

## Chapter 2

# Basics of Stochastic Calculus

Let  $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P})$  be a filtered probability space. We remark again that, unlike in standard literature, we do not assume  $\mathbb{F} = \{\mathcal{F}_t\}_{0 \leq t \leq T}$  satisfy the usual hypothesis. This will be crucial for the fully nonlinear theory in Part III, and for fixed  $\mathbb{P}$  this is a very mild relaxation due to Proposition 1.2.1.

## 2.1 Brownian Motion

### 2.1.1 Definition

**Definition 2.1.1** We say a process  $B : [0, T] \times \Omega \rightarrow \mathbb{R}$  is a (standard) Brownian motion if

- $B_0 = 0$ , a.s.
- For any  $0 = t_0 < \dots < t_n \leq T$ ,  $B_{t_1}, B_{t_1, t_2}, \dots, B_{t_{n-1}, t_n}$  are independent.
- For any  $0 \leq s < t \leq T$ ,  $B_{s,t} \sim N(0, t - s)$ .

Moreover, we call  $B$  an  $\mathbb{F}$ -Brownian motion if  $B \in \mathbb{L}^0(\mathbb{F})$  and

- For any  $0 \leq s < t \leq T$ ,  $B_{s,t}$  and  $\mathcal{F}_s$  are independent.

We note that as in the previous chapter we restrict  $B$  to a finite horizon  $[0, T]$ . But the definition can be easily extended to  $[0, \infty)$ , by first extending the filtration  $\mathbb{F}$  to  $[0, \infty)$ . When necessary, we may interpret  $B$  as a Brownian motion on  $[0, \infty)$  without mentioning it explicitly. Moreover, when there is a need to emphasize the dependence on the probability measure  $\mathbb{P}$  and/or the filtration  $\mathbb{F}$ , we call  $B$  a  $\mathbb{P}$ -Brownian motion or  $(\mathbb{P}, \mathbb{F})$ -Brownian motion. Since  $B$  has independent increments,

clearly  $(B_{t_1}, \dots, B_{t_n})$  have Gaussian distribution, or say  $B$  is a Gaussian process. Moreover, from the definition we can easily compute the finite distribution of  $B$ . Then by the Kolmogorov's Extension Theorem we know that Brownian motion does exist. The following properties are immediate and left to the readers.

**Proposition 2.1.2** *Let  $B$  be a standard Brownian motion. For any  $t_0$  and any constant  $c > 0$ , the processes  $B_t^{t_0} := B_{t_0, t+t_0}$  and  $\hat{B}_t^c := \frac{1}{\sqrt{c}}B_{ct}$  are also standard Brownian motions.*

**Proposition 2.1.3** *A Brownian motion is Markov, and an  $\mathbb{F}$ -Brownian motion is an  $\mathbb{F}$ -martingale.*

In the multidimensional case, we call  $B = (B^1, \dots, B^d)^\top$  a  $d$ -dimensional Brownian motion if  $B^1, \dots, B^d$  are independent Brownian motions. In most cases we do not emphasize the dimension and thus still call it a Brownian motion.

From now on, throughout this chapter,  $B$  is a  $d$ -dimensional  $\mathbb{F}$ -Brownian motion. All our results hold true in multidimensional setting. However, while we shall state the results in multidimensional case, for notional simplicity quite often we will carry out the proofs only in the case  $d = 1$ . The readers may extend the arguments to multidimensional cases straightforwardly.

## 2.1.2 Pathwise Properties

We start with its pathwise continuity. Notice that Brownian motion is defined via its distribution. As mentioned in the paragraph after Theorem 1.2.3, the pathwise properties should be understood for a version of  $B$ .

**Theorem 2.1.4** *For any  $\varepsilon \in (0, \frac{1}{2})$ ,  $B$  is Hölder- $(\frac{1}{2} - \varepsilon)$  continuous, a.s. In particular,  $B$  is continuous, a.s.*

*Proof* For notational simplicity, assume  $d = 1$ . For any  $s < t$ , since  $B_{s,t} \sim N(0, t - s)$ , we have

$$\mathbb{E}[|B_{s,t}|^p] = C_p |t - s|^{\frac{p}{2}}, \text{ for all } p \geq 1.$$

Apply the Kolmogorov's Continuity Theorem 1.2.3, by considering a modification if necessary,  $B$  is Hö- $\gamma$  continuous for  $\gamma := \frac{\frac{p}{2}-1}{p-1}$ . Since  $p$  is arbitrary, one can always find  $p$  large enough so that  $\gamma > \frac{1}{2} - \varepsilon$ . ■

From now on, we shall always consider a continuous version of  $B$ . We next study the quadratic variation of  $B$ . For a time partition  $\pi : 0 = t_0 < \dots < t_n = T$ , denote  $|\pi| := \max_{1 \leq i \leq n} (t_i - t_{i-1})$ . We recall that the *total variation* of a process  $X \in \mathbb{L}^0(\mathbb{F}, \mathbb{R}^d)$  is defined pathwise by: for  $0 \leq a < b \leq T$ ,

$$\bigvee_a^b(X) := \sup_{\pi} \sum_{i=1}^n |X_{a \vee t_{i-1} \wedge b, a \vee t_i \wedge b}|, \text{ in particular, } \bigvee_0^T(X) := \sup_{\pi} \sum_{i=1}^n |X_{t_{i-1}, t_i}|. \quad (2.1.1)$$

**Definition 2.1.5** Let  $X \in \mathbb{L}^0(\mathbb{F}, \mathbb{R}^d)$ . We say  $X$  has quadratic variation if the following limit exists :

$$\langle X \rangle_t := \lim_{|\pi| \rightarrow 0} \sum_{i=1}^n X_{t_{i-1} \wedge t, t_i \wedge t} X_{t_{i-1} \wedge t, t_i \wedge t}^\top, \quad \text{in the sense of convergence in probability.} \quad (2.1.2)$$

In this case we call  $\langle X \rangle$  the quadratic variation process of  $X$ .

Note that  $\langle X \rangle$  takes values in  $\mathbb{S}^d$ , the set of  $d \times d$ -symmetric matrices. Its  $(i, j)$ -th component is:

$$\langle X \rangle_t^{ij} := \lim_{|\pi| \rightarrow 0} \sum_{k=1}^n X_{t_{k-1} \wedge t, t_k \wedge t}^i X_{t_{k-1} \wedge t, t_k \wedge t}^j.$$

We also remark that, unlike total variation, the quadratic variation is not defined in a pathwise manner. It is interesting to understand the pathwise definition of quadratic variation, which we will study in Part III. See also Remark 2.2.6.

**Theorem 2.1.6** It holds that

$$\lim_{|\pi| \rightarrow 0} \mathbb{E} \left[ \left( \sum_{i=1}^n B_{t_{i-1} \wedge t, t_i \wedge t} B_{t_{i-1} \wedge t, t_i \wedge t}^\top - tI_d \right)^2 \right] = 0, \quad \text{and consequently,} \quad \langle B \rangle_t = tI_d.$$

*Proof* For notational simplicity we assume  $d = 1$ , and without loss of generality we prove the theorem only at  $T$ . Fix a partition  $\pi : 0 = t_0 < \dots < t_n = T$ , and denote

$$\Delta t_i := t_i - t_{i-1}, \quad \eta_i := |B_{t_{i-1}, t_i}|^2 - \Delta t_i, \quad i = 1, \dots, n.$$

Then  $\eta_i, i = 1, \dots, n$ , are independent. Since  $B_{t_{i-1}, t_i} \sim N(0, \Delta t_i)$ , we have  $\mathbb{E}[\eta_i] = 0$  and

$$\text{Var}(\eta_i) = \text{Var}(|B_{t_{i-1}, t_i}|^2) = \mathbb{E}[|B_{t_{i-1}, t_i}|^4] - \left( \mathbb{E}[|B_{t_{i-1}, t_i}|^2] \right)^2 = 3(\Delta t_i)^2 - (\Delta t_i)^2 = 2(\Delta t_i)^2.$$

Notice also that  $\sum_{i=1}^n \Delta t_i = T$ . Then

$$\begin{aligned} \mathbb{E} \left[ \left( \sum_{i=1}^n |B_{t_{i-1}, t_i}|^2 - T \right)^2 \right] &= \mathbb{E} \left[ \left( \sum_{i=1}^n \eta_i \right)^2 \right] = \text{Var} \left( \sum_{i=1}^n \eta_i \right) \\ &= \sum_{i=1}^n \text{Var}(\eta_i) = 2 \sum_{i=1}^n (\Delta t_i)^2 \leq 2|\pi| \sum_{i=1}^n \Delta t_i = 2T|\pi| \rightarrow 0, \quad \text{as } |\pi| \rightarrow 0. \end{aligned}$$

Since  $\mathbb{L}^2$  convergence implies convergence in probability, we conclude that  $\langle B \rangle_T = T$ . ■

As a corollary of Theorems 2.1.4 and 2.1.6, we have

**Corollary 2.1.7**  $\bigvee_a^b(B) = \infty$  for any  $0 \leq a < b \leq T$ , a.s.

*Proof* We proceed in two steps, again assuming  $d = 1$ .

