Chapter 4
Backward Stochastic Differential Equations

Let $(\Omega, \mathscr{F}, \mathbb{F}, \mathbb{P})$ be given, and B a d-dimensional Brownian motion. In order to apply the martingale representation theorem, in this chapter we shall always assume

$$\mathbb{F} = \mathbb{F}^B. \tag{4.0.1}$$

While SDE is a nonlinear extension of the stochastic integration, Backward SDE is a nonlinear version of the martingale representation theorem. In fact, both the results and the arguments in this chapter are analogous to those for SDEs, combined with the martingale representation theorem.

Given $\xi \in \mathbb{L}^2(\mathbb{F})$, it induces naturally a martingale $Y_t := \mathbb{E}[\xi|\mathscr{F}_t]$. By the martingale representation theorem, there exists unique $Z \in \mathbb{L}^2(\mathbb{F})$ such that

$$dY_t = Z_t dB_t, \quad \text{or equivalently,} \quad Y_t = \xi - \int_t^T Z_s dB_s. \tag{4.0.2}$$

This is a linear SDE with terminal condition $Y_T = \xi$, and thus is called a *Backward SDE* (BSDE, for short). We emphasize that the solution to a BSDE is a *pair* of $\mathbb{F}$-measurable processes (Y, Z). As we will see more clearly in Section 9.4, the component Z is essentially the derivative of Y with respect to B and thus is uniquely determined by Y (and B). We also emphasize that the presence of Z is crucial to ensure the $\mathbb{F}$-measurability of Y. Indeed, if we consider a SDE with terminal condition in the following form:

$$dY_t = \sigma_t(Y_t) dB_t, \quad Y_T = \xi.$$

Then typically the equation has no $\mathbb{F}$-measurable solution Y. For example, if $\sigma = 0$, then the candidate solution has to be $Y_t = \xi$ for all t, which is not $\mathbb{F}$-measurable unless $\xi \in \mathscr{F}_0$.

J. Zhang, *Backward Stochastic Differential Equations*, Probability Theory and Stochastic Modelling 86, DOI 10.1007/978-1-4939-7256-2_4

In this chapter we consider the following nonlinear BSDE:

$$Y_t = \xi + \int_t^T f_s(Y_s, Z_s)ds - \int_t^T Z_s dB_s, \quad 0 \le t \le T, \ \mathbb{P}\text{-a.s.} \tag{4.0.3}$$

where $Y \in \mathbb{L}^2(\mathbb{F}, \mathbb{R}^{d_2})$, $Z \in \mathbb{L}^2(\mathbb{F}, \mathbb{R}^{d_2 \times d})$ for some dimension d_2. We call f the (nonlinear) generator and ξ the terminal condition of the BSDE. We shall always assume

Assumption 4.0.1

(i) (4.0.1) *holds;*
(ii) $f : [0, T] \times \Omega \times \mathbb{R}^{d_2} \times \mathbb{R}^{d_2 \times d} \to \mathbb{R}^{d_2}$ *is* $\mathbb{F}$*-measurable in all variables;*
(iii) f *is uniformly Lipschitz continuous in* (y, z) *with a Lipschitz constant L;*
(iv) $\xi \in \mathbb{L}^2(\mathscr{F}_T, \mathbb{R}^{d_2})$ *and* $f^0 := f(0, 0) \in \mathbb{L}^{1,2}(\mathbb{F}, \mathbb{R}^{d_2})$.

As in Chapter 3, for notational simplicity we shall assume $d_2 = d = 1$ in most proofs. We remark that, in the standard literature, it is required that $f^0 \in \mathbb{L}^2(\mathbb{F})$. Our condition here is slightly weaker.

4.1 Linear Backward Stochastic Differential Equations

In this section we study the case when f is linear. We first have the following simple result.

Proposition 4.1.1 *Let* $\xi \in \mathbb{L}^2(\mathscr{F}_T, \mathbb{R}^{d_2})$ *and* $f^0 \in \mathbb{L}^{1,2}(\mathbb{F}, \mathbb{R}^{d_2})$*. Then, the following linear BSDE has a unique solution* $(Y, Z) \in \mathbb{S}^2(\mathbb{F}, \mathbb{R}^{d_2}) \times \mathbb{L}^2(\mathbb{F}, \mathbb{R}^{d_2 \times d})$*:*

$$Y_t = \xi + \int_t^T f_s^0 ds - \int_t^T Z_s dB_s \tag{4.1.1}$$

Proof It is obvious that

$$Y_t = \mathbb{E}\Big[\xi + \int_t^T f_s^0 ds \Big| \mathscr{F}_t\Big].$$

Note that

$$\tilde{Y}_t := Y_t + \int_0^t f_s^0 ds = \mathbb{E}\Big[\xi + \int_0^T f_s^0 ds \Big| \mathscr{F}_t\Big]$$

is a square integrable martingale. By the martingale representation theorem, there exists unique $Z \in \mathbb{L}^2(\mathbb{F}, \mathbb{R}^{d_2 \times d})$ such that

$$d\tilde{Y}_t = Z_t dB_t.$$

One can check straightforwardly that the above pair (Y, Z) satisfies (4.1.1), and, from the above derivation, it is the unique solution. ■

We next consider the general linear BSDE with $d_2 = 1$:

$$Y_t = \xi + \int_t^T [\alpha_s Y_s + Z_s \beta_s + f_s^0] ds - \int_t^T Z_s dB_s. \tag{4.1.2}$$

The well-posedness of this BSDE will follow from the general theory. Here we provide a representation formula for its solution.

Proposition 4.1.2 *Let $d_2 = 1$, $\xi \in \mathbb{L}^2(\mathscr{F}_T, \mathbb{R})$, $\alpha \in \mathbb{L}^\infty(\mathbb{F}, \mathbb{R})$, $\beta \in \mathbb{L}^\infty(\mathbb{F}, \mathbb{R}^d)$, and $f^0 \in \mathbb{L}^{1,2}(\mathbb{F}, \mathbb{R})$. If $(Y, Z) \in \mathbb{L}^2(\mathbb{F}, \mathbb{R}) \times \mathbb{L}^2(\mathbb{F}, \mathbb{R}^{1\times d})$ satisfies the linear BSDE* (4.1.2)*, then*

$$Y_t = \Gamma_t^{-1} \mathbb{E}\Big[\Gamma_T \xi + \int_t^T \Gamma_s f_s^0 ds \Big| \mathscr{F}_t\Big], \tag{4.1.3}$$

where

$$\Gamma_t = 1 + \int_0^t \Gamma_s [\alpha_s dt + \beta_s \cdot dB_s], \text{ or say, } \Gamma_t := \exp\Big(\int_0^t \beta_s \cdot dB_s + \int_0^t [\alpha_s - \frac{1}{2}|\beta_s|^2] ds\Big). \tag{4.1.4}$$

Proof Applying Itô formula we have

$$d(\Gamma_t Y_t) = -\Gamma_t f_t^0 dt + \Gamma_t [Y_t \beta_t^\top + Z_t] dB_t.$$

Denote

$$\hat{Y}_t := \Gamma_t Y_t; \quad \hat{Z}_t := \Gamma_t [Y_t \beta_t^\top + Z_t]; \quad \hat{\xi} := \Gamma_T \xi; \quad \hat{f}_t^0 := \Gamma_t f_t^0. \tag{4.1.5}$$

Then, one may rewrite (4.1.2) as

$$\hat{Y}_t = \hat{\xi} + \int_t^T \hat{f}_s^0 ds - \int_t^T \hat{Z}_s dB_s.$$

This is a linear BSDE in the form (4.1.1). By Lemma 2.6.1 and Problem 2.10.7 (i) we see that $\int_0^t \hat{Z}_s dB_s$ is a martingale. Then

$$\hat{Y}_t := \mathbb{E}\Big[\hat{\xi} + \int_t^T \hat{f}_s^0 ds \Big| \mathscr{F}_t\Big],$$

which implies (4.1.3) immediately. ■

4.2 A Priori Estimates for BSDEs

We now investigate the nonlinear BSDE (4.0.3).

