Dept. of Math. University of Oslo
Pure Mathematics
ISBN 82–553–1299–4 No. 18
ISSN 0806–2439 September 2001

On Backward Stochastic Partial Differential Equations

Bernt Øksendal[1,2] Tusheng Zhang[3]

September 26, 2001

Abstract

We prove an existence and uniqueness result for a general class of backward stochastic partial differential equations. This is a type of equations which appear as adjoint equations in the maximum principle approach to optimal control of systems described by stochastic partial differential equations.

AMS Subject Classification: Primary 60H15 Secondary 93E20, 35R60.

1 Introduction

Let $B_t, t \geq 0$ be an m-dimensional Brownian motion on a filtered probability space $(\Omega, \mathcal{F}, \mathcal{F}_t, P)$. Fix $T > 0$ and let $\eta(\omega)$ be an $\mathcal{F}_T$-measurable random variable. Let

$$b : [0,T] \times \mathbf{R}^n \times \mathbf{R}^{n\times m} \to \mathbf{R}^n$$

be a given vector field. Consider the problem to find two $\mathcal{F}_t$-adapted processes $p(t) \in \mathbf{R}^n, q(t) \in \mathbf{R}^{n\times m}$ such that

$$dp(t) = b(t, p(t), q(t))dt + q(t)dB_t, t \in (0,T) \tag{1.1}$$
$$p(T) = \eta \quad a.s. \tag{1.2}$$

This is a backward stochastic (ordinary) differential equation (BSDE). It is called backward because it is the terminal value $p(T) = \eta$ that is given, not the initial value $p(0)$. Still $p(t)$ is required to be $\mathcal{F}_t$-adapted. In general this is only possible if we also are free to choose $q(t)$ (in an $\mathcal{F}_t$-adapted way).

[1] Department of Mathematics, University of Oslo, P.O.Box 1053 Blindern, N–0316 Oslo, Norway. Email: oksendal@math.uio.no
[2] Norwegian School of Economics and Business Administration, Helleveien 30, N–5045 Bergen, Norway
[3] Department of Mathematics, University of Manchester, Oxford Road, Manchester M13 9PL, England, U.K. Email: tzhang@maths.man.ac.uk

The theory of BSDEs is now well developed. See e.g. [EPQ], [MY], [PP] and [YZ] and the references therein.

There are many applications of this theory. Examples include the following:

(i) The problem of finding a replicating portfolio of a given contingent claim in a complete financial market can be transformed into a problem of solving a BSDE.

(ii) The maximum principle method for solving a stochastic control problem involves a BSDE for the adjoint processes $p(t), q(t)$.

For more information about these and other applications of BSDEs we refer to [EPQ] and [YZ] and refences therein.

The purpose of this paper is to study backward stochastic *partial* differential equations (BSPDEs). They are defined in a similar way as BSDEs, but with the basic equation being a stochastic partial differential equation rather than a stochastic ordinary differential equation. More precisely, we will study a class of BSPDEs which includes the following:

$$\begin{aligned} dY(t,x) &= AY(t,x)dt + b(t,x,Y(t,x),Z(t,x))dt \\ &\qquad + Z(t,x)dB_t, (t,x) \in (0,T) \times \mathbf{R}^n \qquad (1.3) \\ Y(T,x) &= \phi(\omega,x) \qquad (1.4) \end{aligned}$$

Here $dY(t,x)$ denotes the Itô differential with respect to t, while A is a partial differential operator with respect to x.

The function $b : [0,T] \times \mathbf{R}^n \times \mathbf{R} \times \mathbf{R} \to \mathbf{R}$ is given and so is the terminal value function $\phi(\omega,x)$. We assume that $\phi(\omega,x)$ is $\mathcal{F}_T$-measurable for all x and that

$$E[\int_{\mathbf{R}^n} \phi(\omega,x)^2 dx] < \infty, \qquad (1.5)$$

where E denotes expectation with respect to P. We are seeking the two processes $Y(t,x)$ and $Z(t,x)$ such that (1.3) and (1.4) hold. The processes $Y(t,x)$ and $Z(t,x)$ are assumed to be $\mathcal{F}_t$-adapted, i.e., $Y(t,x)$ and $Z(t,x)$ are $\mathcal{F}_t$-measurable for all $x \in \mathbf{R}^n$ and we also require that

$$E\Big[\int_{\mathbf{R}^n} \int_0^T \{Y(t,x)^2 + Z(t,x)^2\} dt\, dx\Big] < \infty. \qquad (1.6)$$

Equations of this type are of interest because they appear as adjoint equations in a maximum principle approach to optimal control of stochastic partial differential equations. See [B1] and [Ø] for details.

