Chin. Ann. of Math.
12B: 3(1991), 255—266.

MAXIMUM PRINCIPLE FOR SEMILINEAR STOCHASTIC EVOLUTION SYSTEMS**

PENG SHIGE(彭实戈)* HU YING(胡 瑛)*

Abstract

A maximum principle is proved for semilinear stochastic evolution systems. The main contribution of this work is that in our problem, the infinitesimal generator of the semigroup of the systems need not to be elliptic. This generalizes a result of A. Bensoussan in 1983.

§ 1. Introduction

The stochastic maximum principle for finite dimensional dynamic systems was studied by Kushner[1], Bensoussan[2], Bismut[3], Haussmann[4], and Hu[5].

For distributed parameter systems, Li Xunjing and Yao Yunlong[6] obtained the maximum principle for deterministic systems. But until 1983, little has been done for stochastic distributed systems except for linear quadratic problems[7,8]. In 1983, an important result was obtained by Bensoussan[9] for stochastic distributed parameter systems. But in that work the infinitesimal generator of the semigroup was assumed to be strictly elliptic.

The objective of this paper is to obtain the same conclusion as in [9] without the elliptic assumption. In fact, we treat the case where the infinitesimal generator is of a general C_0–semigroup. In this case, the main difficulty is due to the fact that the Ito's formula is no longer valid. To overcome this difficulty, we introduce a method based mainly on the stochastic Fubini theorem.

The paper is organized as follows: In the next section, we present some necessary preliminaries in this paper. In section III, we discuss the variational inequality of our optimal control systems. In section IV, the ajoint processes and the adjoint equation are derived and treated. The maximum principle is given in the last section.

Manuscript received October 8, 1988.

* Institute of Mathematics, Fudan University, Shanghai, China.

** Partially supported by Chinese National Science Foundation.

§ 2. Preliminaries

2.1 Stochastic integral in Hilbert spaces[10,11]

Let $(\Omega, \mathscr{F}, P)$ be a probability space, equipped with a filtration $\mathscr{F}_t (\mathscr{F} = \mathscr{F}_\infty)$. A Wiener process with values in a separable Hilbert space E is a stochastic process $t \to W(t)$, which is adapted to $\mathscr{F}_t$ such that, for any $e \in E$, $(W(t), e)$ is a real Wiener process, with the correlation function

$$E[(W(t_1), \tilde{e}_1)(W(t_2), \tilde{e}_2)] = (Q\tilde{e}_1, \tilde{e}_2)\min\{t_1, t_2\}, \ \tilde{e}_1, \tilde{e}_2 \in E,$$

where Q is a positive self-adjoint nuclear operator on E. That is to say, there is an orthonormal basis in E denoted by $\{e_n\}$, such that

$$Qe_n = \lambda_n e_n, \ n = 1, 2, \cdots$$

with

$$\lambda_n > 0, \ n = 1, 2, \cdots, \sum_{i=1}^{\infty} \lambda_n < \infty.$$

We assume

$$\mathscr{F}_t = \sigma(W(s); s \leqslant t).$$

For a given separable Hilbert space H, we introduce a space of stochastic processes $L^2_{\mathscr{F}}(0, T; H)$, which is the space of all $\mathscr{F}_t$-adapted processes $b(t)$ with values in E, such that

$$E\int_0 |b(t)|_H dt < \infty.$$

Obviously, $L^2_{\mathscr{F}}(0, T; \mathscr{L}(E; H))$ is a Hilbert space. One can define the Ito's stochastic integral on $L^2_{\mathscr{F}}(0, T; \mathscr{L}(E; H))$

$$\int_0^T b(s)dW(s): L^2_{\mathscr{F}}(0, T; \mathscr{L}(E; H)) \to L^2(\Omega, \mathscr{F}_T, P; H).$$

For details, see [11].

2.2 Nonlinear stochastic equations[12]

Consider $\mathscr{F}_t$-adapted processes $a(\cdot, x)$, $b(\cdot, x)$ parametrized by $x \in H$, such that

$$a(x, \cdot) \in L^2_{\mathscr{F}}(0, T; H), \ b(x, \cdot) \in L^2_{\mathscr{F}}(0, T; \mathscr{L}(E; H)), \ \forall x \in H$$

and such that the following uniform Lipschitz condition hold:

$$|a(x_1, t) - a(x_2, t)|_H \leqslant C|x_1 - x_2|,$$
$$|b(x_1, t) - b(x_2, t)|_{\mathscr{L}(E;H)} \leqslant C|x_1 - x_2|.$$

Let A be a given infinitesimal generator of a strongly continuous semigroup e^{At} on H. We have an existence and uniqueness result for infinite dimensional stochastic differential equations.

