# SOLVING FORWARD-BACKWARD STOCHASTIC DIFFERENTIAL EQUATIONS EXPLICITLY — A FOUR STEP SCHEME

by

Jin Ma and Philip Protter
Purdue University Purdue University

Jiongmin Yong
Fudan University

Technical Report #93-20

Department of Statistics
Purdue University

April 1993

◇

# Solving Forward-Backward Stochastic Differential Equations Explicitly — A Four Step Scheme

Jin Ma[1], Philip Protter[1,2] and Jiongmin Yong[3]

**Abstract.** In this paper we investigate the nature of the adapted solutions to a class of forward-backward stochastic differential equations (SDEs for short) in which the forward equation is non-degenerate. We prove that in this case the adapted solution can always be sought in an "ordinary" sense over an arbitrarily prescribed time duration, via a direct "Four Step Scheme". Using this scheme, we further prove that the backward components of the adapted solution are determined explicitly by the forward component via the solution of a certain quasi-linear parabolic PDE system. Moreover the uniqueness of the adapted solutions (over an arbitrary time duration), as well as the continuous dependence of the solutions on the parameters, can all be proved within this unified framework. Some special cases are studied separately. In particular, we derive a new form of the integral representation of the Clark-Haussmann-Ocone type for functionals (or functions) of diffusions, in which the conditional expectation is no longer needed.

**Keywords.** forward-backward stochastic differential equations, parabolic systems, integral representation, Haussmann-Ocone formula, Malliavin calculus.

[1] Department of Mathematics, Purdue University, West Lafayette, IN 47907-1395.

[2] Supported in part by NSF grant # DMS-9103454.

[3] IMA, University of Minnesota, Minneapolis, MN 55455; on leave from Department of Mathematics, Fudan University, Shanghai 200433, China. This author is partially supported by NSF of China and Fok Ying Tung Education Foundation.

## §1. Introduction.

Let $(\Omega, \mathcal{F}, P; \{\mathcal{F}_t\}_{t\geq 0})$ be a filtered probability space satisfying the *usual conditions*. Assume that a standard $d$-dimensional Brownian motion $\{W_t\}_{t\geq 0}$ is defined on this space. Consider the following forward-backward stochastic differential equations:

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

$$(1.2) \qquad Y_t = g(X_T) + \int_t^T \widehat{b}(s, X_s, Y_s, Z_s)ds + \int_t^T \widehat{\sigma}(s, X_s, Y_s, Z_s)dW_s, \ t \in [0, T],$$

where $(X, Y, Z)$ takes values in $\mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^{m\times d}$ and $b, \widehat{b}, \sigma, \widehat{\sigma}$ and $g$ are smooth functions with appropriate dimensions; $T > 0$ is an arbitrarily prescribed number which stands for 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 $[0, T]$, $P$-almost surely. Such an adapted solution, if it exists, will be called an *ordinary adapted solution* (here the term *ordinary* is inherited from our previous paper [6], in which the adapted solution can have a *relaxed* form). One should note that it is the extra process $Z$ that makes it possible for (1.1) and (1.2) to have an adapted solution (cf. [7,10,12]).

In [7] we studied the solvability of such forward-backward equations over an arbitrarily prescribed time duration $[0, T]$. We showed, by designing an appropriate relaxed stochastic control problem, that the solvability of the forward-backward SDEs (1.1) and (1.2) is equivalent to the non-emptyness of the nodal set of the viscosity solution to a certain Hamilton-Jacobi-Bellman equation. Using this new approach, we proved the solvability and non-solvability of a special class of forward-backward SDEs and we described exactly the nodal set of the corresponding HJB equation. We should note, however, that in general the adapted solution can only be found in a "wider" sense (cf. [7]). More precisely, the component $Z$ is replaced by an adapted measure-valued process and the probability space is subject to change when necessary. Also, we note that the uniqueness of the adapted solution over an arbitrary duration was not studied in [7] since it basically requires the uniqueness of the optimal relaxed control, which is far from obvious. Therefore, the natural questions are: *To what extent can one actually find an "ordinary adapted solution" over an arbitrarily prescribed time duration? Will such an adapted solution be unique?* Also, in

light of the result obtained in [7], we observe that sometimes the backward components $Y$ and $Z$ are determined completely by the forward component $X$ via the nodal surface. On the other hand, in a special case when the forward equation does not depend on the backward components, Pardoux and Peng [11] discovered recently that the components of the adapted solution $(X, Y, Z)$, whenever it exists, are explicitly related via the Malliavin derivatives; and the solution of the backward SDE is closely related to a class of quasilinear parabolic partial differential equations. Thus, one can hope to find an explicit solution (in some sense) for the strongly coupled forward-backward equation (1.1) and (1.2) via a certain quasilinear parabolic PDE system. This paper is devoted to answering these questions.

We will show that for a fairly large class of forward-backward SDEs in which the forward equation is *non-degenerate* (that is, the coefficient $\sigma$ is non-degenerate), there do exist explicit relations between $Y, Z$ and $X$ in terms of a classical solution of a certain parabolic PDE system; and when such relations hold we not only obtain the ordinary adapted solutions of the forward-backward SDEs, but we also find the explicit form of the solutions. We carry out this idea by designing a generic scheme (which we call the "Four Step Scheme" in the sequel) to construct explicitly the adapted solution for forward-backward SDEs. With this scheme we can prove the uniqueness of the adapted solution over an arbitrary interval, which is not obtainable by the contraction mapping theorem (see [1]) and which seems not possible by a pure control theoretic argument like that of [7]. The continuous dependence of the solution on the parameters is also proved within this framework. It is worth noting that solving the parabolic system, which presumably gives the nodal surface of the viscosity solution to the corresponding HJB equation (cf. [7]), is already sufficient for our scheme to work. That is, one does not have to verify whether or not it is really the nodal surface. Thus the technical difficulties are reduced in this special case. Finally, we would like to point out that the non-degeneracy of $\sigma$ is essential for the existence of an adapted solution over an arbitrary time interval $[0, T]$; in fact, Antonelli's counterexample in [1] shows that otherwise the adapted solution may not even exist when the time duration $T$ is large (see also [7] for other non-existence results).

This paper is organized as follows. In §2 we formulate the problem and give some preliminaries. In §3 we study the solvability of the two essential steps in our "Four Step

Scheme". In §4 we give our three main theorems; and in §5 we prove the continuous dependence and differentiability of the adapted solutions with respect to the parameters. In §6 we discuss the applications of our results to an integral representation theorem and compare it with the Clark-Haussmann-Ocone formula.

*Acknowledgement:* The authors wish to thank Fabio Antonelli for many discussions on this subject.

## §2. Formulations of the Problem.

In this paper we will only seek *ordinary adapted solutions* to the forward-backward equations, which we now describe.

Let $(\Omega, \mathcal{F}, P)$ be a probability space carrying a standard $d$-dimensional Brownian motion $W = \{W_t : t \geq 0\}$, and let $\{\mathcal{F}_t\}$ be the $\sigma$-field generated by $W$ (*i.e.*, $\mathcal{F}_t = \sigma\{W_s : 0 \leq s \leq t\}$). We make the usual $P$-augmentation to each $\mathcal{F}_t$ so that $\mathcal{F}_t$ contains all the $P$-null sets of $\mathcal{F}$. Then, $\{\mathcal{F}_t\}$ is right continuous and $\{\mathcal{F}_t\}$ satisfies the *usual hypotheses*. Let us consider the following forward-backward SDEs:

$$(2.1)\quad \begin{cases} X_t = x + \displaystyle\int_0^t b(s, X_s, Y_s, Z_s)ds + \int_0^t \sigma(s, X_s, Y_s)dW_s, \\ Y_t = g(X_T) + \displaystyle\int_t^T \widehat{b}(s, X_s, Y_s, Z_s)ds + \int_t^T \widehat{\sigma}(s, X_s, Y_s, Z_s)dW_s, \end{cases} \qquad t \in [0, T].$$

Here, the processes $X$, $Y$ and $Z$ take values in $\mathbb{R}^n$, $\mathbb{R}^m$ and $\mathbb{R}^{m\times d}$, respectively; and the functions $b$, $\widehat{b}$, $\sigma$, $\widehat{\sigma}$ and $g$ take values in $\mathbb{R}^n$, $\mathbb{R}^m$, $\mathbb{R}^{n\times d}$, $\mathbb{R}^{m\times d}$ and $\mathbb{R}^m$, respectively. In what follows, we use the usual Euclidean norms in $\mathbb{R}^m$ and $\mathbb{R}^n$; and for $z \in \mathbb{R}^{m\times d}$ (resp. $\mathbb{R}^{n\times d}$), we define $|z| = \{\mathrm{tr}\,(zz^T)\}^{1/2}$, where " $^T$ " means transpose. Then, $\mathbb{R}^{m\times d}$ (resp. $\mathbb{R}^{n\times d}$) is a Hilbert space.

**Definition 2.1.** A triple of processes $(X, Y, Z) : [0, T] \times \Omega \to \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^{m\times d}$ is called an ordinary adapted solution of the forward-backward SDEs (2.1), if it is $\{\mathcal{F}_t\}$-adapted and square-integrable, such that it satisfies (2.1) $P$-almost surely.

Since we are looking only for ordinary adapted solutions in this paper, the term "ordinary" will be omitted from now on. Moreover, the adaptedness of the solution enables

us to rewrite (see also [7]) (2.1) in a pure forward differential form:

$$
\begin{cases} dX_t = b(t, X_t, Y_t, Z_t)dt + \sigma(t, X_t, Y_t)dW_t, \\ dY_t = -\widehat{b}(t, X_t, Y_t, Z_t)dt - \widehat{\sigma}(t, X_t, Y_t, Z_t)dW_t, \qquad t \in [0, T], \\ X_0 = x, \qquad Y_T = g(X_T). \end{cases} \tag{2.2}
$$

It is clear that (2.2) is a stochastic two point boundary value problem.

