ds \right)^2 + (1+C) (Y^{*,t})^2 + \\ &+ C \left( \int_t^T |Z_s \sigma(X_s)| ds \right)^2 - 2 \int_t^T Y_s Z_s \cdot dM_s. \end{aligned}$$

Further we introduce the notation

$$U^t = \left( \int_t^T |Z_s \sigma(X_s)|^2 ds \right)^{\frac{1}{2}}, \quad V^t = \int_t^T |Z_s \sigma(X_s)| ds, \quad H^t = \int_t^T |h^\circ_s| ds$$

and take the square root in the last inequality obtaining the following estimate for  $U^t$ ,

$$\sqrt{2}U^t \leq |\xi| + H^t + \sqrt{1 + CY^{*,t}} + \sqrt{C}V^t + \sqrt{2} \left| \int_t^T Y_s Z_s .dM_s \right|^{\frac{1}{2}}.$$

The stochastic integral term is estimated by the inequality of Burkholder-Davis-Gundy with the constant corresponding to  $\frac{p}{2}$ ,

$$\begin{aligned} \left\| \left| \int_t^T Y_s Z_s .dM_s \right|^{\frac{1}{2}} \right\|_p &\leq \sqrt{c_{\frac{p}{2}}} \left\| \left( \int_t^T Y_s^2 |Z_s \sigma(X_s)|^2 ds \right)^{\frac{1}{2}} \right\|_{\frac{p}{2}} \leq \\ &\leq \sqrt{c_{\frac{p}{2}}} \left( E^\nu (Y^{*,t} U^t)^{\frac{p}{2}} \right)^{\frac{1}{p}} \leq \frac{\sqrt{c_{\frac{p}{2}}}}{2} \left( \frac{1}{\varepsilon} \|Y^{*,t}\|_p + \varepsilon \|U^t\|_p \right). \end{aligned}$$

Taking  $\varepsilon$  such that  $\varepsilon \sqrt{\frac{c_{\frac{p}{2}}}{2}} = \sqrt{2} - 1$  one obtains

$$\|U^t\|_p \leq \|\xi\|_p + \|H^t\|_p + \text{const} \|Y^{*,t}\|_p + \sqrt{C} \|V^t\|_p$$

and combining with the relation (\*) one gets

$$\|U^t\|_p \leq \text{const} \left( \|\xi\|_p + \|H^t\|_p \right) + \text{const} \|V^t\|_p.$$

Since  $V^t \leq \sqrt{T-t}U^t$ , one deduces that there exist  $M$  and  $\delta > 0$ , which depend only of  $C$  and  $p$ , such that  $\|U^t\|_p \leq M \left( \|\xi\|_p + \|H^t\|_p \right)$  provided that  $T-t < \delta$ . Finally, iterating this last estimate over the intervals  $[T-k\delta, T-(k-1)\delta]$  one gets summing up

$$\left\| \left( \int_0^T |Z_t \sigma(X_t)|^2 dt \right)^{\frac{1}{2}} \right\|_p \leq K \left( \|\xi\|_p + \left\| \int_0^T |h^\circ_t| dt \right\|_p \right).$$

Combining this inequality with (\*) one deduces the first estimate of the lemma.

In order to deduce the second estimate one uses the first estimate and the preceding lemma obtaining first a particular form

$$\left\| \int_0^T |h'_t(Y_t, Z_t \sigma(X_t))| dt \right\|_p \leq K' \left( \|\xi\|_p + \left\| \int_0^T |h^\circ_t| dt \right\|_p \right),$$

from which one immediately gets the general form.  $\square$

In the next section we will treat the particular case when the process  $Y$  has the form  $Y_t = u(t, X_t)$  and  $\xi = \Phi(X_T)$ ,  $h = h(t, X_t, y, z)$ . In such a situation, if moreover  $\Phi$  and  $h(\cdot, \cdot, 0, 0)$  are bounded one may use the last lemma to obtain the boundedness of  $u$ . Namely, taking the expectation with respect to each measure  $P^x$ ,  $x \in \mathcal{N}^c$  the above lemma leads to

$$u(0, x) \leq K \left( e^{\mu T} \|\Phi\|_\infty + \frac{1}{\mu} (e^{\mu T} - 1) \|h^\circ\|_\infty \right).$$

The following lemma establishes such uniform estimates in the general case considered here.

**Lemma 6.5** *There exists a constant  $K$ , which depends only on  $T$  and  $C$ , such that the solution  $(Y, Z)$  in  $L^2(P^\nu)$  with data  $(\xi, h)$  satisfies the following estimate*

$$e^{2\mu t} |Y_t|^2 + E \left[ \int_t^T e^{2\mu s} |Z_s \sigma(X_s)|^2 ds / \mathcal{F}_t \right] \leq KE [e^{2\mu T} \xi^2 + \left( \int_t^T e^{\mu s} |h^\circ_s| ds \right)^2 / \mathcal{F}_t].$$

In particular, if  $\xi$  and  $h^\circ$  are bounded one has

$$\sup_{0 \leq t \leq T} e^{\mu t} |Y_t| \leq \sqrt{K} \left( e^{\mu T} \|\xi\|_\infty + \frac{e^{\mu T} - 1}{\mu} \|h^\circ\|_\infty \right).$$

**Proof.** We suppose that  $\mu = 0$ . From the quadratic relation (28) one has

$$Y_t^2 + 2E \left[ \int_t^T |Z_s \sigma(X_s)|^2 ds / \mathcal{F}_t \right] \leq E[|\xi|^2 + 2 \int_t^T |Y_s| |h^\circ_s| ds + 2C \int_t^T |Y_s| |Z_s \sigma(X_s)| ds / \mathcal{F}_t].$$

Using the second estimate of Lemma 6.3 one deduces that the left hand side of this inequality is dominated as follows:

$$\begin{aligned} &\leq E[|\xi|^2 + 2|\xi| \left( \int_t^T |h^\circ_s| ds + C \int_t^T |Z_s \sigma(X_s)| ds \right) + \\ &\quad \left( \int_t^T |h^\circ_s| ds + C \int_t^T |Z_s \sigma(X_s)| ds \right)^2 / \mathcal{F}_t] \leq \\ &\leq E[2|\xi|^2 + 6 \left( \int_t^T |h^\circ_s| ds \right)^2 + 6C^2 \left( \int_t^T |Z_s \sigma(X_s)| ds \right)^2 / \mathcal{F}_t]. \end{aligned}$$

Now one uses the inequality  $\left( \int_t^T |Z_s \sigma(X_s)| ds \right)^2 \leq (T-t) \int_t^T |Z_s \sigma(X_s)|^2 ds$  and obtain the first inequality of the lemma provided  $3TC^2 < 1$ . For larger values of  $T$  one splits the interval  $[0, T]$  in intervals of length less than  $\frac{1}{3C^2}$  and iterates the estimate already obtained. The second inequality of the lemma is immediately obtained from the first.  $\square$

Note that in the preceding lemmas one do not make use of the global monotonicity of  $h$  but only of the monotonicity of  $h'$ . The global monotonicity is used in the next lemma which estimates the solution dependence of the data. The proof is completely similar.