*Step 1.* Fix  $0 \leq a < b \leq T$ . For any partition  $\pi$ , denote  $\tilde{t}_i := a \vee t_i \wedge b$  and notice that

$$\sum_{i=1}^n |B_{\tilde{t}_{i-1}, \tilde{t}_i}|^2 \leq \left( \sup_{1 \leq i \leq n} |B_{\tilde{t}_{i-1}, \tilde{t}_i}| \right) \sum_{i=1}^n |B_{\tilde{t}_{i-1}, \tilde{t}_i}| \leq \bigvee_a^b(B) \times \sup_{1 \leq i \leq n} |B_{\tilde{t}_{i-1}, \tilde{t}_i}|$$

Send  $|\pi| \rightarrow 0$ , by Theorems 2.1.4 and 2.1.6 we have

$$\sup_{1 \leq i \leq n} |B_{\tilde{t}_{i-1}, \tilde{t}_i}| \rightarrow 0, \text{ a.s.} \quad \text{and} \quad \sum_{i=1}^n |B_{\tilde{t}_{i-1}, \tilde{t}_i}|^2 \rightarrow \langle B \rangle_b - \langle B \rangle_a = b - a > 0, \text{ in } \mathbb{P}.$$

This clearly implies that

$$\bigvee_a^b(B) = \infty, \quad \text{a.s.}$$

*Step 2.* For any  $0 \leq a < b \leq T$ , by Step 1 we have

$$\mathbb{P}(\mathcal{N}(a, b)) = 0, \quad \text{where} \quad \mathcal{N}(a, b) := \left\{ \omega : \bigvee_a^b(B(\omega)) < \infty \right\}.$$

Denote

$$\mathcal{N} := \bigcup \left[ \mathcal{N}(r_1, r_2) : 0 \leq r_1 < r_2 \leq T, \quad r_1, r_2 \in \mathbb{Q} \right]$$

Then  $\mathbb{P}(\mathcal{N}) = 0$ . Now for any  $\omega \notin \mathcal{N}$ , and for any  $0 \leq a < b \leq T$ , there exist  $r_1, r_2 \in \mathbb{Q}$  such that  $a \leq r_1 < r_2 \leq b$ . Then

$$\bigvee_a^b(B(\omega)) \geq \bigvee_{r_1}^{r_2}(B(\omega)) = \infty.$$

The proof is complete now. ■

**Remark 2.1.8**

- (i) Corollary 2.1.7 implies that  $B$  is nowhere absolutely continuous with respect to  $dt$ . Let  $d = 1$ . We actually have the following so-called *Law of Iterated Logarithm*: for any  $t \in [0, T)$ ,

$$\limsup_{\delta \downarrow 0} \frac{B_{t, t+\delta}}{\sqrt{2\delta \ln \ln \frac{1}{\delta}}} = 1, \quad \liminf_{\delta \downarrow 0} \frac{B_{t, t+\delta}}{\sqrt{2\delta \ln \ln \frac{1}{\delta}}} = -1, \quad \text{a.s.} \quad (2.1.3)$$

This implies that  $B$  is nowhere Hölder- $\frac{1}{2}$  continuous. In particular,  $B$  is nowhere differentiable.

- (ii) The regularity (2.1.3) is right regularity. The left regularity  $B_{t-\delta,t}$  is less clear. Moreover, the null set in (2.1.3) depends on  $t$ . Indeed, the uniform regularities  $\sup_t B_{t,t+\delta}$  and  $\sup_{0 \leq s < t \leq T, t-s \leq \delta} |B_{s,t}|$  are more involved. ■

### 2.1.3 The Augmented Filtration

Let  $\mathbb{F}^B$  denote the filtration generated by  $B$ . We notice that neither  $\mathbb{F}^B$  nor its completed filtration is right continuous. For example,

$$\left\{ \omega \in \Omega : \limsup_{\delta \downarrow 0} \frac{B_\delta(\omega)}{\sqrt{2\delta \ln \ln \frac{1}{\delta}}} = 1 \right\} \in \mathcal{F}_{0+}^B \setminus \mathcal{F}_0^B.$$

However, the augmented filtration, denoted as  $\overline{\mathbb{F}}^B$ , is right continuous. We first have the Blumenthal 0–1 law.

**Theorem 2.1.9** *For any random variable  $X \in \mathbb{L}^0(\mathcal{F}_{0+}^B)$ , we have  $X = \mathbb{E}[X]$ , a.s. Consequently, for any event  $A \in \mathcal{F}_{0+}^B$ , we have  $\mathbb{P}(A) = 0$  or  $1$ .*

*Proof* Let  $X \in \mathbb{L}^0(\mathcal{F}_{0+}^B)$ . For any  $n \geq 1$ , let  $\mathcal{G}_n := \sigma(B_{n^{-1},s}, n^{-1} \leq s \leq T)$ , the  $\sigma$ -field generated by  $\{B_{n^{-1},s}, n^{-1} \leq s \leq T\}$ . Since  $B$  has independent increments,  $\mathcal{G}_n$  and  $\mathcal{F}_{0+}^B \subset \mathcal{F}_{n^{-1}}^B$  are independent. Thus  $\mathbb{E}[X|\mathcal{G}_n] = \mathbb{E}[X]$ , a.s. for all  $n \geq 1$ . On the other hand, denote  $\mathcal{G} := (\bigvee_n \mathcal{G}_n) \vee \mathcal{N}(\mathcal{F})$ . For any  $t > 0$ ,  $B_t = B_{0,t} = \lim_{n \rightarrow \infty} B_{n^{-1},t}$ , a.s. Since  $B_{n^{-1},s} \in \mathbb{L}^0(\mathcal{G}_{n^{-1}}) \subset \mathbb{L}^0(\mathcal{G})$ , we see that  $B_t \in \mathbb{L}^0(\mathcal{G})$  for any  $t > 0$ . Thus  $\mathcal{F}_{0+}^B \subset \mathcal{G}$ . Note that  $\mathcal{G}_n$  is increasing in  $n$ , then by Problem 1.4.2 (iii) we obtain

$$X = \mathbb{E}[X|\mathcal{G}] = \lim_{n \rightarrow \infty} \mathbb{E}[X|\mathcal{G}_n] = \mathbb{E}[X], \quad \text{a.s.}$$

Finally, for any  $A \in \mathcal{F}_{0+}^B$ , set  $X := \mathbf{1}_A$ , we see that  $\mathbb{P}(A) = \mathbf{1}_A$ , a.s. and thus  $\mathbb{P}(A) = 0$  or  $1$ . ■

**Corollary 2.1.10** *The augmented filtration  $\overline{\mathbb{F}}^B$  satisfies the usual hypotheses.*

*Proof* It suffices to show that  $\overline{\mathbb{F}}^B$  is right continuous. Theorem 2.1.9 implies that  $\mathcal{F}_{0+}^B \subset \mathcal{N}(\mathcal{F}) \subset \overline{\mathcal{F}}_0^B$ . Then  $\overline{\mathcal{F}}_{0+}^B = \overline{\mathcal{F}}_0^B$ . Similarly, for any  $t$ , we have  $\overline{\mathcal{F}}_{t+}^B = \overline{\mathcal{F}}_t^B$ . ■

**Remark 2.1.11**

- (i) In this book, we shall use  $\mathbb{F}^B$ . When  $\mathbb{P}$  is given, in most cases this is equivalent to using the augmented filtration  $\overline{\mathbb{F}}^B$  as in standard literature, in the spirit of Proposition 1.2.1.
- (ii) As we will see, all  $\mathbb{F}^B$ -local martingales are continuous. If we consider more general càdlàg martingales, it is more convenient to use right continuous filtration. ■

**2.2 Stochastic Integration****2.2.1 Some Heuristic Arguments**

In this subsection we assume  $d = 1$ . We first recall the Riemann-Stieltjes integral. Let  $A : [0, T] \rightarrow \mathbb{R}$  be a function with bounded variation, and  $b : [0, T] \rightarrow \mathbb{R}$  be continuous. For a partition  $\pi : 0 = t_0 < \dots < t_n = T$ , define the Riemann-Stieltjes partial sum:

$$\sum_{i=0}^{n-1} b(\hat{t}_i) A_{t_i, t_{i+1}} \quad \text{where } \hat{t}_i \in [t_i, t_{i+1}] \text{ is arbitrary.}$$

It is well known that, as  $|\pi| \rightarrow 0$ , the above partial sum converges and the limit is independent of the choices of  $\pi$  and  $\hat{t}_i$ , and thus is defined as the integral of  $b$  with respect to  $A$ :

$$\int_0^T b_t dA_t := \lim_{|\pi| \rightarrow 0} \sum_{i=0}^{n-1} b(\hat{t}_i) A_{t_i, t_{i+1}}. \quad (2.2.1)$$