Example 1.1 Consider the following BSPDE:

$$\begin{aligned} dY(t,x) &= -\tfrac{1}{2}\Delta Y(t,x)dt + Z(t,x)dB_t, (t,x) \in (0,T) \times \mathbf{R}^n \qquad (1.7) \\ Y(T,x) &= \phi(\omega,x) \qquad (1.8) \end{aligned}$$

Here $\Delta Y(t,x) = \sum_{i=1}^n \frac{\partial^2 Y(t,x)}{\partial x_i^2}$ is the Laplacian with respect to x applied to Y, and $\phi(\omega, x)$ satisfies $E[\int_{\mathbf{R}^n} \phi(\omega, x)^2 dx] < \infty$.

In this simple case, we are able to find the solution explicitly. We first use the Itô representation theorem to write, for almost all x,

$$\phi(\omega, x) = h(x) + \int_0^T g(s, x, \omega) dB_s \tag{1.9}$$

where

$$h(x) = E[\phi(\cdot, x)], \tag{1.10}$$

$g(s, x, \cdot)$ is $\mathcal{F}_s$-measurable for all s, x and

$$E[\int_{\mathbf{R}^n} \int_0^T g(s, x, \cdot)^2 ds dx] < \infty. \tag{1.11}$$

Let

$$Q_t f(x) = \int_{\mathbf{R}^n} (2\pi t)^{-\frac{n}{2}} f(y) exp(-\frac{|x-y|^2}{2t}) dy, \qquad t > 0 \tag{1.12}$$

be the transition operator for Brownian motion defined for all measurable $f : \mathbf{R}^n \to \mathbf{R}$ such that the integral converges. Then it is well known that

$$\frac{\partial}{\partial t}(Q_t f(x)) = \tfrac{1}{2}\Delta(Q_t f(x)) \tag{1.13}$$

Now define

$$\begin{aligned} Y(t,x) &= Q_{T-t}(\int_0^t g(s, \cdot, \omega) dB_s + h(\cdot)) \\ &= \int_0^t (Q_{T-t} g(s, \cdot, \omega))(x) dB_s + (Q_{T-t} h)(x) \end{aligned} \tag{1.14}$$

Then

$$dY(t,x) = \left[\int_0^t -\tfrac{1}{2}\Delta(Q_{T-t} g(s, \cdot, \omega))(x) dB_s - \tfrac{1}{2}\Delta Q_{T-t} h(\cdot)(x)\right] dt \tag{1.15}$$

$$+ (Q_{T-t} g(t, \cdot, \omega))(x) dB_t \tag{1.16}$$

$$= -\tfrac{1}{2}\Delta Y(t,x) dt + Z(t,x) dB_t, \tag{1.17}$$

where

$$Z(t,x) = (Q_{T-t} g(t, \cdot, \omega))(x). \tag{1.18}$$

Hence the processes $Y(t,x), Z(t,x)$ given by (1.14) and (1.18) solve the BSPDE (1.7)–(1.8).

In the general case it is not possible to find explicit solutions of a BSPDE. However, in Section 3 we will prove an existence and uniqueness result for a general class of such equations. We will achieve this by regarding the BSPDE of type (1.3)–(1.4) as a special case of a backward stochastic evolution equation for Hilbert space valued processes. This, in turn, is studied by taking finite dimensional projections and then taking the limit. This is the well known Galerkin approximation method which has been used by several authors in other connections. See e.g. [B1], [B2] and [P]. We also refer readers to [PZ] for the general theory of stochastic evolution equations on Hilbert spaces.

The rest of the paper is organized as follows: In Section 2 we give the precise framework. The main result and its proof are given in Section 3.

2 Framework

Let V, H be two separable Hilbert spaces such that V is continuously, densely imbedded in H. Identifying H with its dual we have

$$V \subset H \cong H^* \subset V^*, \tag{2.1}$$

where V^* stands for the topological dual of V. Let A be a bounded linear operator from V to V^* satisfying the following coercivity hypothesis: There exist constants $\alpha > 0$ and $\lambda \geq 0$ such that

$$2\langle Au, u\rangle + \lambda |u|_H^2 \geq \alpha ||u||_V^2 \qquad \text{for all } u \in V, \tag{2.2}$$

where $\langle Au, u\rangle = Au(u)$ denotes the action of $Au \in V^*$ on $u \in V$.