**Lemma 6.6** *Let  $(Y, Z)$  be a solution in  $L^p(P^\nu)$  with data  $(\xi, h)$  (satisfying the relations (26) and (27)). Let  $(\eta, g)$  be a second couple of data satisfying the conditions from Definition 6.2 and let  $(U, V)$  be a solution in  $L^p(P^\nu)$  with data  $(\eta, g)$ . Then one has*

$$\begin{aligned} e^{\mu t} |Y_t - U_t| + \int_t^T e^{\mu s} |h_s(Y_s, V_s \sigma(X_s)) - h_s(U_s, V_s \sigma(X_s)) - \mu(Y_s - U_s)| ds + \\ + \int_t^T e^{\mu s} dO_s \leq e^{\mu T} |\xi - \eta| + \\ + \int_t^T e^{\mu s} |h_s(U_s, V_s \sigma(X_s)) - g_s(U_s, V_s \sigma(X_s))| ds + \\ + C \int_t^T e^{\mu s} |(Z_s - V_s) \sigma(X_s)| ds - \int_t^T e^{\mu s} \text{sign}(Y_s - U_s) (Z_s - V_s) \cdot dM_s, \end{aligned}$$

where  $O = (O_t)$  is the local time of the semimartingale  $Y - U$ . Moreover one has the following estimates in  $L^p$ , with the same constants as in Lemma 6.4

$$\begin{aligned} & \left\| \sup_{0 \leq t \leq T} e^{\mu t} |Y_t - U_t| \right\|_p + \left\| \left( \int_0^T e^{2\mu t} |(Z_t - V_t) \sigma(X_t)|^2 dt \right)^{\frac{1}{2}} \right\|_p + \\ & \leq K \left( e^{\mu T} \|\xi - \eta\|_p + \left\| \int_0^T e^{\mu t} |h_t(U_t, V_t \sigma(X_t)) - g_t(U_t, V_t \sigma(X_t))| dt \right\|_p \right), \\ & \left\| \int_0^T e^{\mu t} |h_t(Y_t, Z_t \sigma(X_t)) - h_t(U_t, V_t \sigma(X_t))| dt \right\|_p \leq \\ & \leq K' \left( e^{\mu T} \|\xi - \eta\|_p + \left\| \int_0^T e^{\mu t} |h_t(U_t, V_t \sigma(X_t)) - g_t(U_t, V_t \sigma(X_t))| dt \right\|_p \right). \end{aligned}$$

In the first inequality of this lemma one may change the term

$$\int_t^T e^{\mu s} |h_s(Y_s, V_s \sigma(X_s)) - h_s(U_s, V_s \sigma(X_s)) - \mu(Y_s - U_s)| ds$$

with the following one

$$\int_t^T e^{\mu s} |h_s(Y_s, Z_s \sigma(X_s)) - h_s(U_s, Z_s \sigma(X_s)) - \mu(Y_s - U_s)| ds.$$

The existence of solutions is ensured by the following theorem.



**Theorem 6.7** *Let  $(\xi, h)$  be the data fixed in the beginning of this subsection and let  $p > 1$  be fixed. Denote, for  $r > 0$ ,*

$$h^{*,r}(s, \omega) = \sup_{|y| \leq r} |h'(s, \omega, y)|$$

and let  $A$  be the set of all points  $x \in \mathcal{N}^c$  with the following properties

$$E^x \left( |\xi| + \int_0^T |h_t^0| dt \right)^p < \infty,$$

$$\int_0^T h_t^{*,r} dt < \infty, \quad P^x - a.s., \quad r > 0.$$

Then there exists a pair  $(Y_t, Z_t)_{0 \leq t \leq T}$  of predictable processes  $Y : [0, T] \times \Omega \rightarrow R$ ,  $Z : [0, T] \times \Omega \rightarrow R^N$  that represents a solution in  $L^p(P^x)$  of the BSDE with data  $(\xi, h)$  for each  $x \in A$ .

If  $\nu$  is a probability measure supported by  $A$  and such that  $E^\nu \left( |\xi| + \int_0^T |h_t^0| dt \right)^p < \infty$ , then an arbitrary solution  $(Y', Z')$  in  $L^p(P^\nu)$  of the BSDE with data  $(\xi, h)$  should coincide with  $(Y, Z)$  in the following sense: on has  $P^\nu - a.s.$   $Y_t = Y'_t$  for any  $t \in [0, T]$  and  $Z_t = Z'_t, dt - a.e.$

**Proof.** The uniqueness assertion follows immediately from Lemma 6.4. For the existence we only consider the case  $\mu = 0$  and will give the proof in three steps.

**FIRST STEP** First we assume that  $\xi$  and  $h$  are bounded. Then the result follows from Theorem 2.2 of [9].

**SECOND STEP** Now we assume that  $\xi$  and  $h^\circ$  are bounded and consider a sequence  $(h^k)$  of bounded functions which approximate  $h$  and satisfy the following conditions

$$1^\circ h^k(\cdot, \cdot, 0, 0) = h(\cdot, \cdot, 0, 0) = h^\circ,$$

$$2^\circ \text{ there exists some constant } c \text{ such that } |h^l - h^\circ| \leq c |h^k - h^\circ| \text{ for } l \leq k.$$

3° for given  $(t, \omega) \in [0, T] \times \Omega$  and  $r, q > 0$  there exists an index  $k_0$  such that  $h^k(t, \omega, y, z) = h(t, \omega, y, z)$  as soon as  $|y| \leq r, |z| \leq q$  and  $k \geq k_0$ .

4° each function  $h^k$  satisfies the conditions (26) and (27) with the same constants as  $h$ .

(For example the sequence  $h^k = ((h - h^\circ) \wedge k) \vee (-k) + h^\circ, k \in \mathbf{N}$ , satisfies these requirements. Observe that the third condition follows from the fact that the map  $(y, z) \rightarrow h(t, \omega, y, z)$  is bounded on  $\{|y| \leq r\} \times \{|z| \leq q\}$ .)

Let  $(Y^k, Z^k)$  be the solution in  $L^p(P^x)$  of the BSDE with data  $(\xi, h^k)$  under all  $P^x, x \in A$ , provided by the preceding step. We are going to prove that, for each  $p' \in (1, p)$ , there exists the limit  $\lim_{k \rightarrow \infty} (Y^k, Z^k) = (Y, Z)$ , in the  $L^{p'}$  sense and that  $(Y, Z)$  is a solution in  $L^p(P^x)$  of the BSDE with data  $(\xi, h)$ , for each  $x \in A$ . In order to simplify the presentation we will analyze first the convergence under an arbitrary fixed measure  $P^x$ , with  $x \in A$ . We leave to the reader the task of discussing the global measurability aspects.