Now assume  $A, b \in \mathbb{L}^0(\mathbb{F})$  such that  $A$  has bounded variation and  $b$  is continuous, a.s. Then clearly we can define the integral pathwise:

$$\left( \int_0^T b_t dA_t \right)(\omega) := \int_0^T b_t(\omega) d(A_t(\omega)).$$

We next discuss stochastic integrals with respect to  $B$ . Let  $\sigma \in \mathbb{L}^0(\mathbb{F})$  be continuous, a.s. We first notice that in this case the limits of the Riemann-Stieltjes partial sum may depend on the choices of  $\hat{t}_i$ . Indeed, let  $\sigma = B$  and set  $\hat{t}_i$  as the left end point and right end point respectively, we have

$$S_L(\pi) := \sum_{i=0}^{n-1} B_{t_i} B_{t_i, t_{i+1}}, \quad S_R(\pi) := \sum_{i=0}^{n-1} B_{t_{i+1}} B_{t_i, t_{i+1}}.$$

Then, by Theorem 2.1.6,

$$S_R(\pi) - S_L(\pi) = \sum_{i=0}^{n-1} |B_{t_i, t_{i+1}}|^2 \rightarrow T, \quad \text{in } \mathbb{P} \text{ as } |\pi| \rightarrow 0.$$

So  $S_R(\pi)$  and  $S_L(\pi)$  cannot converge to the same limit, and therefore, it is important to choose appropriate points  $\hat{t}_i$ . As in standard literature, we shall study the Itô integral, which uses the left end points. The main reason is that, among others, in this case we use  $\sigma_{t_i}$  to approximate  $\sigma$  on the interval  $[t_i, t_{i+1})$  and thus the approximating process  $\sigma^\pi$  defined below is still  $\mathbb{F}$ -measurable:

$$\sigma^\pi := \sum_{i=0}^{n-1} \sigma_{t_i} \mathbf{1}_{[t_i, t_{i+1})}. \quad (2.2.2)$$

**Remark 2.2.1** If we use  $\hat{t}_i := \frac{t_i + t_{i+1}}{2}$ , the corresponding limit is called the Stratonovic Integral. We shall study Itô integral in this book, which has the following advantages:

- The  $\mathbb{F}$ -measurability of the  $\sigma^\pi$  in (2.2.2) is natural in many applications, see, e.g., Section 2.8;
- As we will see soon, the Itô integral has martingale property and thus allows us to use the martingale theory;
- Unlike Stratonovic Integral, the Itô integral does not require any regularity on the integrand  $\sigma$ .

However, Stratonovic Integral is more convenient for pathwise analysis. In particular, under Stratonovic Integral, the chain rule same as the deterministic case remains true. See Problem 2.10.13. ■

## 2.2.2 Itô Integral for Elementary Processes

**Definition 2.2.2** We say  $\sigma \in \mathbb{L}^2(\mathbb{F})$  is an elementary process, denote as  $\sigma \in \mathbb{L}_0^2(\mathbb{F})$ , if there exist a partition  $0 = t_0 < \dots < t_n = T$  such that  $\sigma_t = \sigma_{t_i}$  for all  $t \in [t_i, t_{i+1})$ ,  $i = 0, \dots, n-1$ .

Clearly, for  $\sigma \in L_0^2(\mathbb{F}, \mathbb{R}^d)$ , we may define the stochastic integral in a pathwise manner:

$$\int_0^t \sigma_s \cdot dB_s := \sum_{i=0}^{n-1} \sigma_{t_i} \cdot B_{t_i \wedge t, t_{i+1} \wedge t}, \quad 0 \leq t \leq T. \quad (2.2.3)$$

**Lemma 2.2.3** Let  $\sigma \in \mathbb{L}_0^2(\mathbb{F}, \mathbb{R}^d)$  and denote  $M_t := \int_0^t \sigma_s \cdot dB_s$ .

- (i)  $M$  is an  $\mathbb{F}$ -martingale. In particular,  $\mathbb{E}[M_t] = 0$ .
- (ii)  $M \in \mathbb{L}^2(\mathbb{F})$  and  $N_t := M_t^2 - \int_0^t |\sigma_s|^2 ds$  is a martingale. In particular,

$$\mathbb{E}[|M_t|^2] = \mathbb{E}\left[\int_0^t |\sigma_s|^2 ds\right]. \quad (2.2.4)$$

- (iii) For any  $\sigma^i \in \mathbb{L}_0^2(\mathbb{F}, \mathbb{R}^d)$ ,  $\lambda_i \in \mathbb{L}^\infty(\mathcal{F}_0, \mathbb{R})$ ,  $i = 1, 2$ , we have  $\lambda_1 \sigma^1 + \lambda_2 \sigma^2 \in \mathbb{L}_0^2(\mathbb{F}, \mathbb{R}^d)$  and

$$\int_0^t [\lambda_1 \sigma_s^1 + \lambda_2 \sigma_s^2] \cdot dB_s = \lambda_1 \int_0^t \sigma_s^1 \cdot dB_s + \lambda_2 \int_0^t \sigma_s^2 \cdot dB_s.$$

- (iv)  $M$  is continuous, a.s.

*Proof*

- (i) It suffices to show that, for any  $i$ ,

$$M_t = \mathbb{E}[M_{t_{i+1}} | \mathcal{F}_t], \quad t_i \leq t \leq t_{i+1}.$$

Indeed, note that  $\sigma_{t_i} \in \mathcal{F}_{t_i} \subset \mathcal{F}_t$  and  $B$  has independent increments, then

$$\mathbb{E}[M_{t,t_{i+1}} | \mathcal{F}_t] = \mathbb{E}[\sigma_{t_i} \cdot B_{t,t_{i+1}} | \mathcal{F}_t] = \sigma_{t_i} \cdot \mathbb{E}[B_{t,t_{i+1}} | \mathcal{F}_t] = \sigma_{t_i} \cdot \mathbb{E}[B_{t,t_{i+1}}] = 0.$$

- (ii) The square integrability of  $M$  follows directly from (2.2.4). Then it suffices to show that

$$N_t = \mathbb{E}[N_{t_{i+1}} | \mathcal{F}_t], \quad t_i \leq t \leq t_{i+1}.$$

To illustrate the arguments, in this proof we use multidimensional notations. Note that

$$\begin{aligned} N_{t,t_{i+1}} &= M_{t_{i+1}}^2 - M_t^2 - |\sigma_{t_i}|^2(t_{i+1} - t) = |M_{t,t_{i+1}}|^2 + 2M_t M_{t,t_{i+1}} - |\sigma_{t_i}|^2(t_{i+1} - t) \\ &= [\sigma_{t_i} \sigma_{t_i}^\top] : [B_{t,t_{i+1}} B_{t,t_{i+1}}^\top - (t_{i+1} - t)I_d] + 2M_t \sigma_{t_i} \cdot B_{t,t_{i+1}}. \end{aligned}$$

Then, similar to (i) we have

$$\mathbb{E}[N_{t,t_{i+1}} | \mathcal{F}_t] = [\sigma_{t_i} \sigma_{t_i}^\top] : \mathbb{E}[B_{t,t_{i+1}} B_{t,t_{i+1}}^\top - (t_{i+1} - t)I_d] + 2M_t \sigma_{t_i} \cdot \mathbb{E}[B_{t,t_{i+1}}] = 0.$$

- (iii) and (iv) are obvious. ■

The following estimates are important, and we leave a more general result in Problem 2.10.3 below. Recall the notation  $X^*$  in (1.2.4).



**Lemma 2.2.4 (Doob's Maximum Inequality)** *Let  $\sigma \in \mathbb{L}_0^2(\mathbb{F}, \mathbb{R}^d)$ ,  $M_t := \int_0^t \sigma_s \cdot dB_s$ . Then*

$$\mathbb{E}[|M_T|^2] \leq \mathbb{E}[|M_T^*|^2] \leq 4\mathbb{E}[|M_T|^2]. \quad (2.2.5)$$

*Proof* The left inequality is obvious. We prove the right inequality in two steps.