Theorem 2.1. *For a fixed $x_0 \in H$, there exists a unique process $x(\cdot) \in C_{\mathscr{F}}([0, T]; L^2_{\mathscr{F}}(\Omega, \mathscr{F}, P; H))$ satisfying*

$$x(t)=e^{At}x_0+\int_0^t e^{A(t-s)}a(x(s),\ s)\,ds+\int_0^t e^{A(t-s)}b(x(s),\ s)dW(s).$$

Proof For a given $b>0$, $X=C_{\mathscr{F}}(0,\ T;\ L^2(\Omega,\ \mathscr{F}_T,\ P;\ H))$ is a Banach space under the following norm

$$|x(\cdot)|_x=\max\{e^{-bt}(E|x(t)|_H^2)^{1/2}|t\in[0,\ T]\}.$$

Then, we can introduce a mapping Λ: $X\to X$ as follows

$$(\Lambda x)(t)=e^{At}x_0+\int_0^t e^{A(t-s)}a(x(s),\ s)ds+\int_0^t e^{A(t-s)}b(x(s),\ s)dW(s).$$

One can check easily that there exists a constant $\bar{c}>0$, such that

$$|\Lambda x_1-\Lambda x_2|_x\leqslant\frac{\bar{c}}{b}|x_1-x_2|_x.$$

Thus, for $b>0$ sufficiently large, Λ: $X\to X$ is a contraction mapping. From the fixed point theorem, we derive the theorem.

§ 3. Optimal Control Problem and Variational Inequality

3.1. Optimal control problem

Let U be a separable Hilbert space, called the space of controls. Let U_{ad} be a convex, non-empty subset of U. Let

$$g(x,\ v):\ H\times U\to H,$$
$$\sigma(x,\ v):\ H\times U\to\mathscr{L}(E;\ H).$$

We assume that g and σ are Gauteaux-differentiable and that their derivative g_x, g_v, σ_x, σ_v are continuous and bounded.

An admissible control is an adapted process

$$v(\cdot)\in L^2_{\mathscr{F}}(0,\ T;\ U)\ \text{such that}\ v(t,\ w)\in U,\ \forall(t,\ w)\in[0,\ T]\times\Omega.$$

We denote the set of all admissible controls by $\mathscr{U}_{ad}$. For any admissible control consider the following state equation

$$z(t)=e^{At}y_0+\int_0^t e^{A(t-s)}g(z(s),\ v(s))ds+\int_0^t e^{A(t-s)}\sigma(z(s),\ v(s))dW(s),$$

where $y_0\in H$ is a given initial data. According to Theorem 2.1, the solution of the above state equation is well-defined. Thus we can define the following cost function

$$J(v(\cdot))=E\int_0^T l(z(t),\ v(t))dt+Eh(z(T)),$$

where

$$l(x,\ v):\ H\times U\to\mathbf{R},$$
$$h(x):\ H\to\mathbf{R}.$$

We assume that l and h are Gauteaux-differentiable, and that their derivatives l_x, l_v, h_x are continuous and bounded by

$$|h_x(x)|\leqslant C(1+|x|),$$
$$|l_x(x, v)|+|l_v(x, v)|\leqslant C(1+|x|+|v|).$$

The optimal control problem is to minimize the cost function $J(v(\cdot))$ over the set of admissible controls. The objective of this paper is to obtain a necessary condition of the optimality.

3. 2. The Variational inequality

Let $u(\cdot)$ be an optimal control and let $y(\cdot)$ be the corresponding trajectory

$$y(t)=e^{At}y_0+\int_0^t e^{A(t-s)}g(y(s), u(s))ds+\int_0^t e^{A(t-s)}\sigma(y(s), u(s))dW(s),$$
$$J(u(\cdot))=\min\{J(v(\cdot)); v(\cdot)\in\mathscr{U}_{ad}\}.$$

Then we can obtain the following variational inequality.

Lemma 3.1. *The cost function $J(v(\cdot))$ is Gauteaux-differentiable and the following variational inequality holds*

$$\frac{d}{d\theta}J(u(\cdot)+\theta v(\cdot))\Big|_{\theta=0}=Eh_x(T)z(T)+E\int_0^T[l_x(s))z(s)+l_v(s)v(s)]ds\geqslant 0 \tag{3.1}$$

with

$$l_x(s)=l_x(y(s), u(s)),\ l_v(s)=l_v(y(s), u(s)),$$
$$h_x(T)=h_x(y(T)),$$

where $v(\cdot)$ satisfies $u(\cdot)+v(\cdot)\in\mathscr{U}_{ad}$, and $z(\cdot)$ is the solution of the following linear equation

$$z(t)=\int_0^t e^{A(t-s)}[g_x(s)z(s)+g_v(s)v(s)]\,ds$$
$$+\int_0^t e^{A(t-s)}[\sigma_x(s)z(s)+\sigma_v(s)v(s)]dW(s) \tag{3.2}$$

with

$$g_x(s)=g_x(y(s), u(s)),\ g_v(s)=g_v(y(s), u(s)),$$
$$\sigma_x(s)=\sigma_x(y(s), u(s)),\ \sigma_v(s)=\sigma_v(y(s)\ u(s)).$$

The proof is similar to [2].