First note that, on account of Lemma 6.5 the processes  $|Y^k|$ ,  $k \in \mathbf{N}$ , all are uniformly bounded by a constant say  $r > 0$ . Also, the solutions  $(Y^k, Z^k)$  are uniformly bounded in  $L^p$  by Lemma 6.4,

$$\begin{aligned} & \left\| \sup_{0 \leq t \leq T} |Y_t^k| \right\|_p + \left\| \left( \int_0^T |Z_t^k \sigma(X_t)|^2 dt \right)^{\frac{1}{2}} \right\|_p + \left\| \int_0^T |h_t^k(Y_t^k, Z_t^k \sigma(X_t))| dt \right\|_p \leq \\ & \leq K \left( \|\xi\|_p + \left\| \int_0^T |h_t^\circ| dt \right\|_p \right). \end{aligned} \quad (*)$$

We are now going to show that  $(Y^k, Z^k)$  is Cauchy. We start from the following inequality provided by Lemma 6.6

$$\begin{aligned} & \left\| \sup_{0 \leq t \leq T} |Y_t^k - Y_t^l| \right\|_{p'} + \left\| \left( \int_0^T |(Z_t^k - Z_t^l) \sigma(X_t)|^2 dt \right)^{\frac{1}{2}} \right\|_{p'} + \\ & \left\| \int_0^T |h_t^k(Y_t^k, Z_t^k \sigma(X_t)) - h_t^l(Y_t^l, Z_t^l \sigma(X_t))| dt \right\|_{p'} \leq \\ & \leq K \left( \left\| \int_0^T |h_t^k(Y_t^k, Z_t^k \sigma(X_t)) - h_t^l(Y_t^k, Z_t^k \sigma(X_t))| dt \right\|_{p'} \right). \end{aligned} \quad (**)$$

We use  $p' \in (1, p)$  in this inequality because with such a parameter we will next show that the term in the right hand side tends to zero. For each  $q > 0$  we set  $h^q = h \mathbf{1}_{\{|z| \leq q\}}$  and  $h^{q,k} = h^k \mathbf{1}_{\{|z| \leq q\}}$ . One clearly has

$$|h^q(t, \omega, y, z) - h^\circ(t, \omega)| \leq Cq + h^{*,r}(t, \omega), \quad |y| \leq r.$$

Therefore, on account of property 2° of the approximating sequence  $h^k$  we deduce

$$\left| h_t^{q,l}(Y_t^k, Z_t^k \sigma(X_t)) - h_t^\circ \right| \leq c \left| h_t^{q,k}(Y_t^k, Z_t^k \sigma(X_t)) - h_t^\circ \right| \leq c^2 (Cq + h_t^{*,r}) \quad k \in \mathbf{N}.$$

Then one may use the almost sure integrability property of  $h^{*,r}$ , ensured by the hypothesis, and property 3° from above to deduce the following relation

$$\lim_{l \rightarrow \infty} \int_0^T \sup_{k \geq l} \left| h_t^{q,k}(Y_t^k, Z_t^k \sigma(X_t)) - h_t^{q,l}(Y_t^k, Z_t^k \sigma(X_t)) \right| dt = 0, \quad P^x - a.s.$$

The boundedness in  $L^p$  provided in the relation (\*), implies then

$$\lim_{l \rightarrow \infty} \sup_{k \geq l} \left\| \int_0^T \left| h_t^{q,k}(Y_t^k, Z_t^k \sigma(X_t)) - h_t^{q,l}(Y_t^k, Z_t^k \sigma(X_t)) \right| dt \right\|_{p'} = 0. \quad (\#)$$

Similarly one has

$$\limsup_{l \rightarrow \infty} \sup_{k \geq l} \left\| \int_0^T |h_t^{l,k}(Y_t^k) - h_t^{l,l}(Y_t^k)| dt \right\|_{p'} = 0. \quad (\#\#)$$

Now we write

$$|h_t^k(y, z) - h_t^l(y, z)| \leq |h_t^{q,k}(y, z) - h_t^{q,l}(y, z)| + 2C|z| \mathbf{1}_{\{|z| > q\}} + |h_t^{l,k}(y) - h_t^{l,l}(y)|,$$

where we use the notation  $h^{l,k}(t, \omega, y) = h^k(t, \omega, y, 0) - h^o(t, \omega)$ , for any  $k \in \mathbf{N}$ . This leads to the inequality

$$\begin{aligned} & \left\| \int_0^T |h_t^k(Y_t^k, Z_t^k \sigma(X_t)) - h_t^l(Y_t^k, Z_t^k \sigma(X_t))| dt \right\|_{p'} \leq \\ & \leq \left\| \int_0^T |h_t^{q,k}(Y_t^k, Z_t^k \sigma(X_t)) - h_t^{q,l}(Y_t^k, Z_t^k \sigma(X_t))| dt \right\|_{p'} + \\ & 2C \left\| \int_0^T |Z_t^k \sigma(X_t)| \mathbf{1}_{\{|Z_t^k \sigma(X_t)| > q\}} dt \right\|_{p'} + \left\| \int_0^T |h_t^{l,k}(Y_t^k) - h_t^{l,l}(Y_t^k)| dt \right\|_{p'}. \end{aligned}$$

The first and the third terms in the right hand side of this inequality pass to the limit according to the relations (#) and (\#\#). In order to estimate the middle term of the right hand side we will employ Holder's inequality with respect to the powers  $(2, \frac{2p'}{p-p'}, \frac{2p'}{2p'-p})$ ,

$$\begin{aligned} & \int_0^T |Z_t^k \sigma(X_t)| \mathbf{1}_{\{|Z_t^k \sigma(X_t)| > q\}} dt \leq \\ & \leq T^{\frac{2p'-p}{2p'}} \left( \int_0^T |Z_t^k \sigma(X_t)|^2 dt \right)^{\frac{1}{2}} \left( \int_0^T \mathbf{1}_{\{|Z_t^k \sigma(X_t)| > q\}} dt \right)^{\frac{p-p'}{2p'}}. \end{aligned}$$

This implies a restriction on  $p'$ , namely one should take  $p' > \frac{p}{2}$ , but do not alter our argumentation. After a new use of Holder's inequality one gets

$$\begin{aligned} & E \left( \int_0^T |Z_t^k \sigma(X_t)| \mathbf{1}_{\{|Z_t^k \sigma(X_t)| > q\}} dt \right)^{p'} \leq \\ & \leq T^{\frac{2p'-p}{2}} \left( E \left( \int_0^T |Z_t^k \sigma(X_t)|^2 dt \right)^{\frac{p'}{2}} \right)^{\frac{p'}{p}} \left( E \left( \int_0^T \mathbf{1}_{\{|Z_t^k \sigma(X_t)| > q\}} dt \right)^{\frac{p}{2}} \right)^{\frac{p-p'}{p}}. \end{aligned}$$

>From the estimate (\*) one deduces further that the last quantity is less than a constant multiplied by  $\frac{1}{q^{p-p'}}$ . For large  $q$  this last quantity becomes arbitrarily small. Thus we may conclude that

$$\lim_{l \rightarrow \infty} \sup_{k \geq l} \left\| \int_0^T |h_t^k(Y_t^k, Z_t^k \sigma(X_t)) - h_t^l(Y_t^k, Z_t^k \sigma(X_t))| dt \right\|_{p'} = 0$$

and this in turn implies the following limit relations

$$\lim_{l \rightarrow \infty} \sup_{k \geq l} \left( \left\| \sup_{0 \leq t \leq T} |Y_t^k - Y_t^l| \right\|_{p'} + \left\| \left( \int_0^T |(Z_t^k - Z_t^l) \sigma(X_t)|^2 dt \right)^{\frac{1}{2}} \right\|_{p'} \right) = 0,$$

$$\lim_{l \rightarrow \infty} \sup_{k \geq l} \left\| \int_0^T |h_t^k(Y_t^k, Z_t^k \sigma(X_t)) - h_t^l(Y_t^l, Z_t^l \sigma(X_t))| dt \right\|_{p'} = 0. \quad (\&)$$

>From the first of these relations one deduces that  $(Y^k, Z^k)$  has a limit  $(Y, Z)$  in the following sense

$$\lim_{k \rightarrow \infty} \left( \left\| \sup_{0 \leq t \leq T} |Y_t^k - Y_t| \right\|_{p'} + \left\| \left( \int_0^T |(Z_t^k - Z_t) \sigma(X_t)|^2 dt \right)^{\frac{1}{2}} \right\|_{p'} \right) = 0.$$

Clearly relation (\*) ensures  $\left\| \left( \int_0^T |Z_t \sigma(X_t)|^2 dt \right)^{\frac{1}{2}} \right\|_p < \infty$ .