*Step 1.* We first prove it under an additional assumption:

$$\mathbb{E}[|M_T^*|^2] < \infty. \quad (2.2.6)$$

Given  $\lambda > 0$ , denote

$$\tau_\lambda := \inf\{t \geq 0 : |M_t| \geq \lambda\} \wedge T. \quad (2.2.7)$$

Since  $M$  is continuous, we see that

$$\tau_\lambda \in \mathcal{F}(\mathbb{F}), \quad |M_{\tau_\lambda}| \leq \lambda, \quad \text{and} \quad \{M_T^* \geq \lambda\} = \{|M_{\tau_\lambda}| = \lambda\}. \quad (2.2.8)$$

Moreover, by (2.2.6)  $M$  is a u.i. martingale, then

$$|M_{\tau_\lambda}| = \left| \mathbb{E}[M_T | \mathcal{F}_{\tau_\lambda}] \right| \leq \mathbb{E}[|M_T| | \mathcal{F}_{\tau_\lambda}].$$

This implies

$$\begin{aligned} \mathbb{P}(M_T^* \geq \lambda) &= \mathbb{E}[\mathbf{1}_{\{|M_{\tau_\lambda}| = \lambda\}}] = \mathbb{E}\left[\frac{|M_{\tau_\lambda}|}{\lambda} \mathbf{1}_{\{|M_{\tau_\lambda}| = \lambda\}}\right] \leq \frac{1}{\lambda} \mathbb{E}\left[\mathbb{E}(|M_T| | \mathcal{F}_{\tau_\lambda}) \mathbf{1}_{\{|M_{\tau_\lambda}| = \lambda\}}\right] \\ &= \frac{1}{\lambda} \mathbb{E}\left[\mathbb{E}(|M_T| \mathbf{1}_{\{|M_{\tau_\lambda}| = \lambda\}} | \mathcal{F}_{\tau_\lambda})\right] = \frac{1}{\lambda} \mathbb{E}\left[|M_T| \mathbf{1}_{\{|M_{\tau_\lambda}| = \lambda\}}\right] = \frac{1}{\lambda} \mathbb{E}\left[|M_T| \mathbf{1}_{\{M_T^* \geq \lambda\}}\right]. \end{aligned} \quad (2.2.9)$$

Thus

$$\begin{aligned} \mathbb{E}[|M_T^*|^2] &= 2 \int_0^\infty \lambda \mathbb{P}(M_T^* \geq \lambda) d\lambda \leq 2 \int_0^\infty \mathbb{E}\left[|M_T| \mathbf{1}_{\{M_T^* \geq \lambda\}}\right] d\lambda \\ &= 2 \mathbb{E}\left[\int_0^\infty |M_T| \mathbf{1}_{\{M_T^* \geq \lambda\}} d\lambda\right] = 2 \mathbb{E}\left[|M_T| M_T^*\right] \leq 2 \left(\mathbb{E}[|M_T|^2]\right)^{\frac{1}{2}} \left(\mathbb{E}[|M_T^*|^2]\right)^{\frac{1}{2}}, \end{aligned}$$

where the last inequality thanks to the Hölder's inequality. This implies (2.2.5) immediately.

*Step 2.* In the general case, for each  $n \geq 1$ , let  $\tau_n$  be defined by (2.2.7) and denote

$$\sigma^n := \sigma \mathbf{1}_{[0, \tau_n]}, \quad M_t^n := \int_0^t \sigma_s^n \cdot dB_s, \quad M_t^{n,*} := \sup_{0 \leq s \leq t} |M_s^n|.$$

Since  $M$  is continuous, by Problem 1.4.8 (ii) we see that  $\tau_n$  is increasing, and  $\tau_n = T$  when  $n$  is large enough. Then

$$M_t^n = M_{\tau_n \wedge t}, \quad M_T^{n,*} \leq n, \quad \text{and} \quad M_T^{n,*} \uparrow M_T^*.$$

By Step 1 and (2.2.4) we have

$$\mathbb{E}[|M_T^{n,*}|^2] \leq 4\mathbb{E}[|M_T^n|^2] = 4\mathbb{E}\left[\int_0^{\tau_n} |\sigma_s|^2 ds\right] \leq 4\mathbb{E}\left[\int_0^T |\sigma_s|^2 ds\right] = 4\mathbb{E}[|M_T|^2].$$

Now applying the Monotone Convergence Theorem we obtain (2.2.5).  $\blacksquare$

### 2.2.3 Itô Integral in $\mathbb{L}^2(\mathbb{F})$ and $\mathbb{L}_{loc}^2(\mathbb{F})$

We now extend the Itô stochastic integration to all processes in  $\mathbb{L}^2(\mathbb{F})$ . We first need a lemma.

**Lemma 2.2.5** *For any  $\sigma \in \mathbb{L}^2(\mathbb{F}, \mathbb{R}^d)$ , there exist  $\sigma^n \in \mathbb{L}_0^2(\mathbb{F}, \mathbb{R}^d)$  such that  $\lim_{n \rightarrow \infty} \|\sigma^n - \sigma\|_2 = 0$ .*

*Proof* We proceed in three steps.

*Step 1.* We first assume  $\sigma$  is continuous and bounded. For each  $n$ , define

$$\sigma_t^n := \sum_{i=0}^{n-1} \sigma_{t_i} \mathbf{1}_{[t_i, t_{i+1})} \text{ where } t_i := \frac{i}{n}T, \quad i = 0, \dots, n.$$

Then by the Dominated Convergence Theorem we obtain the result immediately.

*Step 2.* We now assume only that  $|\sigma| \leq C$ . For each  $\delta > 0$ , define  $\sigma_t^\delta := \frac{1}{\delta} \int_{(t-\delta) \vee 0}^t \sigma_s ds$ . Clearly  $|\sigma^\delta| \leq C$ ,  $\sigma^\delta$  is continuous, and by real analysis, in the spirit of Problem 1.4.14, we have  $\lim_{\delta \rightarrow 0} \int_0^T |\sigma_t^\delta - \sigma_t|^2 dt = 0$ , a.s. By the Dominated Convergence Theorem again, we have  $\lim_{\delta \rightarrow 0} \|\sigma^\delta - \sigma\|_2 = 0$ . Now for each  $n$ , there exists  $\delta_n$  such that  $\|\sigma^\delta - \sigma\|_2 \leq \frac{1}{2n}$ . Moreover, by Step 1, there exists  $\sigma^n \in \mathbb{L}_0^2(\mathbb{F}, \mathbb{R}^d)$  such that  $\|\sigma^n - \sigma^\delta\|_2 \leq \frac{1}{2n}$ . This implies  $\|\sigma^n - \sigma\|_2 \leq \frac{1}{n} \rightarrow 0$ , as  $n \rightarrow \infty$ .

*Step 3.* For the general case, for each  $n$ , denote  $\tilde{\sigma}^n := (-n) \vee \sigma \wedge n$ , where the truncation is component wise. Then  $\tilde{\sigma}^n \rightarrow \sigma$  and  $|\sigma^n| \leq |\sigma|$ . Applying the Dominated Convergence Theorem we get  $\lim_{n \rightarrow \infty} \|\tilde{\sigma}^n - \sigma\|_2 = 0$ . Moreover, since  $|\sigma^n| \leq n\sqrt{d}$ , by Step 2 there exists  $\sigma^n \in \mathbb{L}_0^2(\mathbb{F}, \mathbb{R}^d)$  such that  $\|\tilde{\sigma}^n - \sigma^n\|_2 \leq \frac{1}{n}$ . Thus

$$\|\sigma^n - \sigma\|_2 \leq \|\tilde{\sigma}^n - \sigma\|_2 + \|\tilde{\sigma}^n - \sigma^n\|_2 \leq \|\tilde{\sigma}^n - \sigma\|_2 + \frac{1}{n} \rightarrow 0.$$

The proof is complete now.  $\blacksquare$

For the above  $\sigma^n \in \mathbb{L}_0^2(\mathbb{F}, \mathbb{R}^d)$ , we have defined  $M_t^n := \int_0^t \sigma_s^n \cdot dB_s$  by (2.2.3). Applying Lemma 2.2.4 we get

$$\mathbb{E}\left[|(M^n - M^m)_T^*|^2\right] \leq 4\|\sigma^n - \sigma^m\|_2^2 \rightarrow 0, \quad \text{as } m, n \rightarrow \infty.$$

Thus there exists a ( $\mathbb{P}$ -a.s.) unique continuous process  $M \in \mathbb{L}^0(\mathbb{F}, \mathbb{R})$  such that

$$\lim_{n \rightarrow \infty} \mathbb{E}\left[|(M^n - M)_T^*|^2\right] = 0. \quad (2.2.10)$$

Moreover, if there exist another sequence  $\tilde{\sigma}^n \in \mathbb{L}_0^2(\mathbb{F}, \mathbb{R}^d)$  such that  $\lim_{n \rightarrow \infty} \|\tilde{\sigma}^n - \sigma\|_2 = 0$ , then  $\lim_{n \rightarrow \infty} \|\tilde{\sigma}^n - \sigma^n\|_2 = 0$ . This implies that, for  $\tilde{M}_t^n := \int_0^t \tilde{\sigma}_s^n \cdot dB_s$ ,

$$0 \leq \mathbb{E}\left[|(M^n - \tilde{M}^n)_T^*|^2\right] \leq 4\|\sigma^n - \tilde{\sigma}^n\|_2^2 \rightarrow 0, \quad \text{as } n \rightarrow \infty.$$

Thus  $\tilde{M}^n$  also converges to  $M$ . That is, the process  $M$  does not depend on the choices of  $\sigma^n$ . Therefore, we may define  $M$  as the stochastic integral of  $\sigma$ : for each  $t \in [0, T]$ ,

$$\int_0^t \sigma_s \cdot dB_s := \lim_{n \rightarrow \infty} \int_0^t \sigma_s^n \cdot dB_s, \quad \text{where the convergence is in the sense of (2.2.10).} \quad (2.2.11)$$