Remark that A is generally not bounded as an operator from H into H. Let K be another separable Hilbert space. Let $\{B_t, t \geq 0\}$ be a cylindrical Brownian motion with covariance space K on a probability space $(\Omega, \mathcal{F}, P)$, i.e., for any $k \in K$,$\langle B_t, k\rangle$ is a real valued-Brownian motion with $E[\langle B_t, k\rangle^2] = t|k|_K^2$. Denote by $\mathcal{F}_t = \sigma(B_s, s \leq t)$ the natural, completed filteration generated by $\{B_t, t \geq 0\}$. Recall that a linear operator S from K into H is called Hilbert-Schmidt if $\sum_{i=1}^\infty |Sk_i|_H^2 < \infty$ for some orthonormal basis $\{k_i, i \geq 1\}$ of K. $L_2(K, H)$ will denote the Hilbert space of Hilbert-Schmidt operators from K into H equipped with the inner product $\langle S_1, S_2\rangle_{L_2(K,H)} = \sum_{i=1}^\infty \langle S_1 k_i, S_2 k_i\rangle_H$. Let $b(t, y, z, \omega)$ be a measurable mapping from $[0, T] \times H \times L_2(K, H) \times \Omega$ into H such that $b(t, y, z, \omega)$ is $\mathcal{F}_t$-adapted,i.e., $b(t, y, z, \cdot)$ is $\mathcal{F}_t$-measurable for all t, y, z. Suppose we are given an $\mathcal{F}_T$ -measurable, H-valued random variable $\phi(\omega)$. We are looking for two $\mathcal{F}_t$-adapted processes Y_t, Z_t with values in H and $L_2(K, H)$, respectively, such that the following backward stochastic evolution equation holds:

$$dY_t = AY_t dt + b(t, Y_t, Z_t)dt + Z_t dB_t, t \in (0, T) \tag{2.3}$$

$$Y_T = \phi(\omega) \text{ a.s.} \tag{2.4}$$

From now on we assume that the following, (2.5) and (2.6), hold:

There exists a constant $c < \infty$ such that

$$|b(t, y_1, z_1)(\omega) - b(t, y_2, z_2)(\omega)|_H \leq c(|y_1 - y_2|_H + |z_1 - z_2|_{L_2(K,H)}) \quad (2.5)$$

for all $(t, y, z) \in [0, T] \times H \times L_2(K, H)$.

$$E[\int_0^T |b(t, 0, 0)|_H^2 dt] < \infty \quad (2.6)$$

3 Results

We now state and prove the main result of this paper.

Theorem 3.1 *Assume that $E[|\phi|_H^2] < \infty$. Then there exists an unique $H \times L_2(K, H)$-valued progressively measurable process (Y_t, Z_t) such that*

(i) $E[\int_0^T |Y_t|_H^2 dt] < \infty, E[\int_0^T |Z_t|_{L_2(K,H)}^2 dt] < \infty$.

(ii) $\phi = Y_t + \int_t^T AY_s ds + \int_t^T b(s, Y_s, Z_s) ds + \int_t^T Z_s dB_s; \quad 0 \leq t \leq T$.

Proof. We will complete the proof by three steps.

Step 1. Assume that $b(t, y, z, \omega) = b(t, \omega)$ is independent of y and z, and $E[\int_0^T |b(t)|_H^2 dt] < \infty$.

Existence of solution.
Set $D(A) = \{v; v \in V, Av \in H\}$. Then $D(A)$ is a dense subspace of H. Thus we can choose and fix an orthonormal basis $\{e_1, ...e_n, ...\}$ of H such that $e_i \in D(A)$. Set $V_n = span(e_1, e_2, ..., e_n)$. Denote by P_n the projection operator from H into V_n. Put $A_n = P_n A$. Then A_n is a bounded linear operator from V_n to V_n. For the cylindrical Brownian motion B_t, it is well known that the following decomposition holds:

$$B_t = \sum_{i=1}^{\infty} \beta_t^i k_i \quad (3.1)$$

where $\{k_1, k_2, ..., k_i, ...\}$ is an orthonormal basis of K, and $\beta_t^i, i = 1, 2, 3, ...$ are independent standard Brownian motions. Set $B_t^n = (\beta_t^1, ..., \beta_t^n)$. Define $\mathcal{F}_t^n = \sigma(B_s^n, s \leq t)$ completed by the probability measure P, and put $\phi_n = E[P_n\phi|\mathcal{F}_T^n]$ and $b_n(t) = E[P_n b(t)|\mathcal{F}_t^n]$. Consider the following backward stochastic differential equation on the finite dimensional space V_n:

$$dY_t^n = A_n Y_t^n dt + b_n(t) dt + Z_t^n dB_t^n \; ; \qquad t < T \quad (3.2)$$

$$Y_T^n = \phi_n(\omega) \text{ a.s.} \quad (3.3)$$

As A_n is a bounded linear operator from V_n to V_n, it follows by the results of Pardoux and Peng [PP] that (3.2)–(3.3) admits a unique, continuous,