>From relation (&) we deduce that the sequence  $h_t^k(Y_t^k, Z_t^k \sigma(X_t))$ ,  $k \in \mathbf{N}$ , is a Cauchy sequence in  $L^1(dt \otimes dP^x)$ . The property 3° of the approximating sequence  $(h^k)$  allows us to easily identify the limit

$$\lim_{k \rightarrow \infty} h_t^k(Y_t^k, Z_t^k \sigma(X_t)) = h_t(Y_t, Z_t \sigma(X_t)).$$

Finally the above arguments allow one to conclude that  $(Y, Z)$  is an  $L^p(P^x)$  solution corresponding to the data  $(\xi, h)$ .

THIRD STEP Now we can prove the general result. Put

$$\xi^k = (\xi \wedge k) \vee (-k), \quad h^{\circ, k} = (h^\circ \wedge k) \vee (-k), \quad h^k = h - h^\circ + h^{\circ, k}$$

and denote by  $(Y^k, Z^k)$  the solution corresponding to the data  $(\xi^k, h^k)$  provided by the preceding step. The second inequality of Lemma 6.6, again ensures that the sequence  $(Y^k, Z^k)$  is a Cauchy sequence and, consequently has a limit "in  $L^p$ ". Like in the preceding step one deduces that the limit represents the solution we are looking for.  $\square$

**Remark 6.8** *The treatment of the monotonicity presented in this section uses some new estimates for the solution. These estimates reveal that, with  $p = 2$ , the good norm for  $h^\circ$  is  $\left\| \int_0^T |h^\circ_t| dt \right\|_2$  and not  $\left( \int_0^T \|h^\circ_t\|_2^2 dt \right)^{\frac{1}{2}}$ , as was previously used in the literature of BSDE.*

## 6.2 Probabilistic interpretation of weak analytical solutions

We shall now look at the connection between the solutions of BSDE's introduced in this section and PDE's studied in Section 3. In order to do this we have to consider BSDE's over time intervals like  $[s, T]$  with  $0 \leq s \leq T$ , and so we first introduce some notation and terminology related to these intervals. Our notation differs from the standard notation of BSDE's, that is essentially formalized for time inhomogeneous SDE,s. Since the present approach is based on the theory of Markov processes, which is a time homogeneous theory, we have to discuss solutions over the interval  $[s, T]$ , while the process and the coordinate martingales are indexed by a parameter in the interval  $[0, T - s]$ .

Let  $\xi$  be a  $\mathcal{F}_{T-s}$ -measurable random variable and  $h : [s, T] \times \Omega \times R \times R^n \rightarrow R$  a measurable process such that  $(h(s+l, \omega, \cdot, \cdot))_{l \in [0, T-s]}$  is predictable (with respect to  $(\mathcal{F}_l)_{l \in [0, T-s]}$ ). Let  $\nu$  be a probability measure in  $\mathcal{N}^c$  such that  $E^\nu \xi^2 < \infty$ . We say that a pair  $(Y_t, Z_t)_{s \leq t \leq T}$  of processes  $Y : [s, T] \times \Omega \rightarrow R$ ,  $Z : [s, T] \times \Omega \rightarrow R^N$  is a *solution of the BSDE over the interval  $[s, T]$*  under the measure  $P^\nu$  with data  $(\xi, h)$  provided that  $Y$  is continuous and together satisfy the integrability conditions

$$E^\nu \int_s^T |h(t, \cdot, Y_t, Z_t \sigma(X_{t-s}))| dt < \infty,$$

$$E^\nu \int_s^T |Z_t \sigma(X_{t-s})|^2 dt < \infty,$$

and the following equation under  $P^\nu$ ,

$$Y_t = \xi + \int_t^T h(r, \omega, Y_r, Z_r \sigma(X_{r-s})) dr - \int_0^{T-t} Z_{t+l} \cdot dM_l, s \leq t \leq T.$$

Now let us assume that  $f : [0, T] \times R^N \times R \times R^n \rightarrow R$  is again the function appearing in the basic equation (13). Note however that before these functions were considered more as  $L^2$  elements, while now it will be important to look at them as pointwise defined functions. Like in Section 3. we use the notation  $f^\circ(t, x) = f(t, x, 0, 0)$ . The next result interpret probabilistically Theorem 3.3. The proof is inspired from Subsection 3.1.1 of [5].

**Theorem 6.9** *Assume that the function  $f$  satisfies conditions (11), (12) and  $\Phi \in L^2(R^N)$ . Denote by*

$$A = \{(s, x) \in [0, T] \times \mathcal{N}^c / E^x \left( \Phi^2(X_{T-s}) + \int_s^T |f^\circ(t, X_{t-s})|^2 dt \right) < \infty\},$$

$$A_s = \{x \in \mathcal{N}^c / (s, x) \in A\}, s \in [0, T].$$

*Then the set  $A$  is absorbent for the time-space process  $\widehat{X}$  and  $m(A_s^c) = 0$ , for each  $s \in [0, T]$ . There exist universally measurable functions  $(u, \phi)$ ,  $u : A \rightarrow R$ ,  $\phi : A \rightarrow R^N$ , such that, for each  $s \in [0, T]$  and each  $x \in A_s$ , the pair  $(u(t, X_{t-s}), \phi(t, X_{t-s}))_{s \leq t \leq T}$  solves the*

BSDE under  $P^x$  with data  $(\Phi(X_{T-s}), f(t, X_{t-s}, y, z))$  over the interval  $[s, T]$ . The functions  $u, \phi$  also satisfy the following relations, for  $(s, x) \in A$ ,

$$u(s, x) = E^x(\Phi(X_{T-s})) + \int_s^T E^x(f(t, X_{t-s}, u(t, X_{t-s}), \phi\sigma(t, X_{t-s}))dt, \quad (*)$$

$$E^x \left( \sup_{t \in [s, T]} |u(t, X_{t-s})|^2 + \int_s^T |\phi\sigma(t, X_{t-s})|^2 dt \right) \leq$$

$$\leq K E^x \left( |\Phi(X_{T-s})|^2 + \int_s^T |f^\circ(t, X_{t-s})|^2 dt \right).$$

Moreover, the class of  $u1_A$  is an element of  $\widehat{F}$  which is a weak solution of (13) and  $\phi\sigma$  represents a version of  $D_\sigma u$ .