Theorem 4.2.1 *Let Assumption 4.0.1 hold and* $(Y, Z) \in \mathbb{L}^2(\mathbb{F}, \mathbb{R}^{d_2}) \times \mathbb{L}^2(\mathbb{F}, \mathbb{R}^{d_2 \times d})$ *be a solution to BSDE (4.0.3). Then* $Y \in \mathbb{S}^2(\mathbb{F}, \mathbb{R}^{d_2})$ *and there exists a constant C, depending only on T, L, and* d, d_2*, such that*

$$\|(Y, Z)\|^2 := \mathbb{E}\Big[|Y_T^*|^2 + \int_0^T |Z_t|^2 dt\Big] \le C I_0^2, \text{ where } I_0^2 := \mathbb{E}\Big[|\xi|^2 + \big(\int_0^T |f_t^0| dt\big)^2\Big]. \tag{4.2.1}$$

Proof Foy simplicity, we assume $d = d_2 = 1$. We proceed in several steps.

Step 1. We first show that

$$\mathbb{E}[|Y_T^*|^2] \le C \mathbb{E}\Big[\int_0^T [|Y_t|^2 + |Z_t|^2] dt\Big] + C I_0^2 < \infty. \tag{4.2.2}$$

Indeed, note that

$$|Y_t| \le |\xi| + \int_t^T [|f_s^0| + C|Y_s| + C|Z_s|] ds + \Big|\int_t^T Z_s dB_s\Big|.$$

Then,

$$Y_T^* \le C\Big[|\xi| + \int_0^T [|f_t^0| + |Y_t| + |Z_t|] dt + \sup_{0 \le t \le T} \Big|\int_0^t Z_s dB_s\Big|\Big].$$

Applying Burkholder-Davis-Gundy inequality we have

$$\mathbb{E}[|Y_T^*|^2] \le C \mathbb{E}\Big[|\xi|^2 + \Big(\int_0^T |f_t^0| dt\Big)^2 + \int_0^T [|Y_t|^2 + |Z_t|^2] dt\Big],$$

which implies (4.2.2) immediately.

Step 2. We next show that, for any $\varepsilon > 0$,

$$\sup_{0 \le t \le T} \mathbb{E}[|Y_t|^2] + \mathbb{E}\Big[\int_0^T |Z_t|^2 dt\Big] \le \varepsilon [|Y_T^*|^2] + C \varepsilon^{-1} I_0^2. \tag{4.2.3}$$

Indeed, by Itô formula,

$$d|Y_t|^2 = 2Y_t dY_t + |Z_t|^2 dt = -2Y_t f_t(Y_t, Z_t) dt + 2Y_t Z_t dB_t + |Z_t|^2 dt. \tag{4.2.4}$$

Thus,

$$|Y_t|^2 + \int_t^T |Z_s|^2 ds = |\xi|^2 + 2\int_t^T Y_s f_s(Y_s, Z_s) ds + 2\int_t^T Y_s Z_s dB_s. \tag{4.2.5}$$

By (4.2.2) and Problem 2.10.7 (i) we know $\int_0^t Y_s Z_s dB_s$ is a true martingale. Now, taking expectation on both sides of (4.2.5) and noting that $ab \le \frac{1}{2}a^2 + \frac{1}{2}b^2$, we have

$$\mathbb{E}\Big[|Y_t|^2 + \int_t^T |Z_s|^2 ds\Big] = \mathbb{E}\Big[|\xi|^2 + 2\int_t^T Y_s f_s(Y_s, Z_s) ds\Big]$$
$$\le \mathbb{E}\Big[|\xi|^2 + C\int_t^T |Y_s|[|f_s^0| + |Y_s| + |Z_s|] ds\Big]$$
$$\le \mathbb{E}\Big[|\xi|^2 + CY_T^* \int_0^T |f_s^0| ds + C\int_0^T [|Y_s|^2 + |Y_s Z_s|] ds\Big]$$
$$\le \mathbb{E}\Big[|\xi|^2 + CY_T^* \int_0^T |f_s^0| ds + C\int_t^T |Y_s|^2 ds + \frac{1}{2}\int_t^T |Z_s|^2 ds\Big].$$

This leads to

$$\mathbb{E}\Big[|Y_t|^2 + \frac{1}{2}\int_t^T |Z_s|^2 ds\Big] \le \mathbb{E}\Big[C\int_t^T |Y_s|^2 ds + |\xi|^2 + CY_T^* \int_0^T |f_s^0| ds\Big], \quad (4.2.6)$$

which, together with Fubini Theorem, implies that

$$\mathbb{E}[|Y_t|^2] \le \mathbb{E}\Big[|\xi|^2 + CY_T^* \int_0^T |f_s^0| ds\Big] + C\int_t^T E[|Y_s|^2] ds.$$

Applying (backward) Gronwall inequality, we get

$$\mathbb{E}[|Y_t|^2] \le C\mathbb{E}\Big[|\xi|^2 + Y_T^* \int_0^T |f_s^0| ds\Big], \quad \forall t \in [0, T]. \quad (4.2.7)$$

Then, by letting $t = 0$ and plug (4.2.7) into (4.2.6) we have

$$\mathbb{E}\Big[\int_0^T |Z_s|^2 ds\Big] \le C\mathbb{E}\Big[|\xi|^2 + Y_T^* \int_0^T |f_s^0| ds\Big]. \quad (4.2.8)$$

By (4.2.7) and (4.2.8) and noting that $2ab \le \varepsilon a^2 + \varepsilon^{-1} b^2$, we obtain (4.2.3) immediately.

Step 3. Plug (4.2.3) into (4.2.2), we get

$$\mathbb{E}[|Y_T^*|^2] \le C\varepsilon \mathbb{E}[|Y_T^*|^2] + C\varepsilon^{-1} I_0^2.$$

By choosing $\varepsilon = \frac{1}{2C}$ for the constant C above, we obtain

$$\mathbb{E}[|Y_T^*|^2] \le C I_0^2.$$

This, together with (4.2.3), proves (4.2.1). ■

Remark 4.2.2 Similar to Remark 3.2.3, Theorem 4.2.1 remains true if we weaken the Lipschitz condition of Assumption 4.0.1 (iii) to the linear growth condition:

$$|f_t(y,z)| \le |f_t^0| + L[|y| + |z|]. \tag{4.2.9}$$

■

Theorem 4.2.3 *For $i = 1,2$, assume (ξ_i, f^i) satisfy Assumption 4.0.1 and $(Y^i, Z^i) \in \mathbb{L}^2(\mathbb{F}, \mathbb{R}^{d_2}) \times \mathbb{L}^2(\mathbb{F}, \mathbb{R}^{d_2 \times d})$ is a solution to BSDE (4.0.3) with coefficients (ξ_i, f^i). Then*

$$\|(\Delta Y, \Delta Z)\|^2 \le C\mathbb{E}\Big[|\Delta\xi|^2 + \Big(\int_0^T |\Delta f_t(Y_t^1, Z_t^1)|dt\Big)^2\Big], \tag{4.2.10}$$

where

$$\Delta Y := Y^1 - Y^2, \quad \Delta Z := Z^1 - Z^2, \quad \Delta\xi := \xi_1 - \xi_2, \quad \Delta f := f^1 - f^2.$$

Proof Again assume $d = d_2 = 1$. Note that

$$\begin{aligned}\Delta Y_t &= \Delta\xi + \int_t^T [f_s^1(Y_s^1, Z_s^1) - f_s^2(Y_s^2, Z_s^2)]ds - \int_t^T \Delta Z_s dB_s \\ &= \Delta\xi + \int_t^T [\Delta f_s(Y_s^1, Z_s^1) + \alpha_s \Delta Y_s + \beta_s \Delta Z_s]ds - \int_t^T \Delta Z_s dB_s,\end{aligned}$$

where, similar to (3.2.10)

$$\alpha_t := \frac{f_t^2(Y_t^1, Z_t^1) - f_t^2(Y_t^2, Z_t^1)}{\Delta Y_t}\mathbf{1}_{\{\Delta Y_t \ne 0\}}, \quad \beta_t := \frac{f_t^2(Y_t^2, Z_t^1) - f_t^2(Y_t^2, Z_t^2)}{\Delta Z_t}\mathbf{1}_{\{\Delta Z_t \ne 0\}} \tag{4.2.11}$$

are bounded by L. Then, by Theorem 4.2.1 we obtain the result immediately. ■

4.3 Well-Posedness of BSDEs

We now establish the well-posedness of BSDE (4.0.3).

