Chin. Ann. of Math.
16B: 3(1995),279-298.

SOLVABILITY OF FORWARD-BACKWARD SDES AND THE NODAL SET OF HAMILTON-JACOBI-BELLMAN EQUATIONS***

MA JIN* YONG JIONGMIN**

Abstract

The solvability of a class of forward-backward stochastic differential equations (SDEs for short) over an arbitrarily prescribed time duration is studied. The authors design a stochastic relaxed control problem, with both drift and diffusion all being controlled, so that the solvability problem is converted to a problem of finding the nodal set of the viscosity solution to a certain Hamilton-Jacobi-Bellman equation. This method overcomes the fatal difficulty encountered in the traditional contraction mapping approach to the existence theorem of such SDEs.

Keywords Forward-backward stochastic differential equations, Stochastic control, Relaxed control, Viscosity solutions, Nodal set.
1991 MR Subject Classification 35K15, 49L20, 49L25, 60H10, 93E20.

§1. Introduction

This paper studies the solvability (or the existence) of the adapted solutions to a certain class of forward-backward stochastic differential equations. Let $(\Omega, \mathcal{F}, P; \mathcal{F}_t)$ be a filtered probability space satisfying the usual conditions (see §2). Suppose that on this probability space a d-dimensional $\{\mathcal{F}_t\}$-Brownian motion $\{W_t\}_{t\geq 0}$ is given. Consider the following forward-backward stochastic differential equations (SDE for short):

$$X_t = x + \int_0^t b(X_s, Y_s, Z_s)ds + \int_0^t \sigma(X_s, Y_s, Z_s)dW_s, \tag{1.1}$$

$$Y_t = g(X_T) + \int_t^T \widehat{b}(X_s, Y_s, Z_s)ds + \int_t^T \widehat{\sigma}(X_s, Y_s, Z_s)dW_s, \tag{1.2}$$

where (X, Y, Z) takes value in $\mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^{m\times d}$ and $b, \widehat{b}, \sigma, \widehat{\sigma}, g$ are some smooth functions with appropriate dimensions; $T > 0$ is a prescribed constant which is called the time duration. Our objective is to find a triple (X, Y, Z) which is $\{\mathcal{F}_t\}$-adapted, square integrable, such that the equations (1.1)–(1.2) are satisfied on $[0, T]$. One should note that it is the extra process Z that makes it possible for (1.1)–(1.2) to have an adapted solution (cf. [23, 24, 25]).

Manuscript received September 19, 1994.
*Department of Mathematics, Purdue University, West Lafayette, IN 47907.
**Department of Mathematics, Fudan University, Shanghai 200433, China.
***Project supported by the U. S. NSF Grant #DMS-9301516, the National Natural Science Foundation of China and the Trans-Century Training Programme Foundation for the Talents by the State Education Commission of China.

The forward-backward SDEs of this kind was first introduced by Bismut[3] for studying the duality in optimal stochastic control and the stochastic Pontryagin Maximum Principle, in which the adjoint equation is a backward SDE. The SDE has been brought into strong attention recently because of its appeal not only in optimal control theory but also in mathematical finance and partial differential equations (cf. [10, 24] and the references therein). However, all the existing results regarding the existence and uniqueness of an adapted solution require that the product of Lipschitz constants of the coefficients and the time duration T be small enough (see, for example, [1, 24]). The restriction is simply due to the usual scheme of Picard iteration and the contraction mapping theorem. Thus, because of the "forward-backward" nature, one falls into a fatal difficulty when the time duration is large. In fact, [1] provided a counterexample showing that for some special kind of forward-backward SDEs, the adapted solution may fail to exist when the product of the Lipschitz constant and the time duration is larger than one. Therefore, the solvability of such an SDE over an arbitrarily prescribed time interval becomes an interesting issue, and sometimes is even crucial (for instance, when one studies the so-called decoupling problem in optimal stochastic control theory), but so far remains open.

In this paper we reformulate the above "forward-backward" SDE in terms of a martingale problem and consider its solutions in both strong and weak sense. In both cases we allow the underlying probability space to change when necessary; and in the latter case, we even allow the component Z to be a suitable adapted measure-valued process. This relaxation nontrivially contains the ordinary adapted solution as a special case. Our main strategy of attacking the solvability problem is to convert the problem of finding adapted solutions to the forward-backward SDEs to a problem of, roughly speaking, finding the "zero"-point (called the nodal set) of the viscosity solution to a certain Hamilton-Jacobi-Bellman (HJB for short) equation. In fact, this HJB equation is exactly the one that the value function of the properly designed optimal stochastic control problem should satisfy.

Intuitively, our scheme will work if the following two problems can be solved.

Problem 1. For any $x \in \mathbb{R}^n$, there exists a $y \in \mathbb{R}^m$ such that $v(0,x,y) = 0$, where v is the viscosity solution of a certain HJB equation;

Problem 2. (a) For any $x \in \mathbb{R}^n$, $y \in \mathbb{R}^m$, there exists an optimal relaxed control which attains its value function $V(0,x,y)$;

(b) $v(s,x,y) = V(s,x,y)$ for all $(s,x,y) \in [0,T] \times \mathbb{R}^n \times \mathbb{R}^m$.

It turns out that if both Problems 1 and 2 above can be solved, then the optimal trajectory (X,Y), together with the optimal relaxed control process Z, will be an adapted solution to the forward-backward equation (1.1)–(1.2).

The main advantage of this scheme is that both Problems 1 and 2 become attackable. In fact, Problem 2-(a) is an existence theorem of the optimal controls, so if we allow the component Z to take the form as a measure-valued process, it is always solvable in principle when we convert it to a stochastic relaxed control problem. Furthermore, the well-known "Chattering Lemma" (cf. [11, 14]) in the relaxed control context will lead to the solvability of Problem 2-(b). Therefore, in the sense of "weak solvability" of the forward-backward SDEs (see §2 for definition), we need only solve Problem 1, which is basically a problem of the existence of the nodal set for a certain HJB-equation. The notion of the "nodal set" was first introduced in [6] for studying the eigenfunctions of some linear elliptic equations (see [9] also). Recently, this notion was used for the study of the general solutions to elliptic and parabolic equations ([15, 20]). Also, it is further related to the so-called singular set, zero-set and partial regularity of solutions to PDEs (see [4, 5, 13] and references cited therein). Therefore, one should have at least some clue as to how to approach such a problem.

In summary, our scheme overcomes the difficulty encountered in the contraction mapping approach, and provides a novel strategy for us to study the original solvability problem. By studying the nodel set of a certain HJB equation, we are able to prove the solvability and non-solvability of a class of forward-backward SDEs. We hope that this correspondence will also raise some interesting questions in the study of (the existence of) the nodal sets in partial differential equations, via the existence of the adapted solutions to forward-backward SDEs. In [22], the problem is further developed by using the idea discovered in this paper.

This paper is organized as follows. In section 2 we formulate the forward-backward SDEs via the martingale problems and give some definitions. In sections 3 we design the corresponding optimal control problems and briefly review the results related to stochastic relaxed control. In section 4 we give the necessary and sufficient conditions for the solvability of forward-backward SDEs in terms of the solvability of optimal control and the existence of the nodal set of the related HJB-equations. A special class of forward-backward SDEs, raised from mathematical finance, is considered in sections 5 and 6. The weak solvability of such SDEs over an arbitrary time interval, as well as the conditions for non-solvability over the given interval, is studied by using our new approach. Finally, some discussions and concluding remarks are made in section 7.

§2. Problem Formulations

Throughout this paper, we assume that all the probability spaces will be filtered such that the quadruple $(\Omega, \mathcal{F}, P; \mathcal{F}_t)$ satisfies the usual conditions (that is, $\{\mathcal{F}_t\}$ is right-continuous and $\{\mathcal{F}_0\}$ contains all the P-null sets in $\mathcal{F}$). We will often use the notion of a natural extension of a filtered probability space, which can be found in [11,16].

Definition 2.1. *A natural extension of the space $(\Omega, \mathcal{F}, P; \mathcal{F}_t)$ is a filtered probability space $(\widetilde{\Omega}, \widetilde{\mathcal{F}}, \widetilde{P}; \widetilde{\mathcal{F}}_t)$ such that:*

(i) $\widetilde{\Omega} = \Omega \times \Omega'$ *where Ω' is an auxiliary space;*

(ii) $\widetilde{\mathcal{F}}_t \supseteq \mathcal{F}_t$ *where $\mathcal{F}_t$ is also the canonical extension of $\mathcal{F}_t$ to $\widetilde{\Omega}$;*

(iii) *the restriction of $\widetilde{P}$ to Ω is P;*

(iv) *each $(\mathcal{F}_t, P)$ martingale is an $(\widetilde{\mathcal{F}}_t, \widetilde{P})$ martingale.*

Now let U be a compact metric space (we often take U to be the closed subset of $\overline{\mathbb{R}^k}$, the one-point compactification of the Euclidean space $\mathbb{R}^k$). Denote by $\mathcal{P}(U)$ the space of all probability measures on $\mathcal{B}(U)$, the Borel σ-algebra of U. We endow the space $\mathcal{P}(U)$ with the Prohorov metric. Then $\mathcal{P}(U)$ is also a compact metric space. Let $\mathcal{M}([0,T] \times U)$ be the space of all $\mathcal{P}(U)$-valued functions $\mu_t(\cdot)$, $t \ge 0$, such that $\mu_\cdot(A)$ is Borel measurable for all $A \in \mathcal{B}(U)$. Furthermore, if a probability space $(\Omega, \mathcal{F}, P)$ is given, then we denote $\mathbb{M}(\Omega)$ to be the totality of all $\mathcal{M}([0,T] \times U)$-valued random variables $\mu(\cdot)$ defined on $(\Omega, \mathcal{F}, P)$, such that for each $A \in \mathcal{B}(U)$, the process $\mu_\cdot(A, \cdot)$ is $\{\mathcal{F}_t\}$-adapted.