§ 4. Linear Stochastic Differencial Equations and Its Adjoint Processes

The maximum principle relates tightly with the adjoint process and adjoint equation of a linear stochastic differential equation. In this section, we will discuss this problem

4.1 Abstract definition of the adjoint processes

We consider the Hilbert spaces $L^2_{\mathscr{F}}(0, T; H)$ and $L^2_{\mathscr{F}}(0, T; H)^\infty$, the latter being the space of sequences of process in $L^2_{\mathscr{F}}(0, T; H)$

$$\Psi(\cdot)=(\psi_1(\cdot),\ \psi_2(\cdot),\ \cdots),\ \psi_i(\cdot)\in L^2_{\mathscr{F}}(0,\ T;\ H),\ i=1,\ 2,\ \cdots,$$

such that

$$\sum_{i=1}\lambda_i E\int_0^T|\psi_i(s)|^2 ds<\infty,$$

where λ_i $i=1, 2, \cdots$ are the eigenvalues of the nuclear operator Q (see 2.1). This space is endowed the following scalar product

$$\langle\Psi,\ \Psi\rangle=\sum_{i=1}\lambda_i E\int_0^T(\psi_i(s),\ \psi_i(s))_H ds$$

for $\Psi(\cdot)=(\psi_1(\cdot), \psi_2(\cdot), \cdots)$, $\Psi(\cdot)=(\psi_1(\cdot), \psi_2(\cdot), \cdots)\in L^2_{\mathscr{F}}(0, T; H)^\infty$.

Sometimes we denote also

$$\langle\Psi,\ \Psi\rangle=E\int_0^T(\Psi(s),\ \Psi(s))ds.$$

We can give a meaning to the stochastic integral

$$\int_0^T\Psi(s)dW(s)=\sum_{i=1}^{\infty}\int_0^T\Psi_i(s)d(W(s),\ e_i),\ \text{for}\ \Psi(\cdot)\in L^2_{\mathscr{F}}(0,\ T;\ H)^\infty.$$

Indeed

$$\int_0^T\psi_i(s)d(W(s),\ e_i)$$

is an ordinary stochastic integral with values in H and the sum is convergent in $L^2(\Omega, \mathscr{F}, P; H)$ since

$$E\left|\sum_{i=1}\int_0^T\psi_i(s)d(W(s),\ e_i)\right|^2=\sum_{i=1}^{\infty}\lambda_i E\int_0^T|\psi_i(s)|^2 ds.$$

We can then solve the following linear equation

$$x(t)=\int_0^t e^{A(t-s)}[g_x(s)x(s)+\phi(s)]ds+\int_0^t e^{A(t-s)}[\sigma_x(s)x(s)+\Psi(s)]dW(s),\quad (4.1)$$

where $(\phi(\cdot), \Psi(\cdot))\in L^2_{\mathscr{F}}(0, T; H)\times L^2_{\mathscr{F}}(0, T; H)^\infty$ is considered as a 'control'.

It is easy to check that the mapping $(\phi(\cdot), \Psi(\cdot))\to x(\cdot)$ is linear and continuous. According to Riesz representation theorem, we can define in a unique way $(p(\cdot), K(\cdot))\in L^2_{\mathscr{F}}(0, T; H)\times L^2_{\mathscr{F}}(0, T; H)^\infty$ such that

$$\begin{aligned}&E\int_0^T(p(s),\ \phi(s))ds+E\int_0^T(K(s),\ \Psi(s))ds\\&=E(h_x(T),\ x(T))+E\int_0^T(l_x(s)),\ x(s))ds,\\&\forall(\phi(\cdot),\ \Psi(\cdot))\in L^2_{\mathscr{F}}(0,\ T;\ H)\times L^2_{\mathscr{F}}(0,\ T;\ H)^\infty.\end{aligned}\quad (4.2)$$

The process $(p(\cdot), K(\cdot))$ is called adjoint process.

4.2. Approximation

We need to find an equation that solve uniquely the adjoint process $(p(\cdot), K(\ ))$. For this purpose, we use an approximation procedure. Let $\{h_1, h_2, \cdots\}\subset D(A)$ be an orthonormal basis in H. We define

$$V_m=\operatorname{Span}\{h_1,\ h_2,\ \cdots,\ h_m\},$$

$$P_m h=\sum_{i=1}^{m}(h,\ h_i)h_i,\ \forall h\in H.$$

With these notions, we approximate A, g_x, σ_x, $\phi(\cdot)$, $\Psi(\cdot)$ by

$$A_m=P_mAP_m,\ g_m(t)=P_mg_x(t)P_m,\ l_m(t)=P_ml_x(t),\ h_m(T)=P_mh_x(T),$$

$$\sigma_m(t)h_ih=\sum_{i=1}^{m}(\sigma_x(t)h_ih,\ h_j)h_j,\ \forall h\in H,\ (\sigma_m(t)\in\mathscr{L}(V_m;\ \mathscr{L}(E;\ V_m))).$$

$$\phi^m(t)=P_m\phi(t),\ \Psi^m(t)=P_m\Psi(t)=(P_m\psi_1(t),\ P_m\psi_2(t),\ \cdots).$$