Theorem 4.3.1 *Under Assumption 4.0.1, BSDE (4.0.3) has a unique solution $(Y, Z) \in \mathbb{L}^2(\mathbb{F}, \mathbb{R}^{d_2}) \times \mathbb{L}^2(\mathbb{F}, \mathbb{R}^{d_2 \times d})$.*

Proof Uniqueness follows directly from Theorem 4.2.3. In particular, the uniqueness means

$$Y_t^1 = Y_t^2 \text{ for all } t \in [0,T], \ \mathbb{P}\text{-a.s. and } Z_t^1 = Z_t^2, \ dt \times d\mathbb{P}\text{-a.s.} \tag{4.3.1}$$

We now prove the existence by using the Picard iteration. We shall use the local approach similar to that used in the proof of Theorem 3.3.1 and leave the global approach to Exercise. For simplicity we assume $d = d_2 = 1$.

Step 1. Let $\delta > 0$ be a constant which will be specified later, and assume $T \leq \delta$. We emphasize that δ will depend only on the Lipschitz constant L (and the dimensions). In particular, it does not depend on the terminal condition ξ.

Denote $Y_t^0 := 0, Z_t^0 := 0$. For $n = 1, 2, \cdots$, let

$$Y_t^n = \xi + \int_t^T f_s(Y_s^{n-1}, Z_s^{n-1})ds - \int_t^T Z_s^n dB_s. \tag{4.3.2}$$

Assume $(Y^{n-1}, Z^{n-1}) \in \mathbb{L}^2(\mathbb{F}) \times \mathbb{L}^2(\mathbb{F})$. Note that

$$|f_t(Y_t^{n-1}, Z_t^{n-1})| \leq C\Big[|f_t^0| + |Y_t^{n-1}| + |Z_t^{n-1}|\Big].$$

Then, $f_t(Y_t^{n-1}, Z_t^{n-1}) \in \mathbb{L}^{1,2}(\mathbb{F})$. By Proposition 4.1.1, the linear BSDE (4.3.2) uniquely determines $(Y^n, Z^n) \in \mathbb{L}^2(\mathbb{F}) \times \mathbb{L}^2(\mathbb{F})$, and then Theorem 4.2.1 implies further that $(Y^n, Z^n) \in \mathbb{S}^2(\mathbb{F}) \times \mathbb{L}^2(\mathbb{F})$. By induction we have $(Y^n, Z^n) \in \mathbb{S}^2(\mathbb{F}) \times \mathbb{L}^2(\mathbb{F})$ for all $n \geq 0$.

Denote $\Delta Y_t^n := Y_t^n - Y_t^{n-1}, \Delta Z_t^n := Z_t^n - Z_t^{n-1}$. Then,

$$\Delta Y_t^n = \int_t^T [\alpha_s^{n-1}\Delta Y_s^{n-1} + \beta_s^{n-1}\Delta Z_s^{n-1}]ds - \int_t^T \Delta Z_s^n dB_s,$$

where α^n, β^n are defined in a similar way as in (4.2.11) and are bounded by L. Applying Itô formula we have

$$d(|\Delta Y_t^n|^2) = -2\Delta Y_t^n[\alpha_t^{n-1}\Delta Y_t^{n-1} + \beta_t^{n-1}\Delta Z_t^{n-1}]dt + 2\Delta Y_t^n \Delta Z_t^n dB_t + |\Delta Z_t^n|^2 dt.$$

By Problem 2.10.7 (i), $\int_0^t \Delta Y_s^n \Delta Z_s^n dB_s$ is a true martingale. Noting that $\Delta Y_T^n = 0$, we get

$$\begin{aligned}\mathbb{E}\Big[|\Delta Y_t^n|^2 + \int_t^T |\Delta Z_s^n|^2 ds\Big] &= \mathbb{E}\Big[2\int_t^T [\Delta Y_s^n[\alpha_s^{n-1}\Delta Y_s^{n-1} + \beta_s^{n-1}\Delta Z_s^{n-1}]ds\Big] \\ &\leq C\mathbb{E}\Big[\int_0^T |\Delta Y_s^n|[|\Delta Y_s^{n-1}| + |\Delta Z_s^{n-1}|]ds\Big]. \end{aligned} \tag{4.3.3}$$

Thus

$$\begin{aligned}\mathbb{E}\Big[\int_0^T |\Delta Y_t^n|^2 dt\Big] &\leq C\delta\mathbb{E}\Big[\int_0^T |\Delta Y_s^n|[|\Delta Y_s^{n-1}| + |\Delta Z_s^{n-1}|]ds\Big] \\ &\leq C\delta\mathbb{E}\Big[\int_0^T [|\Delta Y_t^n|^2 + |\Delta Y_t^{n-1}|^2 + |\Delta Z_t^{n-1}|^2]dt\Big]\end{aligned}$$

Assume $\delta < \frac{1}{2C}$ for the above constant C and thus $1 - C\delta \le \frac{1}{2}$, then,

$$\mathbb{E}\Big[\int_0^T |\Delta Y_t^n|^2 dt\Big] \le C\delta \mathbb{E}\Big[\int_0^T [|\Delta Y_t^{n-1}|^2 + |\Delta Z_t^{n-1}|^2] dt\Big].$$

Moreover, by setting $t = 0$ in (4.3.3), we have

$$\begin{aligned}\mathbb{E}\Big[\int_0^T |\Delta Z_t^n|^2 dt\Big] &\le C\mathbb{E}\Big[\int_0^T |\Delta Y_s^n|^2 dt\Big] + \frac{1}{8}\mathbb{E}\Big[\int_0^T [|\Delta Y_t^{n-1}|^2 + |\Delta Z_t^{n-1}|^2] dt\Big] \\ &\le \Big[C\delta + \frac{1}{8}\Big]\mathbb{E}\Big[\int_0^T [|\Delta Y_t^{n-1}|^2 + |\Delta Z_t^{n-1}|^2] dt\Big].\end{aligned}$$

Thus

$$\mathbb{E}\Big[\int_0^T [|\Delta Y_t^n|^2 + |\Delta Z_t^n|^2] dt\Big] \le \Big[C\delta + \frac{1}{8}\Big]\mathbb{E}\Big[\int_0^T [|\Delta Y_t^{n-1}|^2 + |\Delta Z_t^{n-1}|^2] dt\Big].$$

Set $\delta := \frac{1}{8C}$ for the above C. Then

$$\mathbb{E}\Big[\int_0^T [|\Delta Y_t^n|^2 + |\Delta Z_t^n|^2] dt\Big] \le \frac{1}{4}\mathbb{E}\Big[\int_0^T [|\Delta Y_t^{n-1}|^2 + |\Delta Z_t^{n-1}|^2] dt\Big].$$

By induction we have

$$\mathbb{E}\Big[\int_0^T [|\Delta Y_t^n|^2 + |\Delta Z_t^n|^2] dt\Big] \le \frac{C}{4^n}, \quad \forall n \ge 1.$$

Now following the arguments in Theorem 3.3.1 one can easily see that there exists $(Y, Z) \in \mathbb{S}^2(\mathbb{F}) \times \mathbb{L}^2(\mathbb{F})$ such that

$$\lim_{n\to\infty} \|(Y_t^n - Y_t, Z_t^n - Z_t)\| = 0.$$

Therefore, by letting $n \to \infty$ in BSDE (4.3.2) we know that (Y, Z) satisfies BSDE (4.0.3).