We note that a U-valued adapted process $\{Z_t\}$ can be identified as an element in $\mathbb{M}(\Omega)$ through the relation $\mu_t(du, \omega) = \delta_{Z_t(\omega)}(du)$. In other words, $\mu_t(\cdot, \omega)$ is a δ-measure supported at $Z_t(\omega) \in U$ (cf. [14]). In this case, for any continuous function $g : U \to \mathbb{R}^n$, we have

$$\int_U g(u)\mu_t(du, \omega) = \int_U g(u)\delta_{Z_t(\omega)}(u)du = g(Z_t(\omega)).$$

We now give some formulations of the forward-backward stochastic equations. Let the probability space $(\Omega, \mathcal{F}, P; \mathcal{F}_t)$ be given such that a d-dimensional Brownian motion W is defined on it. A forward-backward SDE is a pair of SDEs with the following type:

$$X_t = x + \int_0^t b(X_s, Y_s, Z_s)ds + \int_0^t \sigma(X_s, Y_s, Z_s)dW_s, \tag{2.1}$$

$$Y_t = g(X_T) + \int_t^T \widehat{b}(X_s, Y_s, Z_s)ds + \int_t^T \widehat{\sigma}(X_s, Y_s, Z_s)dW_s, \tag{2.2}$$

where X takes value in $\mathbb{R}^n$, Y takes value in $\mathbb{R}^m$ and Z takes value in $U \subseteq \overline{\mathbb{R}^{m\times d}}$, with U being compact as well.

We make the following assumptions on the coefficients $b, \widehat{b}, \sigma, \widehat{\sigma}$ and g:

(A.1) The functions

$$\begin{cases} b : \mathbb{R}^n \times \mathbb{R}^m \times U \to \mathbb{R}^n; \\ \widehat{b} : \mathbb{R}^n \times \mathbb{R}^m \times U \to \mathbb{R}^m; \\ \sigma : \mathbb{R}^n \times \mathbb{R}^m \times U \to \mathbb{R}^{n\times d}; \\ \widehat{\sigma} : \mathbb{R}^n \times \mathbb{R}^m \times U \to \mathbb{R}^{m\times d} \end{cases} \tag{2.3}$$

are continuous and bounded in x, y, z; and are differentiable in x, y, such that all the partial derivatives are uniformly bounded by some constant $C > 0$.

(A.2) The function $g : \mathbb{R}^n \to \mathbb{R}^m$ is continuous.

A triple of adapted process (X, Y, Z) satisfying (2.1) and (2.2) is called an ordinary (or strong) adapted solution. Since such an ordinary solution may not exist when the time duration is large, we therefore look for the weak solution. As usual, we start from the martingale problem formulation of the adapted solutions. Note that we are only looking for adapted solutions, let us assume a priori that (X, Y, Z) is adapted. In this case, we can rewrite, for any $t \geq \tau \geq 0$, the equation (2.2) as

$$Y_t = Y_\tau - \int_\tau^t b(X_s, Y_s, Z_s)ds - \int_\tau^t \widehat{\sigma}(X_s, Y_s, Z_s)dW_s. \tag{2.4}$$

In particular, setting $\tau = 0$ and combining with (2.1), we see that solving (2.1)–(2.2) becomes the following problem: for any $x \in \mathbb{R}^n$, find a probability space $(\Omega, \mathcal{F}, P, \mathcal{F}_t)$, an $\mathbb{R}^m$-valued $\mathcal{F}_0$-measurable random variable Y_0, an $\{\mathcal{F}_t\}$-Brownian motion defined on this space; and an $\{\mathcal{F}_t\}$-adapted, U-valued process Z such that the following SDE with terminal condition:

$$X_t = x + \int_0^t b(X_s, Y_s, Z_s)ds + \int_0^t \sigma(X_s, Y_s, Z_s)dW_s, \tag{2.5}$$

$$Y_t = Y_0 - \int_0^t \widehat{b}(X_s, Y_s, Z_s)ds - \int_0^t \widehat{\sigma}(X_s, Y_s, Z_s)dW_s; \tag{2.6}$$

$$Y_T = g(X_T), \tag{2.7}$$

has an $\{\mathcal{F}_t\}$-adapted solution (X,Y) defined on $(\Omega,\mathcal{F},P)$.

Remark 2.1. One should note that solving (2.5)–(2.7) is not easy in general, since it is essentially a two-point boundary value problem for SDEs. It is quite possible that it does not have any solution of any kind.

Now let $T>0$ be given. Let C_b^2 denote the space of all functions $f:\mathbb{R}^n\times\mathbb{R}^m\to\mathbb{R}$ with bounded continuous partial derivatives up to the second order. For each $f\in C_b^2$, define the differential operator L by:

$$\begin{aligned}L[f](x,y,z)\triangleq\Big\{&\langle f_x,b\rangle-\langle f_y,\widehat{b}\rangle\\&+\frac{1}{2}\operatorname{tr}(\sigma\sigma^T f_{xx})-\operatorname{tr}(\sigma\widehat{\sigma}^T f_{xy})+\frac{1}{2}\operatorname{tr}(\widehat{\sigma}\widehat{\sigma}^T f_{yy})\Big\}(x,y,z).\end{aligned}\tag{2.8}$$

Definition 2.1. *Let $s\in[0,T]$ be given. An eight-tuple $(\Omega,\mathcal{F},P,\mathcal{F}_t,x,X,Y,Z)$ is called a strong adapted solution of* (2.1) *and* (2.2) *on $[s,T]$ if the following hold:*

(i) $(\Omega,\mathcal{F},P;\mathcal{F}_t)$ *is a filtered probability space;*

(ii) X,Y,Z *are $\{\mathcal{F}_t\}$-adapted processes such that*

$$C_t(f,Z)\triangleq f(X_t,Y_t)-f(X_s,Y_s)-\int_s^t L[f](X_r,Y_r,Z_r)dr \tag{2.9}$$

is an $\{\mathcal{F}_t\}$-martingale for $t\in[s,T]$;

(iii) *P-almost surely, it holds that*

$$X_s=x;\qquad Y_T=g(X_T). \tag{2.10}$$

It is easily seen that if (X,Y,Z) is an ordinary adapted solution to (2.5)–(2.7) defined on some probability space $(\Omega,\mathcal{F},P;\mathcal{F}_t)$, then by applying Itô's formula to $f(X_t,Y_t)$, $f\in C_b^2$, we see that $(\Omega,\mathcal{F},P,\mathcal{F}_t,x,X,Y,Z)$ must be a strong adapted solution on $[0,T]$ in the sense of Definition 2.1. (here the starting time can easily be replaced by $s\in[0,T]$). Conversely, if on some probability space $(\Omega,\mathcal{F},P;\mathcal{F}_t)$, there exists an adapted processes (X,Y,Z) that solves the martingale problem (2.9) and (2.10), then possibly on a natural extension of $(\Omega,\mathcal{F},P;\mathcal{F}_t)$, there exists a Brownian motion W such that (2.5)–(2.7) hold (see, for example, [11, Proposition 1.5]). In other words, modulo a possible change of probability space and the underlying Brownian motion, the solvability of (2.1) and (2.2) is equivalent to the solvability of (2.8) and (2.9). Therefore we give the following definition.

Definition 2.2. *Let $T>0$ be any given number. The forward-backward SDE* (2.1) *and* (2.2) *is called strongly solvable over $[0,T]$ if there exists a strong adapted solution in the sense of Definition* 2.1.

We should note here that such a change of formulation does not help us to solve the original problem, since we are still facing a two-point boundary problem. Nevertheless, we are now able to define the weak adapted solution so that our scheme will work.

Definition 2.3. *Let $s\in[0,T]$ be given. An eight-tuple $(\Omega,\mathcal{F},P,\mathcal{F}_t,x,X,Y,\mu)$ is called a weak adapted solution to* (2.1) *and* (2.2) *on $[s,T]$, if the following hold:*

(i) $(\Omega,\mathcal{F},P;\mathcal{F}_t)$ *is a filtered probability space;*

(ii) (X,Y) *is an $\mathbb{R}^n\times\mathbb{R}^m$-valued, $\{\mathcal{F}_t\}$-adapted square integrable process and $\mu\in\mathbb{M}(\Omega)$ such that*

$$C_t(f,\mu)\triangleq f(X_t,Y_t)-f(X_0,Y_0)-\int_0^T\int_U Lf(X_r,Y_r,u)\mu_s(du)dr \tag{2.11}$$

is an $\{\mathcal{F}_t\}$*-martingale for* $t \in [s,T]$*;*

(iii) P*-almost surely, it holds that*

$$X_s = x; \qquad Y_T = g(X_T). \tag{2.12}$$

Definition 2.4. *Let* $T > 0$ *be any given number. The forward-backward SDE* (2.1)–(2.2) *is called weakly solvable over* $[0,T]$ *if there exists a weak adapted solution to* (2.1) *and* (2.2) *on* $[0,T]$.

In what follows, we often distinguish the terms "strongly solvable" and "weakly solvable", but by "non-solvable" we mean neither strongly solvable nor weakly solvable. It is clear that if (2.1) and (2.2) is strongly solvable, then it must be weakly solvable. Indeed, if $(\Omega, \mathcal{F}, P, \mathcal{F}_t, x, X, Y, Z)$ is a strong adapted solution in the sense of Definition 2.1, then by setting $\mu_t(du, \omega) = \delta_{Z_t(\omega)}(du)$, we see that (X, Y, δ_Z) is a weak adapted solution in the sense of Definition 2.3. Moreover, under some restrictive conditions on the data and if the time duration T is small enough (cf. [25]), we know that the forward-backward equation is strongly solvable.