Let us first give a heuristic argument. Suppose that $(X_t, Y_t, Z_t)$ is an adapted solution to (2.1) or equivalently (2.2). In light of the special case studied in [7], we assume that $Y$ and $X$ are related by

$$
Y_t = \theta(t, X_t), \qquad \forall t \in [0, T], \qquad \text{a.s. } P, \tag{2.3}
$$

where $\theta$ is some function to be determined (in [6], $\theta$ is the nodal surface of some HJB-equation). Suppose that all the functions involved are smooth, say at least $C^2$; then by applying Itô's formula, we have for $1 \leq k \leq m$:

$$
\begin{aligned} dY_t^k = d\theta^k(t, X_t) = \{&\theta_t^k(t, X_t) + \langle\, \theta_x^k(t, X_t), b(t, X_t, \theta(t, X_t), Z_t)\,\rangle \\ &+ \frac{1}{2}\mathrm{tr}\,[\theta_{xx}^k(t, X_t)\sigma(t, X_t, \theta(t, X_t))\sigma(t, X_t, \theta(t, X_t))^T]\}dt \\ &+ \langle\, \theta_x^k(t, X_t), \sigma(t, X_t, \theta(t, X_t))dW_t\,\rangle\,. \end{aligned} \tag{2.4}
$$

Comparing (2.4) and (2.2), we see that if $\theta$ is the right choice, then it is necessary that, for $k = 1, \cdots, m$,

$$
\begin{cases} -\widehat{b}^k(t, X_t, \theta(t, X_t)) = \theta_t^k(t, X_t) + \langle\, \theta_x^k(t, X_t), b(t, X_t, \theta(t, X_t), Z_t)\,\rangle \\ \qquad\qquad + \frac{1}{2}\mathrm{tr}\,[\theta_{xx}^k(t, X_t)\sigma(t, X_t, \theta(t, X_t))\sigma(t, X_t, \theta(t, X_t))^T]; \\ \theta(T, X_T) = g(X_T), \end{cases} \tag{2.5}
$$

and

$$
\theta_x(t, X_t)\sigma(t, X_t, \theta(t, X_t)) = -\widehat{\sigma}(t, X_t, Y_t, Z_t). \tag{2.6}
$$

The above arguments suggest that we design the following "Four Step Scheme" to solve the forward-backward SDE (2.1).

**Four Step Scheme:**

*Step 1.* Find a "smooth" (see Remark 2.1) mapping $z : [0,T] \times \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^{m\times n} \to \mathbb{R}^{m\times d}$ satisfying

$$\begin{aligned} p\sigma(t,x,y)+\widehat{\sigma}(t,x,y,z(t,x,y,p)) = 0, \\ \forall (t,x,y,p) \in [0,T] \times \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^{m\times n}. \end{aligned} \tag{2.7}$$

*Step 2.* Using the function $z$ above, solve the following parabolic system for $\theta(t,x)$:

$$\begin{cases} \theta_t^k + \dfrac{1}{2}\mathrm{tr}\,(\theta_{xx}^k \sigma(t,x,\theta)\sigma(t,x,\theta)^T) + \langle\, b(t,x,\theta,z(t,x,\theta,\theta_x)), \theta_x^k \,\rangle \\ \qquad + \widehat{b}^k(t,x,\theta,z(t,x,\theta,\theta_x)) = 0, \qquad k = 1,\cdots,m, \quad (t,x) \in (0,T)\times\mathbb{R}^n, \\ \theta(T,x) = g(x), \qquad x \in \mathbb{R}^n. \end{cases} \tag{2.8}$$

*Step 3.* Using $\theta$ and $z$, solve the following forward SDE:

$$X_t = x + \int_0^t \tilde{b}(s,X_s)ds + \int_0^t \tilde{\sigma}(s,X_s)dW_s, \tag{2.9}$$

where $\tilde{b}(t,x) = b(t,x,\theta(t,x),z(t,x,\theta(t,x),\theta_x(t,x)))$ and $\tilde{\sigma}(t,x) = \sigma(t,x,\theta(t,x))$.

*Step 4.* Set

$$\begin{cases} Y_t = \theta(t,X_t), \\ Z_t = z(t,X_t,\theta(t,X_t),\theta_x(t,X_t)). \end{cases} \tag{2.10}$$

Then if this scheme is realizable, $(X_t,Y_t,Z_t)$ would give an adapted solution of (2.1).

The above Four Step Scheme provides a generic method which of course can be applied to any forward-backward equation (*e.g.*, to those systems in which $\sigma$ depends on $z$ and $z$ can take value in any Euclidean space $\mathbb{R}^\ell$). However, in order to ensure that every step goes through, some restrictions on the data are inevitable. For instance in order for the parabolic system (2.8) to have a classical solution, and for the equation (2.7) to be solvable, we should have at least two reasonable assumptions: (1) the uniform parabolicity of (2.8); (2) the surjectivity of the mapping $\widehat{\sigma}$. We now give the standing assumptions of this paper.

*Standing Assumptions:*

(A1) $d = n$; and the functions $b$, $\widehat{b}$, $\sigma$, $\widehat{\sigma}$ and $g$ are smooth functions taking values in $\mathbb{R}^n$, $\mathbb{R}^m$, $\mathbb{R}^{n\times n}$, $\mathbb{R}^{m\times n}$ and $\mathbb{R}^m$, respectively, and with first order derivatives in $x,y,z$ being bounded by some constant $L > 0$.

(A2) The function $\sigma$ satisfies

$$\sigma(t,x,y)\sigma(t,x,y)^T \geq \nu(|y|)I, \qquad \forall (t,x,y) \in [0,T]\times \mathbb{R}^n \times \mathbb{R}^m, \tag{2.11}$$

for some positive continuous function $\nu(\cdot)$.

(A3) For each fixed $(t,x,y,z) \in [0,T]\times\mathbb{R}^n\times\mathbb{R}^m\times\mathbb{R}^{m\times n}$, the linear map $\widehat{\sigma}_z(t,x,y,z) \in \mathcal{L}(\mathbb{R}^{m\times n})$ (the space of all linear transforms on $\mathbb{R}^{m\times n}$) is invertible with the inverse $\widehat{\sigma}_z(t,x,y,z)^{-1}$ satisfying

$$\begin{aligned} &\|\widehat{\sigma}_z(t,x,y,z)^{-1}\|_{\mathcal{L}(\mathbb{R}^{m\times n})} \leq \lambda(|y|), \\ &\qquad\qquad (t,x,y,z) \in [0,T]\times\mathbb{R}^n\times\mathbb{R}^{m\times n}\times\mathbb{R}^{m\times n}, \end{aligned} \tag{2.12}$$

for some continuous function $\lambda(\cdot)$. Moreover, for any $(t,x,y) \in [0,T]\times\mathbb{R}^n\times\mathbb{R}^m$,

$$\{\widehat{\sigma}(t,x,y,z) \mid z \in \mathbb{R}^{m\times n}\} = \mathbb{R}^{m\times n}; \tag{2.13}$$

and there exists a positive continuous function $\kappa(\cdot)$, such that

$$\sup\{|z| \mid \widehat{\sigma}(t,x,y,z) = 0\} \leq \kappa(|y|), \qquad \forall (t,x,y) \in [0,T]\times\mathbb{R}^n\times\mathbb{R}^m. \tag{2.14}$$

(A4) There exists a function $\mu$ and constants $C > 0$ and $\alpha \in (0,1)$, such that $g$ is bounded in $C^{2+\alpha}(\mathbb{R}^m)$ and for all $(t,x,y,z) \in [0,T]\times\mathbb{R}^n\times\mathbb{R}^m\times\mathbb{R}^{n\times m}$,

$$|\sigma(t,x,y)| \leq \mu(|y|), \tag{2.15}$$

$$|b(t,x,y,0)| \leq \mu(|y|), \tag{2.16}$$

$$|\widehat{b}(t,x,0,z)| \leq C. \tag{2.17}$$

*Remark 2.1.* Throughout this paper, by "smooth" we mean that the involved functions possess partial derivatives of all necessary orders. We prefer not to indicate the exact order of smoothness for the sake of simplicity of presentation. Also, the boundedness of the first order derivatives in $x, y, z$ requires only the usual uniform Lipschitz condition in

these variables, which is close to necessary in order to have global well-posedness for any differential equations.

## §3. Solvability of (2.7) and (2.8).

It is readily seen that among all steps in our Four Step Scheme, the first two (*i.e.*, the solvability of (2.7) and (2.8)) are essential. Thus we devote this section to these two steps, which can also be viewed as the preliminaries of our main theorems in the following section. The first proposition concerns the solvability of (2.7).