Step 2. We now prove the existence for arbitrary T. Let $\delta > 0$ be the constant in Step 1. Consider a partition $0 = t_0 < \cdots < t_n = T$ such that $t_{i+1} - t_i \le \delta$, $= 0, \cdots, n-1$. Define $Y_{t_n} := \xi$, and for $i = n-1, \cdots, 0$ and $t \in [t_i, t_{i+1})$, let (Y_t, Z_t) be the solution to the following BSDE on $[t_i, t_{i+1}]$:

$$Y_t = Y_{t_{i+1}} + \int_t^{t_{i+1}} f_s(Y_s, Z_s) ds - \int_t^{t_{i+1}} Z_s dB_s, \quad t \in [t_i, t_{i+1}].$$

Since $t_{i+1} - t_i \le \delta$, by Step 1 the above BSDE is well posed. Moreover, since n is finite here, we see that $(Y, Z) \in \mathbb{L}^2(\mathbb{F}) \times \mathbb{L}^2(\mathbb{F})$, and thus they are a global solution on the whole interval $[0, T]$. ■

Remark 4.3.2 Assume f satisfies Assumption 4.0.1, $\tau \in \mathscr{T}(\mathbb{F})$, and $\xi \in \mathbb{L}^2(\mathscr{F}_\tau)$. Consider the following BSDE

$$Y_t = \xi + \int_t^T \tilde{f}_s(Y_s, Z_s)ds - \int_t^T Z_s dB_s, \quad \text{where} \quad \tilde{f}_s(y, z) := f_s(y, z)\mathbf{1}_{[0,\tau]}(s). \tag{4.3.4}$$

One can easily see that $\tilde{f}$ also satisfies Assumption 4.0.1, and thus the above BSDE has a unique solution. Since $\xi \in \mathscr{F}_\tau$, we see immediately that $Y_s := \xi, Z_s := 0$ satisfy (4.3.4) for $s \in [\tau, T]$. Therefore, we may rewrite (4.3.4) as

$$Y_t = \xi + \int_t^\tau f_s(Y_s, Z_s)ds - \int_t^\tau Z_s dB_s, \quad 0 \le t \le \tau, \tag{4.3.5}$$

and it is also well posed. ■

4.4 Basic Properties of BSDEs

As in Section 3.4, we start with the comparison result, in the case $d_2 = 1$.

Theorem 4.4.1 (Comparison Theorem) *Let $d_2 = 1$. Assume, for $i = 1, 2$, (ξ_i, f^i) satisfies Assumption 4.0.1 and $(Y^i, Z^i) \in \mathbb{S}^2(\mathbb{F}, \mathbb{R}) \times \mathbb{L}^2(\mathbb{F}, \mathbb{R}^{1\times d})$ is the unique solution to the following BSDE:*

$$Y_t^i = \xi_i + \int_t^T f_s^i(Y_s^i, Z_s^i)ds - \int_t^T Z_s^i dB_s. \tag{4.4.1}$$

Assume further that $\xi_1 \le \xi_2$, $\mathbb{P}$-a.s., and $f^1(y, z) \le f^2(y, z)$, $dt \times d\mathbb{P}$–a.s. that for any (y, z). Then,

$$Y_t^1 \le Y_t^2, \quad 0 \le t \le T, \quad \mathbb{P}\text{-a.s.} \tag{4.4.2}$$

Proof Denote

$$\Delta Y_t := Y_t^1 - Y_t^2; \quad \Delta Z_t := Z_t^1 - Z_t^2; \quad \Delta\xi := \xi_1 - \xi_2, \quad \Delta f := f^1 - f^2.$$

Then,

$$\begin{aligned} \Delta Y_t &= \Delta\xi + \int_t^T [f_s^1(Y_s^1, Z_s^1) - f_s^2(Y_s^2, Z_s^2)]ds - \int_t^T \Delta Z_s dB_s \\ &= \Delta\xi + \int_t^T [\alpha_s \Delta Y_s + \Delta Z_s \beta_s + \Delta f_s(Y_s^2, Z_s^2)]ds - \int_t^T \Delta Z_s dB_s, \end{aligned}$$

where α and β are bounded. Define Γ by (4.1.4). By (4.1.3) we have

$$\Delta Y_t = \Gamma_t^{-1}\mathbb{E}\Big[\Gamma_T \Delta\xi + \int_t^T \Gamma_s \Delta f_s(Y_s^2, Z_s^2)ds \Big| \mathscr{F}_t\Big]. \tag{4.4.3}$$

Similar to (3.4.1), by Problem 1.4.6 (ii) we have

$$f^1(y,z) \le f^2(y,z) \text{ for all } (y,z), \quad dt \times d\mathbb{P}\text{-a.s.}$$

This implies that $\Delta f(Y^2, Z^2) \le 0$, $dt \times d\mathbb{P}$-a.s. Since $\Gamma \ge 0$ and $\Delta\xi \le 0$, then (4.4.2) follows from (4.4.3) immediately. ■

Remark 4.4.2 In the Comparison Theorem we require the process Y to be scalar. The comparison principle for general multidimensional BSDEs is an important but very challenging subject. See Problem 4.7.5 for some simple result. ■

We next establish the stability result.

Theorem 4.4.3 (Stability) *Let (ξ, f) and (ξ_n, f^n), $n = 1, 2, \cdots$, satisfy Assumption 4.0.1 with the same Lipschitz constant L, and $(Y, Z), (Y^n, Z^n) \in \mathbb{S}^2(\mathbb{F}, \mathbb{R}^{d_2}) \times \mathbb{L}^2(\mathbb{F}, \mathbb{R}^{d_2 \times d})$ be the solution to the corresponding BSDE* (4.0.3). *Denote*

$$\Delta Y^n := Y^n - Y, \quad \Delta Z^n := Z^n - Z; \quad \Delta\xi_n := \xi_n - \xi, \quad \Delta f^n := f^n - f.$$

Assume

$$\lim_{n\to\infty} \mathbb{E}\Big[|\Delta\xi_n|^2 + \Big(\int_0^T |\Delta f_t^n(0,0)|dt\Big)^2\Big] = 0, \tag{4.4.4}$$

and that $\Delta f^n(y,z) \to 0$ in measure $dt \times d\mathbb{P}$, for all (y,z). Then,

$$\lim_{n\to\infty} \|(\Delta Y^n, \Delta Z^n)\| = 0. \tag{4.4.5}$$

Proof First, by (4.2.10) we have

$$\begin{aligned}
&\|(\Delta Y^n, \Delta Z^n)\|^2 \le C\mathbb{E}\Big[|\Delta\xi_n|^2 + \Big(\int_0^T |\Delta f_t^n(Y_t, Z_t)|dt\Big)^2\Big] \\
&\le C\mathbb{E}\Big[|\Delta\xi_n|^2 + \Big(\int_0^T |\Delta f_t^n(0,0)|dt\Big)^2 + \Big(\int_0^T |\Delta f_t^n(Y_t, Z_t) - \Delta f_t^n(0,0)|dt\Big)^2\Big].
\end{aligned} \tag{4.4.6}$$

By Problem 1.4.6 (iii), $\Delta f^n(Y, Z) \to 0$, in measure $dt \times d\mathbb{P}$. Note that

$$|\Delta f_n(t, Y_t, Z_t) - \Delta f_n(t, 0, 0)| \le C[|Y_t| + |Z_t|].$$

Applying the dominated convergence Theorem we have

$$\lim_{n\to\infty} \mathbb{E}\Big[\Big(\int_0^T |\Delta f_t^n(Y_t, Z_t) - \Delta f_t^n(0,0)|dt\Big)^2\Big] = 0.$$

This, together with (4.4.4) and (4.4.6), leads to the result. ■

We conclude this section by extending the well-posedness result to $\mathbb{L}^p(\mathbb{F})$ for $p \geq 2$.

Theorem 4.4.4 *Assume Assumption 4.0.1 holds and* $\xi \in \mathbb{L}^p(\mathscr{F}_T, \mathbb{R}^{d_2})$, $f^0 \in \mathbb{L}^{1,p}(\mathbb{F}, \mathbb{R}^{d_2})$ *for some* $p \geq 2$. *Let* $(Y, Z) \in \mathbb{S}^2(\mathbb{F}, \mathbb{R}^{d_2}) \times \mathbb{L}^2(\mathbb{F}, \mathbb{R}^{d_2 \times d})$ *be the unique solution to BSDE* (4.0.3). *Then,*

$$\mathbb{E}\Big[|Y_T^*|^p + \Big(\int_0^T |Z_t|^2 dt\Big)^{\frac{p}{2}}\Big] \leq C_p I_p^p, \text{ where } I_p^p := \mathbb{E}\Big[|\xi|^p + \Big(\int_0^T |f_t^0| dt\Big)^p\Big]. \tag{4.4.7}$$

Proof As in Theorem 3.4.3 we proceed in two steps. Again assume $d = d_2 = 1$ for simplicity.