Our purpose is to investigate the solvability of (2.1) and (2.2) on any time interval $[0,T]$. We shall give a necessary condition for the solvability (both strongly and weakly) of (2.1) and (2.2) in the sense of Definitions 2.2 and 2.4; and we show that this condition is also sufficient if only the weak solvability is concerned. In the accompany paper [22], we study the strong solvability of such forward-backward SDEs via a slightly different approach.

§3. Formulation of the Control Problems

In this section we design some optimal control problems that will be important for our future discussion, the main reference related to these formulations is [11].

First, let us define a differential operator L_0 similar to L defined by (2.4): for some bounded continuous functions $b : \mathbb{R}^k \times U \to \mathbb{R}^k$, $\sigma : \mathbb{R}^k \times U \to \mathbb{R}^{k\times d}$, define

$$L_0[f](x,u) = \Big\{ \langle f_x, b \rangle + \frac{1}{2}\mathrm{tr}(\sigma\sigma^T f_{xx}) \Big\}(x,u). \tag{3.1}$$

Now let us define the control problems. Assume that $G : \mathbb{R}^k \to \mathbb{R}$ is a continuous function.

Definition 3.1 (Strong Problem). *Let* $(s,x) \in [0,T] \times \mathbb{R}^k$ *be given. A strong control problem is to find a filtered probability space* $(\Omega^*, \mathcal{F}^*, P^*, \mathcal{F}^*_t)$ *and a* U*-valued process* $\{u^*_t\}$ *that is progressively measurable with respect to* $\{\mathcal{F}^*_t\}$*, and an* $\mathbb{R}^k$*-valued adapted process* $\{\xi^*_t\}$ *with continuous paths, such that*

(i) *for each* $f \in C^2_b(\mathbb{R}^k)$

$$C_t(f,\mu) = f(\xi^*_t) - f(\xi^*_s) - \int_s^t L_0[f](\xi^*_r, u^*_r)dr \tag{3.2}$$

is a P^**-martingale for* $t \geq s$*;*

(ii) *for* $t \leq s$, $\xi^*_t = x$ P^**-a. s.*

(iii) *the six-tuple* $\mathcal{A}^{o,*} \triangleq (\Omega^*, \mathcal{F}^*, P^*, \mathcal{F}^*_t, u^*_t, \xi^*_t)$ *is a minimizer of the cost function*

$$J(s,x;\mathcal{A}^o) \triangleq E^P(G(X^*_T)) \tag{3.3}$$

among all the possible choices of the six-tuples $\mathcal{A}^o = (\Omega, \mathcal{F}, P, \mathcal{F}_t, u, \xi)$ *satisfying* (i) *and* (ii).

A six-tuple $\mathcal{A}^S = (\Omega, \mathcal{F}, P, \mathcal{F}_t, u, \xi)$ satisfying (i) and (ii) in Definition 3.1 is called an admissible strong control and the minimizer $\mathcal{A}^{S,*}$ is called an optimal strong control. The value function of the strong control problem is defined by

$$V^S(s,x) = \inf_{\mathcal{A}^S} J(s,x;\mathcal{A}^S). \tag{3.4}$$

Definition 3.2 (Relaxed Problem). *Let $(s,x) \in [0,T] \times \mathbb{R}^k$ be given. A relaxed control problem is to find a filtered probability space $(\Omega^*, \mathcal{F}^*, P^*, \mathcal{F}^*_t)$ and a $\mathcal{P}(U)$-valued process $\{\mu^*_t\} \in \mathbb{M}(\Omega^*)$ and an $\mathbb{R}^k$-valued adapted process $\{\xi^*_t\}$ with continuous paths, such that*

(i) *for each $f \in C^2_b(\mathbb{R}^k)$*

$$C_t(f,\mu) = f(\xi^*_t) - f(\xi^*_s) - \int_s^t \int_U L_0[f](\xi^*_r, u)\mu^*_r(du)dr \tag{3.5}$$

is a P^*-martingale for $t \ge s$;

(ii) *for $t \le s$, $\xi^*_t = x$ P^*-a. s.*

(iii) *the six-tuple $\mathcal{A}^* \triangleq (\Omega^*, \mathcal{F}^*, P^*, \mathcal{F}^*_t, u^*_t, \xi^*_t)$ is a minimizer of the cost function*

$$J(s,x;\mathcal{U}) \triangleq E^P(G(X^*_T))$$

among all the possible choices of the six-tuples $\mathcal{A} = (\Omega, \mathcal{F}, P, \mathcal{F}_t, \mu, \xi)$ satisfying (i) *and* (ii).

A six-tuple $\mathcal{A} = (\Omega, \mathcal{F}, P, \mathcal{F}_t, \mu, \xi)$ satisfying (i) and (ii) is called an admissible relaxed control and the minimizer $\mathcal{A}^*$ is called an optimal relaxed control. The value function of the relaxed control problem is defined by

$$V^R(s,x) = \inf_{\mathcal{A}} J(s,x;\mathcal{A}).$$

It is clear that an admissible strong control can be identified as an admissible relaxed control via the relation $\mu_t(du,\omega) = \delta_{Z_t(\omega)}(u)du$. Thus, if we define the set of all admissible strong controls by $\mathcal{U}^o$ and that of all admissible relaxed controls by $\mathcal{U}^R$, then we can imbed $\mathcal{U}^o$ into $\mathcal{U}^R$ in an obvious way. Therefore, in general we have

$$V^R(s,x) \le V^S(s,x), \qquad (s,x) \in [0,T] \times \mathbb{R}^k. \tag{3.6}$$

The following theorem combines the results in [11], which will be essential in our future discussions.

Theorem 3.1. (1) *For any given $(s,x) \in [0,T] \times \mathbb{R}^k$, there exists a relaxed control $\mathcal{A}^* = (\Omega, \mathcal{F}, P, \mathcal{F}_t, X, \mu, s, x) \in \mathcal{U}^R$ such that*

$$J(s,x;\mathcal{A}^*) = V^R(s,x).$$

(2) *The strong problem and the relaxed problem have the same value function. In other words, it holds that $V^R(s,x) = V^S(s,x)$ for all $(s,x) \in [0,T] \times \mathbb{R}^k$.*

Proof. Note that our assumption (A.1) guarantees the boundedness and the uniform Lipschitz property of the coefficients, so it can be checked that the so-called "weak uniqueness for controlled equations" (cf. [11, Definition 4.4]) holds. Therefore, (1) follows from Theorem 3.4 in [11] and (2) follows from Theorem 4.11 in [11].

Recall that (see also [7], [11,§6]), the value function $V(t,x) = V^S(s,x) = V^R(s,x)$ is the unique viscosity solution of the following HJB equation:

$$\begin{cases} V_t + \inf\limits_{u \in U}\{\langle V_x, b(x,u)\rangle + \dfrac{1}{2}\mathrm{tr}\,(\sigma(x,u)\sigma(x,u)^T V_{xx})\} = 0, \\ V\,\big|_{t=T} = G(x). \end{cases} \tag{3.7}$$

Here, V_x and V_{xx} stand for the gradient and the Hessian of V in x. For the definition and basic results of viscosity solutions of HJB equations, see [7].

Finally, we design the strong and relaxed control problems related to the forward-backward SDEs (2.1) and (2.2). Let $\mathbb{R}^k$ in the Definitions 3.1 and 3.2 be replaced by $\mathbb{R}^{n\times m}$, and denote the generic element of $\mathbb{R}^{n\times m}$ by (x,y), where $x\in\mathbb{R}^n$ and $y\in\mathbb{R}^m$. Let $\xi\triangleq(X,Y)$; also define $\tilde b=(b,-\widehat b)$ and $\tilde\sigma=(\sigma,-\widehat\sigma)$ in a similar way. Define $L_0=L$ where L is the differential operator defined by (2.4), and define $G(x,y)=|g(x)-y|^2$. Then $G:\mathbb{R}^{n\times m}\to\mathbb{R}$ is continuous and a strong (resp. relaxed) admissible control will be of the form $\mathcal{A}^R=(\Omega,\mathcal{F},P,\mathcal{F}_t,x,y,X,Y,Z)$ (resp. $\mathcal{A}=(\Omega,\mathcal{F},P,\mathcal{F}_t,x,y,X,Y,\mu)$) such that (i)–(iii) in Definition 3.1 (resp. Definition 3.2) are satisfied. It is readily seen that the HJB equation (3.7) now takes the following form:

$$\begin{cases} V_t+H(x,V_x,V_y,V_{xx},V_{xy},V_{yy})=0,\\ V(T,x,y)=|g(x)-y|^2,\end{cases}\tag{3.8}$$

where H is given by

$$\begin{aligned}&H(x,V_x,V_y,V_{xx},V_{xy},V_{yy})\\ =&\inf_{u\in U}\{\langle V_x,b(x,y,u)\rangle-\langle V_y,\widehat b(x,y,u)\rangle\\ &+\frac{1}{2}\mathrm{tr}\,(\sigma(x,y,u)\sigma(x,y,u)^TV_{xx})-\mathrm{tr}\,(\sigma(x,y,u)\widehat\sigma(x,y,u)^TV_{xy})\\ &+\frac{1}{2}\mathrm{tr}\,(\widehat\sigma(x,y,u)\widehat\sigma(x,y,u)^TV_{yy})\}.\end{aligned}\tag{3.9}$$

In order to build up the relations between the solvability of the forward-backward SDEs and the property of the viscosity solutions of the HJB equations in the next section, the following notion of nodal set of a continuous function $W(t,x)$ is important.