Thus we can approximate (4.1) by

$$x_m(t)=\int_0^t e^{A_m(t-s)}[g_m(s)x_m(s)+\phi^m(s)]ds$$
$$+\int_0^t e^{A_m(t-s)}[\sigma_m(s)x_m(s)+\Psi^m(s)]dW(s).$$

Indeed, as $m\to\infty$, we have the following convergence result.

Lemma 4.2. *We have, as $m\to\infty$,*

$$x_m(\cdot)\to x(\cdot)\ in\ L^2_{\mathscr{F}}(0,\ T;\ H)\cap C_{\mathscr{F}}(0,\ T;\ L^2(\Omega,\ \mathscr{F},\ P;\ H)),$$

where $x(\cdot)$ is the solution of the linear equation (4.1).

The proof is similar to that of Bensoussan[9].

Now we take arbitrary $\phi^m(\cdot)\in L^2_{\mathscr{F}}(0,\ T;\ V_m)$, $\psi^m(\cdot)\in L^2_{\mathscr{F}}(0,\ T;\ V_m)$, such that

$$E\int_0^T\sum_{i=1}^{\infty}\lambda_i|\psi^m(s)|^2ds<\infty,$$

and define in a unique way

$$p^m(\cdot)\in L^2_{\mathscr{F}}(0,\ T;\ V_m),\ K^m(\cdot)\in L^2_{\mathscr{F}}(0,\ T;\ V_m),$$

such that

$$E\int_0^T\sum_{i=1}^{\infty}\lambda_i|K^m(s)|^2ds<\infty,$$

as follows

$$E\int_0^T(p^m(s),\ \phi^m(s))ds+\sum_{i=1}^{\infty}\lambda_iE\int_0^T(K^m(s),\ \Psi^m(s))ds$$
$$=E(h_x(T),\ x_m(T))+E\int_0^T(l_x(s)),\ x_m(s))ds, \tag{4.3}$$
$$\forall(\phi(\cdot),\ \Psi(\cdot))\in L^2_{\mathscr{F}}(0,\ T;\ H)\times L^2_{\mathscr{F}}(0,\ T;\ H)^\infty.$$

We have the following convergence result.

Lemma 4.2. *As $m\to\infty$, we have*

$$p^m(\cdot)\to p(\cdot),\ weakly\ in\ L^2_{\mathscr{F}}(0,\ T;\ H),$$
$$K^m(\cdot)\to K(\cdot),\ weakly\ in\ L^2_{\mathscr{F}}(0,\ T;\ H)^\infty.$$

Proof Since

$$\phi^m(t)=P_m\phi(t),\ \Psi^m(t)=P_m\Psi(t), \tag{4.4}$$

from Ito's formula we derive

$$E|x^m(t)|^2\leqslant C_1\int_0^t E|x^m(s)|^2ds+C_2\Big(\int_0^T E|\phi(s)|^2ds+\int_0^T E|\Psi(s)|^2ds\Big).$$

It follows from Gronwall inequality that

$$E|x^m(t)|^2 \leqslant C_3\Big(\int_0^T E|\phi(s)|^2 ds + \int_0^T E|\Psi(s)|^2 ds\Big).$$

Therefore necessarily

$$|p^m(\cdot)|_{L^2_{\mathscr{F}}(0,T;H)} \leqslant \text{Const.},$$
$$|K^m(\cdot)|_{L^2_{\mathscr{F}}(0,T;H)^\infty} \leqslant \text{Const.}$$

Thus there exists a subsequence, still denoted by $(p^m(\cdot), K^m(\cdot))$, such that

$$(p^m(\cdot), K^m(\cdot)) \to (\bar{p}(\cdot), \bar{K}(\cdot)) \text{ weakly in } L^2_{\mathscr{F}}(0, T; H) \times L^2_{\mathscr{F}}(0,T; H)^\infty.$$

Passing limit in (4.3), we have

$$E\int_0^T (\bar{p}(s), \phi(s))ds + \sum_{i=1}^\infty \lambda_i E\int_0^T (\bar{K}(s), \Psi(s))ds$$
$$= E(h_x(T), x(T)) + E\int_0^T (l_x(s)), x(s))ds,$$
$$\forall (\phi(\cdot), \Psi(\cdot)) \in L^2_{\mathscr{F}}(0, T; H) \times L^2_{\mathscr{F}}(0, T; H)^\infty.$$

Thus, from Lemma 4.1, we have

$$(\bar{p}(\cdot), \bar{K}(\cdot)) = (p(\cdot), K(\cdot)).$$

The proof is complete.