**Proposition 3.1.** *Suppose that (A1), (A2) and (2.12) hold. Then (2.7) admits a unique smooth solution* $z : [0,T] \times \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^{m\times n}$ *if and only if (2.13) holds.*

*In particular, (2.7) is solvable if the following holds:*

$$\lim_{|z|\to\infty} \frac{\mathrm{tr}\,(\widehat{\sigma}(t,x,y,z)z^T)}{|z|} = +\infty, \qquad \forall (t,x,y) \in [0,T] \times \mathbb{R}^n \times \mathbb{R}^m. \tag{3.1}$$

*In addition, if (2.14) holds, then, the solution* $z(t,x,y,p)$ *of (2.7) satisfies*

$$\begin{aligned} |z(t,x,y,p)| \leq & \kappa(|y|) + \lambda(|y|)|\sigma(t,x,y)|\,|p|, \\ & \forall (t,x,y,p) \in [0,T] \times \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^{m\times n}. \end{aligned} \tag{3.2}$$

*Proof.* Recall that a solution of (2.7) is a mapping $z : [0,T] \times \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^{m\times n}$ satisfying

$$\begin{aligned} & p\sigma(t,x,y) + \widehat{\sigma}(t,x,y,z(t,x,y,p)) = 0, \\ & \qquad \forall (t,x,y,p) \in [0,T] \times \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^{m\times n}. \end{aligned} \tag{3.3}$$

Since all the functions involved are smooth and $\widehat{\sigma}_z(t,x,y,z)$ is invertible as an element of $\mathcal{L}(\mathbb{R}^{m\times n})$, whenever a solution $z(t,x,y,p)$ of (3.3) exists, it must be smooth. Moreover, such a solution is unique due to (2.12). Indeed, suppose for some $(t,x,y)$, there exist $z_1, z_2 \in \mathbb{R}^{m\times n}$ with $z_1 \neq z_2$, such that $\widehat{\sigma}(t,x,y,z_1) = \widehat{\sigma}(t,x,y,z_2)$. Let

$$\varphi(r) = \widehat{\sigma}(t,x,y,rz_1 + (1-r)z_2), \qquad r \in [0,1].$$

Since $\varphi(0) = \varphi(1)$, there exists some $r \in (0,1)$ such that $\varphi'(r) = 0$. In other words,

$$\langle\, \widehat{\sigma}_z(t,x,y,rz_1 + (1-r)z_2), z_1 - z_2 \,\rangle = 0,$$

contradicting (2.12). It is evident that if (2.13) holds, then such a function $z$ will exist. Conversely, because of (2.11), for any fixed $(t,x,y)$, the range of the matrix function $p \mapsto p\sigma(t,x,y)$ is all of $\mathbb{R}^{m\times n}$. Thus, (2.13) has to hold if (3.3) has a solution $z(t,x,y,p)$. Thus we proved the first part of the proposition.

Now noting that $\mathbb{R}^{m\times n}$ under the norm $|z| = (\mathrm{tr}(zz^T))^{1/2}$ is isometric to $\mathbb{R}^{mn}$, the condition (3.1) implies that the map $z \mapsto \widehat{\sigma}(t,x,y,z)$ is surjective (cf. [2, Theorem I.3.3]) for each $(t,x,y)$. This gives (2.13) and hence (2.7) is solvable.

Finally, it follows immediately from (3.3) and (2.12) that, for $(t,x,y,p) \in [0,T] \times \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^{m\times n}$,

$$\left\| \frac{\partial z}{\partial p}(t,x,y,p) \right\|_{\mathcal{L}(\mathbb{R}^{n\times m})} \leq \lambda(|y|)|\sigma(t,x,y)|, \tag{3.4}$$

whence (3.2) follows from (2.14) and (3.4). □

We now turn to the solvability of (2.8). Resolving this step relies heavily on the theory of parabolic systems. Our main references are [5] and [14]. Let us first try to apply the result of [5]. Consider the following initial boundary value problem:

$$\begin{cases} \theta_t^k + \sum_{i,j=1}^n a_{ij}(t,x,\theta)\theta_{x_ix_j} + \sum_{i=1}^n b_i(t,x,\theta,z(t,x,\theta,\theta_x))\theta_{x_i}^k \\ \qquad\qquad + \widehat{b}^k(t,x,\theta,z(t,x,\theta,\theta_x)) = 0, \\ \qquad\qquad 1 \leq k \leq m, (t,x) \in [0,T] \times B_R, \\ \theta\,|_{\partial B_R} = g(x), \qquad |x| = R, \\ \theta(T,x) = g(x), \qquad x \in B_R, \end{cases} \tag{3.5}$$

where $B_R$ is the ball centered at the origin with raduis $R > 0$ and

$$\begin{cases} (a_{ij}(t,x,y)) = \dfrac{1}{2}\sigma(t,x,y)\sigma(t,x,y)^T, \\ (b_1(t,x,y,z),\cdots,b_n(t,x,y,z))^T = b(t,x,y,z), \\ (\widehat{b}^1(t,x,y,z),\cdots,\widehat{b}^m(t,x,y,z))^T = \widehat{b}(t,x,y,z). \end{cases}$$

Suppose (A1)–(A3) hold, then by Proposition 3.1, the solution $z(t,x,y,p)$ of (2.7) exists and is smooth. We now give a lemma, which is an analogue of [5, Ch. VII, Theorem 7.1].

**Lemma 3.2.** *Suppose that all the functions $a_{ij}$, $b_i$, $\widehat{b}^k$ and $g$ are smooth. Suppose also that for all $(t,x,y) \in [0,T] \times \mathbb{R}^n \times \mathbb{R}^m$ and $p \in \mathbb{R}^{m\times n}$, it holds that*

$$\nu(|y|)I \leq \big(a_{ij}(t,x,y)\big) \leq \mu(|y|)I, \tag{3.6}$$

$$|b(t,x,y,z(t,x,y,p))| \leq \mu(|y|)(1+|p|), \tag{3.7}$$

$$\left|\frac{\partial}{\partial x_\ell} a_{ij}(t,x,y)\right| + \left|\frac{\partial}{\partial y^k} a_{ij}(t,x,y)\right| \leq \mu(|y|), \tag{3.8}$$

*for some continuous functions $\mu(\cdot)$ and $\nu(\cdot)$, with $\nu(r) > 0$;*

$$|\widehat{b}(t,x,y,z(t,x,y,p))| \leq [\varepsilon(|y|) + P(|p|,|y|)](1+|p|^2), \tag{3.9}$$

*where $P(|p|,|y|) \to 0$, as $|p| \to \infty$ and $\varepsilon(|y|)$ is small enough;*

$$\sum_{k=1}^{m} \widehat{b}^k(t,x,y,z(t,x,y,p))y^k \leq L(1+|y|^2), \tag{3.10}$$

*for some constant $L > 0$. Finally, suppose that $g$ is bounded in $C^{2+\alpha}(\mathbb{R}^n)$ for some $\alpha \in (0,1)$. Then (3.5) admits a unique classical solution $\theta(t,x)$.* □

Using Lemma 3.2, we can now prove the solvability of (2.8) under our standing assumptions.

**Proposition 3.3.** *Let (A1)–(A4) hold. Then (2.8) admits a unique classical solution $\theta(t,x)$ which is bounded and $\theta_t(t,x)$, $\theta_x(t,x)$ and $\theta_{xx}(t,x)$ are bounded as well.*

*Proof.* We need check only that all the required conditions in Lemma 3.2 are satisfied. First, by Proposition 2.3 we know that there exists a smooth function $z(t,x,\theta,p)$ satisfying (3.3) and (3.2). By (2.15), we further have

$$\begin{aligned} |z(t,x,y,p)| &\leq \kappa(|y|) + \lambda(|y|)\mu(|y|)|p|, \\ &\forall (t,x,y,p) \in [0,T] \times \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^{m\times n}. \end{aligned} \tag{3.11}$$

Now, we see that (3.6) and (3.8) follow from (A1), (2.11) and (2.15); (3.7) follows from (A1), (2.16) and (3.11); and (3.9)–(3.10) follow from (A1) and (2.17). Therefore by Lemma 3.2

there exists a unique bounded solution $\theta(t,x;R)$ of (3.5) for which $\theta_t(t,x;R)$, $\theta_x(t,x;R)$ and $\theta_{xx}(t,x;R)$ together with $\theta(t,x)$ are bounded uniformly in $R>0$. Using a diagonalization argument one further shows that there exists a subsequence $\theta(t,x,R)$ which converges to $\theta(t,x)$ as $R\to\infty$, and $\theta(t,x)$ is a classical solution of (2.8), with $\theta_t(t,x)$, $\theta_x(t,x)$ and $\theta_{xx}(t,x)$ all being bounded.

Finally, noting that all the functions together with the possible solutions are smooth with required bounded partial derivatives, the uniqueness follows from a standard argument using Gronwall's inequality. □

*Remark 3.4.* Note that the solution $z(t,x,y,p)$ of (2.7) is not bounded in general. Thus, (3.10) almost implies that $\widehat{b}(t,x,y,z)$ is bounded for fixed $(t,x,y)$ uniformly in $z$. This leads to our assumption (2.17) in the present framework. This assumption could be relaxed if we had some more information about $\widehat{b}(t,x,y,z)$ and the function $z(t,x,y,p)$.

## §4. Main Theorems.

In this section we state and prove our main theorems concerning the existence and uniqueness of the (ordinary) adapted solution to the forward-backward SDEs (2.1). By slightly changing the conditions on the data, we can derive different forms of the results. We shall therefore consider three cases.

*1. The General Case:*

**Theorem 4.1.** *Let (A1)–(A4) hold. Then the forward-backward SDE (2.1) admits a unique adapted solution $(X,Y,Z)$ which has the expression (2.10) with $z(t,x,y,p)$ and $\theta(t,x)$ being the solutions of (2.7) and (2.8).*

*Proof.* By Proposition 3.1 we know that there exists a unique smooth function $z(t,x,y,p)$ satisfying (3.3). Next by Proposition 3.3, one can find a classical solution $\theta(t,x)$ of the uniform parabolic system (2.8). Now we consider the forward SDE (2.9). Since under our conditions both $\widetilde{b}(t,x)$ and $\widetilde{\sigma}(t,x)$ are uniformly Lipschitz continuous in $x$, we see that for any $x\in\mathbb{R}^n$, (2.9) has a unique strong solution. Then, by defining $Y_t$ and $Z_t$ via (2.10) and applying Itô's formula, we can easily check that (2.4) is satisfied. Hence, $(X,Y,Z)$ is a solution of (2.1).