**Remark 2.2.6** We emphasize that the convergence in (2.2.11) is in  $\mathbb{L}^2$ -sense, and thus the above definition of stochastic integral is not in a pathwise manner. That is, given  $\sigma(\omega)$  and  $B(\omega)$ , in general we cannot determine  $\left(\int_0^t \sigma_s \cdot dB_s\right)(\omega)$ . The theory on pathwise stochastic integration is important and challenging, see some discussion along this line in Sections 2.8.3 and 12.1.1, and Problem 2.10.14. ■

By the uniform convergence in (2.2.10), it follows immediately that

**Theorem 2.2.7** *Let  $\sigma \in \mathbb{L}^2(\mathbb{F}, \mathbb{R}^d)$  and  $M_t := \int_0^t \sigma_s \cdot dB_s$ . All the results in Lemmas 2.2.3 and 2.2.4 still hold true.*

We finally extend the stochastic integration to all processes  $\sigma \in \mathbb{L}_{loc}^2(\mathbb{F}, \mathbb{R}^d)$ . For  $n \geq 1$ , define

$$\tau_n := \inf\{t \geq 0 : \int_0^t |\sigma_s|^2 ds \geq n\} \wedge T, \quad \sigma_t^n := \sigma_t \mathbf{1}_{[0, \tau_n)}(t).$$

Then  $\sigma^n \in \mathbb{L}^2(\mathbb{F}, \mathbb{R}^d)$ ,  $\tau_n$  is increasing and  $\tau_n = T$  for  $n$  large enough, a.s. Denote  $M_t^n := \int_0^t \sigma_s^n \cdot dB_s$ . One can easily check that, for  $n < m$ ,

$$M_t^n = M_t^m \quad \text{for } t \leq \tau_n.$$

Thus we may define

$$\int_0^t \sigma_s \cdot dB_s := M_t^n \text{ for } t \leq \tau_n. \quad (2.2.12)$$

So  $M_t := \int_0^t \sigma_s \cdot dB_s$  is well defined for all  $t \in [0, T]$ . By Theorem 2.2.7 it is obvious that

**Theorem 2.2.8** *For any  $\sigma \in \mathbb{L}_{loc}^2(\mathbb{F}, \mathbb{R}^d)$ ,  $M_t := \int_0^t \sigma_s \cdot dB_s$  is a continuous local martingale.*

## 2.3 The Itô Formula

The Itô formula is the extension of the chain rule in calculus to stochastic calculus, and plays a key role in stochastic calculus. In particular, it will be crucial to build the connection between the martingale theory and partial differential equations, see, e.g., Section 5.1 below.

### 2.3.1 Some Heuristic Arguments

Assume  $A \in \mathbb{L}^0(\mathbb{F}, \mathbb{R})$  has bounded variation, a.s. and  $f \in C^1(\mathbb{R})$  is a deterministic function. The standard chain rule tells that

$$df(A_t) = f'(A_t) dA_t. \quad (2.3.1)$$

The following simple example shows that the above formula fails if we replace  $A$  with the Brownian motion  $B$  and thus  $dA$  becomes stochastic integration  $dB$ .

**Example 2.3.1** *Let  $d = 1$  and set  $f(x) := x^2$ . Then*

$$|B_T|^2 = 2 \int_0^T B_t dB_t + T.$$

*Proof* For any partition  $\pi : 0 = t_0 < \dots < t_n = T$ , we have

$$|B_T|^2 = \sum_{i=0}^{n-1} \left[ |B_{t_{i+1}}|^2 - |B_{t_i}|^2 \right] = \sum_{i=0}^{n-1} \left[ |B_{t_i, t_{i+1}}|^2 + 2B_{t_i} B_{t_i, t_{i+1}} \right].$$

Send  $|\pi| \rightarrow 0$ , we have

$$\sum_{i=0}^{n-1} |B_{t_i, t_{i+1}}|^2 \rightarrow T \text{ in } \mathbb{L}^2(\mathcal{F}_T). \quad (2.3.2)$$

Moreover, denote  $B_t^\pi := \sum_{i=0}^{n-1} B_{t_i} \mathbf{1}_{[t_i, t_{i+1})}$ . Then  $B^\pi \in \mathbb{L}_0^2(\mathbb{F})$ , and one can easily check that

$$\lim_{|\pi| \rightarrow 0} \|B^\pi - B\|_2 = 0.$$

This implies that

$$\sum_{i=0}^{n-1} B_{t_i} B_{t_i, t_{i+1}} \rightarrow \int_0^T B_t dB_t, \quad \text{in } \mathbb{L}^2(\mathcal{F}_T),$$

which, together with (2.3.2), proves the result. ■

Note that

$$f'(B_t) = 2B_t, \quad f''(B_t) = 2, \quad \langle B \rangle_t = t.$$

Then Example 2.3.1 implies

$$f(B_T) - f(B_0) = \int_0^T f'(B_t) dB_t + \frac{1}{2} \int_0^T f''(B_t) d\langle B \rangle_t. \quad (2.3.3)$$

This is a special case of the Itô formula. We see that there is a correction term  $\frac{1}{2} \int_0^T f''(B_t) d\langle B \rangle_t$  for stochastic integrations. We prove the general case in the next subsection.

### 2.3.2 The Itô Formula

In this subsection we focus on one-dimensional case. The multidimensional case will be introduced in detail in the next subsection. Let  $b \in \mathbb{L}_{loc}^1(\mathbb{F})$ ,  $\sigma \in \mathbb{L}_{loc}^2(\mathbb{F})$ , and denote

$$X_t = X_0 + \int_0^t b_s ds + \int_0^t \sigma_s dB_s \quad \text{and} \quad \langle X \rangle_t := \int_0^t |\sigma_s|^2 ds. \quad (2.3.4)$$

**Theorem 2.3.2 (Itô Formula)** *Let  $f \in C^{1,2}([0, T] \times \mathbb{R}, \mathbb{R})$ . Then*

$$\begin{aligned} df(t, X_t) &= \partial_t f(t, X_t) dt + \partial_x f(t, X_t) dX_t + \frac{1}{2} \partial_{xx} f(t, X_t) d\langle X \rangle_t \\ &= \left[ \partial_t f + \partial_x f b_t + \frac{1}{2} \partial_{xx} f |\sigma_t|^2 \right](t, X_t) dt + \partial_x f(t, X_t) \sigma_t dB_t. \end{aligned} \quad (2.3.5)$$

Or equivalently,

$$f(t, X_t) = f(0, X_0) + \int_0^t \left[ \partial_t f + \partial_x f b_s + \frac{1}{2} \partial_{xx} f |\sigma_s|^2 \right](s, X_s) ds + \int_0^t \partial_x f(s, X_s) \sigma_s dB_s. \quad (2.3.6)$$

*Proof* We first note that, since  $X$  is continuous and  $f \in C^{1,2}$ , for  $\varphi = \partial_t f, \partial_x f, \partial_{xx} f$ , we know that  $\varphi(t, X_t)$  is continuous and thus  $\sup_{0 \leq t \leq T} |\varphi(t, X_t)| < \infty$ , a.s. This implies that

$$\left[ \partial_t f + \partial_x f b + \frac{1}{2} \partial_{xx} f |\sigma|^2 \right](\cdot, X) \in \mathbb{L}_{loc}^1(\mathbb{F}), \quad \partial_x f(\cdot, X) \sigma \in \mathbb{L}_{loc}^2(\mathbb{F}),$$

and thus the right side of (2.3.6) is well defined.

Without loss of generality, we prove (2.3.6) only for  $t = T$ . We proceed in several steps.

*Step 1.* We first assume that  $b_t = b_0, \sigma_t = \sigma_0$  are  $\mathcal{F}_0$ -measurable and bounded, and  $f$  is smooth enough with all related derivatives bounded.

For an arbitrary partition  $\pi : 0 = t_0 < \dots < t_n = T$ , we have

$$f(T, X_T) - f(0, X_0) = \sum_{i=0}^{n-1} \left[ f(t_{i+1}, X_{t_{i+1}}) - f(t_i, X_{t_i}) \right]. \quad (2.3.7)$$