$\mathcal{F}_t^n$- adapted solution (Y_t^n, Z_t^n), where $Y_t^n \in \mathbf{R}^n \cong V_n$, $Z_t^n \in \mathbf{R}^n \otimes \mathbf{R}^n \cong L_2(K_n, V_n)$, $K_n = \text{span}(k_1, k_2, ..., k_n)$. Next we are going to show that the sequence (Y_t^n, Z_t^n) admits a convergent subsequence. Using Itô's formula, we find that

$$E[|Y_t^n|_H^2] = E[|\phi_n|_H^2] - 2E[\int_t^T \langle Y_s^n, P_n A Y_s^n \rangle ds] - 2E[\int_t^T < Y_s^n, b_n(s) > ds] - E[\int_t^T |Z_s^n|^2_{L_2(K_n,V_n)} ds], \tag{3.4}$$

where $|Z_s^n|^2_{L_2(K_n,V_n)} = \sum_{i,j=1}^n (Z_s^n(i,j))^2$ stands for the Hilbert-Schmidt norm. It follows from (2.2) that

$$E[|Y_t^n|_H^2] \leq E[|\phi|_H^2] - \alpha E[\int_t^T ||Y_s^n||_V^2 ds] + \lambda E[\int_t^T |Y_s^n|_H^2 ds] + E[\int_t^T |Y_s^n|_H^2 ds] + E[\int_t^T |b_n(s)|_H^2 ds] - E[\int_t^T |Z_s^n|^2_{L_2(K_n,V_n)} ds] \tag{3.5}$$

Hence,

$$E[|Y_t^n|_H^2] + \alpha E[\int_t^T ||Y_s^n||_V^2 ds] + E[\int_t^T |\bar{Z}_s^n|^2_{L_2(K,H)} ds] \leq E[|\phi|_H^2] + (\lambda+1) E[\int_t^T |Y_s^n|_H^2 ds] + E[\int_t^T |b(s)|_H^2 ds]$$

where $\bar{Z}_s^n = Z_s^n \bar{P}_n$, and $\bar{P}_n$ is the projection from K into $K_n = span(k_1, ..., k_n)$. Therefore,

$$E[|Y_t^n|_H^2] \leq E[|\phi|_H^2] + (\lambda+1) E[\int_t^T |Y_s^n|_H^2 ds] + E[\int_t^T |b(s)|_H^2 ds] \tag{3.6}$$

Set $\bar{Y}_t^n = \int_t^T |Y_s^n|_H^2$. Then (3.6) implies that

$$-\frac{d(e^{(\lambda+1)t} \bar{Y}_t^n)}{dt} \leq e^{(\lambda+1)t} (E[|\phi|_H^2] + E[\int_t^T |b(s)|_H^2 ds]) \tag{3.7}$$

Hence,

$$\int_0^T E[|Y_s^n|_H^2] ds \leq C(E[|\phi|_H^2] + E[\int_0^T |b(s)|_H^2 ds]), \tag{3.8}$$

where C is an appropriate constant. This together with (3.5) yields that

$$\sup_n \{ \int_0^T E[|Y_s^n|_H^2] ds \} < \infty \tag{3.9}$$

$$\sup_n \{ \int_0^T E[||Y_s^n||_V^2] ds \} < \infty \tag{3.10}$$

$$\sup_n \{ \int_0^T E[|\bar{Z}_s^n|^2_{L_2(K,H)}] ds \} < \infty \tag{3.11}$$

For a separable Hilbert space L, we denote by $M^2([0,T],L)$ the Hilbert space of progressively measurable, square integrable, L-valued processes equipped with the inner product $<a,b>_M = E[\int_0^T <a_t,b_t>_L dt]$. By the weak compactness of a Hilbert space, it follows from (3.10) and (3.11) that a subsequence $\{n_k, k\geq 1\}$ can be selected so that $Y_\cdot^{n_k}, k\geq 1$ converges weakly to some limit Y in $M^2([0,T],V)$, and $\bar{Z}_\cdot^{n_k}, k\geq 1$ converges weakly to some limit Z in $M^2([0,T],L_2(K,H))$. Let us prove that (a version of)(Y,Z) is a solution to the backward stochastic evolution equation (2.3) and (2.4). For $n\geq i\geq 1$, we have that

$$d\langle Y_t^n, e_i\rangle = \langle P_n A Y_t^n, e_i\rangle dt + \langle b_n(t), e_i\rangle dt + \langle \bar{Z}_t^n dB_t, e_i\rangle \tag{3.12}$$

$$= \langle A Y_t^n, e_i\rangle dt + \langle b_n(t), e_i\rangle dt + \langle \bar{Z}_t^n dB_t, e_i\rangle \tag{3.13}$$