Step 1. We first assume $Y \in \mathbb{L}^{\infty,p}(\mathbb{F}), Z \in \mathbb{L}^{2,p}(\mathbb{F})$ and prove (4.4.7). Applying Itô formula we have

$$\begin{aligned} d|Y_t|^2 &= -2Y_t f_t(Y_t, Z_t)dt + |Z_t|^2 dt + 2Y_t Z_t dB_t; \\ d(|Y_t|^p) &= d(|Y_t|^2)^{\frac{p}{2}} = -p|Y_t|^{p-2} Y_t f_t(Y_t, Z_t)dt + \frac{1}{2}p(p-1)|Y_t|^{p-2}|Z_t|^2 dt + p|Y_t|^{p-2} Y_t Z_t dB_t. \end{aligned} \tag{4.4.8}$$

Following the arguments in Theorem 4.2.1 Steps 1 and 2 one can easily show that, for any $\varepsilon > 0$,

$$\begin{aligned} &\mathbb{E}\Big[|Y_T^*|^p\Big] \leq C_p \sup_{0 \leq t \leq T} \mathbb{E}[|Y_t|^p] + C_p \mathbb{E}\Big[\int_0^T |Y_t|^{p-2}|Z_t|^2 dt\Big] + C_p I_p^p; \\ &\sup_{0 \leq t \leq T} \mathbb{E}[|Y_t|^p] + \mathbb{E}\Big[\int_0^T |Y_t|^{p-2}|Z_t|^2 dt\Big] \leq \varepsilon \mathbb{E}\Big[|Y_T^*|^p\Big] + C_p \varepsilon^{-1} I_p^p. \end{aligned}$$

Then, by choosing $\varepsilon > 0$ small enough we obtain

$$\mathbb{E}\Big[|Y_T^*|^p\Big] \leq C_p I_p^p. \tag{4.4.9}$$

Next, by (4.4.8) we see that

$$\begin{aligned} \int_0^T |Z_t|^2 dt &= |\xi|^2 - |Y_0|^2 + 2\int_0^T Y_t f_t(Y_t, Z_t)dt - 2\int_0^T Y_t Z_t dB_t \\ &\leq C|Y_T^*|^2 + C\int_0^T |Y_t|[|f_t^0| + |Y_t| + |Z_t|]dt + C\Big|\int_0^T Y_t Z_t dB_t\Big| \\ &\leq C|Y_T^*|^2 + C\Big(\int_0^T |f_t^0| dt\Big)^2 + \frac{1}{2}\int_0^T |Z_t|^2 dt + C\Big|\int_0^T Y_t Z_t dB_t\Big|. \end{aligned}$$

Then, by (4.4.9) and Burkholder-Davis-Gundy inequality,

$$\begin{aligned}
&\mathbb{E}\Big[\Big(\int_0^T |Z_t|^2 dt\Big)^{\frac{p}{2}}\Big] \le C_p I_p^2 + C_p \mathbb{E}\Big[|Y_T^*|^p + \Big|\int_0^T Y_t Z_t dB_t\Big|^{\frac{p}{2}}\Big] \\
&\le C_p I_p^2 + C_p \mathbb{E}\Big[\Big(\int_0^T |Y_t Z_t|^2 dt\Big)^{\frac{p}{4}}\Big] \le C_p I_p^2 + C_p \mathbb{E}\Big[|Y_T^*|^{\frac{p}{2}}\Big(\int_0^T |Z_t|^2 dt\Big)^{\frac{p}{4}}\Big] \\
&\le C_p I_p^2 + C_p \mathbb{E}[|Y_T^*|^p] + \frac{1}{2}\mathbb{E}\Big[\Big(\int_0^T |Z_t|^2 dt\Big)^{\frac{p}{2}}\Big] \le C_p I_p^2 + \frac{1}{2}\mathbb{E}\Big[\Big(\int_0^T |Z_t|^2 dt\Big)^{\frac{p}{2}}\Big].
\end{aligned}$$

This leads to the desired estimate for Z, and together with (4.4.9), proves further (4.4.7).

Step 2. In the general case, we shall use the space truncation arguments in Theorem 3.4.3. We note that the time truncation does not work well here because it will involve Y_{τ_n} which still lacks desired integrability. For each $n \ge 1$, denote $\xi_n := (-n) \vee \xi \wedge n, f_n := (-n) \vee f \wedge n$. Clearly (ξ_n, f^n) satisfy all the conditions of this theorem with the same Lipschitz constant L, and

$$(\xi_n, f^n) \to (\xi, f), \quad |\xi_n| \le |\xi|, |f^n| \le |f|, \quad |\xi_n| \le n, |f^n| \le n, \quad \text{for all } (t, \omega, y, z).$$

Let $(Y^n, Z^n) \in \mathbb{S}^2(\mathbb{F}) \times \mathbb{L}^2(\mathbb{F})$ be the unique solution to BSDE (4.0.3) with coefficients (ξ_n, f^n). Then

$$Y_t^n = \mathbb{E}\Big[\xi_n + \int_t^T f_s^n(Y_s^n, Z_s^n) ds \Big| \mathscr{F}_t\Big], \quad \int_0^t Z_s^n dB_s = Y_t^n - Y_0^n + \int_0^t f_s^n(Y_s^n, Z_s^n) ds$$

are bounded. By the Burkholder-Davis-Gundy inequality, this implies further that $Z^n \in \mathbb{L}^{2,p}(\mathbb{F})$. Then it follows from Step 1 that

$$\mathbb{E}\Big[|(Y^n)_T^*|^p + \Big(\int_0^T |Z_t^n|^2 dt\Big)^{\frac{p}{2}}\Big] \le C_p \mathbb{E}\Big[|\xi_n|^p + \Big(\int_0^T |f_t^n(0,0)| dt\Big)^p\Big] \le C_p I_p^p.$$

Now similar to the arguments in Theorem 4.2.1, (4.4.7) follows from Theorem 4.4.3 and Fatou lemma. ■

4.5 Some Applications of BSDEs

The theory of BSDEs has wide applications in many fields, most notably in mathematical finance, stochastic control theory, and probabilistic numerical methods for nonlinear PDEs. We shall discuss its connection with PDE rigorously in the next chapter. In this section we present the first two types of applications in very simple settings and in a heuristic way, just to illustrate the idea.

4.5.1 Application in Asset Pricing and Hedging Theory

Consider the Black-Scholes model in Section 2.8. Assume a self-financing portfolio (λ, h) hedges ξ. By (2.8.8) and (2.8.7) we have:

$$\begin{aligned} dV_t &= \Big[\lambda_t r e^{rt} + h_t S_t \mu\Big]dt + h_t S_t \sigma dB_t \\ &= \Big[r(V_t - h_s S_t) + h_t S_t \mu\Big]dt + h_t S_t \sigma dB_t. \end{aligned} \tag{4.5.1}$$

Denote

$$Y_t := V_t, \quad Z_t := \sigma S_t h_t. \tag{4.5.2}$$

Then (4.5.1) leads to

$$dY_t = \Big[r[Y_t - \frac{Z_t}{\sigma S_t}] + \frac{\mu Z_t}{\sigma S_t}\Big]dt + Z_t dB_t, \quad Y_T = \xi, \quad \mathbb{P}\text{-a.s.} \tag{4.5.3}$$

This is a linear BSDE. Once we solve it, we obtain that:

$$Y \text{ is the price of the option } \xi \text{ and Z induces the hedging portfolio: } h_t = \frac{Z_t}{\sigma S_t}. \tag{4.5.4}$$

We remark that BSDE (4.5.3) is under the market measure $\mathbb{P}$. In this approach, there is no need to talk about the risk neutral measure.