Denote  $\Delta t_{i+1} := t_{i+1} - t_i$  and note that  $X_{t_i, t_{i+1}} = b_0 \Delta t_{i+1} + \sigma_0 B_{t_i, t_{i+1}}$ . Then, by Taylor expansion,

$$\begin{aligned} f(t_{i+1}, X_{t_{i+1}}) - f(t_i, X_{t_i}) &= f(t_i + \Delta t_{i+1}, X_{t_i} + X_{t_i, t_{i+1}}) - f(t_i, X_{t_i}) \\ &= \partial_t f(t_i, X_{t_i}) \Delta t_{i+1} + \partial_x f(t_i, X_{t_i}) X_{t_i, t_{i+1}} \\ &\quad + \frac{1}{2} \partial_{tt} f(t_i, X_{t_i}) |\Delta t_{i+1}|^2 + \partial_{tx} f(t_i, X_{t_i}) \Delta t_{i+1} X_{t_i, t_{i+1}} + \frac{1}{2} \partial_{xx} f(t_i, X_{t_i}) |X_{t_i, t_{i+1}}|^2 + R_{i+1}^\pi \\ &= \left[ \partial_t f + b_0 \partial_x f + \frac{1}{2} \partial_{xx} f |\sigma_0|^2 \right](t_i, X_{t_i}) \Delta t_{i+1} + \sigma_0 \partial_x f(t_i, X_{t_i}) B_{t_i, t_{i+1}} \\ &\quad + \frac{1}{2} \partial_{xx} f(t_i, X_{t_i}) |\sigma_0|^2 [|B_{t_i, t_{i+1}}|^2 - \Delta t_{i+1}] + I_{i+1}^\pi \end{aligned} \quad (2.3.8)$$

where

$$\begin{aligned} I_{i+1}^\pi &:= \frac{1}{2} \left[ \partial_{tt} f + \partial_{xx} f |b_0|^2 \right](t_i, X_{t_i}) |\Delta t_{i+1}|^2 + \left[ \partial_{tx} f + b_0 \sigma_0 \partial_{xx} f \right](t_i, X_{t_i}) \Delta t_{i+1} B_{t_i, t_{i+1}} + R_{i+1}^\pi, \\ \text{and } |R_{i+1}^\pi| &\leq C[|\Delta t_{i+1}|^3 + |X_{t_i, t_{i+1}}|^3] \leq C[|\Delta t_{i+1}|^3 + |B_{t_i, t_{i+1}}|^3]. \end{aligned}$$

Send  $|\pi| \rightarrow 0$ . First, applying the Dominated Convergence Theorem we have:

$$\begin{aligned} &\sum_{i=0}^{n-1} \left[ \partial_t f + b_0 \partial_x f + \frac{1}{2} \partial_{xx} f |\sigma_0|^2 \right](t_i, X_{t_i}) \Delta t_{i+1} \\ &\rightarrow \int_0^T \left[ \partial_t f + b_0 \partial_x f + \frac{1}{2} |\sigma_0|^2 \partial_{xx} f \right](t, X_t) dt, \quad \text{in } \mathbb{L}^2(\mathcal{F}_T). \end{aligned} \quad (2.3.9)$$

Next, applying the Dominated Convergence Theorem again we have

$$\mathbb{E}\left[\sum_{i=0}^{n-1} \int_{t_i}^{t_{i+1}} |\partial_{xx}f(t, X_t) - \partial_{xx}f(t_i, X_{t_i})|^2 dt\right] \rightarrow 0,$$

and thus

$$\sum_{i=0}^{n-1} \sigma_0 \partial_{xx}f(t_i, X_{t_i}) B_{t_i, t_{i+1}} \rightarrow \int_0^T \sigma_0 \partial_{xx}f(t, X_t) dB_t, \quad \text{in } \mathbb{L}^2(\mathcal{F}_T). \quad (2.3.10)$$

Moreover, note that, for any  $p \geq 1$  and some constant  $c_p > 0$ ,

$$\mathbb{E}[|B_{t_i, t_{i+1}}|^p] = c_p |\Delta t_{i+1}|^{\frac{p}{2}}.$$

Then

$$\begin{aligned} \mathbb{E}\left[\left(\sum_{i=0}^{n-1} B_{t_i, t_{i+1}}^2\right)^2\right] &\leq C \mathbb{E}\left[\sum_{i=0}^{n-1} (|\Delta t_{i+1}|^2 + \Delta t_{i+1} |B_{t_i, t_{i+1}}| + |B_{t_i, t_{i+1}}|^3)\right] \\ &\leq C \sum_{i=0}^{n-1} |\Delta t_{i+1}|^{\frac{3}{2}} \leq C |\pi|^{\frac{1}{2}} \rightarrow 0. \end{aligned} \quad (2.3.11)$$

Finally, by Example 2.3.1 we see that

$$|B_{t_i, t_{i+1}}|^2 - \Delta t_{i+1} = 2 \int_{t_i}^{t_{i+1}} B_{t_i, t} dB_t.$$

Clearly

$$\begin{aligned} \mathbb{E}\left[\int_0^T \left|\sum_{i=0}^{n-1} \partial_{xx}f(t_i, X_{t_i}) B_{t_i, t} \mathbf{1}_{[t_i, t_{i+1})}\right|^2 dt\right] &= \mathbb{E}\left[\sum_{i=0}^{n-1} \int_{t_i}^{t_{i+1}} \left|\partial_{xx}f(t_i, X_{t_i}) B_{t_i, t}\right|^2 dt\right] \\ &\leq C \sum_{i=0}^{n-1} \int_{t_i}^{t_{i+1}} (t - t_i) dt = C \sum_{i=0}^{n-1} |\Delta t_{i+1}|^2 \leq C |\pi| \rightarrow 0. \end{aligned}$$

Then

$$\sum_{i=0}^{n-1} \partial_{xx}f(t_i, X_{t_i}) [B_{t_i, t_{i+1}}^2 - \Delta t_{i+1}] = \int_0^T \sum_{i=0}^{n-1} \partial_{xx}f(t_i, X_{t_i}) B_{t_i, t} \mathbf{1}_{[t_i, t_{i+1})} dB_t \rightarrow 0, \text{ in } \mathbb{L}^2(\mathcal{F}_T). \quad (2.3.12)$$

Plug (2.3.9)–(2.3.12) into (2.3.7) and (2.3.8), we prove (2.3.6).

*Step 2.* Assume that  $b_t = b_0$ ,  $\sigma_t = \sigma_0$  are  $\mathcal{F}_0$ -measurable and bounded, and  $f \in C^{1,2}$  with all related derivatives bounded. Let  $f^n$  be a smooth mollifier of  $f$ , see Problem 1.4.14. Then  $f^n$  is smooth with all the related derivatives bounded with a constant  $C_n$  which may depend on  $n$ , and for  $\varphi = \partial_t f, \partial_x f, \partial_{xx} f$ ,

$$\varphi^n \rightarrow \varphi \quad \text{and} \quad |\varphi^n| \leq C \quad \text{where } C \text{ is independent of } n.$$

By Step 1, we have

$$f^n(T, X_T) = f^n(0, X_0) + \int_0^T \left[ \partial_t f^n + \partial_x f^n b_0 + \frac{1}{2} \partial_{xx} f^n |\sigma_0|^2 \right](t, X_t) dt + \int_0^T \partial_x f^n(t, X_t) \sigma_0 dB_t.$$

Send  $n \rightarrow \infty$ , we prove (2.3.6) for  $f$  immediately.

*Step 3.* Assume  $b = \sum_{i=0}^{n-1} b_{t_i} \mathbf{1}_{[t_i, t_{i+1})} \in \mathbb{L}_0^2(\mathbb{F})$ ,  $\sigma = \sum_{i=0}^{n-1} \sigma_{t_i} \mathbf{1}_{[t_i, t_{i+1})} \in \mathbb{L}_0^2(\mathbb{F})$  are bounded, and  $f \in C^{1,2}$  with all related derivatives bounded. Applying Step 2 on  $[t_i, t_{i+1}]$  one can easily see that

$$f(t_{i+1}, X_{t_{i+1}}) = f(t_i, X_{t_i}) + \int_{t_i}^{t_{i+1}} \left[ \partial_t f + \partial_x f b_{t_i} + \frac{1}{2} \partial_{xx} f |\sigma_{t_i}|^2 \right](t, X_t) dt + \int_{t_i}^{t_{i+1}} \partial_x f(t, X_t) \sigma_{t_i} dB_t.$$

Sum over all  $i$  we obtain the result.

*Step 4.* Assume  $b \in \mathbb{L}^1(\mathbb{F})$ ,  $\sigma \in \mathbb{L}^2(\mathbb{F})$ , and  $f \in C^{1,2}$  with all related derivatives bounded. Analogous to Lemma 2.2.5, one can easily show that there exist bounded  $b^n, \sigma^n \in \mathbb{L}_0^2(\mathbb{F})$  such that

$$\lim_{n \rightarrow \infty} \|b^n - b\|_1 = 0, \quad \lim_{n \rightarrow \infty} \|\sigma^n - \sigma\|_2 = 0.$$

Denote

$$X_t^n := X_0 + \int_0^t b_s^n ds + \int_0^t \sigma_s^n dB_s,$$

and note that

$$(X^n - X)_T^* \leq \int_0^T |b_t^n - b_t| dt + \sup_{0 \leq t \leq T} \left| \int_0^t [\sigma_s^n - \sigma_s] dB_s \right|.$$

Then by (2.2.11) we have

$$\lim_{n \rightarrow \infty} \mathbb{E} \left[ (X^n - X)_T^* \right] = 0, \quad \text{and thus } (X^n - X)_T^* \rightarrow 0 \text{ in probability.}$$

By Step 3, we have

$$f(T, X_T^n) = f(0, X_0) + \int_0^T \left[ \partial_t f + \partial_x f b_t^n + \frac{1}{2} \partial_{xx} f |\sigma_t^n|^2 \right](t, X_t^n) dt + \int_0^T \partial_x f(t, X_t^n) \sigma_t^n dB_t.$$



Send  $n \rightarrow \infty$ . Note that

$$\begin{aligned} & \mathbb{E} \left[ \int_0^T |\partial_x f(t, X_t^n) \sigma_t^n - \partial_x f(t, X_t) \sigma_t|^2 dt \right] \\ & \leq C \mathbb{E} \left[ \int_0^T [|\sigma_t^n - \sigma_t|^2 + |\partial_x f(t, X_t^n) - \partial_x f(t, X_t)|^2 |\sigma_t|^2] dt \right] \rightarrow 0, \end{aligned}$$

thanks to the Dominated Convergence Theorem. Then

$$\int_0^T \partial_x f(t, X_t^n) \sigma_t^n dB_t \rightarrow \int_0^T \partial_x f(t, X_t) \sigma_t dB_t \quad \text{in } \mathbb{L}^2(\mathcal{F}_T).$$

Similarly,

$$\int_0^T \left[ \partial_t f + \partial_x f b_t^n + \frac{1}{2} \partial_{xx} f |\sigma_t^n|^2 \right](t, X_t^n) dt \rightarrow \int_0^T \left[ \partial_t f + \partial_x f b_t + \frac{1}{2} \partial_{xx} f |\sigma_t|^2 \right](t, X_t) dt, \quad \text{in } \mathbb{L}^1(\mathcal{F}_T).$$

We thus obtain the result.