Let $h(t)$ be an absolutely continuous function from $[0,T]$ to $\mathbf{R}$ with $h'(\cdot)\in L^2([0,T])$ and $h(0)=0$. By the Itô formula,

$$\langle Y_T^n, e_i\rangle h(T) \tag{3.14}$$

$$= \int_0^T h(t)\langle A Y_t^n, e_i\rangle dt + \int_0^T h(t)\langle b_n(t), e_i\rangle dt \tag{3.15}$$

$$+ \int_0^T h(t) d\langle \int_0^t \bar{Z}_s^n dB_s, e_i\rangle + \int_0^T \langle Y_t^n, e_i\rangle h'(t)dt. \tag{3.16}$$

Replacing n by n_k in (3.16) and letting $k\to\infty$ to obtain

$$\langle \phi, e_i\rangle h(T) \tag{3.17}$$

$$= \int_0^T h(t)\langle A Y_t, e_i\rangle dt + \int_0^T h(t)\langle b(t), e_i\rangle dt \tag{3.18}$$

$$+ \int_0^T h(t) d\langle \int_0^t Z_s dB_s, e_i\rangle + \int_0^T \langle Y_t, e_i\rangle h'(t)dt. \tag{3.19}$$

From (3.16) to (3.19), we have used the fact that the linear mapping G from $M^2([0,T],L_2(K,H))$ into $L^2(\Omega)$ defined by

$$G(Z) = \int_0^T h(t) d\langle \int_0^t Z_s dB_s, e_i\rangle = \sum_{j=1}^{\infty}\int_0^T h(t)\langle Z_t(k_j), e_i\rangle d\beta_t^j$$

is continuous. So, the convergence of (3.16) to (3.19) takes place weakly in $L^2(\Omega)$. Fix $t\in(0,T)$ and choose , for $n\geq 1$,

$$h_n(s) = \begin{cases} 1, & s\geq t+\frac{1}{2n}, \\ 1-\frac{1}{n}(t+\frac{1}{2n}-s), & t-\frac{1}{2n}\leq s\leq t+\frac{1}{2n}, \\ 0, & s\leq t-\frac{1}{2n} \end{cases}$$

With $h(\cdot)$ replaced by $h_n(\cdot)$ in (3.19), it follows that

$$\langle \phi, e_i\rangle = \int_0^T h_n(s)\langle A Y_s, e_i\rangle ds + \int_0^T h_n(s)\langle b(s), e_i\rangle ds \tag{3.20}$$

$$+ \int_0^T h_n(s) d\langle \int_0^s Z_u dB_u, e_i\rangle + \frac{1}{n}\int_{t-\frac{1}{2n}}^{t+\frac{1}{2n}} \langle Y_s, e_i\rangle ds. \tag{3.21}$$

Sending n to infinity in (3.21) we get that

$$\langle \phi, e_i\rangle = \int_t^T \langle AY_s, e_i\rangle ds + \int_t^T \langle b(s), e_i\rangle ds \tag{3.22}$$

$$+ \int_t^T d\langle \int_0^s Z_u dB_u, e_i\rangle + \langle Y_t, e_i\rangle. \tag{3.23}$$

for almost all $t \in [0, T]$ (with respect to Lebesgue measure).

As i is arbitrary, this implies that

$$\phi = \int_t^T AY_s ds + \int_t^T b(s)ds + \int_t^T Z_s dB_s + Y_t. \tag{3.24}$$

for almost all $t \in [0, T]$ (with respect to Lebesgue measure).

For $t \in [0, T]$, define

$$\hat{Y}_t = \phi - \int_t^T AY_s ds - \int_t^T b(s)ds - \int_t^T Z_s dB_s$$

Then we see that $(\hat{Y}_t, Z_t)$ also satisfies (ii) in the Theorem 3.1 with Y replaced by $\hat{Y}$ for all $t \in [0, T]$. Hence, $(\hat{Y}_t, Z_t)$ is a solution to the equations (2.3) and (2.4).