The following lemma extends martingale representation theorem to infinite dimensional space.

Lemma 4.3. *Let $X(\cdot) \in L^2_{\mathscr{F}}(0, T; V_m)$ be an $\mathscr{F}_t$-martingale with $EX(t) = 0$, $\forall t$. Then there exists $G_i(\cdot) \in L^2_{\mathscr{F}}(0, T; V_m)$, $i = 1, 2, \cdots$, with*

$$\sum_{i=1}^\infty \lambda_i E\int_0^T |G_i(s)|^2 ds < \infty,$$

such that

$$X(t) = \sum_{i=1}^\infty \int_0^t G_i(s) d(W(s), e_i).$$

Proof We denote

$$\mathscr{F}_t^k = \bigvee_{i=1}^k \sigma((W(s), e_i); s \leqslant t).$$

Since

$$\mathscr{F}_t = \bigvee_{i=1}^\infty \sigma((W(s), e_i); s \leqslant t),$$

we have

$$\mathscr{F}_t^k \uparrow \mathscr{F}_t.$$

Let $X^k(t) = E(X(t) | \mathscr{F}_t^k)$. Then $X^k(t)$ is $\mathscr{F}_t$-martingale and is square integrable. According to martingale representation theorem[13], we can write

$$X^k(t) = \sum_{i=1}^k \int_0^t G_i(s) d(W(s), e_i)$$

with $G_i \in L^2_{\mathscr{F}}(0, T; V_m)$. Obviourslу

$$\sum_{i=1}^\infty \lambda_i E\int_0^T |G_i(s)|^2 ds < \infty.$$

According to [14]

$$E|X^k(T)-X(T)|^2\to 0.$$

Thus, as $k, l\to\infty$

$$E|X^k(T)-X^l(T)|^2\to 0.$$

It follows that

$$\sum_{i=1}^{\infty}\lambda_i E\int_0^T|G_i(s)-G_i(s)|^2ds\to 0.$$

So there exist $G_i(\cdot)\in L^2_{\mathscr{F}}(0, T; V_m)$, $i=1, 2, \cdots$, such that

$$\sum_{i=1}^{\infty}\lambda_i E\int_0^T|G_i(s)-G_i(s)|^2ds\to 0, \text{ as } k\to\infty.$$

This implies

$$X(t)=\sum_{i=1}^{\infty}\int_0^t G_i(s)d(W(s), e_i).$$

The proof is complete.

Corollary 4.4. *Let $X(\cdot)\in L^2_{\mathscr{F}}(0, T; H)$ be an $\mathscr{F}_t$-martingale with $EX(t)=0$, $\forall t$. Then there exists $K(\cdot)\in L^2_{\mathscr{F}}(0, T; H)^\infty$, such that*

$$X(t)=\int_0^t K(s)dW(s).$$

We can first derive the adjoint equation for $(p^m(\cdot), K^m(\cdot))$, introduced in (4.3).

Lemma 4.5. *$(p^m(\cdot), K^m(\cdot))$ is the unique solution of the following adjoint equation*

$$-dp^m(t)=A_m^*p^m(t)dt+(g_m^*p^m(t)+\sigma_m^*K^m(t)+l_m(t))dt-K^m(t)dW(t),$$
$$p^m(T)=h_m(T). \tag{4.5}$$

The proof can be found in Bensoussan[2], [9].

4. 4. Convergence and adjoint equation

Now we will pass limit in (4.5) to derive the adjoint equation in H. We can assert

Theorem 4.6. *The adjoint process $(p(\cdot), K(\cdot))$ derived from (4.2) solves*

$$p(t)=e^{A^*(T-t)}h_x(T)+\int_t^T e^{A^*(s-t)}[g_x^*p(s)$$
$$+\sigma_x^*K(s)+l_x(s)]\,ds-\int_t^T e^{A^*(s-t)}K(s)dW(s). \tag{4.6}$$

Proof. From the extended representation theorem, for any given $f(\cdot)\in L^2(\Omega\times[0. T], \mathscr{F}_T\times\mathscr{B}([0, T]), dP\times dt; \mathbf{R})$, we can write

$$\bar{f}(t)=E(\bar{f}(t)|\mathscr{F}_t)+\int_t^T\sum_{i=1}^{\infty}\bar{G}_i(s, t)d(W(s), e_i)$$

with

$$\bar{G}_i(\cdot, t)\in L^2_{\mathscr{F}}(0, T; \mathscr{L}(E; R)), i=1, 2, \cdots,$$

$$\sum_{i=1}^{\infty}\lambda_i E\int_0^T|\bar{G}_i(s,\ t)|^2ds<\infty.$$

For any given $h\in H$, we denote

$$f(t)=\bar{f}(t)h,\ f_1(t)=E(f(t)\,|\,\mathscr{F}_t),$$
$$G(s,\ t)=\bar{G}(s,\ t)h.$$