*Step 5.* We now show the general case, namely  $b \in \mathbb{L}_{loc}^1(\mathbb{F})$ ,  $\sigma \in \mathbb{L}_{loc}^2(\mathbb{F})$  and  $f \in C^{1,2}$ . For each  $n \geq 1$ , define

$$\tau_n := \inf \left\{ t \geq 0 : \int_0^t |b_s| ds + \int_0^t |\sigma_s|^2 ds + |X_t| + \int_0^t |\partial_x f(s, X_s) \sigma_s|^2 ds \geq n \right\} \wedge T, \quad (2.3.13)$$

and denote

$$b^n := b \mathbf{1}_{[0, \tau_n]}, \quad \sigma^n := \sigma \mathbf{1}_{[0, \tau_n]}, \quad X^n := X_{\tau_n \wedge \cdot},$$

and  $f^n \in C^{1,2}$  with bounded derivatives such that

$$f^n(t, x) = f(t, x), \quad \text{for all } 0 \leq t \leq T, |x| \leq n.$$

Then

$$X_t^n = X_0 + \int_0^t b_s^n ds + \int_0^t \sigma_s^n dB_s \quad \text{and} \quad |X^n| \leq n.$$

By Step 4, we have

$$\begin{aligned} f^n(T, X_T^n) &= f^n(0, X_0) + \int_0^T \left[ \partial_t f^n + \partial_x f^n b_t^n + \frac{1}{2} \partial_{xx} f^n |\sigma_t^n|^2 \right](t, X_t^n) dt \\ &\quad + \int_0^T \partial_x f^n(t, X_t^n) \sigma_t^n dB_t. \end{aligned}$$

This is equivalent to

$$f(T, X_{\tau_n}) = f(0, X_0) + \int_0^T \left[ \partial_t f + \partial_x f b_t^n + \frac{1}{2} \partial_{xx} f |\sigma_t^n|^2 \right] (t, X_t^n) dt + \int_0^T \partial_x f(t, X_t^n) \sigma_t^n dB_t. \quad (2.3.14)$$

Recall (2.2.12) for stochastic integration in  $\mathbb{L}_{loc}^2(\mathbb{F})$  and notice that (2.3.13) include the term  $\int_0^t |\partial_x f(s, X_s) \sigma_s|^2 ds$ , then

$$\int_0^T \partial_x f(t, X_t^n) \sigma_t^n dB_t = \int_0^{\tau_n} \partial_x f(t, X_t^n) \sigma_t^n dB_t = \int_0^{\tau_n} \partial_x f(t, X_t) \sigma_t dB_t.$$

Plug this into (2.3.14) and send  $n \rightarrow \infty$ . Note that, for  $n$  large enough,  $\tau_n = T$ ,  $b^n = b$ ,  $\sigma^n = \sigma$ ,  $X^n = X$ , a.s. This implies that (2.3.6) holds a.s.  $\blacksquare$

### 2.3.3 Itô Formula in Multidimensional Case

Let  $B = (B^1, \dots, B^d)^\top$  be a  $d$ -dimensional  $\mathbb{F}$ -Brownian Motion,  $b^i \in \mathbb{L}_{loc}^1(\mathbb{F})$ ,  $\sigma^{ij} \in \mathbb{L}_{loc}^2(\mathbb{F})$ ,  $1 \leq i \leq d_1$ ,  $1 \leq j \leq d$ . Set  $b := (b^1, \dots, b^{d_1})^\top$  and  $\sigma := (\sigma^{ij})_{1 \leq i \leq d_1, 1 \leq j \leq d}$  which take values in  $\mathbb{R}^{d_1}$  and  $\mathbb{R}^{d_1 \times d}$ , respectively. Let  $X = (X^1, \dots, X^{d_1})^\top$  satisfy

$$dX_t^i := b_t^i dt + \sum_{j=1}^d \sigma_t^{ij} dB_t^j, \quad i = 1, \dots, d_1; \quad \text{or equivalently,} \quad dX_t = b_t dt + \sigma_t dB_t. \quad (2.3.15)$$

Denote

$$\langle X \rangle_t := \int_0^t \sigma_s \sigma_s^\top ds \text{ taking values in } \mathbb{S}^{d_1}. \quad (2.3.16)$$

We have the following multidimensional Itô formula whose proof is analogous to that of Theorem 2.3.2 and is omitted.

**Theorem 2.3.3** Assume  $f : [0, T] \times \mathbb{R}^{d_1} \rightarrow \mathbb{R}$  is in  $C^{1,2}$ . Then

$$\begin{aligned} df(t, X_t) &= \partial_t f(t, X_t) dt + \partial_x f(t, X_t) dX_t + \frac{1}{2} \partial_{xx} f(t, X_t) : d\langle X \rangle_t \\ &= \left[ \partial_t f + \partial_x f b_t + \frac{1}{2} \partial_{xx} f : (\sigma_t \sigma_t^\top) \right] (t, X_t) dt + \partial_x f(t, X_t) \sigma_t dB_t \\ &= \left[ \partial_t f + \sum_{i=1}^{d_1} \partial_{x_i} f b_t^i + \frac{1}{2} \sum_{i,j=1}^{d_1} \sum_{k=1}^d \partial_{x_i x_j} f \sigma_t^{i,k} \sigma_t^{j,k} \right] (t, X_t) dt + \sum_{i=1}^{d_1} \sum_{j=1}^d \partial_{x_i} f(t, X_t) \sigma_t^{ij} dB_t^j. \end{aligned} \quad (2.3.17)$$

Throughout the book, we take the convention that  $\partial_x f = (\partial_{x_1} f, \dots, \partial_{x_{d_1}} f)$  is a row vector, and we note that  $\partial_{xx} f$  takes values in  $\mathbb{S}^{d_1}$ .

### 2.3.4 An Extended Itô Formula

For future purpose, we need to extend the Itô formula to the case where the drift term  $b_t dt$  is replaced with a bounded variational process  $A$ . For simplicity, we state the result only for the case  $d_1 = 1$ , but one may easily generalize it to multidimensional cases. Let  $B$  be a  $d$ -dimensional  $\mathbb{F}$ -Brownian Motion,  $\sigma \in \mathbb{L}_{loc}^2(\mathbb{F}, \mathbb{R}^{1 \times d})$ ,  $A \in \mathbb{L}^0(\mathbb{F}, \mathbb{R})$  is continuous in  $t$  and  $\bigvee_0^T A < \infty$ , a.s. Denote

$$dX_t := \sigma_t dB_t + dA_t, \quad \langle X \rangle_t := \int_0^t \sigma_s \sigma_s^\top ds. \quad (2.3.18)$$

We have the following extended Itô formula whose proof is left to the readers in Problem 2.10.4.

**Theorem 2.3.4** *Assume  $f : [0, T] \times \mathbb{R} \rightarrow \mathbb{R}$  is in  $C^{1,2}$ . Then*

$$\begin{aligned} df(t, X_t) &= \partial_t f(t, X_t) dt + \partial_x f(t, X_t) dX_t + \frac{1}{2} \partial_{xx} f(t, X_t) : d\langle X \rangle_t \\ &= \left[ \partial_t f + \frac{1}{2} \partial_{xx} f(\sigma_t \sigma_t^\top) \right](t, X_t) dt + \partial_x f(t, X_t) \sigma_t dB_t + \partial_x f(t, X_t) dA_t, \end{aligned} \quad (2.3.19)$$

where the last term is understood in the sense of (2.2.1).

## 2.4 The Burkholder-Davis-Gundy Inequality

As an application of the Itô formula, we prove the following important inequality due to Burkholder-Davis-Gundy. For any  $p > 0$  and  $\sigma \in \mathbb{L}^{2,p}(\mathbb{F}, \mathbb{R}^d) \subset \mathbb{L}_{loc}^2(\mathbb{F}, \mathbb{R}^d)$ , define  $M_t := \int_0^t \sigma_s \cdot dB_s$  and  $M^*$  by (1.2.4).