Note that BSDE (4.5.3) is linear, which can be solved explicitly. In particular, for the special example we are presenting, Y_0 can be computed via the well-known Black-Scholes formula. To motivate nonlinear BSDEs, let us assume in a more practical manner that the lending interest rate r_1 is less than the borrowing interest rate r_2. That is, the self-financing condition (4.5.1) should be replaced by

$$dV_t := \Big[r_1(V_t - h_t S_t)^+ - r_2(V_t - h_t S_t)^-\Big]dt + h_t dS_t, \tag{4.5.5}$$

and therefore, BSDE (4.5.3) becomes a nonlinear one:

$$dY_t = \Big[r_1(Y_t - \frac{Z_t}{\sigma S_t})^+ - r_2(Y_t - \frac{Z_t}{\sigma S_t})^- + \frac{\mu Z_t}{\sigma S_t}\Big]dt + Z_t dB_t, \quad Y_T = \xi. \tag{4.5.6}$$

Nonlinear BSDEs typically do not have explicit formula. We shall discuss its numerical method in the next chapter.

4.5.2 Applications in Stochastic Control

Consider a controlled SDE:

$$X_t^k = x + \int_0^t b(s, X_s^k, k_s)ds + \int_0^t \sigma(s, X_s^k, k_s)dB_s, \quad 0 \le t \le T, \ \mathbb{P}\text{-a.s.} \tag{4.5.7}$$

Here B, X, b, σ take values in $\mathbb{R}^d$, $\mathbb{R}^{d_1}$,$\mathbb{R}^{d_1}$, and $\mathbb{R}^{d_1 \times d}$, respectively, and $k \in \mathscr{K}$ are admissible controls. We assume k takes values in certain Polish space $\mathbb{K}$ and is $\mathbb{F}$-measurable. Our goal is the following stochastic optimization problem (with superscript ^{S} indicating strong formulation in contrast to the weak formulation in (4.5.12) below):

$$V_0^S := \sup_{k \in \mathscr{K}} J_S(k) \quad \text{where} \quad J_S(k) := \mathbb{E}^{\mathbb{P}}\Big[g(X_T^k) + \int_0^T f(t, X_t^k, k_t)dt\Big], \tag{4.5.8}$$

where f and g are 1-dimensional and thus J_S and V_0^S are scalars.

If we follow the standard stochastic maximum principle, the above problem will lead to a forward-backward SDE, which is the main subject of Chapter 8 and is in general not solvable. We thus transform the problem to weak formulation as follows. We remark that the weak formulation, especially when there is diffusion control (namely σ depends on k), will be our main formulation for stochastic control problems and will be explored in details in Part III. Here we just present some very basic ideas. For this purpose, we assume

Assumption 4.5.1

(i) b, σ, f, g are deterministic, Borel measurable in all variables, and bounded (for simplicity);
(ii) $\sigma = \sigma(t, x)$ does not contain the control k, and is uniformly Lipschitz in x;
(iii) There exists a bounded $\mathbb{R}^d$-valued function $\theta(t, x, k)$ such that $b(t, x, k) = \sigma(t, x)\theta(t, x, k)$.

We note that, when $d = d_1$ and $\sigma \in \mathbb{S}^d$ is invertible, it is clear that $\theta(t, x, k) = \sigma^{-1}(t, x)b(t, x, k)$ and is unique.

Let X be the unique solution to the following SDE (without control):

$$X_t = x + \int_0^t \sigma(s, X_s)dB_s, \quad 0 \le t \le T, \ \mathbb{P}\text{-a.s.} \tag{4.5.9}$$

For each $k \in \mathscr{K}$, recall the notations in Section 2.6 and denote

$$\theta_t^k := \theta(t, X_t, k_t), \quad B_t^k := B_t - \int_0^t \theta_s^k ds, \quad M^k := M^{\theta^k}, \quad \mathbb{P}^k := \mathbb{P}^{\theta^k}. \tag{4.5.10}$$

Under Assumption 4.5.1 (iii), θ^k is bounded and thus it follows from the Girsanov Theorem that B^k is a $\mathbb{P}^k$-Brownian motion. Since $\mathbb{P}^k$ is equivalent to $\mathbb{P}$, then (4.5.9) leads to

$$X_t = x + \int_0^t b(s, X_s, k_s)ds + \int_0^t \sigma(s, X_s)dB^k_s, \quad 0 \le t \le T, \ \mathbb{P}^k\text{-a.s.} \tag{4.5.11}$$

Compare (4.5.11) with (4.5.7), we modify (4.5.8) as

$$V_0 := \sup_{k \in \mathscr{K}} J(k), \quad \text{where} \quad J(k) := \mathbb{E}^{\mathbb{P}^k}\Big[g(X_T) + \int_0^T f(t, X_t, k_t)dt\Big]. \tag{4.5.12}$$

This is the stochastic optimization problem under weak formulation (with drift control only).

Remark 4.5.2

(i) In strong formulation (4.5.8), $\mathbb{P}$ is fixed and one controls the state process X^k, while in weak formulation (4.5.8), the state process X is fixed and one controls the probability $\mathbb{P}^k$, or more precisely controls the distribution of X.
(ii) Although formally (4.5.11) looks very much like (4.5.7), the $\mathbb{P}^k$-distribution of k is different from the $\mathbb{P}$-distribution of k, then the joint $\mathbb{P}^k$-distribution of (B^k, k, X) is different from the joint $\mathbb{P}$-distribution of (B, k, X^k). Consequently, for given $k \in \mathscr{K}$, typically $J(k) \neq J_S(k)$.
(iii) In most interesting applications, it holds that $V_0^S = V_0$. However, in general it is possible that they are not equal. Nevertheless, in this section we investigate V_0. This is partially because the optimization problem (4.5.12) is technically easier, and more importantly because the weak formulation is more appropriate in many applications, as we discuss next.
(iv) As discussed in Section 2.8.3, in many applications one can actually observe the state process X, rather than the noise B. So it makes more sense to assume the control k depends on X, instead of on B (or ω). That is, weak formulation is more appropriate than strong formulation in many applications, based on the information one observes. In this case, of course, we shall either restrict $\mathscr{K}$ to $\mathbb{F}^X$-measurable processes or assume $\mathbb{F}^X = \mathbb{F}^B$ (e.g., when $d = d_1$ and $\sigma > 0$).
(v) Even when $V_0^S = V_0$, it is much more likely to have the existence of optimal control in weak formulation than in strong formulation. See Remark 4.5.4 below. ■

We now solve (4.5.12). For each $k \in \mathscr{K}$, applying Theorem 2.6.6, the martingale representation theorem under Girsanov setting, one can easily see that the following linear BSDE under $\mathbb{P}^k$ has a unique solution (Y^k, Z^k):

$$Y^k_t = g(X_T) + \int_t^T f(s, X_s, k_s)ds - \int_t^T Z^k_s dB^k_s, \quad \mathbb{P}^k\text{-a.s.} \tag{4.5.13}$$

Clearly $J(k) = Y_0^k$. By (4.5.10) and noting that $\mathbb{P}^k$ and $\mathbb{P}$ are equivalent, we may rewrite (4.5.13) as

$$Y_t^k = g(X_T) + \int_t^T \Big[f(s, X_s, k_s) + Z_s^k \theta(s, X_s, k_s)\Big] ds - \int_t^T Z_s^k dB_s, \quad \mathbb{P}\text{-a.s.} \tag{4.5.14}$$

Define the Hamiltonians:

$$H^*(t, x, z) := \sup_{k \in \mathbb{K}} H(t, x, z, k), \quad \text{where} \quad H(t, x, z, k) := f(t, x, k) + z\theta(t, x, k). \tag{4.5.15}$$

By Assumption 4.5.1 (iii) and (i), H^* is uniformly Lipschitz continuous in z and $H^*(t, x, 0)$ is bounded. Then the following BSDE has a unique solution (Y^*, Z^*):

$$Y_t^* = g(X_T) + \int_t^T H^*(s, X_s, Z_s^*) ds - \int_t^T Z_s^* dB_s, \quad \mathbb{P}\text{-a.s.} \tag{4.5.16}$$

We have the following main result for this subsection.