Definition 3.3. *Let $W(t,x)$ be a continuous function. The nodal set of W, denoted by $\mathcal{N}(W)$, is the set*

$$\mathcal{N}(W)=\{(t,x)\bigm|W(t,x)=0\}.\tag{3.10}$$

As pointed out in the introduction, this notion was first introduced in [6] for eigenfunctions of elliptic differential operators. Recently, this notion was also used in the study of general solutions of elliptic and parabolic equations. Here, we apply it to the viscosity solutions of HJB equations.

§4. Necessary and Sufficient Conditions for the Solvability

In this section we study the relationship between the solvability of forward-backward SDEs and the nodal set of the viscosity solutions of a certain HJB equations of the corresponding optimal control problems.

The first theorem concerns the necessary condition for the solvability of the forward-backward equation (2.1) and (2.2).

Theorem 4.1. *Let* (A.1) *and* (A.2) *hold. Suppose that for some $T>0$ and $x\in\mathbb{R}^n$, the forward-backward SDE* (2.1)–(2.2) *is either strongly or weakly solvable on* $[0,T]$. *Then it is necessary that the nodal set $\mathcal{N}(v)$ of v contains the point $(0,x,y)$ for some $y\in\mathbb{R}^m$, where*

v is the unique viscosity solutions of the following HJB equation:

$$\begin{cases} v_t + H(x, v_x, v_y, v_{xx}, v_{xy}, v_{yy}) = 0, \\ v(T, x, y) = (y - g(x))^2, \end{cases} \tag{4.1}$$

where H is given by

$$\begin{aligned} &H(x, v_x, v_y, v_{xx}, v_{xy}, v_{yy}) \\ &= \inf_{u\in U}\{\langle v_x, b(x, y, u)\rangle - \langle v_y, \widehat{b}(x, y, u)\rangle \\ &+ \frac{1}{2}\mathrm{tr}\,(\sigma(x, y, u)\sigma(x, y, u)^T v_{xx}) - \mathrm{tr}\,(\sigma(x, y, u)\widehat{\sigma}(x, y, u)^T v_{xy}) \\ &+ \frac{1}{2}\mathrm{tr}\,(\widehat{\sigma}(x, y, u)\widehat{\sigma}(x, y, u)^T v_{yy})\}. \end{aligned} \tag{4.2}$$

Proof. First, comparing the Definitions 3.1 and 2.1, we see that if the forward-backward SDEs (2.1) and (2.2) are strongly solvable over $[0, T]$ for some $x \in \mathbb{R}^n$, and if we let $(\Omega, \mathcal{F}, P, \mathcal{F}_t)$ be the probability space which carries the adapted strong solution (X, Y, Z) of the martingale problem (2.8) and (2.9), then we must have $E^P|g(X_T) - Y_T|^2 = 0$.

Next, let $\pi \in \mathcal{P}(\mathbb{R}^m)$ be the distribution of Y_0 and P_y be the regular conditional probability of P, given $Y_0 = y$. Denote $\mathcal{A}^y = (\Omega, \mathcal{F}, P_y, \mathcal{F}_t, x, y, X, Y, \mu)$. Then it is clear that $\mathcal{A}^y \in \mathcal{U}^R$ and for π-a.e. $y \in \mathbb{R}^m$, $X_0 = x$ and $Y_0 = y$, P_y-a. s. Therefore,

$$0 = E^P[g(X_T) - Y_T]^2 = \int_{\mathbb{R}^m} E^{P_y}[g(X_T) - Y_T]^2\pi(dy) = \int_{\mathbb{R}^m} J(0, x, y; \mathcal{A}^y)\pi(dy). \tag{4.9}$$

Since $J(0, x, y; \mathcal{A}^y) \geq 0$ for any $y \in \mathbb{R}^m$, we have $J(0, x, y, \mathcal{A}^y) = 0$, for π-a.e. $y \in \mathbb{R}^m$. Finally, since $V(0, x, y) \geq 0$ for all $(x, y) \in \mathbb{R}^{n+m}$, we obtain $V^S(0, x, y) = 0$, for π-a.e. $y \in \mathbb{R}^m$.

Finally, by Theorem 3.1 and the argument following it, we see that the value functions of the strong control problem and the relaxed control problem are the same and it is the unique viscosity solution of the corresponding HJB equation. Thus the above argument is also true if (2.1) and (2.2) is weakly solvable and the conclusion follows.

Intuitively, by introducing the above stochastic control problem, we see that if the viscosity solution of the HJB equation (4.1) and (4.2) satisfies that for any $x \in \mathbb{R}^n$, there is a $y \in \mathbb{R}^m$ such that $v(0, x, y) = 0$; and the optimal strong control (resp. relaxed control) exists, then the forward-backward SDEs (2.1) and (2.2) should be strongly (resp. weakly) solvable and the optimal strong (resp. relaxed) control is the strong (resp. weak) solution. In other words, the necessary condition in Theorem 4.1 becomes sufficient. The following theorem validates this idea.

Theorem 4.2. *Let* (A.1)–(A.2) *hold. Then, the forward-backward SDE* (2.1)–(2.2) *is weakly solvable over* $[0, T]$ *if and only if for any* $x \in \mathbb{R}^n$ *there exists a* $y \in \mathbb{R}^m$ *such that* $v(0, x, y) = 0$*, where* v *is the viscosity solution of* (4.1)–(4.2)*. In other words, the nodal set* $\mathcal{N}(v)$ *of* v *contains the point* $(0, x, y)$ *for some* $y \in \mathbb{R}^m$.

Proof. The necessity follows from Theorem 4.1. We need only show the sufficiency. To begin with, let $v(t, x, y)$ denote the unique continuous viscosity solution of HJB-equation (4.11)–(4.12). By Theorem 3.1, we know that the value function of the relaxed control problem and that of the strong control problem coincide and equal to v. Thus by the assumption, for any $x \in \mathbb{R}^n$, we can choose $y \in \mathbb{R}^m$ such that $v(0, x, y) = 0$. Now by Theorem 3.1 again, we can find an optimal relaxed control $\mathcal{A}^*_{x,y} = (\Omega^*, \mathcal{F}^*, P^*, \mathcal{F}^*_t, x, y, X^*, Y^*, \mu^*) \in \mathcal{U}^R$ such that

$$E^P|g(X^*_T) - Y^*_T|^2 = J(0, x, y; \mathcal{A}^*) = V(0, x, y) = 0. \tag{4.10}$$

Namely $g(X^*_T) = Y^*_T$, P-a.s. By Definition 2.3, $(\Omega^*, \mathcal{F}^*, P^*, \mathcal{F}^*_t, x, y, X^*, Y^*, \mu^*)$ is a weak solution to the forward-backward equation on $[0, T]$.

§5. Solvability of a Class of Forward-Backward SDEs

In this section, we try to solve a class of forward-backward SDEs by using the results of previous sections. Suppose a probability space $(\Omega, \mathcal{F}, P; \mathcal{F}_t)$ and an $\{\mathcal{F}_t\}$-Brownian motion W is given. Consider the following forward-backward SDE:

$$\begin{cases} X_t = x + \displaystyle\int_0^t b(X_s, Y_s)ds + \int_0^t \sigma(X_s, Y_s)dW_s, \\ Y_t = g(X_T) + \displaystyle\int_t^T \widehat{b}(X_s, Y_s)ds + \int_t^T \widehat{\sigma}(X_s, Y_s, Z_s)dW_s, \end{cases} \quad t \in [0, T], \tag{5.1}$$

where U is a compact subset of $\overline{\mathbb{R}^{n\times m}}$.

For simplicity, let us assume that all the processes and functions appeared in (5.1) are scalar valued; the higher dimensional case is basically the same (see Remark 6.6). We claim that if the filtration $\{\mathcal{F}_t\}$ is actually generated by the Brownian motion W, then the SDE of the type (5.1) is actually equivalent to the following form which is of special interest in mathematical finance (see, for example, [10] and the references therein):

$$\begin{cases} X_t = x + \displaystyle\int_0^t b(X_s, Y_s)ds + \int_0^t \sigma(X_s, Y_s)dW_s, \\ Y_t = E\left\{ g(X_T) + \displaystyle\int_t^T \widehat{b}(X_s, Y_s)ds \,\middle|\, \mathcal{F}_t \right\}, \end{cases} \quad t \in [0, T]. \tag{5.2}$$

To verify our claim, we first note that if (X, Y, Z) is an adapted solution of (5.1) in the usual sense ([23,24]), then by simply taking the condition expectation $E\{\cdot|\mathcal{F}_t\}$ on both sides of the second equation for each $t \in [0, T]$, we see immediately that (X, Y) is an adapted solution of (5.2). Conversely, suppose that (5.2) has an adapted solution (X, Y) which is square integrable; we shall prove that there exists an $\{\mathcal{F}_t\}$-adapted, square integrable process Z such that (X, Y, δ_Z) is a solution of a forward-backward SDE of the type (5.1). To see this, consider the square integrable martingale

$$M_t = E\left\{ g(X_T) + \int_0^T \widehat{b}(X_s, Y_s)ds \,\middle|\, \mathcal{F}_t \right\}, \ t \geq 0. \tag{5.3}$$

By the martingale representation theorem (see, for example, [16, p.182]), there exists a square integrable, $\{\mathcal{F}_t\}$-adapted process Z such that

$$M_t = M_0 + \int_0^t Z_s dW_s, \qquad \forall t \geq 0. \tag{5.4}$$

Since

$$\begin{aligned} &g(X_T) + \int_0^T \widehat{b}(X_s, Y_s)ds = M_0 + \int_0^T Z_s dW_s, \\ &Y_t = M_t - \int_0^t \widehat{b}(X_s, Y_s)ds, \end{aligned} \tag{5.5}$$

a simple computation leads to that

$$Y_t = g(X_T) + \int_t^T \widehat{b}(X_s, Y_s)ds - \int_t^T Z_s dW_s, \tag{5.6}$$

which is exactly of the form (5.1) with $\widehat{\sigma}(x, y, z) \equiv z$. It is obvious that our equation (5.1) is nontrivially more general.