Uniqueness:
Let (Y_t, Z_t) and $(\bar{Y}_t, \bar{Z}_t)$ be two solutions of the equation (2.3). Then

$$\int_t^T A(Y_s - \bar{Y}_s)ds + \int_t^T (Z_s - \bar{Z}_s)dB_s + (Y_t - \bar{Y}_t) = 0 \tag{3.25}$$

Applying Itô's formula, we get

$$0 = |Y_t - \bar{Y}_t|_H^2 + 2\int_t^T \langle Y_s - \bar{Y}_s, dM_s\rangle \tag{3.26}$$

$$+ 2\int_t^T \langle A(Y_s - \bar{Y}_s), Y_s - \bar{Y}_s\langle ds + \int_t^T |Z_s - \bar{Z}_s|^2_{L_2(K,H)} ds \tag{3.27}$$

where $M_t = \int_0^t (Z_s - \bar{Z}_s)dB_s$. By (2.2),we get that

$$\begin{aligned} E[|Y_t - \bar{Y}_t|_H^2] &= -2\int_t^T E[\langle A(Y_s - \bar{Y}_s), Y_s - \bar{Y}_s\rangle]ds - E[\int_t^T |Z_s - \bar{Z}_s|^2_{L_2(K,H)} ds] \\ &\le -\alpha \int_t^T E[||Y_s - \bar{Y}_s||_V^2]ds + \lambda \int_t^T E[|Y_s - \bar{Y}_s|_H^2]ds \\ &\le \lambda \int_t^T E[|Y_s - \bar{Y}_s|_H^2]ds. \end{aligned}$$

By a Gronwall type inequality, it follows that $E[|Y_t - \bar{Y}_t|_H^2] = 0$, which proves the uniqueness.

Step 2. Assume that $b(t, y, z)(\omega) = b(t, z)(\omega)$ is independent of y.

Set $Z_t^0 = 0$. Denote by (Y_t^n, Z_t^n) the unique solution of the backward stochastic evolution equation:

$$dY_t^n = AY_t^n dt + b(t, Z_t^{n-1})dt + Z_t^n dB_t \tag{3.28}$$

$$Y_T^n = \phi(\omega). \tag{3.29}$$

The existence of such a solution (Y_t^n, Z_t^n) has been proved in step 1. Putting $M_t^n = \int_0^t Z_s^n dB_s$, and by Itô's formula we get that

$$0 = |Y_T^{n+1} - Y_T^n|_H^2 \tag{3.30}$$

$$= |Y_t^{n+1} - Y_t^n|_H^2 + 2\int_t^T \langle A(Y_s^{n+1} - Y_s^n), Y_s^{n+1} - Y_s^n \rangle ds \tag{3.31}$$

$$+ 2\int_t^T \langle b(t, Z_s^n) - b(t, Z_s^{n-1}), Y_s^{n+1} - Y_s^n \rangle ds \tag{3.32}$$

$$+ 2\int_t^T \langle Y_s^{n+1} - Y_s^n, d(M_s^{n+1} - M_s^n) \rangle + \int_t^T |Z_s^{n+1} - Z_s^n|_{L_2(K,H)}^2 ds \tag{3.33}$$

In virtue of (2.2), for $\varepsilon > 0$,

$$E[|Y_t^{n+1} - Y_t^n|_H^2] + E[\int_t^T |Z_s^{n+1} - Z_s^n|_{L_2(K,H)}^2 ds] \tag{3.34}$$

$$= -2E[\int_t^T \langle A(Y_s^{n+1} - Y_s^n), Y_s^{n+1} - Y_s^n \rangle ds] \tag{3.35}$$

$$- 2E[\int_t^T \langle b(t, Z_s^n) - b(t, Z_s^{n-1}), Y_s^{n+1} - Y_s^n \rangle ds] \tag{3.36}$$

$$\leq \lambda E[\int_t^T |Y_s^{n+1} - Y_s^n|_H^2 ds] - \alpha E[\int_t^T ||Y_s^{n+1} - Y_s^n||_V^2 ds] \tag{3.37}$$

$$+ \varepsilon E[\int_t^T |b(t, Z_s^n) - b(t, Z_s^{n-1})|_H^2 ds] + \frac{1}{\varepsilon} E[\int_t^T |Y_s^{n+1} - Y_s^n|_H^2 ds] \tag{3.38}$$

Choose $\varepsilon < \frac{1}{2c}$, where c is the Lipschitz constant in (2.5). It follows from (3.38) that

$$E[|Y_t^{n+1} - Y_t^n|_H^2] + E[\int_t^T |Z_s^{n+1} - Z_s^n|_{L_2(K,H)}^2 ds] + \alpha E[\int_t^T ||Y_s^{n+1} - Y_s^n||_V^2 ds]$$

$$\leq (\lambda + \frac{1}{\varepsilon}) E[\int_t^T |Y_s^{n+1} - Y_s^n|_H^2 ds] + \tfrac{1}{2} E[\int_t^T |Z_s^{n+1} - Z_s^n|_{L_2(K,H)}^2 ds] \tag{3.39}$$