**Proof.** The properties of  $A$  are ensured by condition (11) and by the strong Markov property.

Let us construct the functions  $u, \phi$ . For  $s \in [0, T]$ , denote by  $(Y_t^s, Z_t^s)_{s \leq t \leq T}$  the solution of the BSDE over the interval  $[s, T]$ , under  $P^x, x \in A_s$  with data  $(\Phi(X_{T-s}), f(t, X_{t-s}, y, z))$ . By the uniqueness part of Theorem 6.7 one may deduce that

$$Y_t^{s+r} \circ \theta_r = Y_t^s, \quad t \in [s+r, T], \quad P^x - a.s.,$$

$$(Z_t^{s+r} \sigma(X_{t-s-r})) \circ \theta_r = Z_t^s \sigma(X_{t-s}), \quad dt \times P^x - a.e.,$$

for each fixed  $r \in [0, T-s)$  and all measures  $P^x, x \in A_s$ . In particular, if we define

$$u(s, x) =: E^x(Y_s^s),$$

we will have

$$u(t, X_{t-s}) = E^{X_{t-s}}(Y_t^t) = E^x(Y_t^t \circ \theta_{t-s} / \mathcal{F}_{t-s}) = E^x(Y_t^s / \mathcal{F}_{t-s}) = Y_t^s \quad P^x - a.s, \forall x \in A_s.$$

Denote  $W_l(s, \omega) = (W_{1,l}(s, \omega), \dots, W_{n,l}(s, \omega)) = Z_{l+s}^s \sigma(X_l)(\omega)$ , for  $(s, \omega) \in [0, T] \times \Omega$  and  $l \in [0, t-s)$ . One has  $W_l(r+s, \theta_r(\omega)) = W_{l+r}(s, \omega)$ ,  $dl \times P^x$  -a.e. In terms of the time-space Markov process  $\widehat{X}$  introduced in the preceding section, this writes  $W_l(\widehat{\theta}_r(s, \omega)) = W_{l+r}(s, \omega)$ . Therefore  $t \rightarrow A_{k,t}(s, \omega) = \int_0^{t \wedge T} W_{k,l}(s, \omega) dl$  represents an additive functional for the time-space process  $\widehat{X}$ . By definition this functional is absolutely continuous with respect to  $dt$  so Theorem 66.2 in [12] asserts that there exists an universally measurable function  $\psi_k : [0, T] \times R^N \rightarrow R$ , such that  $\psi_k(t, X_{t-s}(\omega)) = W_{k,t-s}(s, \omega)$ ,  $dt \times P^x$  -a.e. We define

$$\phi =: \psi\tau$$

where  $\tau$  is given in the beginning of the paper. Note that  $\phi\sigma(t, X_{t-s}) = Z_t^s(\sigma\tau\sigma)(X_{t-s}) = Z_t^s\sigma(X_{t-s})$ , so that we conclude

$$Z_t^s\sigma(X_{t-s}) = \phi\sigma(t, X_{t-s}) \quad dt \times P^x \text{ a.e.}, \forall x \in A_s.$$

Now one easily deduces relation (\*) from the statement. In particular, one deduces that  $t \rightarrow u(t, X_{t-s})$  is continuous  $P^x - a.s.$  for each  $x \in A_s$ . This implies the equality of the processes  $u(\cdot, X_{\cdot-s}) = Y^s$ , completing the proof of the fact that  $u(\cdot, X_{\cdot-s}), \phi(\cdot, X_{\cdot-s})$  solves the BSDE over the time interval  $[s, T]$ . The remainder proof presents no more difficulties.  $\square$

The above result suggests us the following definition.

**Definition 6.10** *Let  $A$  be an absorbent set with respect to the space-time process  $\widehat{X}$ . A pair of universally measurable functions  $(u, \phi), u : A \rightarrow R, \phi : A \rightarrow R^N$ , is called solution of equation (13) in the sense of BSDE's on  $A$  provided that, for each  $s \in [0, T]$  and each  $x \in A_s$ , the pair  $(u(t, X_{t-s}), \phi(t, X_{t-s}))_{s \leq t \leq T}$  solves the BSDE under  $P^x$  with data  $(\Phi(X_{T-s}), f(t, X_{t-s}, y, z))$  over the interval  $[s, T]$ .*

Obviously, the uniqueness ensured by Lemma 6.4 implies that if  $(u', \phi')$  is another solution in the sense of BSDE's on  $A$ , then  $u \equiv u'$  and  $\phi\sigma = \phi'\sigma$  modulo sets of zero potential with respect to  $\widehat{X}$  (i.e.  $\phi\sigma(t, X_{t-s}) = \phi'\sigma(t, X_{t-s}), dt \times P^x - a.s.$ , for any  $(s, x) \in A$ ). So, unlike the generalized gradient  $D_\sigma u$  of an arbitrary function in  $\widehat{F}$ , which is determined up to  $dt \times m$ -negligible sets, the gradient  $\phi\sigma$  of a solution in the sense of BSDE's is more precisely determined. The preceding theorem ensures existence of solutions in the sense of BSDE's on a maximal set  $A$  and describes their basic properties. Clearly, if  $\Phi$  and  $f^\circ$  are bounded one has  $A = [0, T] \times \mathcal{N}^c$ . Relation (\*) from the statement is a kind of punctual mild equation. As we have seen, the mild equation (15) has a meaning on  $L^2$  elements and it is equivalent to the weak equation (14). The punctual mild equation (\*) has a stronger sense.

However, compared with the notion of solution in the sense of BSDE's, relation (\*) alone is weaker, because it does not determine the gradient as precisely as BSDE's do. In the case of the linear mild equation, for example, the gradient  $D_\sigma u$  would not appear at all in (\*). Therefore one can not deduce that a pair  $(u, \phi)$  which satisfies the punctual mild equation is necessarily a solution in the sense of BSDE.

## 7 Comparison theorem and a control problem

We shall first solve explicitly the linear BSDE

$$Y_t = \xi + \int_t^T (\alpha_s + \beta_s Y_s + \langle \gamma_s, Z_s \sigma(X_s) \rangle) ds - \int_t^T Z_s . dM_s \quad (29)$$

where  $\alpha, \beta, \gamma^i, i = 1, n$ , are adapted processes,  $\gamma = (\gamma^i)$ , the process  $\beta$  is assumed to be bounded from above, the processes  $\gamma^i$  are bounded and the following integrability conditions are satisfied

$$\int_0^T |\beta_t| dt < \infty, P^x - a.s., x \in \mathcal{N}^c, \quad (*)$$

$$\int_0^T |\alpha_t| dt, \xi \in \cap_{x \in \mathcal{N}^c} L^2(P^x).$$

This linear equation satisfies the conditions from the hypothesis of Theorem 6.7 with  $p = 2$ . So the solution exists and satisfies the estimates from Lemma 6.4. The next lemma gives a formula for the solutions in  $L^2$ .

**Lemma 7.1** *The solution in  $L^2(P^x)$  of the linear equation (29) is given, under  $P^x, x \in \mathcal{N}^c$ , by*

$$Y_t = E^x(\xi \Gamma_{t,T} + \int_t^T \Gamma_{t,s} \alpha_s ds / \mathcal{F}_t)$$

where  $\Gamma_{t,s} = \exp(N_s - N_t - \frac{1}{2}(\langle N \rangle_s - \langle N \rangle_t) + \int_s^t \beta_r dr)$  with  $N_t = \frac{1}{2} \int_0^t \gamma_s \tau(X_s) .dM_s^i$  and  $\tau$  is the matrix defined at the beginning of the paper.

**Proof.** Because  $\exp(2N_t - \frac{1}{2}\langle 2N \rangle_t)$  is a supermartingale and  $\langle N \rangle$  is bounded it follows that  $\exp(2N_t - \langle N \rangle_t)$  is integrable. From this one deduces that  $\sup_{t \leq s \leq T} \Gamma_{t,s}$  is square integrable and that  $\int_t^T N_{t,s} |\alpha_s| ds$  is  $P^x$ -integrable, for each  $x \in \mathcal{N}^c$ . Therefore the expression in the right hand side of the formula asserted by the lemma makes sense.

Now we note that, because of the linearity of our problem, there is no loss of generality if we assume, and so we will do in this proof, that  $\xi$  and  $\alpha$  are nonnegative.