We now use our scheme designed in the previous sections to solve the SDE (5.1). Let Z_t take values in some compact set $U \subset \mathbb{R}$, which will be determined later. Then, we formulate the relaxed control problem with the infinitesimal generator

$$\begin{aligned} L_0[f](x, y, z) = \Big\{ &\langle f_x, b(x, y) \rangle - \langle f_y, \widehat{b}(x, y) \rangle \\ &+ \frac{1}{2}\sigma(x, y)^2 f_{xx} - \sigma(x, y)\widehat{\sigma}(x, y, z) f_{xy} + \frac{1}{2}\widehat{\sigma}(x, y, z)^2 \Big\}, \end{aligned} \tag{5.7}$$

for any $f \in C_b^2(\mathbb{R}^{n\times m})$; and the cost functional

$$J(s, x, y; \mu) = E|Y_T - g(X_T)|^2. \tag{5.8}$$

Namely, for any $s \in [0, T]$, we are to find an eight-tuple $(\Omega, \mathcal{F}, P, \mathcal{F}_t, x, y, X, Y, \mu) \in \mathcal{U}^R$ satisfying: (1) for any $f \in C_b^2(\mathbb{R}^{n\times m})$,

$$C_t(f, \mu) = f(X_t, Y_t) - f(X_s, Y_s) - \int_s^t \int_U L_0[f](X_r, Y_r, u)\mu_r(du)dr$$

is a P-martingale; (2) $P\{X_t = x, Y_t = y : t \le s\} = 1$; (3) the cost functional (5.8) is minimized. We define the value function $V(s, x, y)$ as

$$V(s, x, y) = \inf_{\mu} J(s, x, y; \mu). \tag{5.9}$$

Then, we know that $V(s, x, y)$ is the unique viscosity solution of the following HJB-equation:

$$\begin{cases} V_t + H(x, y, V_x, V_y, V_{xx}, V_{xy}, V_{yy}) = 0, \\ V(T, x, y) = (y - g(x))^2, \end{cases} \tag{5.10}$$

where

$$\begin{aligned} H(x, y, V_x, V_y, V_{xx}, V_{xy}, V_{yy}) = {}& b(x, y)V_x - \widehat{b}(x, y)V_y + \frac{1}{2}\sigma(x, y)^2 V_{xx} \\ &+ \inf_{z\in U}\{-\sigma(x, y)\widehat{\sigma}(x, y, z)V_{xy} + \frac{1}{2}\widehat{\sigma}(x, y, z)^2 V_{yy}\}. \end{aligned} \tag{5.11}$$

Theorem 4.3 tells us that the weak solvability of (5.1) is equivalent to the following problem: For each $x \in \mathbb{R}$, find a $y \in \mathbb{R}$, such that

$$V(0, x, y) = 0. \tag{5.12}$$

Recall that (see [6, 9, 15, 20]) the nodal set $\mathcal{N}(V)$ of V is defined to be

$$\mathcal{N}(V) = \{(s, x, y) \in [0, \infty) \times \mathbb{R}^2 \mid V(s, x, y) = 0\}. \tag{5.13}$$

Thus, (5.1) is solvable if and only if the nodal set $\mathcal{N}(V)$ of V intersects each set of form $\{(0,x)\} \times \mathbb{R}$. Here, we should note that the value function actually depends on the time duration $T > 0$. Now, let us set

$$v(t,x,y) = V(T-t,x,y), \qquad (t,x,y) \in [0,T] \times \mathbb{R} \times \mathbb{R}. \tag{5.14}$$

Then, v is the unique viscosity solution of the following (see [7]):

$$\begin{cases} v_t - \dfrac{1}{2}\sigma(x,y)^2 v_{xx} - b(x,y)v_x + \widehat{b}(x,y)v_y \\ \quad - \inf\limits_{z\in U}\{-\sigma(x,y)\widehat{\sigma}(x,y,z)v_{xy} + \dfrac{1}{2}\widehat{\sigma}(x,y,z)^2 v_{yy}\} = 0, \\ v(0,x,y) = (y-g(x))^2. \end{cases} \tag{5.15}$$

The advantage of (5.15) is that this problem is posed on $(0,\infty)\times\mathbb{R}\times\mathbb{R}$, and the solvability of (5.1) over any $[0,T]$ is equivalent to the following statement: For any $(t,x)\in[0,\infty)\times\mathbb{R}$, there exists a $y\in\mathbb{R}$, such that

$$v(t,x,y) = 0. \tag{5.16}$$

Clearly, one way to do this is to find a function $\theta : [0,\infty)\times\mathbb{R}\to\mathbb{R}$, such that

$$v(t,x,\theta(t,x)) = 0, \qquad \forall (x,t) \in [0,\infty)\times\mathbb{R}; \tag{5.17}$$

or equivalently,

$$\{(t,x,\theta(t,x)) \bigm| (t,x)\in[0,\infty)\times\mathbb{R}\} \subset \mathcal{N}(v). \tag{5.18}$$

Sometimes, such a hypersurface $y=\theta(t,x)$ is called a nodal surface of V (see [6]). In the rest of this section, we are going to give some intuitive arguments of constructing such a nodal set. The existence of such a nodal set implies the solvability of the forward-backward SDE (5.1).

Suppose $v(t,x,y)$ is a classical solution of (5.15) and $\theta(t,x)$ is an undetermined smooth function with

$$\theta(0,x) = g(x), \qquad x\in\mathbb{R}. \tag{5.19}$$

We define

$$w(t,x) = v(t,x,\theta(t,x)), \qquad (t,x)\in(0,\infty)\times\mathbb{R}. \tag{5.20}$$

Then,

$$w(0,x) = 0, \qquad x\in\mathbb{R}. \tag{5.21}$$

On the other hand, by (5.20), we have, at $(t,x,\theta(t,x))$, that

$$\begin{cases} w_t = v_t + v_y\theta_t, \\ w_x = v_x + v_y\theta_x, \\ w_{xx} = v_{xx} + 2v_{xy}\theta_x + v_{yy}\theta_x^2 + v_y\theta_{xx}. \end{cases} \tag{5.22}$$

Then, by the equation in (5.15), we obtain

$$\begin{aligned} 0 = {} & w_t - \frac{1}{2}\sigma^2 w_{xx} - bw_x - (\theta_t - \frac{1}{2}\sigma^2\theta_{xx} - b\theta_x - \widehat{b})v_y \\ & - \inf_{z\in U}\{(\widehat{\sigma}+\sigma\theta_x)(-\sigma v_{xy} + \frac{1}{2}(\widehat{\sigma}-\sigma\theta_x)v_{yy})\}. \end{aligned} \tag{5.23}$$

Here, v_y, v_{xy} and v_{yy} are evaluated at $(t, x, \theta(t,x))$, $b, \widehat{b}$ and σ at $(x, \theta(t,x))$ and $\widehat{\sigma}$ at $(t, x, \theta(t,x), z)$. Now, we take the function $\theta(t,x)$ to be the classical solution of the following problem: (assuming, for the time being, such a solution exists)

$$\begin{cases} \theta_t - \dfrac{1}{2}\sigma(x,\theta)^2\theta_{xx} - b(x,\theta)\theta_x - \widehat{b}(x,\theta) = 0, \\ \theta(0,x) = g(x). \end{cases} \tag{5.24}$$

Further, we assume that

$$0 \in \{\widehat{\sigma}(x, \theta(t,x), z) + \sigma(x, \theta(t,x))\theta_x(t,x) \bigm| z \in U\}. \tag{5.25}$$

Then, from (5.23)–(5.25), we see that

$$w_t - \frac{1}{2}\sigma^2 w_{xx} - b w_x \le 0. \tag{5.26}$$

Hence, by (5.21) and maximum principle ([19]),

$$w(t,x) \le 0, \qquad (t,x) \in (0,\infty) \times \mathbb{R}. \tag{5.27}$$

However, by definition we know that $w(t,x)$ is nonnegative. Hence, we obtain

$$v(t,x,\theta(t,x)) \equiv w(t,x) = 0, \qquad (t,x) \in (0,\infty) \times \mathbb{R}. \tag{5.28}$$

Thus, $\theta(t,x)$ is a nodal surface of v.

§6. Nodal Sets of HJB Equations

In this section, we will make the heuristic arguments given in the previous section rigorously.