We denote also

$$\hat{l}_m(s)=g_m^*p^m(s)+\sigma_m^*K^m(s)+l_m(s),$$
$$\hat{l}(s)=g_x^*p(s)+\sigma_x^*K(s)+l_x(s).$$

From Fubini theorem, we have

$$\begin{aligned}I_m&\triangleq E\int_0^T\Big\langle\int_t^T e^{A^*(s-t)}\hat{l}_m(s)ds,\ f(t)\Big\rangle dt\\&=E\int_0^T\Big\langle\hat{l}_m(s),\int_0^s e^{A(s-t)}f(t)\,dt\Big\rangle ds\\&=E\int_0^T\Big\langle\hat{l}_m(s),\int_0^s e^{A(s-t)}\Big[f_1(t)\\&\quad+\int_t^s G(r,\ t)dW(r)+\int_s^T G(r,\ t)dW(r)\Big]dt\Big\rangle ds\\&=E\int_0^T\Big\langle\hat{l}_m(s),\ \int_0^s e^{A(s-t)}\Big[f_1(t)+\int_t^s G(r,\ t)dW(r)\Big]dt\Big\rangle ds.\end{aligned}$$

Since $\hat{l}_m(\cdot)\to\hat{l}(\cdot)$ weakly in $L^2_{\mathscr{F}}(0,\ T;\ H)$, we have

$$\begin{aligned}I_m&\to E\int_0^T\Big\langle\hat{l}(s)ds,\int_s^T e^{A(s-t)}\Big[f_1(t)+\int_t^s G(r,\ t)\ dW(r)\Big]dt\Big\rangle ds\\&=E\int_0^T\Big\langle\int_r^T e^{A^*(s-t)}\hat{l}(s)ds,\ f(t)\Big\rangle dt.\end{aligned}$$

We have also

$$\begin{aligned}E\int_0^T\langle p^m(t),\ f(t)\rangle dt&=E\int_0^T\langle p^m(t),\ f_1(t)\rangle dt\\&\to E\int_0^T\langle p(t),\ f_1(t)\rangle dt\\&=E\int_0^T\langle p(t),\ f(t)\rangle dt,\end{aligned}$$

$$E\int_0^T\langle e^{A^*(T-t)}h(T),\ f(t)\rangle dt\to E\int_0^T\langle e^{A^*(T-t)}h(T),\ f(t)\rangle dt.$$

It follows that

$$\begin{aligned}Q_m&=E\int_0^T\langle p^m(t)-e^{A^*(T-t)}h_m(T)-\int_t^T e^{A^*(s-t)}\hat{l}_m(s)ds,\ f(t)\rangle dt\\&\to E\int_0^T\Big\langle p(t)-e^{A^*(T-t)}h_x(T)-\int_t^T e^{A^*(s-t)}\hat{l}(s)ds,\ f(t)\Big\rangle dt\end{aligned}\tag{4.7}$$

On the other hand, since $(p^m,\ K^m)$ solves (4.5), we have

$$\begin{aligned}Q_m&=-E\int_0^T\Big\langle\int_t^T e^{A^*(s-t)}K^m(s)dW(s),\ f_1(t)+\int_t^T G(s,\ t)dW(s)\Big\rangle dt\\&=-E\int_0^T\int_t^T\langle e^{A^*(s-t)}K^m(s),\ G(s,\ t)\rangle dsdt\end{aligned}$$

$$\to -E\int_0^T\int_t^T \langle e^{A^*(s-t)}K(s),\ G(s,t)\rangle ds dt$$

$$= -E\int_0^T\Big\langle \int_t^T e^{A^*(s-t)}K(s)dW(s),\ f(t)\Big\rangle dt. \tag{4.8}$$

This with (4.7) implies

$$E\int_0^T\Big\langle p(t) - e^{A^*(T-t)}h(T) - \int_t^T e^{A^*(s-t)}l(s)ds + \int_t^T e^{A^*(s-t)}K(s)dW(s),\ f(t)\Big\rangle dt = 0,$$

$$\forall f(\cdot)\in L^2(\Omega\times[0.T],\ \mathscr{F}_T\times\mathscr{B}([0,T]),\ dP\times dt;\ \mathbf{R}).$$

Thus $(p(\cdot), K(\cdot))$ solves (4.6). The proof is complete.

4.5. Uniqueness

The following theorem gives a uniqueness result for the adjoint equation (4.6).

Theorem 4.7. *There is a unique pair*

$$(p(\cdot), K(\cdot))\in L^2_{\mathscr{F}}(0,T;H)\times L^2_{\mathscr{F}}(0,T;H)^\infty$$

that solves the adjoint equation (4.6).