It remains to show the uniqueness. First, we claim that any solution $(X, Y, Z)$ of (2.1) must be of the form we constructed using the Four Step Scheme. Let $(X, Y, Z)$ be any solution of (2.1). We define

$$\widetilde{Y}_t = \theta(t, X_t), \quad \widetilde{Z}_t = z(t, X_t, \theta(t, X_t), \theta_x(t, X_t)). \tag{4.1}$$

By applying Itô's formula to $\theta(t, X_t)$ and noting the definition of $z(t, x, y, p)$ and $\theta(t, x)$, one shows as in §2 that $(\widetilde{Y}, \widetilde{Z})$ is an adapted solution of the following (pure) backward SDE:

$$Y_t = g(X_T) + \int_t^T \widehat{b}(s, X_s, Y_s, Z_s)ds + \int_t^T \widehat{\sigma}(s, X_s, Y_s, Z_s)dW_s, \qquad t \in [0, T]. \tag{4.2}$$

Therefore both $(Y, Z)$ and $(\widetilde{Y}, \widetilde{Z})$ satisfy (4.2), with the same $X_t$. We now prove that it is necessary that $Y_t = \widetilde{Y}_t$ and $Z_t = \widetilde{Z}_t$. In fact, applying Itô's formula and taking expectations, we have

$$\begin{aligned} E|Y_t - \widetilde{Y}_t|^2 &= 2\int_t^T E\left\langle Y_s - \widetilde{Y}_s, \widehat{b}(s, X_s, Y_s, Z_s) - \widehat{b}(s, X_s, \widetilde{Y}_s, \widetilde{Z}_s)\right\rangle ds \\ &\quad - \int_t^T E|\widehat{\sigma}(s, X_s, Y_s, Z_s) - \widehat{\sigma}(s, X_s, \widetilde{Y}_s, \widetilde{Z}_s)|^2 ds \end{aligned} \tag{4.3}$$

By (A3) we see that for any fixed $(t, x, y)$ (we suppress these three arguments below)

$$\begin{aligned} |z - \widetilde{z}| &= |\widehat{\sigma}^{-1}\widehat{\sigma}(z) - \widehat{\sigma}^{-1}\widehat{\sigma}(z)| \\ &= \left|\int_0^1 \widehat{\sigma}_z^{-1}(\widehat{\sigma}(\widetilde{z}) + \tau(\widehat{\sigma}(z) - \widehat{\sigma}(\widetilde{z})))d\tau(\widehat{\sigma}(z) - \widehat{\sigma}(\widetilde{z}))\right| \\ &\le \lambda(|y|)|\widehat{\sigma}(z) - \widehat{\sigma}(\widetilde{z})|, \qquad \forall z, \widetilde{z} \in \mathbb{R}^{n\times m}. \end{aligned} \tag{4.4}$$

Thus, we have

$$\begin{aligned} &2\int_t^T E\left\langle Y_s - \widetilde{Y}_s, \widehat{b}(s, X_s, Y_s, Z_s) - \widehat{b}(s, X_s, \widetilde{Y}_s, \widetilde{Z}_s)\right\rangle ds \\ &\le 2L\int_t^T E[|Y_s - \widetilde{Y}_s|(|Y_s - \widetilde{Y}_s| + |Z_s - \widetilde{Z}_s|)]ds \\ &\le 2L\int_t^T \{E|Y_s - \widetilde{Y}_s|^2 + E[|Y_s - \widetilde{Y}_s|\lambda(|\widetilde{Y}_s|)|\widehat{\sigma}(s, X_s, \widetilde{Y}_s, Z_s) - \widehat{\sigma}(s, X_s, \widetilde{Y}_s, \widetilde{Z}_s)|]\}ds \\ &\le \int_t^T \{CE|Y_s - \widetilde{Y}_s|^2 + CE[|Y_s - \widetilde{Y}_s||\widehat{\sigma}(s, X_s, Y_s, Z_s) - \widehat{\sigma}(s, X_s, \widetilde{Y}_s, \widetilde{Z}_s)|]\}ds \\ &\le \int_t^T \{CE|Y_s - \widetilde{Y}_s|^2 + \frac{1}{2}E|\widehat{\sigma}(s, X_s, Y_s, Z_s) - \widehat{\sigma}(s, X_s, \widetilde{Y}_s, \widetilde{Z}_s)|^2\}ds. \end{aligned} \tag{4.5}$$

Here, we have used the fact that $\widetilde{Y}_t = \theta(t, X_t)$ is unformly bounded. Then, (4.3) yields

$$
\begin{aligned}
&E|Y_t - \widetilde{Y}_t|^2 + \int_t^T E|\widehat{\sigma}(s, X_s, Y_s, Z_s) - \widehat{\sigma}(s, X_s, \widetilde{Y}_s, \widetilde{z}_s)|^2 ds \\
&\leq C \int_t^T E|Y_s - \widetilde{Y}_s|^2 ds.
\end{aligned} \tag{4.6}
$$

By using Gronwall's inequality, we conclude that

$$
Y_t = \widetilde{Y}_t, \qquad \widehat{\sigma}(t, X_t, Y_t, Z_t) = \widehat{\sigma}(t, X_t, \widetilde{Y}_t, \widetilde{Z}_t), \qquad \text{a.s. } P. \tag{4.7}
$$

An argument similar to that in Proposition 3.1 shows that $Z_t = \widetilde{Z}_t$, a.s. $P.$; Thus any solution of (2.1) must have the form that we have constructed, proving our claim.

Finally, let $(X, Y, Z)$ and $(\widetilde{X}, \widetilde{Y}, \widetilde{Z})$ be any two solutions of (2.1). By the previous argument we have

$$
\begin{aligned}
&Y_t = \theta(t, X_t), \quad Z_t = z(t, X_t, \theta(t, X_t), \theta_x(t, X_t)), \\
&\widetilde{Y}_t = \theta(t, \widetilde{X}_t), \quad \widetilde{Z}_t = z(t, \widetilde{X}_t, \theta(t, \widetilde{X}_t), \theta_x(t, \widetilde{X}_t)).
\end{aligned} \tag{4.8}
$$

Hence $X_t$ and $\widetilde{X}_t$ satisfy exactly the same forward SDE (2.9) with the same initial state $x$. Thus we must have $X_t = \widetilde{X}_t$, $\forall t \in [0, T]$, a.s. $P.$, which in turn shows that $Y_t = \widetilde{Y}_t$, $Z_t = \widetilde{Z}_t$, $\forall t \in [0, T]$, a.s. $P.$ by (4.8). The proof is now complete. □

*2. Special Case I: $\widehat{b}$ has linear growth in $z$.*

Although Theorem 4.1 gives a general solvability result of the forward-backward SDE (2.1), the condition (A4) is rather restrictive; for instance, the case that the coefficient $\widehat{b}(t, x, y, z)$ is linearly growing in $z$ is excluded. This case, however, is very important for applications in optimal stochastic control theory. For example in the Pontryagin maximum principle for optimal stochastic control, the adjoint equation is of the form that the corresponding $\widehat{b}$ is affine in $z$. Thus we would like to discuss this case separately.

In order to relax the condition (A4), we compensate by considering the following special forward-backward SDE:

$$
\begin{cases}
X_t = x + \displaystyle\int_0^t b(s, X_s, Y_s, Z_s) ds + \int_0^t \sigma(s, X_s) dW_s, \\
Y_t = g(X_T) + \displaystyle\int_t^T \widehat{b}(s, X_s, Y_s, Z_s) ds + \int_t^T \widehat{\sigma}(s, X_s, Y_s, Z_s) dW_s,
\end{cases} \qquad t \in [0, T]. \tag{4.9}
$$

We assume that $\sigma$ is independent of $Y_t$ and $Z_t$, but we allow $\widehat{b}$ to have linear growth in $z$. In this case, the parabolic system looks like the following (compare with (2.8)):

$$
\begin{cases}
\theta_t^k + \dfrac{1}{2}\mathrm{tr}\,(\theta_{xx}^k \sigma(t,x)\sigma(t,x)^T) + \langle\, b(t,x,\theta,z(t,x,\theta,\theta_x)), \theta_x^k \,\rangle \\
\qquad\qquad + \widehat{b}^k(t,x,\theta,z(t,x,\theta,\theta_x)) = 0, \\
\qquad\qquad\qquad 1 \le k \le m, (t,x) \in [0,T] \times \mathbb{R}^n, \\
\theta(T,x) = g(x), \qquad x \in \mathbb{R}^n.
\end{cases} \tag{4.10}
$$

Since now $\widehat{b}$ has linear growth in $z$, the result of [5] does not apply. We use the result of [14] instead. To this end, let us rewrite the above parabolic system in divergence form:

$$
\begin{cases}
\theta_t^k + \displaystyle\sum_{i,j=1}^n (a_{ij}(t,x)\theta_{x_i})_{x_j} = f^k(t,x,\theta,\theta_x), \qquad (t,x) \in [0,T] \times \mathbb{R}^m, \\
\qquad\qquad\qquad\qquad\qquad\qquad k = 1, \cdots, m, \\
\theta(T,x) = g(x), \qquad x \in \mathbb{R}^n,
\end{cases} \tag{4.11}
$$

where

$$
\begin{cases}
\big(a_{ij}(t,x)\big) = \dfrac{1}{2}\sigma(t,x)\sigma(t,x)^T, \\
f^k(t,x,y,p) = \displaystyle\sum_{i,j=1}^n a_{ijx_j}(t,x)p_i^k - \sum_{i=1}^n b_i(t,x,y,z(t,x,y,p))p_i^k \\
\qquad\qquad - \widehat{b}^k(t,x,y,z(t,x,y,p)).
\end{cases} \tag{4.12}
$$