First, we recall some standard notations. For any bounded or unbounded region $G \subseteq \mathbb{R}^n$, we let $C(\bar{G})$ be the set of all bounded continuous functions defined on $\bar{G}$ ($\overline{\mathbb{R}^n} = \mathbb{R}^n$) with the norm

$$\|w\|_{C(\overline{G})} = \max_{x \in \overline{G}} |w(x)|, \qquad \forall w \in C(\overline{G}).$$

Then, we let $C^2(\overline{G})$ be the set of all bounded twice continuously differentiable functions defined on $\overline{G}$ with the norm

$$\|w\|_{C^2(\overline{G})} = \|w\|_{C(\overline{G})} + \|w_x\|_{C(\overline{G})} + \|w_{xx}\|_{C(\overline{G})}, \qquad \forall w \in C^2(\overline{G}).$$

Here w_x and w_{xx} stand for the gradient and the Hessian of w, respectively. For $\alpha \in (0,1)$, we define $C^{2+\alpha}(\overline{G})$ to be the set of all elements in $C^2(\overline{G})$ such that the second partial derivatives are Hölder continuous with the exponent α. The norm in $C^{2+\alpha}(\overline{G})$ is defined to be

$$\|w\|_{C^{2+\alpha}(\overline{G})} = \|w\|_{C^2(\overline{G})} + \sup_{x \ne x', x, x' \in \overline{G}} \frac{|w_{xx}(x) - w_{xx}(x')|}{|x - x'|^\alpha}.$$

Next, for any $T > 0$ and any bounded or unbounded region $G \subseteq \mathbb{R}^n$, denote $Q_T = (0,T) \times G$. Let $C^{2+\alpha, 1+\alpha/2}(\overline{Q}_T)$ be the space of all functions $\theta(t,x)$ which are differentiable in t and twice differentiable in x with θ_t and θ_{xx} being $\alpha/2$- and α-Hölder continuous in $(t,x) \in \overline{Q}_T$, respectively. In $C^{2+\alpha, 1+\alpha/2}(\overline{Q}_T)$, we define the norm to be

$$\begin{aligned} \|\theta\|_{C^{2+\alpha, 1+\frac{\alpha}{2}}(\overline{Q}_T)} &= \|\theta\|_{C(\overline{Q}_T)} + \|\theta_t\|_{C(\overline{Q}_T)} + \|\theta_x\|_{C(\overline{Q})} + \|\theta_{xx}\|_{C(\overline{Q})} \\ &\quad + \sup_{(t,x),(t'x') \in \overline{Q}, (t,x) \ne (t',x')} \frac{|\theta_t(t,x) - \theta_t(t',x')| + |\theta_{xx}(t,x) - \theta_{xx}(t',x')|}{(|x - x'|^2 + |t - t'|)^{\alpha/2}}. \end{aligned}$$

Now, let us make some hypotheses.

(H1) Functions $b, \widehat{b}, \sigma$ are $C^2(\mathbb{R}^2)$ and g is $C^{2+\alpha}(\mathbb{R})$ (for some $\alpha \in (0,1)$) with

$$\|b\|_{C^2(\mathbb{R}^2)} + \|\widehat{b}\|_{C^2(\mathbb{R}^2)} + \|\sigma\|_{C^2(\mathbb{R}^2)} + \|g\|_{C^{2+\alpha}(\mathbb{R})} \le C. \tag{6.1}$$

Moreover, there exists a constant $\nu > 0$, such that

$$\sigma(x,y)^2 \ge \nu, \qquad \forall (x,y) \in \mathbb{R}^2. \tag{6.2}$$

(H2) Function $\widehat{\sigma}$ is continuous. For each $z \in \mathbb{R}$, $\widehat{\sigma}(\cdot,\cdot,z)$ is in $C^2(\mathbb{R}^2)$ with

$$\|\widehat{\sigma}(\cdot,\cdot,z)\|_{C^2(\mathbb{R}^2)} \le C_R, \qquad \forall |z| \le R. \tag{6.3}$$

(H3) Function $\widehat{\sigma}$ satisfies

$$\{\widehat{\sigma}(x,y,z) \bigm| z \in \mathbb{R}\} = \mathbb{R}, \qquad \forall (x,y) \in \mathbb{R}^2. \tag{6.4}$$

Remark 6.1. The regularity of $b, \widehat{b}, \sigma$ and $\widehat{\sigma}$ might be relaxed. In order to do this, some arguments of [32] should be adopted. Also, the regularity of g can be relaxed as well. In this case, the solution θ of (5.24) will be less regular near $t = 0$. We also point out that (H1)–(H3) imply the state equation (5.7) satisfies (A.1) of §3 (for $z \in U$ with U being a compact set in $\mathbb{R}$).

The following result concerns the well-posedness of (5.24). A proof can be found in [8,19].

Lemma 6.1. *Let* (H1) *hold. Then,* (5.24) *admits a unique solution* $\theta(t,x)$ *in* $C^{2+\alpha,1+\alpha/2}$ $([0,\infty)\times\mathbb{R})$. *In particular, for any* $T > 0$,

$$\sup_{x\in\mathbb{R}, t\in[0,T]} |\theta_x(t,x)| < \infty. \tag{6.5}$$

Now, we come up with the first main result of this section.

Theorem 6.1. *Let* (H1)–(H3) *hold. Let* $v(t,x,y)$ *be the viscosity solutions of* (5.15) *and* $\theta(t,x)$ *be the solution of* (5.24). *Then, the nodal set* $\mathcal{N}(v)$ *of* $v(t,x,y)$ *is given by*

$$\mathcal{N}(v) = \{(t,x,\theta(t,x)) \bigm| (t,x) \in [0,\infty)\times\mathbb{R}\}. \tag{6.6}$$

Proof. For any $\varepsilon > 0$, we introduce the following problem:

$$\begin{cases} v_t^\varepsilon - \dfrac{1}{2}\sigma(x,y)^2 v_{xx}^\varepsilon - b(x,y)v_x^\varepsilon + \widehat{b}(x,y)v_y^\varepsilon \\ \qquad - \inf\limits_{z\in U}\{-\sigma(x,y)\widehat{\sigma}(x,y,z)v_{xy}^\varepsilon + \dfrac{1}{2}\widehat{\sigma}(x,y,z)^2 v_{yy}^\varepsilon\} - \varepsilon v_{yy}^\varepsilon = 0, \\ v^\varepsilon(0,x,y) = (y - g(x))^2. \end{cases} \tag{6.7}$$

Here, we take U to be a compact set in $\mathbb{R}$, such that

$$\begin{gathered} \inf_{z\in U} \widehat{\sigma}(x,\theta(x,t),z) \le \sigma(x,\theta(t,x))\theta_x(t,x) \le \sup_{z\in U} \widehat{\sigma}(x,\theta(t,x),z), \\ \forall (t,x) \in [0,T]\times\mathbb{R}. \end{gathered} \tag{6.8}$$

This is possible due to (H3) and the boundedness of $\sigma(x,y)$ and $\theta_x(t,x)$ (see (6.1) and (6.5)). Then, we know that (6.7) is a nondegenerate fully nonlinear parabolic equation. From [18] (see [12] also), we know that there exists a classical solution v^ε of (6.7). On the other hand, this v^ε is the value function of the optimal stochastic control problem similar to (5.7)–(5.8) in which a term $\sqrt{\varepsilon/2}\,dW_r'$ is added in the second equation of (5.7) with W' being another Brownian motion independent of W. Thus, by [17], we can find a continuous function $K(t,x,y) > 0$, independent of $\varepsilon > 0$, such that

$$v_{yy}^\varepsilon(t,x,y) \le K(t,x,y), \qquad \forall (t,x,y) \in [0,\infty)\times\mathbb{R}^2, \quad \varepsilon > 0. \tag{6.9}$$

Now, we set

$$w^\varepsilon(t,x)=v^\varepsilon(t,x,\theta(t,x)),\qquad (t,x)\in[0,\infty)\times\mathbb{R}. \tag{6.10}$$

Then, similar to (5.22), we have

$$\begin{aligned}
&w_t^\varepsilon-\frac{\sigma(x,\theta(t,x))^2}{2}w_{xx}^\varepsilon-b(x,\theta(t,x))w_x^\varepsilon\\
&=\inf_{z\in U}\{(\widehat{\sigma}(x,\theta(t,x),z)+\sigma(x,\theta(t,x))\theta_x(x,t))[-\sigma(x,\theta(t,x))v_{xy}^\varepsilon(t,x,\theta(t,x))\\
&\quad+\frac{1}{2}\big(\widehat{\sigma}(x,\theta(t,x),z)-\sigma(x,\theta(t,x))\theta_x(t,x)\big)v_{yy}^\varepsilon(t,x,\theta(t,x))]\}+\varepsilon v_{yy}^\varepsilon(x,t,\theta(t,x))\\
&\le\varepsilon K(t,x,\theta(x,t)).
\end{aligned}\tag{6.11}$$

Here, we have used the facts (6.8) and (6.9). Hence, the function $w^\varepsilon(t,x)$ defined by (6.10) satisfies the following (in the classical sense and thus in the viscosity sense):

$$\begin{cases} w_t^\varepsilon-\dfrac{\sigma(x,\theta(t,x))^2}{2}w_{xx}^\varepsilon-b(x,\theta(t,x))w_x^\varepsilon\le\varepsilon K(t,x,y),\\ w^\varepsilon(0,x)=(g(x)-\theta(0,x))^2=0.\end{cases}\tag{6.12}$$

On the other hand, by [7], we know that $v^\varepsilon(t,x,y)$ converges to the unique viscosity solution $v(t,x,y)$ of (5.15) uniformly in any compact sets. Thus, we see that $w^\varepsilon(t,x)$ converges to $w(t,x)=v(t,x,\theta(t,x))$ uniformly in any compact sets. Then, by [7] again, this $w(t,x)$ is a viscosity solution of

$$\begin{cases} w_t-\dfrac{\sigma(x,\theta(t,x))^2}{2}w_{xx}-b(x,\theta(t,x))w_x\le 0,\\ w(0,x)=0.\end{cases}\tag{6.13}$$

Therefore, by a comparison theorem, we must have $w(t,x)\le 0$. But, we know that $v(t,x,y)$ is nonnegative. Hence, we have

$$\mathcal{N}(v)\supseteq\{(t,x,\theta(t,x))\bigm|(t,x)\in[0,\infty)\times\mathbb{R}\}. \tag{6.14}$$