Theorem 4.5.3 *Under Assumption 4.5.1, we have*

$$V_0 = Y_0^*. \tag{4.5.17}$$

Moreover, if there exists a Borel measurable function $I : [0, T] \times \mathbb{R}^{d_1} \times \mathbb{R}^d \to \mathbb{K}$ *such that*

$$H^*(t, x, z) = H(t, x, z, I(t, x, z)). \tag{4.5.18}$$

Then

$$k_t^* := I(t, X_t, Z_t^*) \textit{ is an optimal control.} \tag{4.5.19}$$

Proof First, applying comparison theorem, we have $Y_0^k \le Y_0^*$ for all $k \in \mathscr{K}$, and thus $V_0 \le Y_0^*$. On the other hand, for any $\varepsilon > 0$, by standard measurable selection there exists a Borel measurable function $I^\varepsilon : [0, T] \times \mathbb{R}^{d_1} \times \mathbb{R}^d \to \mathbb{K}$ such that

$$H^*(t, x, z) \le H(t, x, z, I^\varepsilon(t, x, z)) + \varepsilon.$$

Denote $k_t^\varepsilon := I^\varepsilon(t, X_t, Z_t^*)$, and thus $H^*(t, X_t, Z_t^*) \le H(t, X_t, Z_t^*, k_t^\varepsilon) + \varepsilon$. Note that

$$Y_t^{k^\varepsilon} = g(X_T) + \int_t^T H(s, X_s, Z_s^{k^\varepsilon}, k_s^\varepsilon) ds - \int_t^T Z_s^{k^\varepsilon} dB_s.$$

Denote $\Delta Y^\varepsilon := Y^* - Y^{k^\varepsilon}$, $\Delta Z^\varepsilon := Z^* - Z^{k^\varepsilon}$. Then

$$\begin{aligned}\Delta Y_t^\varepsilon &= \int_t^T \Big[H^*(s, X_s, Z_s^*) - H(s, X_s, Z_s^*, k_s^\varepsilon) + \Delta Z_s^\varepsilon \theta(s, X_s, k_s^\varepsilon)\Big]ds - \int_t^T \Delta Z_s^\varepsilon dB_s \\ &= \int_t^T \Big[H^*(s, X_s, Z_s^*) - H(s, X_s, Z_s^*, k_s^\varepsilon)\Big]ds - \int_t^T \Delta Z_s^\varepsilon dB_s^{k^\varepsilon} \le \varepsilon(T-t) - \int_t^T \Delta Z_s^\varepsilon dB_s^{k^\varepsilon}.\end{aligned}$$

This implies that $\Delta Y_0^\varepsilon \le T\varepsilon$. Since $\varepsilon > 0$ is arbitrary, we obtain $Y_0^* \le V_0$, and hence the equality holds.

Finally, under (4.5.18) it is clear that $Y^* = Y^{k^*}$, which implies (4.5.19) immediately. ■

Remark 4.5.4 We emphasize that the optimal control k^* in (4.5.19) is optimal in weak formulation, but not necessarily in strong formulation. To illustrate the main idea, let us consider a special case: $d = d_1 = 1$, $\sigma = 1$, $x = 0$, and then $X = B$. Since k^* is $\mathbb{F}^B$-measurable, so we may write $k^* = k^*(B) = k^*(X)$. Assume $V_0^S = V_0$, then the above k^* provides an optimal control in strong formulation amounts to say the following SDE admits a strong solution:

$$X_t = x + \int_0^t b(s, X_s, k_s^*(X_\cdot))ds + \int_0^t \sigma(s, X_s)dB_s, \quad \mathbb{P}\text{-a.s.} \tag{4.5.20}$$

We remark that, in this special case here, actually one can show that $k_t^* = k^*(t, X_t^*)$ depends only on X_t^*. However, k^* may be discontinuous in X, and thus it is difficult to establish a general theory for the strong solvability of SDE (4.5.20). Moreover, one may easily extend Theorem 4.5.3 to the path dependent case, namely b, f, and/or g depend on the paths of X. In this case k^* may also depend on the paths of X^* and thus (4.5.20) becomes path dependent. Typically this SDE does not have a strong solution, see a counterexample in Wang & Zhang [231] which is based on Tsirelson's [229] counterexample. Consequently, the optimization problem (4.5.8) (or its extension to path dependent case) in strong formulation may not have an optimal control. ■

4.6 Bibliographical Notes

The linear BSDE was first proposed by Bismut [16], motivated from applications in stochastic control, and the well-posedness of nonlinear BSDEs was established by the seminal paper Paradox & Peng [167]. There is an excellent exposition on the basic theory and applications of BSDEs in El Karoui, Peng, & Quenez [81], and Peng [182] provides a detailed survey on the theory and its further developments. Another application which independently leads to the connection with BSDE is the recursive utility proposed by Duffie and Epstein [69, 70]. We also refer to some book chapters El Karoui & Mazliak [80], Peng [175], Yong & Zhou [242], Pham

[190], Cvitanic & Zhang [52], Touzi [227], as well as the recent book Pardoux & Rascanu [170] on theory and applications of BSDEs. In particular, many materials of this and the next chapter follow from the presentation in [52].

We note that the materials in this chapter are very basic. There have been various extensions of the theory, with some important ones presented in the next chapter and Part II. The further extension to fully nonlinear situation is the subject of Part III. Besides those and among many others, we note that Lepeltier & San Martin [135] studied BSDEs with non-Lipschitz continuous generators, Tang & Li [223] studied BSDEs driven by jump processes, Fuhrman & Tessitore [95] studied BSDEs in infinite dimensional spaces, Darling & Pardoux [54] studied BSDEs with random terminal time, Buckdahn, Engelbert, & Rascanu [24] studied weak solutions of BSDEs, and Pardoux & Peng [169] studied backward doubly SDEs which provides a representation for solutions to (forward) stochastic PDEs. Moreover, we note that Hu & Peng [110] provided some general result concerning comparison principle for multidimensional BSDEs, and Hamadene & Lepeltier [103] extended the stochastic optimization problem to a zero-sum stochastic differential game problem, again in weak formulation. Another closely related concept is the g-expectation developed by Peng [176, 179], see also Coquet, Hu, Memin, & Peng [38], Chen & Epstein [30], and Delbaen, Peng, & Rosazza Gianin [53]. This is a special type of the nonlinear expectation which we will introduce in Chapter 10.

4.7 Exercises

Problem 4.7.1 Similar to Problem 3.7.2, this problem consider the decoupling strategy for multidimensional linear BSDE. For simplicity, we consider the following linear BSDE with $d = 1$ and $d_2 = 2$:

$$Y_t^i = \xi_i + \int_0^t \Big[\sum_{j=1}^{2}[\alpha_s^{ij} Y_s^j + \beta_s^{ij} Z_s^j] + \gamma_s^i\Big] ds + \int_0^t Z_s^i dB_s, \quad i = 1, 2. \quad (4.7.1)$$

Here $\xi_i \in \mathbb{L}^2(\mathscr{F}_T, \mathbb{R})$, $\alpha^{ij}, \beta^{ij} \in \mathbb{L}^\infty(\mathbb{F}, \mathbb{R})$, and $\gamma^i \in \mathbb{L}^{1,2}(\mathbb{F}, \mathbb{R})$. Show that there exists a process Γ such that $\overline{Y} := Y^1 + \Gamma Y^2$ solves a one-dimensional BSDE, whose coefficients may depend on Γ. ■

Problem 4.7.2

(i) Provide an alternative proof for Theorem 4.3.1 by using the global approach similar to that used in the proof of Theorem 3.3.1. (Hint: first provide a priori estimate for $\|(Y, Z)\|_\lambda^2 := \sup_{0\le t\le T} \mathbb{E}[e^{\lambda t}|Y_t|^2] + \mathbb{E}\Big[\int_0^T e^{\lambda t}|Z_t|^2 dt\Big]$ for some $\lambda > 0$ large enough.)