We first treat the case where the process  $\beta$  is bounded. The proof is analogous to the proof in the standard case (see e.g. [5]). Denote by  $\Gamma_t = \Gamma_{0,t}$  and  $\bar{Y}_t = Y_t \Gamma_t + \int_0^t \alpha_s \Gamma_s ds$ , where  $Y$  has the expression in the statement, and note that  $\bar{Y}_t = E^x(\bar{Y}_T / \mathcal{F}_t)$ . Then, by the representation theorem 5.4, we may write  $\bar{Y}_t = \int_0^t \bar{Z}_s .dM_s$  with  $\bar{Z} = (\bar{Z}_i)$ , where  $\bar{Z}_i, i = 1, N$  are some predictable processes. We also note that  $d\Gamma_t = \Gamma_t(dN_t + \beta_t dt)$  and  $d\Gamma_t^{-1} = -\Gamma_t^{-1}(dN_t + \beta_t dt - d\langle N \rangle_t)$ . Since  $Y_t = (\bar{Y}_t - \int_0^t \alpha_s \Gamma_s ds) \Gamma_t^{-1}$  we get

$$\begin{aligned} dY_t &= -Y_t(dN_t + \beta_t dt - d\langle N \rangle_t) + \Gamma_t^{-1} \bar{Z}_t .dM_t - \alpha_t dt - d\langle \bar{Z} .dM, \Gamma^{-1} dN \rangle_t = \\ &= (\Gamma_t^{-1} \bar{Z}_t - Y_t \frac{1}{2} \gamma_t \tau(X_t)) .dM_t - Y_t \beta_t dt - \alpha_t dt - \\ &\quad - 2 \left\langle (\Gamma_t^{-1} \bar{Z}_t - Y_t \frac{1}{2} \gamma_t \tau(X_t)) \sigma(X_t), (\frac{1}{2} \gamma_t \tau(X_t)) \sigma(X_t) \right\rangle dt. \end{aligned}$$

We denote  $Z_t = \Gamma_t^{-1} \bar{Z}_t - \frac{1}{2} Y_t (\tau \gamma_t)$ . The last term in the above equality is

$$- \langle Z_t \sigma(X_t), \gamma_t \tau \sigma(X_t) \rangle dt = - \langle Z_t \sigma(X_t), \gamma_t \rangle dt = - \langle \gamma_t, Z_t \sigma(X_t) \rangle dt,$$

where we have used the relation  $\sigma \sigma^* \tau^* = \sigma \tau \sigma = \sigma$  (see the appendix). So we get

$$dY_t = Z_t .dM_t - Y_t \beta_t dt - \alpha_t dt - \langle \gamma_t, Z_t \sigma(X_t) \rangle dt,$$

that is the equation (29).



When  $\beta$  is not bounded we will define an approximating sequence as follows:  $\beta^k = \beta \vee (-k), k \in \mathbf{N}$ . Let us denote by  $\Gamma_{t,s}^k, \Gamma_t^k, \bar{Y}^k, \bar{Z}^k, Y^k, Z^k$  the processes corresponding to  $\beta^k$  according to the above notation. One easily sees that the sequence  $\Gamma_{t,s}^k, k \in \mathbf{N}$ , is decreasing and hence the sequence of processes defined by the formula from the statement is nonnegative and decreases:  $\lim_{k \rightarrow \infty} Y^k = Y$ . The sequence of random variables  $\bar{Y}_T^k = \xi \Gamma_T^k + \int_0^T \alpha_t \Gamma_t^k dt, k \in \mathbf{N}$ , is also decreasing to  $\bar{Y}_T$ , and consequently the processes representing them,  $\bar{Z}^k$ , converge locally

$$E \left( \bar{Y}_\tau^k - \bar{Y}_\tau \right)^2 = E \left[ \int_0^\tau \left| \left( \bar{Z}_t^k - \bar{Z}_t \right) \sigma(X_t) \right|^2 dt \right] \rightarrow 0,$$

where  $\tau$  is the first time when  $\bar{Y}^0$  exceeds a certain, arbitrary, value. Then clearly one has the local convergence of the local martingales:

$$\int_0^t \left( \bar{\Gamma}_s^k \right)^{-1} \bar{Z}_s^k . dM_s \rightarrow \int_0^t \left( \bar{\Gamma}_s \right)^{-1} \bar{Z}_s . dM_s.$$

>From the relation that express  $Z$  in terms of  $\bar{Z}$ , used in the preceding step of the demonstration, one concludes that the following convergence also holds locally,

$$\int_0^t Z_s^k . dM_s \rightarrow \int_0^t Z_s . dM_s.$$

The estimates of Lemma 6.4 hold in the limit for  $(Y, Z)$ . In order to verify that this pair is an  $L^2$  solution it remains to verify that the relation (29) pass to the limit *a.s.* The terms that contain  $Z$  pass to the limit by the arguments already discussed. The convergence of the term  $\int_t^T Y_s^k \beta_s^k ds$  is ensured by the inequality

$$\int_0^T Y_t^k |\beta_t^k| dt \leq \left( \sup_{0 \leq t \leq T} Y_t^0 \right) \int_0^T |\beta_t| dt < \infty, a.s.$$

This finishes the proof.  $\square$

We shall now use this result in order to prove the following comparison result.

**Theorem 7.2** *Consider two couples of data  $(\xi, f)$  and  $(\bar{\xi}, \bar{f})$  which satisfy the conditions (26) and (27) and admit solutions in  $L^2(P^\nu)$  where  $\nu$  is supported by  $\mathcal{N}^c$ ; denote by  $(Y, Z)$  and  $(\bar{Y}, \bar{Z})$  the corresponding solutions. Assume that*

$$\xi \geq \bar{\xi}, P^\nu - a.s. \text{ and } f(s, \omega, Y_s, Z_s \sigma(X_s)) \geq \bar{f}(s, \omega, Y_s, Z_s \sigma(X_s)), ds \times dP^\nu - a.e.$$

*Then  $Y \geq \bar{Y}, ds \times dP^\nu - a.e.$  The comparison is strict in the sense that, on the event  $Y_t = \bar{Y}_t$ , one has  $\xi = \bar{\xi}, f(s, \omega, Y_s, Z_s \sigma(X_s)) = \bar{f}(s, \omega, \bar{Y}_s, \bar{Z}_s \sigma(X_s))$ , and  $Y_s = \bar{Y}_s$ , for every  $s \geq t$ ,  $P^\nu - a.s.$*

**Proof.** The proof is the same as in [9]. We denote

$$\begin{aligned}\Delta Y &= Y - \bar{Y}, \quad \Delta Z = Z - \bar{Z}, \quad \Delta \xi = \xi - \bar{\xi} \\ \delta f_s &= f(s, \omega, Y_s, Z_s \sigma(X_s)) - \bar{f}(s, \omega, Y_s, Z_s \sigma(X_s)) \\ \delta_y f_s &= (\bar{f}(s, \omega, Y_s, Z_s \sigma(X_s)) - \bar{f}(s, \omega, \bar{Y}_s, Z_s \sigma(X_s))) / \Delta Y_s\end{aligned}$$

and, for each  $i = 1, \dots, n$ ,

$$\begin{aligned}\delta_{z_i} f_s &= (\bar{f}(s, \omega, \bar{Y}_s, V_s^i) - \bar{f}(s, \omega, \bar{Y}_s, V_s^{i-1})) / (\Delta Z_s \sigma(X_s))_i \quad \text{with} \\ V_s^i &= ((Z_s \sigma(X_s))_1, \dots, (Z_s \sigma(X_s))_i, (\bar{Z}_s \sigma(X_s))_{i+1}, \dots, (\bar{Z}_s \sigma(X_s))_n).\end{aligned}$$

We take  $\delta_y f_s = 0$  (respectively  $\delta_{z_i} f_s = 0$ ) if  $\Delta Y_s = 0$  (respectively  $(\Delta Z_s \sigma(X_s))_i = 0$ ).