From [14], we know that for any $T > 0$, (4.11) will have a unique classical solution, global in time, provided the following conditions hold:

$$
\nu I \le \big(a_{ij}(t,x)\big) \le \mu I, \qquad \forall (t,x) \in [0,T] \times \mathbb{R}^n, \tag{4.13}
$$

$$
\sum_{k=1}^m y^k f^k(t,x,y,p) \le \varepsilon_0 |p|^2 + C(1+|y|^2), \qquad \forall (t,x,y,p) \in [0,T] \times \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^{n \times m}, \tag{4.14}
$$

where $\nu, \mu, C, \varepsilon_0$ are constants with $\varepsilon_0$ being small enough. (To fit the framework of [10], we have taken $H = |y|^2$, $c^k \equiv 0$ and $r^k \equiv 0$, $k = 1, \cdots, m$. See [14] for details). Therefore, we need the following assumption:

(A5) There exist positive constants $\nu, \mu$, such that

$$\nu I \leq \sigma(t,x)\sigma(t,x)^T \leq \mu I, \qquad \forall (t,x) \in [0,T] \times \mathbb{R}^n, \tag{4.15}$$

$$|b(t,x,0,0)|,\ |\widehat{b}(t,x,0,0)| \leq \mu, \qquad \forall (t,x,) \in [0,T] \times \mathbb{R}^n, \tag{4.16}$$

Moreover, (A2)–(A3) hold with $\lambda$ being independent of $y$ and $\kappa(|y|)$ having no more than linear growth in $|y|$.

**Theorem 4.2.** *Suppose that (A1) and (A5) hold. Then (4.8) admits a unique adapted solution* $(X,Y,Z)$.

*Proof.* In the present case, similar to Proposition 3.1, there exists a unique smooth function $z(t,x,y,p)$ satisfying

$$\begin{aligned} &p\sigma(t,x)+\widehat{\sigma}(t,x,y,z(t,x,y,p)) = 0, \\ &\qquad \forall (t,x,y,p) \in [0,T] \times \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^{m\times n}. \end{aligned} \tag{4.17}$$

Moreover, we have the estimate (see (3.2))

$$\begin{aligned} &|z(t,x,y,p)| \leq C(1+|y|+|p|), \\ &\qquad \forall (t,x,y,p) \in [0,T] \times \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^{n\times m}. \end{aligned} \tag{4.18}$$

Therefore conditions (4.13) and (4.14) hold, which will lead to the existence and uniqueness of classical solutions of (4.11) or (4.10). Next, applying an argument similar to the one we used in the proof of Theorem 4.1, we can show that there exists a unique adapted solution $(X,Y,Z)$ of (4.8). □

*Remark 4.3.* It is not hard to see that since $\widehat{b}(t,x,y,z)$ is uniformly Lipschitz continuous in $(y,z)$ (see (A1)), (4.16) gives that

$$\begin{aligned} &|\widehat{b}(t,x,y,z)| \leq C(1+|y|+|p|), \\ &\qquad \forall (t,x,y,z) \in [0,T] \times \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^{m\times n}. \end{aligned} \tag{4.19}$$

In other words, the function $\widehat{b}$ is allowed to have a linear growth in $(y,z)$.

*3. Special Case II:* $m = 1$.

Unlike the special case I, this is the case in which the existence and uniqueness result can be derived for a more general system than (2.1). The main reason is that in this case, the function $\theta(t,x)$ is scalar valued, and the theory of quasilinear parabolic equations is much more complete than that for parabolic systems. Consequently, the corresponding results for the forward-backward SDEs will allow more complicated nonlinearities. Remember that in the present case, the backward component is one dimensional, but the forward part is still $n$ dimensional. This is exactly the case when we discuss the integral representation formula in the next section.

We can now consider more general forward-backward SDEs:

$$\text{(4.20)}\quad \begin{cases} X_t = x + \int_0^t b(s, X_s, Y_s, Z_s)ds + \int_0^t \sigma(s, X_s, Y_s, Z_s)dW_s, \\ Y_t = g(X_T) + \int_t^T \widehat{b}(s, X_s, Y_s, Z_s)ds + \int_t^T \widehat{\sigma}(s, X_s, Y_s, Z_s)^T dW_s, \end{cases} \qquad t \in [0,T].$$

Here $W$ is an $n$-dimensional standard Brownian motion, $b$, $\sigma$, $\widehat{b}$, $\widehat{\sigma}$ and $g$ take values in $\mathbb{R}^n$, $\mathbb{R}^{n\times n}$, $\mathbb{R}$, $\mathbb{R}^n$ and $\mathbb{R}$, respectively. We use the convention that all vectors are column vectors (thus we have $\widehat{\sigma}(t, X_t, Y_t, Z_t)^T$ in the above). We note that (4.20) is of the same form as (1.1)–(1.2); and the difference between (4.20) and (2.1) is that we now allow $\sigma$ to depend on $Z$. In what follows we will try to use our Four Step Scheme given in §2 to solve (4.20). To this end, we first need to solve the following equation for $z$:

$$\text{(4.21)}\qquad \sigma(t,x,y,z)^T p + \widehat{\sigma}(t,x,y,z) = 0.$$

Let us make the following assumptions. We keep the assumptions (A1), (A3) with $m = 1$ and replace (A2) by the following:

(A2)' There exist positive continuous functions $\mu(\cdot)$, $\nu(\cdot)$, such that

$$\text{(4.22)}\qquad \begin{aligned} \nu(|y|)I \le & \sigma(t,x,y,z)\sigma(t,x,y,z)^T \le \mu(|y|)I, \\ & \forall (t,x,y,z) \in [0,t]\times \mathbb{R}^n \times \mathbb{R} \times \mathbb{R}^n. \end{aligned}$$

We have the following result.

**Proposition 4.4.** *Let (A1), (A2)' and (A3) hold. Then there exists a unique smooth function $z(t,x,y,p)$ that solves (4.21) and satisfies (3.2).*

The proof is the same as that of Proposition 3.1.

**Theorem 4.5.** *Let (A1), (A2)' and (A3) hold. Let the function $g$ be bounded in $C^{2+\alpha}(\mathbb{R}^n)$ for some $\alpha \in (0,1)$ and suppose that there exists a constant $C > 0$ such that for all $(t,x) \in [0,T] \times \mathbb{R}^n$,*

$$|b(t,x,0,0)| + |\widehat{b}(t,x,0,0)| + |\sigma(t,x,0,0)| + |\widehat{\sigma}(t,x,0,0)| \leq C. \tag{4.23}$$

*Then, there exists a unique adapted solution $(X,Y,Z)$ of (4.20).*

The proof is omitted here because it is similar to that of Theorem 4.1. We should note that the well-posedness of (2.8) in the present case ($m = 1$) follows from [5, Ch.V, Theorem 8.1]. We see that the condition (4.23) together with (A1) means that all the functions $b$, $\widehat{b}$, $\sigma$ and $\widehat{\sigma}$ are allowed to have linear growth in $y$ and $z$.

## §5. Dependence of Solutions on Parameters.

In this section we study the dependence of the solution of forward-backward SDEs (2.1) on parameters, which will be useful in applications. We show that all the tasks can be accomplished under a unified framework—the Four Step Scheme. Consider forward-backward SDE with a parameter:

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

where $\alpha$ is a parameter taking value in a metric space, say $\Lambda$. The solution of (5.1), whenever it exists, will be denoted by $(X(\alpha),Y(\alpha),Z(\alpha))$. Let us make the following basic assumption:

(A6) The functions $b(\alpha,\cdot,\cdot,\cdot,\cdot)$, $\widehat{b}(\alpha,\cdot,\cdot,\cdot,\cdot)$, $\sigma(\alpha,\cdot,\cdot,\cdot)$, $\widehat{\sigma}(\alpha,\cdot,\cdot,\cdot,\cdot)$ and $g(\alpha,\cdot)$ satisfy the standing assumptions (A1)–(A4), uniformly in $\alpha \in \Lambda$.

It is readily seen that under the assumption (A6), for each $\alpha \in \Lambda$ and any $T > 0$, the adapted solution $(X(\alpha),Y(\alpha),Z(\alpha))$ of (5.1) exists and is unique on $[0,T]$. Furthermore, let us introduce the notion of $\mathcal{L}$- (resp. $\mathcal{L}(B)$-) continuity and differentiability of N. V. Krylov (cf. [4]).