Hence,

$$-\frac{d}{dt}(e^{(\lambda+\frac{1}{\varepsilon})t}E[\int_t^T |Y_s^{n+1}-Y_s^n|_H^2 ds] \tag{3.40}$$

$$+e^{(\lambda+\frac{1}{\varepsilon})t}E[\int_t^T |Z_s^{n+1}-Z_s^n|_{L_2(K,H)}^2 ds] \tag{3.41}$$

$$+\alpha e^{(\lambda+\frac{1}{\varepsilon})t}E[\int_t^T ||Y_s^{n+1}-Y_s^n||_V^2 ds] \tag{3.42}$$

$$\leq \tfrac{1}{2}e^{(\lambda+\frac{1}{\varepsilon})t}E[\int_t^T |Z_s^n-Z_s^{n-1}|_{L_2(K,H)}^2 ds] \tag{3.43}$$

From here, following a similar proof as in [PP] we will show that (Y^n, Z^n) converges to some limit (Y, Z) in the product space of $M^2([0,T],V)$ and $M^2([0,T],L_2(K,H))$.

Let $\beta = \lambda + \frac{1}{\varepsilon}$. Integrating both sides in (3.43) we get that

$$\begin{aligned} E[\int_0^T |Y_s^{n+1}-Y_s^n|_H^2 ds] + \int_0^T E[\int_t^T |Z_s^{n+1}-Z_s^n|_{L_2(K,H)}^2 ds]e^{\beta t}dt \\ +\alpha\int_0^T E[\int_t^T ||Y_s^{n+1}-Y_s^n||_V^2 ds]e^{\beta t}dt \\ \leq \tfrac{1}{2}\int_0^T E[\int_t^T |Z_s^n-Z_s^{n-1}|_{L_2(K,H)}^2 ds]e^{\beta t}dt \end{aligned} \tag{3.44}$$

In particular,

$$\begin{aligned} \int_0^T E[\int_t^T |Z_s^{n+1}-Z_s^n|_{L_2(K,H)}^2 ds]e^{\beta t}dt \\ \leq \tfrac{1}{2}\int_0^T E[\int_t^T |Z_s^n-Z_s^{n-1}|_{L_2(K,H)}^2 ds]e^{\beta t}dt \end{aligned} \tag{3.45}$$

This implies that

$$\int_0^T E[\int_t^T |Z_s^{n+1}-Z_s^n|_{L_2(K,H)}^2 ds]e^{\beta t}dt \leq (\tfrac{1}{2})^n C$$

for some constant C. Thus, it follows from (3.44) that

$$E[\int_0^T |Y_s^{n+1}-Y_s^n|_H^2 ds] \leq (\frac{1}{2})^n C \tag{3.46}$$

Hence, we conclude from (3.39) that

$$E[\int_0^T |Z_s^{n+1}-Z_s^n|_{L_2(K,H)}^2 ds] \leq (\tfrac{1}{2})^n C\beta + \tfrac{1}{2}E[\int_0^T |Z_s^n-Z_s^{n-1}|_{L_2(K,H)}^2 ds] \tag{3.47}$$

Using the above inequality repeatedly gives

$$E[\int_0^T |Z_s^{n+1} - Z_s^n|^2_{L_2(K,H)} ds] \le (\tfrac{1}{2})^n \left(nC\beta + E[\int_0^T |Z_s^1|^2_{L_2(K,H)} ds] \right) \quad (3.48)$$

Combining (3.39) and (3.47) we have that

$$E[\int_0^T ||Y_s^{n+1} - Y_s^n||_V^2 ds] \le (\frac{1}{2})^n \{(n+1)C\beta + E[\int_t^T |Z_s^1|^2_{L_2(K,H)} ds]\} \quad (3.49)$$

It follows now from (3.48) and (3.49) that the the sequence (Y_t^n, Z_t^n), $n \ge 1$ converges in $M^2([0,T], V) \times M^2([0,T], L_2(K,H))$ to some limit (Y_t, Z_t). Letting $n \to \infty$ in (3.28), we see that (Y_t, Z_t) satisfies

$$Y_t + \int_t^T AY_s ds + \int_t^T b(s, Z_s) ds + \int_t^T Z_s dB_s = \phi \quad (3.50)$$

i.e., (Y_t, Z_t) is a solution to equation (2.3).