Proof Let $(p'(\cdot), K'(\cdot))$ be another solution. We denote

$$(\hat p(\cdot), \hat K(\cdot)) = (p(\cdot)-p'(\cdot),\ K(\cdot)-K'(\cdot)).$$

Then $(\hat p(\cdot), \hat K(\cdot))$ solves

$$\hat p(t) = \int_t^T e^{A^*(s-t)}[g_x^*\hat p(s) + \sigma_x^*\hat K(s)]ds - \int_t^T \hat K(s)dW(s). \tag{4.9}$$

Thus

$$\hat p(t) = \int_t^T e^{A^*(s-t)}E[g_x^*\hat p(s) + \sigma_x^*\hat K(s)\,|\,\mathscr{F}_t]ds.$$

Using once again the martingale representation theorem

$$E[g_x^*\hat p(s) + \sigma_x^*\hat K(s)\,|\,\mathscr{F}_t] = g_x^*\hat p(s) + \sigma_x^*\hat K(s) - \int_t^s K_1(r,s)dW(r) \tag{4.10}$$

with

$$K_1(\cdot, s)\in L^2_{\mathscr{F}}(0,s;\mathscr{L}(E,H)).$$

It follows that

$$\int_t^T e^{A^*(r-t)}\hat K(r)\,dW(r) = \int_t^T\int_r^s e^{A^*(s-t)}K_1(r,s)ds dW(r).$$

Thus

$$e^{A^*(r-t)}\hat K(r) = \int_r^T e^{A^*(s-t)}K_1(r,s)ds.$$

Particularly, when $t=r$

$$\hat K(r) = \int_r^T e^{A^*(s-r)}K_1(r,s)ds. \tag{4.11}$$

On the other hand, let

$$\hat\phi(t) = \phi(t) + g_x(t)x(t),\quad \hat\Psi(t) = \Psi(t) + \sigma_x(t)x(t),$$

where $x(t)$ solves (4.1), or

$$x(t) = \int_0^t e^{A(t-s)}\hat\phi(s)ds + \int_0^t e^{A(t-s)}\hat\Psi(s)dW(s).$$

We have

$$E\int_0^T\langle\hat p(t),\phi(t)\rangle dt+E\int_0^T\langle\hat K(t),\ \Psi(t)\rangle dt$$
$$=E\int_0^T\langle\hat p(t),\ \hat\phi(t)-g_x(t)x(t)\rangle dt+E\int_0^T\langle\hat K(t),\ \hat\Psi(t)-\sigma_x(t)x(t)\rangle dt$$
$$=E\int_0^T\langle\hat p(t),\ \hat\phi(t)\rangle dt+E\int_0^T\langle\hat K(t),\ \hat\Psi(t)\rangle dt$$
$$-E\int_0^T\Big\langle g_x^*(t)\hat p(t)+\sigma_x^*(t)\hat K(t),\int_0^t e^{A(t-s)}\hat\phi(s)\Big\rangle ds+\int_0^t e^{A(t-s)}\hat\Psi(s)dW(s)\Big\rangle dt$$
$$=E\int_0^T\Big\langle p(t)-\int_t^T e^{A^*(s-t)}(g_x^*p(s)+\sigma_x^*K(s))ds,\ \hat\phi(t)\Big\rangle dt+E\int_0^T\langle\hat K(t),\ \hat\Psi(t)\rangle dt$$
$$-E\int_0^T\Big\langle g_x^*(s)\hat p(s)+\sigma_x^*(s)\hat K(s),\ \int_0^s e^{A(s-t)}\hat\Psi(r)dW(r)\Big\rangle ds.$$

From (4.9), the first integral on the right hand is equal to zero. Substitute (4.10) in the third integral of the right hand with $t=0$,

$$E\int_0^T\langle p(t),\ \phi(t)\rangle dt+E\int_0^T\langle K(t),\ \Psi(t)\rangle dt$$
$$=E\int_0^T\langle\hat K(t),\ \hat\Psi(t)\rangle dt-E\int_0^T\Big\langle\int_0^s K_1(r,\ s)dW(r),\ \int_0^s e^{A(s-r)}\hat\Psi(r)dW(r)\Big\rangle ds$$
$$=E\int_0^T\langle\hat K(t),\ \hat\Psi(t)\rangle dt-E\int_0^T\int_0^s\langle e^{A^*(s-r)}K_1(r,\ s),\ \hat\Psi(r)\rangle dr ds$$
$$=E\int_0^T\langle\hat K(r)-\int_r^T e^{A^*(s-r)}K_1(r,\ s),\ \Psi(r)\rangle ds\,dr.$$

From (4.11), we derive finally

$$E\int_0^T\langle\hat p(t),\ \phi(t)\rangle dt+E\int_0^T\langle\hat K(t),\ \Psi(t)\rangle dt=0,$$
$$\forall(\phi(\cdot),\ \Psi(\cdot))\in L^2_{\mathscr F}(0,\ T;\ H)\times L^2_{\mathscr F}(0,\ T\ H)^\infty.$$

Thus $(\hat p(\cdot),\ \hat K(\cdot))=0$, a.e. a.s.. The proof is complete.