**Definition 5.1.** An $\mathbb{R}^\ell$-valued process $\{\xi_t(\alpha) : t \geq 0\}$ with parameter $\alpha \in \Lambda$ is called $\mathcal{L}$- (resp. $\mathcal{L}(B)$-) continuous at $\alpha_0 \in \Lambda$ if

$$\lim_{\alpha \to \alpha_0} E\left\{\int_0^T |(\xi_t(\alpha) - \xi_t(\alpha_0)|^2 dt\right\} = 0. \tag{5.2}$$

(resp.

$$\lim_{\alpha \to \alpha_0} E\left\{\sup_{0 \leq t \leq T} |\xi_t(\alpha) - \xi_t(\alpha_0)|^2\right\} = 0.) \tag{5.3}$$

$\xi(\alpha)$ is called $\mathcal{L}$- (resp. $\mathcal{L}(B)$-) differentiable at $\alpha_0$ if for all $l \in \mathbb{R}^\ell$, the limit

$$\begin{aligned}&(\mathcal{L})\text{-}\lim_{r \to 0} \frac{1}{r}[\xi(\alpha_0 + rl) - \xi(\alpha_0)] \\ &\left(\text{resp. } (\mathcal{L}(B))\text{-}\lim_{r \to 0} \frac{1}{r}[\xi(\alpha_0 + rl) - \xi(\alpha_0)]\right)\end{aligned} \tag{5.4}$$

exists. In this case, the limit process is called the $\mathcal{L}$- (resp. $\mathcal{L}(B)$-) derivative of $\xi(\alpha)$.

Our main result of this section is the following.

**Theorem 5.2.** *Suppose that (A6) holds. If for each $t \in [0,T]$, the functions $b$, $\widehat{b}$, $\sigma$, $\widehat{\sigma}$ and $g$ together with their first order partial derivatives in $x$, $y$, $z$ are continuous with respect to $\alpha$; and $x(\alpha)$ is continuous at $\alpha_0 \in \Lambda$, then the adapted solutions $(X_t(\alpha), Y_t(\alpha), Z_t(\alpha))$ of (5.1) is $\mathcal{L}(B)$- (whence $\mathcal{L}$-) continuous at $\alpha_0$.*

*Moreover, if $x(\alpha)$, $g(\alpha, x)$ are $i$ times continuously differentiable in $\alpha$ and $x$, and for fixed $t$, the functions $b$, $\widehat{b}$, $\sigma$ and $\widehat{\sigma}$ are all $i+1$ times continuously differentiable with respect to the variables $\alpha$, $x$, $y$ and $z$; and all the derivatives (up to order $i+1$) above are at most polynomial growth in $x$, $y$ and $z$, uniformly in $t$ and $\alpha$, then the solution $(X(\alpha), Y(\alpha), Z(\alpha))$ is $i$ times $\mathcal{L}$-differentiable with respect to $\alpha$.*

*Proof.* (i) *Continuous dependence.* We shall check our main scheme step by step. First of all, if $\sigma(\cdot, t, x, y)$ and $\widehat{\sigma}(\cdot, s, x, y, z)$ are continuous at $\alpha_0$, then obviously, the implicit function $z(s, x, y, p; \alpha)$ of the equation

$$p\sigma(\alpha, t, x, y) + \widehat{\sigma}(\alpha, t, x, y, z(t, x, y, p; \alpha)) = 0, \tag{5.5}$$

which always exists by (A6), will be continuous (in $\alpha$) at $\alpha_0 \in \Lambda$.

Next consider the parabolic system with parameter:

$$(5.6)\qquad \begin{cases} \theta_t^k + \frac{1}{2}\mathrm{tr}\,(\theta_{xx}^k \sigma(\alpha,t,x,\theta)\sigma(\alpha,t,x,\theta)^T) + \langle\, b(\alpha,t,x,\theta,z(t,x,\theta,\theta_x;\alpha)), \theta_x^k \,\rangle \\ \qquad\qquad + \widehat{\sigma}^k(\alpha,t,x,\theta,z(t,x,\theta,\theta_x;\alpha)) = 0, \qquad k = 1,\cdots,m, \\ \theta(T,x) = g(\alpha,x), \qquad x \in \mathbb{R}^n. \end{cases}$$

Using the assumptions of the theorem it easy to see that the solutions of (5.6), denoted by $\theta(t,x;\alpha)$, will be continuous in $\alpha$ at $\alpha_0$ for fixed $(t,x) \in [0,T] \times \mathbb{R}^n$. Moreover, it is not hard to check that $\theta_x(t,x;\alpha)$ is continuous at $\alpha_0$ as well.

Now we turn to the third step. Consider the forward SDE:

$$(5.7)\qquad X_t = x(\alpha) + \int_0^t \tilde{b}(\alpha,s,X_s)ds + \int_0^t \tilde{\sigma}(\alpha,s,X_s)dW_s,$$

where

$$(5.8)\qquad \begin{cases} \tilde{b}(\alpha,t,x) = b(\alpha,t,x,\theta(t,x;\alpha),z(t,x,\theta(t,x;\alpha),\theta_x(t,x;\alpha);\alpha)); \\ \tilde{\sigma}(\alpha,t,x) = \sigma(\alpha,t,x,\theta(t,x;\alpha)). \end{cases}$$

Let us denote the solution of (5.7) by $X(\alpha)$. The assumption of the theorem and the results from the last two steps show that both $\tilde{b}$ and $\tilde{\sigma}$ are continuous in $\alpha$ at $\alpha_0$ for fixed $(t,x) \in [0,T]$. Hence by a continuous dependence theorem (cf. [4]), we have

$$(5.9)\qquad \mathcal{L}(B)\text{-}\lim_{\alpha\to\alpha_0} X_t(\alpha) = X_t(\alpha_0).$$

Finally, recall from Theorem 4.1 that the adapted solution of (5.1) must have the form

$$(5.10)\qquad Y_t(\alpha) = \theta(t,X_t(\alpha);\alpha); \quad Z_t(\alpha) = z(t,X_t(\alpha),\theta(t,X_t(\alpha);\alpha),\theta_x(t,X_t(\alpha);\alpha);\alpha),$$

the conclusion follows immediately from (5.9).

(ii) *Differentiability.* We again follow the Four Step Scheme. First, by our assumption, the solution $z(t,x,\theta,p;\alpha)$ will be $i$ times differentiable with respect to $x$, $\theta$, $p$ and $\alpha$ for each fixed $t$. Second, the solution $\theta(t,x;\alpha)$ (together with $\theta_x(t,x;\alpha)$) of the parabolic system (5.6) will also be $i$ times differentiable for fixed t, with respect to $x$ and $\alpha$. Therefore the functions $\tilde{b}$ and $\tilde{\sigma}$ defined by (5.8) are all $i$ times continuously differentiable with respect to $\alpha$ and $x$; and their derivatives (up to order $i$) are at most polynomialy growing in $x$,

uniformly in $t$ and $\alpha$. Thus, by [4, Theorem 2.8.4], the solution $X(\alpha)$ of (5.6) is $i$ times $\mathcal{L}$- (in fact $\mathcal{L}(B)$-) differentiable. The conclusion then follows easily from the expression (5.10) for $Y(\alpha)$ and $Z(\alpha)$. The proof is now complete. □

## §6. An Integral Representation Formula.

In this section we consider a special case: $\widehat{b} \equiv 0$ and $\widehat{\sigma} \equiv -z$. We shall derive an integral representation formula for functions (or functionals) of diffusions via our explicit solution of forward-backward SDEs. The forward-backward equations (2.1) now take the form:

$$\begin{cases} X_t = x + \int_0^t b(s, X_s, Y_s, Z_s)ds + \int_0^t \sigma(s, X_s, Y_s)dW_s, \\ Y_t = g(X_T) - \int_t^T Z_s dW_s, \quad t \in [0,T], \end{cases} \tag{6.1}$$

From our Four Step Scheme, we know that if we define

$$z(t,x,y,p) = p\sigma(t,x,y), \qquad \forall (t,x,y,p) \in [0,T] \times \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^{m\times n}, \tag{6.2}$$

and let $\theta(t,x)$ be the classical solution of the following parabolic system:

$$\begin{cases} \theta_t^k + \frac{1}{2}\mathrm{tr}\,[\theta_{xx}^k \sigma(t,x,\theta)\sigma(t,x,\theta)^T] + \langle\, b(t,x,\theta,z(t,x,\theta,\theta_x)), \theta_x^k \,\rangle = 0, \\ \qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad k = 1, \cdots, m; \\ \theta(T,x) = g(x), \end{cases} \tag{6.3}$$

then we can derive the (unique) adapted solution of (6.1) by first solving the following forward SDE:

$$X_t = x + \int_0^t \tilde{b}(s, X_s)ds + \int_0^t \tilde{\sigma}(s, X_s)dW_s, \tag{6.4}$$

where

$$\begin{cases} \tilde{b}(t,x) = b(t,x,\theta(t,x),\theta_x(t,x)\sigma(t,x,\theta(t,x))); \\ \tilde{\sigma}(t,x) = \sigma(t,x,\theta(t,x)), \end{cases} \tag{6.5}$$

and then setting $Y_t = \theta(t, X_t)$ and $Z_t = \theta_x(t, X_t)\sigma(t, X_t, \theta(t, X_t))$. Further, letting $t = 0$ in the second (backward) equation in (6.1), we get

$$Y_0 = g(X_T) - \int_0^T \theta_x(s, X_s)\sigma(s, X_s, \theta(s, X_s))dW_s. \tag{6.6}$$

Note from (6.1) that

$$Y_t = E\{g(X_T)|\mathcal{F}_t\}, \quad t \in [0,T], \tag{6.7}$$

and since $\mathcal{F}_0$ is a trivial $\sigma$-field, $Y_0$ is non-random, hence $Y_0 = EY_0 = Eg(X_T)$ and (6.6) can be rewritten as

$$g(X_T) = Eg(X_T) + \int_0^T \theta_x(s,X_s)\sigma(s,X_s,\theta(s,X_s))dW_s. \tag{6.8}$$