Uniqueness

Let (Y_t, Z_t) , $(\bar{Y}_t, \bar{Z}_t)$ be two solutions to (3.50). By Itô's formula , as in (3.27) we have

$$E[|Y_t - \bar{Y}_t|_H^2] + E[\int_t^T |Z_s - \bar{Z}_s|^2_{L_2(K,H)} ds] \quad (3.51)$$

$$= -2E[\int_t^T < A(Y_s - \bar{Y}_s), Y_s - \bar{Y}_s > ds] \quad (3.52)$$

$$- 2E[\int_t^T \langle b(t, Z_t) - b(t, \bar{Z}_t), Y_s - \bar{Y}_s \rangle ds] \quad (3.53)$$

$$\le \lambda E[\int_t^T |Y_s - \bar{Y}_s|_H^2 ds] - \alpha E[\int_t^T ||Y_s - \bar{Y}_s||_V^2 ds] \quad (3.54)$$

$$+ \tfrac{1}{2} E[\int_t^T |Z_t - \bar{Z}_t|^2_{L_2(K,H)} ds] + cE[\int_t^T |Y_s - \bar{Y}_s|_H^2 ds] \quad (3.55)$$

Consequently,

$$E[|Y_t - \bar{Y}_t|_H^2] \le (\lambda + c) E[\int_t^T |Y_s - \bar{Y}_s|_H^2 ds] \quad (3.56)$$

By Gronwall's inequality ,

$$Y_t = \bar{Y}_t$$

which further implies $Z_t = \bar{Z}_t$ by (3.55).

Step 3. General case $b(t, x, y)(\omega)$.
Let $Y_t^0 = 0$. Define , for $n \geq 1$, (Y_t^{n+1}, Z_t^{n+1}) to be the solution of the equation:

$$\begin{aligned} dY_t^{n+1} &= AY_t^{n+1}dt + b(t, Y_t^n, Z_t^{n+1})dt + Z_t^{n+1}dB_t \qquad (3.57) \\ Y_T^{n+1} &= \phi \qquad (3.58) \end{aligned}$$

The existence of (Y_t^{n+1}, Z_t^{n+1}) is contained in step 2.

Similarly as in step 2 we can show that (Y_t^{n+1}, Z_t^{n+1}) converges to some limit (Y_t, Z_t), and moreover (Y_t, Z_t) is the unique solution to equation (2.3). We omit the details .

Example 3.2 Let $H = L^2(\mathbf{R}^d)$, and set

$$V = H_2^1(\mathbf{R}^d) = \{u \in L^2(R^d); \nabla u \in L^2(\mathbf{R}^d \to \mathbf{R}^d)\}$$

Denote by $a(x) = (a_{ij}(x))$ a matrix-valued function on $\mathbf{R}^d$ satisfing the uniform ellipticity condition:

$$\frac{1}{c}I_d \leq a(x) \leq cI_d \qquad \text{for some constant } c \in (0, \infty).$$

Let $f(x)$ be a vector field on $\mathbf{R}^d$ with $f \in L^p(\mathbf{R}^d)$ for some $p > d$. Define

$$Au = -div(a(x)\nabla u(x)) + f(x) \cdot \nabla u(x)$$

Then (2.2) is fulfilled for (H, V, A). Thus, for any choice of cylindrical Brownian motion B, any drift coefficient $b(t, y, z, \omega)$ satisfying (2.5) and (2.6) and terminal random variable ϕ, the main result in Section 3 applies.

Acknowledgement

This work was initiated when both of us attended a workshop on SPDEs in Warwick in May, 2001. T. Zhang also gave a talk on this work in the SPDE conference in Warwick in July 2001. We would like to thank the organizers Jo Kennedy, Roger Tribe and Jerzy Zabczyk, as well as David Elworthy, for the invitation and hospitality. We also thank the Mathematics Institute at University of Warwick for providing the stimulating research environment.

References

[B1] A.Bensoussan: Maximum principle and dynamic programming approaches of the optimal control of partially observed diffusions. Stochastics **9** (1983),169-222.

[B2] A.Bensoussan: Stochastic maximum principle for systems with partial information and application to the separation principle. In M.Davis and R.Elliott (editors): Applied Stochastic Analysis. Gordon and Breach 1991, pp 157-172.

[EPQ] N. El Karoui, S. Peng and M.C. Queuez: Backward stochastic differential equations in finance. Mathematical Finance **7** (1997) 1–71.

[MY] J. Ma and J. Yong: Forward-Backward Stochastic Differential Equations and Their Applications. Springer LNM 1702, Springer-Verlag 1999.

[Ø] B. Øksendal: Optimal control of stochastic partial differential equations and application to partial observation control. Preprint, University of Oslo 17/2001.

[P] E. Pardoux: Stochastic partial differential equations and filtering of diffussion processes. Stochastics **3** (1979) 127–167.

[PP] E. Pardoux and S. Peng: Adapted solutions of backward stochastic differential equations. Systems and Control Letters **14** (1990) 55–61.

[PZ] G.D. Prato and J. Zabczyk: Stochastic equations in infinite dimensions. Cambridge University Press, 1992.

[YZ] J.Yong and X.Y.Zhou: Stochastic Controls. Springer-Verlag 1999.