(ii) Provide another proof for Theorem 4.3.1 by using contraction mapping. That is, define a mapping $F : \mathbb{L}^2(\mathbb{F}, \mathbb{R}^{d_2}) \times \mathbb{L}^2(\mathbb{F}, \mathbb{R}^{d_2\times d}) \to \mathbb{L}^2(\mathbb{F}, \mathbb{R}^{d_2}) \times \mathbb{L}^2(\mathbb{F}, \mathbb{R}^{d_2\times d})$

by $F(Y, Z) := (\tilde{Y}, \tilde{Z})$, where

$$\tilde{Y}_t := \xi + \int_t^T f_s(Y_s, Z_s)ds - \int_t^T Z_s dB_s.$$

Show that F is a contraction mapping under the norm $\|(Y, Z)\|_\lambda^2$ for $\lambda > 0$ large enough. ■

Problem 4.7.3 Show that the result of Theorem 4.3.1 still holds if, in Assumption 4.0.1, the Lipschitz continuity of f in y is replaced with the following slightly weaker monotonicity condition:

$$[f_t(y_1, z) - f_t(y_2, z)] \cdot [y_1 - y_2] \le L|y_1 - y_2|^2, \quad \forall (t, \omega), y_1, y_2, z.$$

■

Problem 4.7.4 Let f satisfy Assumption 4.0.1 (i), (ii), (iv), and the linear growth condition (4.2.9).

(i) If $\xi \in \mathbb{L}^2(\mathscr{F}_T, \mathbb{R}^{d_2})$ and $(Y, Z) \in \mathbb{L}^2(\mathbb{F}, \mathbb{R}^{d_2}) \times \mathbb{L}^2(\mathbb{F}, \mathbb{R}^{d_2 \times d})$ is a solution to BSDE (4.0.3). Show that (Y, Z) satisfies the a priori estimate (4.2.1).

(ii) Let $(Y^n, Z^n) \in \mathbb{L}^2(\mathbb{F}, \mathbb{R}^{d_2}) \times \mathbb{L}^2(\mathbb{F}, \mathbb{R}^{d_2 \times d})$ be a solution to BSDE (4.0.3) with terminal condition $\xi_n \in \mathbb{L}^2(\mathscr{F}_T, \mathbb{R}^{d_2})$. Assume $\lim_{n\to\infty} \mathbb{E}\Big[|\xi_n - \xi|^2 + \int_0^T |Y_t^n - Y_t|^2 dt\Big] = 0$ for some $\xi \in \mathbb{L}^2(\mathscr{F}_T, \mathbb{R}^{d_2})$ and $Y \in \mathbb{L}^2(\mathbb{F}, \mathbb{R}^{d_2})$. Show that there exists $Z \in \mathbb{L}^2(\mathbb{F}, \mathbb{R}^{d_2 \times d})$ such that $\lim_{n\to\infty} \mathbb{E}\Big[\int_0^T |Z_t^n - Z_t|^2 dt\Big] = 0$ and (Y, Z) is a solution to BSDE (4.0.3) with terminal condition ξ.

(iii) Assume $d_2 = 1$, $\xi \in \mathbb{L}^2(\mathscr{F}_T, \mathbb{R})$, and f is continuous in (y, z). Show that BSDE (4.0.3) has a solution $(Y, Z) \in \mathbb{S}^2(\mathbb{F}, \mathbb{R}) \times \mathbb{L}^2(\mathbb{F}, \mathbb{R}^{1\times d})$.

(iv) Under the conditions in (iii), find a counterexample such that the BSDE has multiple solutions. ■

Problem 4.7.5

(i) Find a counterexample for comparison principle of multidimensional BSDEs. To be precise, let $d_2 = 2, d = 1$, (ξ, f, Y, Z) be as in (4.0.3), and $(\tilde{\xi}, \tilde{f}, \tilde{Y}, \tilde{Z})$ be another system. We want $\xi_i \le \tilde{\xi}_i$ and $f^i \le \tilde{f}^i$, $i = 1, 2$, but it does not hold that $Y^i \le \tilde{Y}^i$, $i = 1, 2$.

(ii) Prove the comparison for the following special multidimensional BSDE. Let (ξ, f) and $(\tilde{\xi}, \tilde{f})$ satisfy Assumption 4.0.1, and (Y, Z), $(\tilde{Y}, \tilde{Z})$ be the corresponding solution to BSDE (4.0.3). Assume

$$\xi_i \le \tilde{\xi}_i, \quad f^i \le \tilde{f}^i, \quad i = 1, \cdots, d_2.$$

Moreover, for $i = 1, \cdots, d_2$, assume f^i does not depend on z_j and is increasing in y_j for all $j \neq i$. Show that

$$Y^i_t \leq \tilde{Y}^i_t, \quad 0 \leq t \leq T, \quad \mathbb{P}\text{-a.s.}, \quad i = 1, \cdots, d_2.$$

■

Problem 4.7.6 This problem extends the optimization problem in Subsection 4.5.2 to a game problem, still in weak formulation. Assume $\mathbb{K} = \mathbb{K}_1 \times \mathbb{K}_2$, its elements are denoted as $k = (k_1, k_2)$, and denote $\mathscr{K}_1, \mathscr{K}_2$ in obvious sense. Assume Assumption 4.5.1 holds true. Denote

$$\overline{H}(t,x,z) := \inf_{k_1 \in \mathbb{K}_1} \sup_{k_2 \in \mathbb{K}_2} H(t,x,z,u), \quad \underline{H}(t,x,z) := \sup_{k_2 \in \mathbb{K}_2} \inf_{k_1 \in \mathbb{K}_1} H(t,x,z,u), \tag{4.7.2}$$

and let $(\overline{Y}, \overline{Z})$, $(\underline{Y}, \underline{Z})$ denote the solution to the following BSDEs:

$$\begin{aligned} \overline{Y}_t &= g(X_T) + \int_t^T \overline{H}(s, X_s, \overline{Z}_s)ds - \int_t^T \overline{Z}_s dB_s, \\ \underline{Y}_t &= g(X_T) + \int_t^T \underline{H}(s, X_s, \underline{Z}_s)ds - \int_t^T \underline{Z}_s dB_s, \end{aligned} \quad \mathbb{P}\text{-a.s.} \tag{4.7.3}$$

(i) Show that

$$\overline{Y}_0 = \inf_{k_1 \in \mathscr{K}_1} \sup_{k_2 \in \mathscr{K}_2} Y_0^{k_1,k_2}, \quad \underline{Y}_0 = \sup_{k_2 \in \mathscr{K}_2} \inf_{k_1 \in \mathscr{K}_1} Y_0^{k_1,k_2} \tag{4.7.4}$$

Moreover, if the following Isaacs condition holds:

$$\overline{H} = \underline{H} =: H^*, \tag{4.7.5}$$

then the game value exists, namely

$$\inf_{k_1 \in \mathscr{K}_1} \sup_{k_2 \in \mathscr{K}_2} Y_0^{k_1,k_2} = \sup_{k_2 \in \mathscr{K}_2} \inf_{k_1 \in \mathscr{K}_1} Y_0^{k_1,k_2} = Y_0^*, \tag{4.7.6}$$

where Y^* is the solution to BSDE (4.5.16) with the generator H^* defined by (4.7.5).

(ii) Assume further that there exists Borel measurable functions $I_1(t,x,z) \in \mathbb{K}_1$ and $I_2(t,x,z) \in \mathbb{K}_2$ such that, for all (t,x,z) and all $(k_1,k_2) \in \mathbb{K}_1 \times \mathbb{K}_2$,

$$H(t,x,z,k_1,I_2(t,x,z)) \geq H(t,x,z,I_1(t,x,z),I_2(t,x,z)) \geq H(t,x,z,I_1(t,x,z),k_2). \tag{4.7.7}$$

Then Isaacs condition (4.7.5) holds and the game has a saddle point:

$$k_t^{1,*} := I_1(t, X_t, Z_t^*), \quad k_t^{2,*} := I_2(t, X_t, Z_t^*), \tag{4.7.8}$$

where H^*, Y^*, Z^* are as in (i). Here the saddle point, also called equilibrium, means:

$$Y_0^{k^1, k^{2,*}} \geq Y_0^* \geq Y_0^{k^{1,*}, k^2}, \quad \forall (k^1, k^2) \in \mathscr{K}_1 \times \mathscr{K}_2. \tag{4.7.9}$$

■

http://www.springer.com/978-1-4939-7254-8

Backward Stochastic Differential Equations
From Linear to Fully Nonlinear Theory
Zhang, J.
2017, XVI, 388 p., Hardcover
ISBN: 978-1-4939-7254-8