We see that (6.8) provides an integral representation for $g(X_T)$. A more general formula is the Clark-Haussmann-Ocone formula, which is (in this case):

$$g(X_T) = Eg(X_T) + \int_0^T E\{D_s g(X_T)|\mathcal{F}_s\}dW_s, \tag{6.9}$$

where $D$ is the "Malliavin derivative" operator (cf. [8] or [9]). To see the relationship between (6.8) and (6.9), note that since $X$ is the solution of (6.4), which is actually a diffusion, we can calculate $E\{D_s g(X_T)|\mathcal{F}_s\}$ as follows:

$$D_s g(X_T) = g_x(X_T)D_s X_T,$$

and

$$D_s X_t = \tilde{\sigma}(s,X_s) + \int_s^t \tilde{b}_x(r,X_r)D_s X_r dr + \int_s^t \tilde{\sigma}_x(r,X_r)D_s X_r dW_r,$$

for $0 \le s \le t \le T$. Thus fixing $s$ and letting $u_t = D_s X_t$, we have that $u$ is the solution of a linear SDE. For simplicity, we now assume that $m = n = 1$. Then it is known (cf. [12]) that $u_t$ has an explicit form:

$$u_t = \mathcal{E}(Z)_t \tilde{\sigma}(s,X_s), \tag{6.10}$$

where $\mathcal{E}(Z)_t$ denotes the Doléans-Dade stochastic exponential of $Z$ and

$$Z_t = \int_s^t \tilde{b}_x(r,X_r)dr + \int_s^t \tilde{\sigma}_x(r,X_r)dW_r.$$

In other words,

$$\begin{aligned} \mathcal{E}(Z)_t &= \exp\{Z_t - \frac{1}{2}[Z,Z]_t\} \\ &= \exp\left\{\int_s^t \tilde{\sigma}_x(r,X_r)dW_r + \int_s^t \left(\tilde{b}_x(r,X_r) - \frac{1}{2}\tilde{\sigma}_x^2(r,X_r)\right)dr\right\}, \end{aligned} \tag{6.11}$$

Therefore (in the case $m = n = 1$) we have

$$E\{D_s g(X_T)|\mathcal{F}_s\} = E\{g_x(X_T)D_s X_T|\mathcal{F}_s\} = E\{g_x(X_T)\mathcal{E}(Z)_T|\mathcal{F}_s\}\tilde{\sigma}(s, X_s). \tag{6.12}$$

Furthermore, if we let $\varphi(x; s, T)$ denote the flow of $X$ from $s$ to $T$ given $X_s = x$ (cf. [11, §V-7]), then

$$E\{g_x(X_T)\mathcal{E}(Z)_T|\mathcal{F}_s\} = H(X_s, s, T),$$

where

$$H(x, s, T) = E\left\{ g_x(\varphi(x; s, T)) \exp\left[ \int_s^T \tilde{\sigma}_x(r, \varphi(x; s, r))dW_r \right.\right.$$
$$\left.\left. + \int_s^T \left( \tilde{b}_x(r, \varphi(x; s, r)) - \frac{1}{2}\tilde{\sigma}_x^2(r, \varphi(x; s, r)) \right) dr \right] \right\},$$

and we are able to conclude (for $m = n = 1$):

$$g(X_T) = Eg(X_T) + \int_0^T H(X_s, s, T)\tilde{\sigma}(s, X_s)dW_s \tag{6.13}$$

directly from the Ocone formula. The function $H(x, s, T)$ is however not amenable either to closed form expression nor to simple computations. But if we compare (6.13) and (6.8), then we have $H(X_s, s, T) = \theta_x(s, X_s)$.

For general $m$ and $n$, we have that $D_s X_t$ solves a linear system of SDEs, and hence it has a closed form solution (see [13, p.271]), and we can again obtain information from the Ocone formula, but it will be more complicated (a simpler argument without using Malliavian derivatives is given at the end of this section). The next theorem shows that in this case, our formula is simpler.

**Theorem 6.1.** *Under assumptions (A1)–(A4), let $X$ solve (6.8). Then*

$$g(X_T) = Eg(X_T) + \int_0^T E\{D_s g(X_T)|\mathcal{F}_s\}dW_s$$
$$= Eg(X_T) + \int_0^T \theta_x(s, X_s)\tilde{\sigma}(s, X_s)dW_s.$$

*Consequently, $E\{D_s g(X_T)|\mathcal{F}_s\} = \tilde{\sigma}(s, X_s)\theta_x(s, X_s)$, $dP \otimes dt$-almost surely.*

*Proof.* The first equality is the Ocone formula (cf [9]); the second is (6.8), and subtraction gives

$$\int_0^T E\{D_s g(X_T)|\mathcal{F}_s\}dW_s = \int_0^T \tilde{\sigma}(s, X_s)\theta_x(s, X_s)dW_s.$$

Subtracting one stochastic integral from the other yields the zero local martingale, hence

$$\int_0^T \left(E\{D_s g(X_T)|\mathcal{F}_s\} - \tilde{\sigma}(s, X_s)\theta_x(s, X_s)\right)^2 ds = 0,$$

and so we have the third equality $dP \otimes dt$-almost surely. □

We can simplify the assumptions of Theorem 6.1:

**Theorem 6.2.** *Let $W$ be an $n$-dimensional Brownian motion, let $b$, $\sigma$, and $g$ be $C^1$ with bounded first derivatives and suppose that $g$ is bounded. Suppose further that the function $\sigma$ is invertible for each $(t,x)$ and that $\sigma^{-1}(t,x)$ is bounded. Let $X$ be the solution to the SDE:*

$$X_t = x + \int_0^t b(s, X_s)ds + \int_0^t \sigma(s, X_s)dW_s. \tag{6.14}$$

*Then it holds that*

$$g(X_T) = Eg(X_T) + \int_0^T \theta_x(s, X_s)\sigma(s, X_s)dW_s. \tag{6.15}$$

*where $\theta$ is the classical solution of the following PDE:*

$$\begin{cases} \theta_t + \frac{1}{2}\mathrm{tr}[\theta_{xx}\sigma(t,x)\sigma(t,x)^T] + \langle\, b(t,x), \theta_x \,\rangle = 0, \\ \theta(T,x) = g(x), \end{cases} \tag{6.16}$$

*Proof.* Let us consider the following uncoupled forward-backward SDE:

$$\begin{cases} X_t = x + \int_0^t b(s, X_s)ds + \int_0^t \sigma(s, X_s)dW_s, \\ Y_t = g(X_T) - \int_t^T Z_s dW_s, \quad t \in [0,T]. \end{cases} \tag{6.17}$$

We again use our Four Step Scheme to get the explicit solution of (6.17). To do this, note that now $m = 1$, so we need only to apply Theorem 4.5; but in this case it is easily seen that the conditions of Theorem 4.5 are reduced to ones given in the theorem, therefore the result is a direct consequence of the argument at the beginning of this section. □

To end this section, we discuss a little bit about the process $p_t \triangleq \theta_x(t, X_t)$. Such a process is of independent interest; for example, in stochastic control theory. The following proposition gives an important property of $p$.

**Proposition 6.3** *There exists an $\mathbb{R}^{n\times n}$-valued adapted process $\{K_t : t \geq 0\}$ such that $(p, K)$ is the unique adapted solution of the following backward SDE:*

$$p_t = g_x(X_T) + \int_t^T [b_x(s, X_s)^T p_s + \sigma_x(s, X_s) \otimes K_s] ds - \int_t^T K_s dW_s, \tag{6.18}$$

*where $\sigma_x \otimes K \triangleq (\mathrm{tr}\,(\sigma_{x_1} K), \cdots, \mathrm{tr}\,(\sigma_{x_n} K))^T$. In particular, if the function $\theta$ is $C^3$, then $K_t = \theta_{xx}(t, X_t)\sigma(t, X_t)$ for $t \geq 0$.*

*Proof.* We first assume that $\theta$ is $C^3$. Taking one more derivative in the $x$ variable for equation (6.16) and doing a little computation, we get, for every $\ell = 1, \cdots, n$:

$$\begin{cases} \theta_{tx_\ell} + \sum_{i=1}^n \theta_{x_i x_\ell} b^i(t,x) + \sum_{i=1}^n b^i_{x_\ell}(t,x)\theta_{x_i} \\ \qquad\qquad + \frac{1}{2}\sum_{i,j=1}^n \theta_{x_i x_j x_\ell} \left( \sum_{k=1}^n \sigma^{jk}(t,x)\sigma^{ki}(t,x) \right) \\ \qquad\qquad + \sum_{i,j=1}^n \theta_{x_i x_j} \left( \sum_{k=1}^n \sigma^{jk}(t,x)\sigma^{ki}_{x_\ell}(t,x) \right) = 0, \\ \theta(T,x) = g(x). \end{cases} \tag{6.19}$$

On the other hand, if we apply Itô's formula to $\theta_{x_\ell}$ from $t$ to $\tau$ $(0 \leq t \leq \tau)$, then we have for every $\ell = 1, \cdots, n$:

$$\begin{aligned} \theta_{x_\ell}(\tau, X_\tau) = \theta_{x_\ell}(t, X_t) + \int_t^\tau \Bigg\{ & \theta_{x_\ell t}(s, X_s) + \sum_{i=1}^n \theta_{x_\ell x_i}(s, X_s) b^i(s, X_s) \\ & + \frac{1}{2}\sum_{i,j=1}^n \theta_{x_\ell x_i x_j}(s, X_s) \left( \sum_{k=1}^n \sigma^{jk}(s, X_s)\sigma^{ki}(s, X_s) \right) \Bigg\} ds \\ & + \int_t^\tau \sum_{k=1}^n \left( \sum_{i=1}^n \theta_{x_\ell x_i}(s, X_s)\sigma^{ik}(s, X_s) \right) dW_s^k. \end{aligned} \tag{6.20}$$