Using the assumptions (26) and (27) one may easily check that  $\delta_{z_i} f_s$  are bounded and  $\delta_y f_s \leq \mu$ . We also notice that  $(\Delta Y, \Delta Z)$  solve the linear BSDE

$$\Delta Y_t = \Delta \xi + \int_t^T (\delta f_s + (\delta_y f_s) \Delta Y_s + \langle \delta_{z_i} f_s, \Delta Z_s \sigma(X_s) \rangle) ds - \int_t^T \Delta Z_s \cdot dM_s$$

and so

$$\Delta Y_t = E^x(\Delta \xi \Gamma_T \Gamma_t^{-1} + \int_t^T \Gamma_s \Gamma_t^{-1} \delta_s f ds / F_t) \geq 0$$

with  $\Gamma$  given in the previous lemma (one takes  $\alpha = \delta f$ ,  $\beta = \delta_y f$  and  $\gamma = \delta_{z_i} f$ ). This proves the assertions of the theorem.  $\square$

We shall now consider a control problem associated to the Markov process  $X$ . An admissible control is a process  $\theta(t, \omega)$  which is progressively measurable with respect to the filtration  $(\mathcal{F}_t)_{t \geq 0}$  and which takes values in a compact subset  $K$  of some metric space. We denote by  $\Theta$  the class of admissible controls.

A bounded measurable function  $b : [0, T] \times R^N \times K \rightarrow R^N$  is given and we suppose that it is continuous with respect of the last variable. For a given admissible control  $\theta$  we define  $N_t^\theta = \int_0^t b_s(X_s, \theta_s) \cdot dM_s$ ,  $\Gamma_t^\theta = \exp(N_t^\theta - \frac{1}{2} \langle N^\theta \rangle_t)$ , and  $P^{\theta, x} = \Gamma^\theta \cdot P^x$ . The payoff function of the control problem is defined as

$$J^\theta(x) = E^{\theta, x}(g(X_T) + \int_0^T h(s, X_s, \theta_s) ds),$$

where  $g$  and  $h$  are bounded measurable functions and  $h$  is continuous in  $\theta$ . One wants to minimize the payoff function, that is to calculate the value function

$$J^*(x) = \inf_{\theta \in \Theta} J^\theta(x)$$

and to find an optimal control  $\theta^*$ , that is an admissible control such that  $J^*(x) = J^{\theta^*}(x)$ .

In what follows we shall restrict our analysis to points  $x \in \mathcal{N}^c$ . We shall now calculate  $J^\theta(x)$  by solving the BSDE

$$Y_t^\theta = g(X_T) + \int_t^T H_s(X_s, Z_s^\theta \sigma(X_s), \theta_s) - \int_t^T Z_s^\theta \cdot dM_s \quad (30)$$

where  $H : [0, T] \times R^N \times R^n \times K$  is the Hamiltonian defined by

$$H(s, x, z, \theta) = h(s, x, \theta) + 2 \langle b(s, x, \theta) \sigma, z \rangle.$$

Let us prove that

$$J^\theta(x) = Y_0^{\theta, x}$$

where  $Y_0^{\theta, x}$  is the initial value of the solution of the preceding equation (30) under  $P^x$ . Indeed, one has

$$\begin{aligned} \Gamma_T^\theta Y_T^\theta &= Y_0^\theta + \int_0^T Y_s^\theta \Gamma_s^\theta dN_s^\theta - \int_0^T \Gamma_s^\theta H_s(X_s, Z_s^\theta \sigma(X_s), \theta_s) ds + \\ &+ \int_0^T \Gamma_s^\theta Z_s^\theta \cdot dM_s + 2 \int_0^T \Gamma_s^\theta \langle b \sigma(X_s), Z_s^\theta \sigma(X_s) \rangle ds. \end{aligned}$$

Taking the expectation with respect to  $P^x$  one gets

$$E^{\theta, x} g(X_T) = E^x \Gamma_T^\theta Y_T^\theta = E^x Y_0^\theta - E^x \int_0^T \Gamma_s^\theta h_s(X_s, \theta_s) ds = Y_0^{\theta, x} - E^{\theta, x} \int_0^T h_s(X_s, \theta_s) ds,$$

that is the asserted relation.

In order to calculate the value function and to produce the optimal control we have to solve the following BSDE

$$Y_t^* = g(X_T) + \int_t^T H^*(s, X_s, Z_s^* \sigma(X_s)) - \sum_{i=1}^N \int_t^T Z_{i,s}^* dM_s^i, \quad (31)$$

where

$$H^*(s, x, z) = \inf_{\theta \in K} H(s, x, z, \theta).$$

It is easy to check that  $z \rightarrow H^*(s, x, z)$  is Lipschitz continuous, so that there exists a unique solution  $(Y^*, Z^*)$  of equation (31). To be more specific, we denote by  $(Y_t^{*, x}, Z_t^{*, x})_{0 \leq t \leq T}$  the solution of the equation under  $P^x$ , so that the initial value of the solution is a constant:  $Y_0^{*, x} = E^x Y_0^*$ .

Since  $H$  is continuous as a function of  $\theta$  and  $K$  is a compact set the infimum is attained at a point  $\theta^*$  and, by a measurable selection theorem one may choose a measurable function  $(s, x, z) \rightarrow \theta^*(s, x, z)$  which realizes the infimum. We construct the optimal control in the following way:  $\theta_s^* := \theta^*(s, \omega) = \theta^*(s, X_s(\omega), Z_s^*(\omega))$ .

**Corollary 7.3** *Under the above hypotheses  $J^*(x) = Y_0^{*, x}$  and  $\theta^*$  is an optimal control.*

**Proof.** Note that

$$H^*(s, X_s, Z_s^* \sigma(X_s)) = \inf_{\theta \in K} H(s, X_s, Z_s^* \sigma(X_s), \theta) = H(s, X_s, Z_s^* \sigma(X_s), \theta_s^*).$$

It follows that  $Y^* = Y^{\theta^*}$ . On the other hand, for every admissible control  $\theta$ , we have  $H(s, X_s, Z_s^\theta \sigma(X_s), \theta_s) \geq H(s, X_s, Z_s^{\theta^*} \sigma(X_s), \theta_s^*)$ , and consequently, by the comparison theorem,  $J^\theta(x) = Y_0^{\theta, x} \geq Y_0^{\theta^*, x} = J^{\theta^*}(x)$ . It follows that  $J^*(x) = Y_0^{*, x}$  and  $\theta^*$  is an optimal control.  $\square$

We now may interpret the solution of the Hamilton Jacobi Bellman equation as the value function of the above control problem. The HJB equation is

$$(\partial_t + L)u + H^*(t, x, \nabla u \sigma) = 0, \quad u(T, x) = g.$$

We have proved in the previous section that this equation admits a unique solution on  $\mathcal{N}^c$  in the sense of punctual mild equation and it satisfies  $u(0, x) = Y_0^{*, x} = J^*(x)$ .