On the other hand, let us set $\psi(t,x,y)=\frac{1}{2}(y-\theta(t,x))^2e^{-\alpha t}$, $(t,x,y)\in[0,\infty)\times\mathbb{R}^2$, where α is some constant to be determined. Then an easy computation leads to that

$$\begin{aligned}
&\psi_t-\frac{\sigma(x,y)^2}{2}\psi_{xx}-b(x,y)\psi_x+\widehat{b}(x,t)\psi_y\\
&\quad-\inf_{z\in U}[-\sigma(x,y)\widehat{\sigma}(x,y,z)\psi_{xy}+\frac{\widehat{\sigma}(x,y,z)^2}{2}\psi_{yy}]\\
&=e^{-\alpha t}\Big\{(\theta-y)\big[-\frac{\alpha}{2}(\theta-y)+\theta_t-\frac{\sigma(x,y)^2}{2}\theta_{xx}-b(x,y)\theta_x-\widehat{b}(x,t)\big]\\
&\quad-\frac{1}{2}\theta_x^2\sigma(x,y)^2-\inf_{z\in U}[\sigma(x,y)\widehat{\sigma}(x,y,z)\theta_x+\frac{\widehat{\sigma}(x,y,z)^2}{2}]\Big\}.
\end{aligned}\tag{6.15}$$

Note that

$$\begin{aligned}
&-\inf_{z\in U}[\sigma(x,y)\widehat{\sigma}(x,y,z)\theta_x+\frac{\widehat{\sigma}(x,y,z)^2}{2}]\\
&=\sup_{z\in U}\big\{\frac{1}{2}\sigma(x,y)^2\theta_x^2-[\sigma(x,y)\theta_x+\widehat{\sigma}(x,y,z)]^2\big\}\\
&\le\frac{1}{2}\sigma(x,y)^2\theta_x^2,
\end{aligned}$$

and that θ satisfies (5.24), we see that the right hand side of (6.15) is no greater than

$$\begin{aligned} e^{-\alpha t}\Big\{(\theta-y)\big[-\frac{\alpha}{2}(\theta-y)-\frac{(\sigma(x,y)^2-\sigma(x,\theta)^2)}{2}\theta_{xx} \\ -(b(x,y)-b(x,\theta))\theta_x-(\widehat{b}(x,y)-\widehat{b}(x,\theta))\big]\Big\}. \end{aligned} \tag{6.16}$$

The Lipschitz condition and the boundedness of the coefficients immediately yield that the expression in (6.16) is negative if α is sufficiently large. Therefore we have

$$\begin{aligned} &\psi_t-\frac{\sigma(x,y)^2}{2}\psi_{xx}-b(x,y)\psi_x+\widehat{b}(x,t)\psi_y \\ &-\inf_{z\in U}[-\sigma(x,y)\widehat{\sigma}(x,y,z)\psi_{xy}+\frac{\widehat{\sigma}(x,y,z)^2}{2}\psi_{yy}]\le 0. \end{aligned}$$

Note that $\varphi(0,x,y)=(y-g(x))^2$, we conclude that $\psi(t,x,y)$ is a viscosity subsolution of (5.15) for sufficiently large α. Thus by a comparison theorem we obtain

$$(y-\theta(t,x))^2\le v(t,x,y)e^{\alpha t},\qquad \forall(t,x,y)\in[0,\infty)\times\mathbb{R}^2, \tag{6.17}$$

which leads to

$$\mathcal{N}(v)\subseteq\{(t,x,\theta(t,x))\bigm|(t,x)\in[0,\infty)\times\mathbb{R}\}. \tag{6.18}$$

Hence, (6.6) follows from (6.14) and (6.18).

The above result together with the results of §4 will give the following

Theorem 6.2. *Let* (H1)–(H3) *hold. Then, for any $T>0$ and $x\in\mathbb{R}$, the forward-backward SDE* (5.1) *is solvable.*

We note that since the optimal (relaxed) controls are not necessarily unique, we do not have the uniqueness of the solutions of (5.1).

Next result is related to the condition (6.4).

Theorem 6.3. *Let* (H1)-(H2) *hold. Suppose one of the following holds:*

(i) *The continuity of $\widehat{\sigma}$ in y is uniform in all its arguments. Moreover, the following holds:*

$$\inf_{x,z\in\mathbb{R}}[\sigma(x,g(x))g'(x)+\widehat{\sigma}(x,g(x),z)]^2>0; \tag{6.19}$$

(ii) *The following holds:*

$$\inf_{x,y,z\in\mathbb{R}}[\sigma(x,y)g'(x)+\widehat{\sigma}(x,y,z)]^2>0. \tag{6.20}$$

Then, there exists a $T_0>0$, such that for any $T\in(0,T_0]$, (5.1) *is not solvable on* $[0,T]$.

Proof. First, we let (i) holds. Then, there exist a $\delta>0$ and $T_0>0$, such that

$$\inf_{t\in[0,T_0],x\in\mathbb{R}}[\sigma(x,y)\theta_x(t,x))+\widehat{\sigma}(x,y,z)]^2\ge\delta,\qquad\forall|y-g(x)|\le\delta. \tag{6.21}$$

Then, we can find a function $h\in C^\infty(\mathbb{R}^2)$ such that

$$\begin{cases} 0\le h(x,y)\le 1, & (x,y)\in\mathbb{R}^2, \\ h(x,g(x))=1, & \forall x\in\mathbb{R}, \\ h(x,y)=0, & \forall|y-g(x)|\ge\delta. \end{cases} \tag{6.22}$$

Now, we define

$$\psi(t,x,y)=\varepsilon th(x,y)+(y-g(x))^2,\qquad (t,x,y)\in[0,\infty)\times\mathbb{R}^2, \tag{6.23}$$

with $\varepsilon > 0$ being undetermined. Similar to (6.16), we can obtain

$$\begin{aligned}
&\psi_t - \frac{\sigma(x,y)^2}{2}\psi_{xx} - b(x,y)\psi_x + \widehat{b}(x,t)\psi_y \\
&- \inf_{z\in U}[-\sigma(x,y)\widehat{\sigma}(x,y,z)\psi_{xy} + \frac{\sigma(x,y)^2}{2}\psi_{yy} \\
= &\varepsilon h(x,y) - \frac{\sigma(x,y)^2}{2}\varepsilon h_{xx}(x,y)t - b(x,y)\varepsilon h_x(x,y)t + \widehat{b}(x,y)\varepsilon h_y(x,y)t \\
&- \inf_{z\in U}\{[\sigma(x,y)\theta_x(t,x) + \widehat{\sigma}(x,y,z)]^2 - \sigma(x,y)\widehat{\sigma}(x,y,z)\varepsilon h_{xy}(x,y)t \\
&+ \frac{\widehat{\sigma}(x,y,z)}{2}\varepsilon h_{yy}(x,y)t\} \le 0,
\end{aligned} \tag{6.24}$$

provided $\varepsilon > 0$ is sufficiently small and $(t,x) \in [0,T_0] \times \mathbb{R}$. Also, we have $\psi(0,x,y) = (y-g(x))^2$. Hence, it follows that

$$\varepsilon h(x,y)t + (y-\theta(t,x))^2 \le v(t,x,y), \qquad (t,x,y) \in [0,T_0] \times \mathbb{R}^2. \tag{6.25}$$

This implies that

$$\mathcal{N}(v) \bigcap \{(0,T_0] \times \mathbb{R}^2\} = \emptyset. \tag{6.26}$$

Hence, (5.1) has no solutions on any $[0,T]$ with $T \in (0,T_0]$.

Now, in the case (ii) holds, we can similarly prove (6.26) by taking $\psi(t,x,y) = \varepsilon t + (y - g(x))^2$ for some sufficiently small $\varepsilon > 0$.

All the above results are sort of global. We next look at some local results. To this end, we let $x \in \mathbb{R}$. Define

$$T_x = \sup\{t > 0 \bigm| \inf_{z\in\mathbb{R}}[\sigma(x,\theta(t,x))\theta_x(t,x) + \widehat{\sigma}(x,\theta(t,x),z)]^2 = 0\}, \tag{6.27}$$

with $\theta(t,x)$ being the solution of (5.24). The following result gives a local version of Theorem 6.1.

Theorem 6.4. *Let* (H1) *and* (H2) *hold. Then,*

$$\mathcal{N}(v) \bigcap \{[0,T_x] \times \mathbb{R}^2\} = \{(t,x,\theta(t,x)) \bigm| t \in [0,T_x]\ \}, \qquad \forall x \in \mathbb{R}. \tag{6.28}$$

Consequently, for any $x \in \mathbb{R}$, (5.1) *is solvable on $[0,T_x]$ and there exists a $\delta_0 > 0$, such that* (5.1) *is not solvable on $[0,T_x+\delta]$ for any $\delta \in (0,\delta_0]$.*

The proof is clear.

Remark 6.2. Although we only discussed the case in which all the functions involved in (5.1) are scalar valued, it is not hard to see that our arguments are good for higher dimensions. In this case, some assumptions should be accordingly changed. For example, (6.2) and (6.4) should be replaced respectively by

$$\sigma(x,y)\sigma(x,y)^T \ge \nu I, \qquad \forall (x,y) \in \mathbb{R}^n \times \mathbb{R}^m, \tag{6.29}$$

and

$$\{\widehat{\sigma}(x,y,z) \bigm| z \in \mathbb{R}^\ell\} = \mathbb{R}^m, \qquad \forall (x,y) \in \mathbb{R}^n \times \mathbb{R}^m, \tag{6.30}$$

provided we assume x, y and z to be in $\mathbb{R}^n$, $\mathbb{R}^m$ and $\mathbb{R}^\ell$. We omit the exact statements here.