§ 5. Maximum Principle

Now we can assert the so called stochastic maximum principle for the optimal control problem described in 3.1.

Theorem 5.1. *Let $u(\cdot)$ be the optimal control and let $y(\cdot)$ be the corresponding trajectory. Then, there exists a unique pair $(p(\cdot),\ K(\cdot))$ in $L^2_{\mathscr F}(0,\ T;\ H)^\infty$, that solves the follwing adjoint equation*

$$p(t)=e^{A^*(T-t)}h_x(y(T))+\int_t^T e^{A^*(s-t)}[g_x^*(y(s),\ u(s))p(s)$$
$$+\sigma_x^*(y(s),\ u(s))K(s)+l_x(y(s),\ u(s))]ds$$
$$-\int_t^T e^{T^*(s-t)}K(s)dW(s),\tag{5.1}$$

such that the following maximum principle holds:

$$\langle l_v(y(t), u(t)) + g_v^*(y(t), u(t))p(t) + \sum_{i=1}^{\infty} \lambda_i \sigma_{iv}^*(y(t), u(t))K(t), v - u(t)\rangle \geqslant 0, \tag{5.2}$$

$$\forall v \in U_{ad}, \text{ a.e., a. s..}$$

Proof From the variational inequality (3.1), we have

$$Eh_x(y(T))z(T) + E\int_0^T [l_x(y(s), u(s))z(s) + l_v(y(s), u(s))v(s)]ds \geqslant 0, \tag{5.3}$$

$$\forall v(\cdot) \text{ such that } u(\cdot) + v(\cdot) \in \mathscr{U}_{ad},$$

where $z(\cdot)$ solves (3.2). But from (4.2),

$$Eh_x(y(T))z(T) + E\int_0^T l_x(y(s), u(s))z(s)ds$$

$$= E\int_0^T \langle p(s), g_v(y(s), u(s))v(s)\rangle ds + E\int_0^T \langle K(s), \sigma_v(y(s), u(s))\rangle ds,$$

where, according to Theorems 4.6 and 4.7, $(p(\cdot). K(\cdot))$ solves uniquely (5.1). The above relation with (5.3) implies

$$E\int_0^T \Big\langle l_v(y(s), u(s)) + g_v^*(y(s), u(s))p(s) + \sum_{i=1}^{\infty} \lambda_i \sigma_{iv}^*(y(s), u(s))K(s), v(s) - u(s)\Big\rangle \geqslant 0,$$

$$\forall v(\cdot) \in \mathscr{U}_{ad}.$$

And then (5.2) follows.

References

[1] Kushner, H. J., Necessary conditions for continous parameter stochastic optimazation problems, *SIAM J, Control*, **10**(1972), 550—565.

[2] Bensoussan, A., Lecture on stochastic control, Lecture Notes in Math. No. 972, Nonlinear Filtering and Stochastic Control, Proceedings, Cortona, 1981.

[3] Bismut, J. M.. An introductory approach to duality in optimal stochastic control, *SIAN*, *Rev.* **20**: 1(1978).

[4] Haussmann, U. G., General necessary conditions for optimal control of stochastic system, *Math. Programming Study*, **6**(1976), 34—48.

[5] Hu, Y., Maximum principle of optimal control for Markov processes, Proceedings of Symposium on Control Theory and Applications, PRC, Sept., 1987 (submitted to Acta Mathematica Sinica).

[6] Li Xunjing & Yao Yunlong, Maximum principle of distributed Parameter Systems with Time lags, Lecture Notes in Information and Control Sciences, No. 75, Springer-Verlag, 1985.

[7] Kushner, H. J., On the optimal control of the system governed by a linear parabolic equation with white noise inputs, *SIAM J. Control*, **6** (1968), 596—614.

[8] Bensoussan, A., Controle Optimal Stochastique de Systemes Gouvernes pardes Equations aux Derivees Partielles, Renti Conti di Matematica, Rome, 1969.

[9] Bensoussan, A., Stochastic maximum principle for distributed parameter system, *J. of the Franklin Institute*, **315**: **516**, (1983), 387—406.

[10] Balakrishnan, A. V, Applied Functional Analysis, Springer-Verlag, New York, 1976.

[11] Curtain, R. F. & Pritchard, A. J., Infinite Dimensional Linear Syatems Theory, Lecture Notes in Control and Information Sciences.

[12] Ichikawa, A., Stabililty of semilinear stochastic evolution equations, *J. Math. Anal. Appl.*, **90**(1982), 12—44.

[13] Kunita, H. & Watanabe, S., On square integrable martingles, *Nagaga Math. Journal*, **30** (1967)

[14] Kunita, H., Stochastic flows and stochastic differencial equations (to appear).