Using (6.19) and denoting $K_t^{\ell k} = \sum_{i=1}^n \theta_{x_\ell x_i}(t, X_t)\sigma^{ik}(t, X_t)$, we obtain from (6.20) that

$$
\begin{aligned}
\theta_{x_\ell}(\tau, X_\tau) &= \theta_{x_\ell}(t, X_t) \\
&\quad - \int_t^\tau \left( \sum_{i=1}^n \theta_{x_i} b^i_{x_\ell} + \sum_{i,j=1}^n \theta_{x_i x_j} \sum_{k=1}^n \sigma^{jk}\sigma^{ki}_{x_\ell} \right)(s, X_s)ds \\
&\quad + \int_t^\tau \sum_{k=1}^n \left( \sum_{i=1}^n \theta_{x_\ell x_i}(s, X_s)\sigma^{ik}(s, X_s) \right) dW_s^k \\
&= \theta_{x_\ell}(t, X_t) - \int_t^\tau \left( \sum_{i=1}^n \theta_{x_i} b^i_{x_\ell}(s, X_s) + \sum_{i,k=1}^n K_s^{ik}\sigma^{ki}_{x_\ell}(s, X_s) \right) ds \\
&\quad + \int_t^\tau \sum_{k=1}^n K_s^{\ell k} dW_s^k, \qquad \ell = 1, \cdots, n.
\end{aligned} \tag{6.21}
$$

Now setting $p_t^\ell = \theta_{x_\ell}(t, X_t)$ and $\tau = T$, and rewriting (6.21) in vector form, we obtain (6.18) immediately.

To show the general case, let $t \in [0, T]$ be fixed and define $Q^\omega = P(\cdot|\mathcal{F}_t)(\omega)$, $\omega \in \Omega$. For any rational $\xi \in \mathbb{R}^n$, and ($P$-almost) every $\omega \in \Omega$, define a diffusion process $X(\xi)$ (on the probability space $(\Omega, \mathcal{F}, Q^\omega)$) by

$$
X_\tau(\xi) = \xi + \int_t^\tau b(s, X_s(\xi))ds + \int_t^\tau \sigma(s, X_s(\xi))dW_s, \qquad t \le \tau \le T. \tag{6.22}
$$

By a standard argument, one shows that for $P -$ a.e. $\omega \in \Omega$,

$$
E^{Q^\omega}\left\{ \sup_{\tau \in [t,T]} |X_\tau(\xi) - X_\tau|^2 \right\} \le C|\xi - X_t(\omega)|^2. \tag{6.23}
$$

Moreover, if we define $\zeta_\tau(\xi) = X_\tau(\xi) - X_\tau$, $t \le \tau \le T$, then $\zeta(\xi)$ satisfies

$$
\zeta_\tau(\xi) = (\xi - X_t) + \int_t^\tau b_x(s, X_s)\zeta_s(\xi)ds + \sum_{i=1}^n \int_t^\tau \sigma_x^{(i)}(s, X_s)\zeta_s(\xi)dW_s^i + \varepsilon(\tau, \xi), \tag{6.24}
$$

for $t \le \tau \le T$, where $\sigma^{(i)}$ is the $i$-th column of $\sigma$; and (6.23) would lead to

$$
E^{Q^\omega}\left\{ \sup_{\tau \in [t,T]} |\varepsilon(\tau, \xi)|^2 \right\} = o(|\xi - X_t(\omega)|^2), \tag{6.25}
$$

for $P$-a.e. $\omega \in \Omega$. We now consider the backward SDE (6.18) on the probability space $(\Omega, \mathcal{F}, Q^\omega)$ for $\tau \in [t, T]$. It is known (cf. [10]) that it has a unique adapted solution under

our assumptions. We shall prove that $\theta_x(t, X_t(\omega)) = p_t(\omega)$, $t \in [0,T]$, for $P$-a.e. $\omega \in \Omega$. To this end, first note that an application of Itô's formula to $\langle\, p_\tau, \zeta_\tau(\xi)\,\rangle$ from $t$ to $T$, together with (6.25), leads to

$$E^{Q^\omega}\{\langle\, p_T, \zeta_T(\xi)\,\rangle\} = \langle\, p_t, \xi - X_t\,\rangle(\omega) + o(|\xi - X_t|(\omega)), \qquad P\text{-a.e. } \omega \in \Omega. \tag{6.26}$$

Next, for $P$-a.e. $\omega \in \Omega$, we have

$$\begin{aligned} \theta(t,\xi) - \theta(t, X_t(\omega)) &= E\{g(X_T(\xi)) - g(X_T)|\mathcal{F}_t\}(\omega) = E^{Q^\omega}\{g(X_T(\xi)) - g(X_T)\} \\ &= E^{Q^\omega}\{\langle\, g_x(X_T), \zeta_T(\xi)\,\rangle\} + o(|\xi - X_t|(\omega)) \\ &= \langle\, p_t, \xi - X_t\,\rangle(\omega) + o(|\xi - X_t|(\omega)). \end{aligned} \tag{6.27}$$

Therefore, by the definition of the derivative, we see that for $P$-a.e. $\omega \in \Omega$, $\theta_x(t, X_t(\omega)) = p_t(\omega)$ (a similar argument can also be found in [15]). The proposition is proved. □

As a final remark, we wish to point out that (6.26) actually gives an *adjoint relation* between the backward SDE (6.18) and the linear forward SDE

$$d\zeta_t = b_x(t, X_t)\zeta_t + \sum_{i=1}^{n} \sigma_x^{(i)}(t, X_t)\zeta_t dW_t^i \tag{6.28}$$

More precisely, if $\zeta$ and $p$ are the (adapted) solutions of (6.18) and (6.28) respectively, then

$$\langle\, p_T, \zeta_T\,\rangle = \langle\, p_t, \zeta_t\,\rangle + m_T - m_t, \tag{6.29}$$

where $m$ stands for some $\{\mathcal{F}_t\}$-martingale. Therefore, upon taking a conditional expectation $E\{\cdot|\mathcal{F}_t\}$ on both sides of (6.29), we obtain that

$$E\{\langle\, p_T, \zeta_T\,\rangle|\mathcal{F}_t\} = \langle\, p_t, \zeta_t\,\rangle, \qquad t \in [0,T], \qquad \text{a.s. } P. \tag{6.30}$$

Note that $p_T = g_x(X_T)$ and that (6.30) is true for any solution $\zeta$ of (6.28). If we let $\Phi(\tau, t)$ $(t \le \tau \le T)$ be the fundamental matrix of the linear SDE (6.28) satisfying $\Phi(t,t) = I$, then (6.30) leads to

$$E\{g_x(X_T)^T\Phi(T,t)|\mathcal{F}_t\} = p_t = \theta_x(t, X_t), \qquad t \in [0,T], \qquad \text{a.s. } P.$$

Noting that if $m = n = 1$, $\Phi$ is actually the stochastic exponential as defined by (6.11), and comparing with (6.12), (6.13) and Theorem 6.1, we see that this again proves the Haussmann formula in our special case.

## References


[1] F. Antonelli, *Backward-Forward Stochastic Differential Equations,* to appear in Ann. Applied Proba. (1992).

[2] K. Deimling, *Nonlinear Functional Analysis,* Springer-Verlag, Berlin, 1985.

[3] U. G. Haussmann, *On the Integral Representation of Functionals of Itô's Processes, Stochastics,* Vol.3 (1979), 17–27.

[4] N. V. Krylov, *Controlled Diffusion Processes,* Springer-Verlag, New York-Heidelerg-Berlin (1980).

[5] O. A. Ladyzenskaja, V. A. Solonnikov and N. N. Ural'ceva, *Linear and Quasilinear Equations of Parabolic Type,* Amer. Math. Soc., Providence, R.I., 1968.

[6] I. Karatzas, D. L. Ocone and J. Li, *An Extension of Clark's Formula, Stochastics and Stochastic Reports,* Vol. 37 (1991), 127–131.

[7] J. Ma and J. Yong, *Solvability of Forward-Backward Equations and the Nodal Set of Hamilton-Jacobi-Equations,* Preprint (1993).

[8] D. Nualart, *Noncausal Stochastic Integrals and Calculus, Stochastic Analysis and Related Topics,* Lecture Notes in Math. (1986), 80–129.

[9] D. Ocone, *Malliavin's Calculus and Stochastic Integral Representations of Functionals of Diffusion Processes, Stochastics,* Vol.12 (1984), 161–185.

[10] E. Pardoux and S. Peng, *Adapted Solution of a Backward Stochastic Differential Equation,* Systems & Control Letters, **14** (1990), 55–61.

[11] E. Pardoux and S. Peng, *Backward Stochastic Differential Equations and Quasilinear Parabolic Partial Differential Equations, Stochastic Partial Differential Equations and*

*Their Applications,* B. L. Bozovskii & R. S. Sowers eds, LNCIS **176**, Springer (1992), 200–217.

[12] S. Peng, *Adapted Solution of Backward Stochastic Equations and Related Partial Differential Equations,* Preprint.

[13] P. Protter, *Stochastic Integration and Differential Equations, A new Approach.* Springer (1990).

[14] M. Wiegner, *Global solutions to a class of strongly coupled parabolic systems, Math. Ann.*, 292 (1992), 711–727.

[15] X. Y. Zhou, *A Unified Treatment of Maximum Principle and Dynamic Programming in Stochastic Controls, Stochastics and Stochastics Reports,* Vol 36 (1991), 137–161.