§7. Discussions

In this section we discuss some interesting implications of our result obtained in the previous sections.

1. Discussion of the Hypotheses (H1)—(H3)

In order that the conclusion of Theorem 6.2 be true, we assumed several conditions on the data. Among them, (6.1), (6.3) are regularity conditions, and as we pointed out in the remark after (H1)—(H3), these conditions can be relaxed if the methodology of [24] is adopted. The condition (6.4) is in some sense standard as long as a backward equation is involved. In fact, even in the pure backward case (cf. [23, 24]), this condition was also posed so that the process Z can be solved via the martingale representation theorem. From Theorem 6.4, we see that condition (6.4) is very close to a necessary condition for the solvability of (5.1) on any $[0,T]$ and any $x \in \mathbb{R}$. However, the seemingly artificial non-degeneracy condition of σ (6.2), which guarantees that the PDE (5.24) has a classical solution, seems unremovable. A counterexample can be found in [1], in which $\sigma \equiv 0$ and the forward-backward equation with the type of (5.2) is not solvable for $T \geq 1$. Therefore, in order to make the forward-backward equation (5.1) solvable, it is necessary that the forward diffusion process is "random enough". We should say that the nondegeneracy of $\sigma(x,y)$ and the condition (6.4) represent the essential solvability feature of our forward-backward SDE (5.1). As a matter of fact, if, say, σ and $\widehat{\sigma}$ are both identically zero, then (5.1) is reduced to a two-point boundary value problem. We know that in general it may have no solutions ([2]).

2. The Existence of Ordinary Adapted Solutions

As we pointed out in §1, the only reason that we use the wider-sense (or relaxed) solution for the forward-backward SDEs is to guarantee that the optimal control exists, which is essential to our scheme. Therefore, it would be nice to know when the ordinary solution of the forward-backward equation exists. The result of [24] shows that under some restrictive conditions on data, the forward-backward equation (5.1) and (5.2) has a unique ordinary solution when T is small enough. Using the equivalent relations (Theorems 6.2 and 6.3), we see that this implies that the Problem (C) will have an ordinary optimal control which is even unique. However, when T is large, whether the relaxed control problem (C) will have an ordinary optimal control is a quite challenging problem in general. But in some special cases, it is still workable. Let us take the forward-backward equation (5.1), with an assumption that $\widehat{\sigma}(x,y,z) \equiv z$, as an example.

Recall from (6.1) that g is bounded. Then it is easily seen that there exists a compact set $U \subset \mathbb{R}$ such that for each $(t,x) \in [0,T] \times \mathbb{R}$, $0 \in \{z + \sigma(t,\theta(t,x))\theta_x(t,x) | z \in U\}$. From the proof of Theorem 6.3, this is sufficient for us to conclude that $v(t,x,\theta(t,x)) = 0$ for any $(t,x) \in [0,T] \times \mathbb{R}$.

Now let us take this compact set U to set up an ordinary control problem with the state equation (5.7) with $\widehat{\sigma}(x,y,z) \equiv z$, and cost functional (5.8). Note that now σ is linear in z, so one can show by a similar technique as we used in §3 and §4 that the ordinary optimal control exists (modulo a change of probability space). Therefore the SDE (5.1) will have an ordinary adapted solution (X,Y,Z), with Z taking value in a compact set! Such a result seems hard to be obtained by using the martingale representation theorem. It is evident that this scheme will also work for all forward-backward SDEs that are linear in Z, as long as a suitable compact set can be found a priori from the study of Problem (N).

3. Solvability of SDEs and Controllability

Another interesting implication of our results, which does not seem possible to argue by using a contraction mapping theorem, is the non-solvability of the forward-backward SDEs. Theorem 6.3 provides a non-existence result which basically says that it is possible that for some $T > 0$, (5.1) is solvable over $[0, T]$, but not solvable over any $[0, S]$ with $S < T$. In other words, even if the time duration is small, (5.1) can still be unsolvable if the coefficients are not well-matched. Therefore, it describes a deeper feature of the solvability of forward-backward SDEs that is so far undiscovered. One can check that all the known cases (cf. e.g., [1, 24]) in which an adapted solution exists are actually in the complement of this case (see condition (6.20)). This phenomenon is, however, quite natural if we look at it from other points of view; for instance, as a boundary value problem of linear ordinary differential equations (see [2], for example); or as a controllability problem. It is known (cf. [21]) that it is possible that the considered systems is completely controllable but not small time controllable. Therefore, there may be some kind of waiting time before hitting a target becoming possible.

4. The Nodal Sets

To our best knowledge, in most of the literatures concerning nodal sets of the solutions to PDEs, people mainly focused on the estimates of the upper bound of the size of nodal sets in terms of Hausdorff measure (see [9,15,20] and also [4,5,13]). In this paper, we actually addressed the problem in terms of the non-emptyness and the shape of a nodal set. We have constructed such a nodal set for the viscosity solution to a certain class of HJB equations. We also believe that the study of the nodal sets for solutions to general nonlinear degenerate elliptic and parabolic PDEs would be very interesting; and some more investigation will be made along this line in our future publications.

Acknowledgements. We would like to thank Professor N. V. Krylov for pointing out a mistake in the first draft of this paper; our thanks are also due to Professor E. Pardoux for bringing the reference [11] into our attention.

References

[1] Antonelli, F., Backward-forward stochastic differential Equations, *Ann. Appl. Prob.,* **3** (1993), 777-793.

[2] Bailey, P. B., Shampine, L. F. & Waltman, P. E., Nonlinear two point boundary value problems, Academic Press, New York, 1968.

[3] Bismut, J. M., An introductory approach to duality in optimal stochastic control, *SIAM Rev.,* **20** (1978), 62-78.

[4] Caffarelli, L. A. & Friedman, A., Partial regularity of the zero-set of semilinear and superlinear elliptic equations, *J. Diff. Eqn.,* **60** (1985), 420-433.

[5] Cannarsa, P. & Soner, H. M., On the singularities of the viscosity solutions to Hamiliton-Jacobi-Bellman equations, *Indiana Univ. Math. J.,* **36** (1987), 501-524.

[6] Courant, R. & Hilbert, D., Methods of mathematical physics, Vol.1, Interscience, New York, 1953.

[7] Crandall, M. G., Ishii, H. & Lions, P. L., User's guide to viscosity solutions of second order partial differential equations, *Bull. Amer. Math. Soc. (New Series)*, **27** (1992), 1-67.

[8] Dong, G., Nonlinear partial differential equations of second order, Translation of Mathematical Monographs Vol.95, AMS, Providence, 1991.

[9] Donnelly, H. & Fefferman, C., Nodal sets of eigenfunctions on Riemannian manifolds, *Invent. Math.,* **93** (1988), 161-183.

[10] Duffie, D., Geoffard, P. Y. & Skiadas, C., Efficient and equilibrium allocations with stochastic differential utility, (Preprint).

[11] Karoui, N. E., Huu Nguyen, D. & Jeanblanc-Picqué, M., Compactification methods in the control of degenerate diffusions: existence of an optimal control, *Stochastics,* **20** (1987), 169-219.

[12] Evans. L. C. & Lenhart, S., The parabolic Bellman equation, *Nonlinear Anal. TMA,* **5**(1981), 765-773.

[13] Fleming, W. H., The Cauchy problem for a nonlinear first order differential equation, *J. Diff. Eqn.,* **5**(1969), 515-530.

[14] Fleming, W. H., Generalized solutions in optimal stochastic control, Differential games and control theory II, *Lect. Notes in Pure and Appl. Math.,* **30**, Dekker, 1977.

[15] Hardt, R. & Simon, L., Nodal sets for solutions of elliptic equations, *J. Diff. Geom.,* **30**(1989), 505-522.

[16] Ikeda, N. & Watanabe, S., Stochastic differential equations and diffusion processes, North-Holland, 1989.

[17] Krylov, N. V., Controlled diffusion processes, Springer-Verlag, 1980.

[18] Krylov, N. V., Nonlinear elliptic and parabolic equations of the second order, Reidel, Dordrecht, Holland, 1987.

[19] Ladyzenskaja, O. A., Solonnikov, V. A. & Ural'ceva, N. N., Linear and quasilinear equations of parabolic type, Amer. Math. Soc., Providence, R.I., 1968.

[20] Lin, F. H,, Nodal sets of solutions of elliptic and parabolic equations, *Comm. Pure Appl. Math.,* **44**(1991), 287-308.

[21] Lions, J. L., Exact controllability, stabilization and perturbations for distributed systems, *SIAM Rev.,* **30**(1988), 1-68.

[22] Protter, J. M. P. & Yong, J., Solving forward-backward stochastic differential equations explicitly — a four step scheme, *Probability Theory* & *Related Fields,* **98**(1994), 339-359.

[23] Pardoux, E. & Peng, H., Adapted solution of a backward stochastic differential equation, *Systems* & *Control Letters,* **14**(1990), 55-61.

[24] Pardoux, E. & Peng, S., Backward stochastic differential equations and quasilinear parabolic partial differential equations, Stochastic Partial Differential Equations and Their Applications, B. L. Rozovskii & R. S. Sowers eds, LNCIS, **176**, Springer, 1992.

[25] Peng, S., Adapted solution of backward stochastic equations and related partial differential equations (Preprint).

[26] Safonov, M. S., On the classical solution of nonlinear elliptic equations of second order, *Math. USSR Izv,* **33**(1989), 597-612.