## 8 Appendix

Let us denote by  $M(N, n)$  the set of all real matrices with  $N$  rows and  $n$  columns. Each such matrix represents a linear operator from  $\mathbf{R}^n$  to  $\mathbf{R}^N$ . On the other hand, the set  $M(N, n)$  is isomorphic to the vector space  $\mathbf{R}^{N \times n}$ , so that it has a natural topology and particularly one may speak about Borel measurability.

**Lemma 8.1** *For given natural numbers  $n$  and  $N$ , there exists a measurable function  $f : M(N, n) \rightarrow M(n, N)$  such that, for each  $E \in M(N, n)$  the matrix  $F = f(E)$  has the following properties:  $1^0$   $EF$  represents the operator of orthogonal projection onto the image space  $E\mathbf{R}^n (\subset \mathbf{R}^N)$ ,  $2^0$   $FE$  represents the operator of orthogonal projection onto the space  $(\ker E^* E)^\perp (\subset \mathbf{R}^n)$ , that is onto the orthogonal complement of the kernel of  $E$ . In particular  $EFE = E$ ,  $EF = F^* E^*$ ,  $FE = E^* F^*$  and  $\|EF\| = \|FE\| = 1$ , with the operator norm.*

**Proof.** The proof of this lemma follows from an analysis of the polar decomposition of the matrix  $E$ . Recall first that the series

$$\sqrt{1+x} = 1 + \sum_{k=0}^{\infty} (-1)^k \frac{(2k-1)!}{(2k)!} \frac{1}{2k+2} x^{k+1}$$

has convergence radius 1, so that for  $A \in M(N, N)$  with  $\|A\| < 1$ , one may define

$$(I+A)^{\frac{1}{2}} = I + \sum_{k=0}^{\infty} (-1)^k \frac{(2k-1)!}{(2k)!} \frac{1}{2k+2} A^k.$$

Then, if  $A$  is non- negative definite, we may write

$$A + \varepsilon I = (\|A\| + \varepsilon) \left[ I - (\|A\| + \varepsilon)^{-1} (\|A\| \cdot I - A) \right],$$

and, because  $\| \| A \| \cdot I - A \| \leq \| A \|$ , one may use the above series to compute  $(A + \varepsilon I)^{\frac{1}{2}}$ . Therefore, the square root is obtained as a measurable function of  $A$ ,

$$A^{\frac{1}{2}} = \lim_{k \rightarrow \infty} \left( A + \frac{1}{k} I \right)^{\frac{1}{2}}.$$

If  $E \in M(N, n)$  set  $A = E^* E$ ,  $B = A^{\frac{1}{2}}$  and note that  $U = A\mathbf{R}^n = B\mathbf{R}^n$ , the image space is the same. Let  $P$  the matrix representing the orthogonal projection from  $\mathbf{R}^n$  onto  $U$ . Then one has  $\ker E = \ker A = \ker B = U^\perp$  and  $PB = BP = B$ .

The following matrix

$$C = \lim_{\varepsilon \rightarrow 0} A^{\frac{1}{2}} (A + \varepsilon I)^{-1},$$

is well defined and satisfies the relations  $PC = CP = C$ ,  $BC = CB = P$ . This can easily be seen by choosing an orthonormal base of eigenvectors of  $A$  and performing a change of base in  $\mathbf{R}^n$  so that all these matrices become diagonal matrices. Further we set  $T = EC$  and observe that one has  $TB = E$ , which is the well-known polar decomposition of the matrix  $E$ . One immediately checks that  $T$  is an isometry from  $U$  to  $V = E\mathbf{R}^n = T\mathbf{R}^n \subset \mathbf{R}^N$ . The matrix  $Q$ , corresponding to the orthogonal projection from  $\mathbf{R}^N$  to  $V$ , satisfies  $QT = T$ . One also has  $TP = T$  and  $PT^* = T^*$ ,  $T^*Q = T^*$ . Using these one successively verifies that  $1^0 T^*T = P$ ,  $2^0 T^*$  is an isometry from  $V$  to  $U$  and  $3^0 TT^* = Q$ . We conclude then that the matrix  $F = CT^*$  satisfies the relations  $EF = Q$  and  $FE = P$ . Clearly  $F$  is expressed as a measurable function of  $E$ .  $\square$

## References

- [1] V. Bally, A. Matoussi: Stochastic PDE's and Doubly Stochastic Backward Differential Equations. Preprint, 1999, submitted.
- [2] G. Barles, E. Lesigne: SDE, BSDE and PDE. In "Backward Stochastic Differential Equations", N. El Karoui and L. Mazliak, editors, Pitman Research Notes in Math. 364, Longman, 1997.
- [3] R.M. Blumenthal and R.K. Gettoor: Markov Processes and Potential Theory, *Academic Press*, New York-London, 1966.
- [4] M. Fukushima, Y. Oshima and M. Takeda: Dirichlet Forms and Symmetric Markov Processes, *Walter de Gruyter*, Berlin- New York, 1994.
- [5] N. El Karoui: Backward stochastic differential equations a general introduction, in vol. Backward Stochastic Differential Equations, Pitman Res. Notes in Math. Ser. 364 *Longman*, 1997.
- [6] N. El Karoui, S. Peng, M.C. Quenez: Backward stochastic differential equations in finance. *Math. finance*, 1994.

- [7] O.A.Ladyzenskaya, N.N.Solonikov, V.A.Ural'tseva: Linear and Quasilinear Equations of Parabolic Type, *Nauka*, Moskow, 1967.
- [8] J.L.Lions: Equations Differentielles Operationnelles, Springer, Berlin-Gottingen-Heidelberg, 1961.
- [9] E.Pardoux: BSDEs, weak convergence and homogenization of semilinear PDEs, in vol. Nonlinear Analysis, Differential Equations and Control, 505-549, F.H. Clarke and R.J. Stern (eds.) Kluwer Academic Publishers, 1999.
- [10] E.Pardoux and S.Peng: Adapted solution of a backward stochastic differential equation, *Systems and Control Letters*, bf 14, 55-61, 1990.
- [11] E. Pardoux, S.Peng: Backward stochastic differential equations and quasilinear parabolic partial differential equations and their applications. (Lect. Notes Control Inf. Sci. ,vol 176.pp.200-217) Berlin Heidelberg New York: Springer 1992.
- [12] M.Sharpe: General Theory of Markov Processes, *Academic Press*, Boston- San Diego- New York- Berkeley- London- Sydney- Tokyo- Toronto,1988.



---

Unité de recherche INRIA Rocquencourt

Domaine de Voluceau - Rocquencourt - BP 105 - 78153 Le Chesnay Cedex (France)

Unité de recherche INRIA Lorraine : LORIA, Technopôle de Nancy-Brabois - Campus scientifique  
615, rue du Jardin Botanique - BP 101 - 54602 Villers-lès-Nancy Cedex (France)

Unité de recherche INRIA Rennes : IRISA, Campus universitaire de Beaulieu - 35042 Rennes Cedex (France)

Unité de recherche INRIA Rhône-Alpes : 655, avenue de l'Europe - 38330 Montbonnot-St-Martin (France)

Unité de recherche INRIA Sophia Antipolis : 2004, route des Lucioles - BP 93 - 06902 Sophia Antipolis Cedex (France)

---

Éditeur

INRIA - Domaine de Voluceau - Rocquencourt, BP 105 - 78153 Le Chesnay Cedex (France)

<http://www.inria.fr>

ISSN 0249-6399