**Theorem 2.4.1 (Burkholder-Davis-Gundy Inequality)** *For any  $p > 0$ , there exist universal constants  $0 < c_p < C_p$ , depending only on  $p$  and  $d$ , such that*

$$c_p \mathbb{E} \left[ \left( \int_0^T |\sigma_t|^2 dt \right)^{\frac{p}{2}} \right] \leq \mathbb{E} [|M_T^*|^p] \leq C_p \mathbb{E} \left[ \left( \int_0^T |\sigma_t|^2 dt \right)^{\frac{p}{2}} \right]. \quad (2.4.1)$$

*Proof* We again assume  $d = 1$ . The case  $p = 2$  is exactly the Doob's maximum inequality in Theorem 2.2.7 and Lemma 2.2.4. Note that  $\langle M \rangle_t = \int_0^t \sigma_s^2 ds$ . Following the truncation arguments in Step 2 of Lemma 2.2.4, we may assume without loss of generality that

$$M_T^* \text{ and } \langle M \rangle_T \text{ are bounded.} \quad (2.4.2)$$

However, we shall emphasize that the constants  $C_p, c_p$  in the proof below will not depend on this bound. We proceed in several steps.

*Step 1.* We first prove the left inequality by using the right inequality. Apply Itô formula, we have

$$d|M_t|^2 = |\sigma_t|^2 dt + 2M_t \sigma_t dB_t. \quad (2.4.3)$$

Then

$$\langle M \rangle_T = \int_0^T |\sigma_t|^2 dt = M_T^2 - M_0^2 - 2 \int_0^T M_t \sigma_t dB_t.$$

Thus, by the right inequality and noting that  $ab \leq \frac{1}{2}[a^2 + b^2]$ , we have

$$\begin{aligned} \mathbb{E}[\langle M \rangle_T^{\frac{p}{2}}] &\leq C_p \mathbb{E}[|M_T^*|^p] + C_p \mathbb{E}\left[\left|\int_0^T M_t \sigma_t dB_t\right|^{\frac{p}{2}}\right] \leq C_p \mathbb{E}[|M_T^*|^p] + C_p \mathbb{E}\left[\left(\int_0^T |M_t \sigma_t|^2 dt\right)^{\frac{p}{4}}\right] \\ &\leq C_p \mathbb{E}[|M_T^*|^p] + C_p \mathbb{E}\left[|M_T^*|^{\frac{p}{2}} \langle M \rangle_T^{\frac{p}{4}}\right] \leq C_p \mathbb{E}[|M_T^*|^p] + \frac{1}{2} \mathbb{E}\left[\langle M \rangle_T^{\frac{p}{2}}\right]. \end{aligned}$$

This, together with (2.4.2), implies the left inequality.

*Step 2.* We next prove the right inequality for  $p \geq 2$ . By the same arguments as in (2.2.9), we have

$$\begin{aligned} \mathbb{E}[|M_T^*|^p] &= p \int_0^\infty \lambda^{p-1} \mathbb{P}(M_T^* \geq \lambda) d\lambda \leq p \int_0^\infty \lambda^{p-2} \mathbb{E}\left[|M_T| \mathbf{1}_{\{M_T^* \geq \lambda\}}\right] d\lambda \\ &= \mathbb{E}\left[p \int_0^\infty \lambda^{p-2} |M_T| \mathbf{1}_{\{M_T^* \geq \lambda\}} d\lambda\right] = \mathbb{E}\left[p |M_T| \int_0^{M_T^*} \lambda^{p-2} d\lambda\right] = \frac{p}{p-1} \mathbb{E}\left[|M_T| |M_T^*|^{p-1}\right]. \end{aligned}$$

Note that  $p$  and  $\frac{p}{p-1}$  are conjugates. Then by Hölder inequality we have

$$\mathbb{E}[|M_T^*|^p] \leq \frac{p}{p-1} \left(\mathbb{E}[|M_T|^p]\right)^{\frac{1}{p}} \left(\mathbb{E}[|M_T^*|^p]\right)^{\frac{p-1}{p}}.$$

This, together with (2.4.2), implies

$$\mathbb{E}[|M_T^*|^p] \leq \left(\frac{p}{p-1}\right)^p \mathbb{E}[|M_T|^p]. \quad (2.4.4)$$

On the other hand, by (2.4.3) and applying the Itô formula, we have

$$d(|M_t|^p) = d([|M_t|^2]^{\frac{p}{2}}) = \frac{1}{2} p(p-1) |M_t|^{p-2} |\sigma_t|^2 dt + |M_t|^{p-2} M_t \sigma_t dB_t.$$

By (2.4.2), clearly  $|M|^{p-2} M \sigma \in \mathbb{L}^2(\mathbb{F})$ . Then

$$\mathbb{E}\left[\int_0^T |M_t|^{p-2} M_t \sigma_t dB_t\right] = 0.$$

Thus

$$\mathbb{E}[|M_T|^p] = \frac{1}{2}p(p-1)\mathbb{E}\left[\int_0^T |M_t|^{p-2}|\sigma_t|^2 dt\right] \leq C_p \mathbb{E}\left[|M_T^*|^{p-2}\langle M \rangle_T\right].$$

Note that  $\frac{p}{p-2}$  and  $\frac{p}{2}$  are conjugates. Applying Hölder inequality again we obtain

$$\mathbb{E}[|M_T|^p] \leq C_p \mathbb{E}\left[|M_T^*|^{p-2}\langle M \rangle_T\right] \leq C_p \left(\mathbb{E}[|M_T^*|^p]\right)^{\frac{p-2}{p}} \left(\mathbb{E}\left[\langle M \rangle_T^{\frac{p}{2}}\right]\right)^{\frac{2}{p}}.$$

This, together with (2.4.4) and (2.4.2), implies the right inequality in (2.4.1) immediately.

*Step 3.* We finally prove the right inequality for  $0 < p < 2$ . Note that

$$\mathbb{E}\left[\int_0^T |\langle M \rangle_t^{\frac{p-2}{4}} \sigma_t|^2 dt\right] = \mathbb{E}\left[\int_0^T \langle M \rangle_t^{\frac{p-2}{2}} d\langle M \rangle_t\right] = \frac{2}{p} \mathbb{E}[\langle M \rangle_T^{\frac{p}{2}}] < \infty.$$

Then  $N_t := \int_0^t \langle M \rangle_s^{\frac{p-2}{4}} \sigma_s dB_s$  is a square integrable martingale and  $\mathbb{E}[N_T^2] = \frac{2}{p} \mathbb{E}[\langle M \rangle_T^{\frac{p}{2}}]$ . Apply Itô formula, we have

$$M_t = \int_0^t \langle M \rangle_s^{\frac{2-p}{4}} dN_s = \langle M \rangle_t^{\frac{2-p}{4}} N_t - \int_0^t N_s d\langle M \rangle_s^{\frac{2-p}{4}}$$

Note that  $\langle M \rangle$  is increasing in  $t$ . Then

$$M_T^* \leq \langle M \rangle_T^{\frac{2-p}{4}} N_T^* + \int_0^T |N_s| d\langle M \rangle_s^{\frac{2-p}{4}} \leq C N_T^* \langle M \rangle_T^{\frac{2-p}{4}}.$$

Note that  $\frac{2}{p}$  and  $\frac{2}{2-p}$  are conjugates. Applying the Hölder inequality and then the Doob's maximum inequality Lemma 2.2.4, we have

$$\begin{aligned} \mathbb{E}[|M_T^*|^p] &\leq C_p \mathbb{E}\left[|N_T^*|^p \langle M \rangle_T^{\frac{p(2-p)}{4}}\right] \leq C_p \left(\mathbb{E}[|N_T^*|^2]\right)^{\frac{p}{2}} \left(\mathbb{E}[\langle M \rangle_T^{\frac{p}{2}}]\right)^{\frac{2-p}{2}} \\ &\leq C_p \left(\mathbb{E}[|N_T|^2]\right)^{\frac{p}{2}} \left(\mathbb{E}[\langle M \rangle_T^{\frac{p}{2}}]\right)^{\frac{2-p}{2}} = C_p \left(\mathbb{E}[\langle M \rangle_T^{\frac{p}{2}}]\right)^{\frac{p}{2}} \left(\mathbb{E}[\langle M \rangle_T^{\frac{p}{2}}]\right)^{\frac{2-p}{2}} = C_p \mathbb{E}[\langle M \rangle_T^{\frac{p}{2}}]. \end{aligned}$$

This completes the proof. ■

**Corollary 2.4.2** *Let  $\sigma \in \mathbb{L}^{2,1}(\mathbb{F}, \mathbb{R}^d) \subset \mathbb{L}_{loc}^2(\mathbb{F}, \mathbb{R}^d)$ . Then  $M_t := \int_0^t \sigma_s \cdot dB_s$  is a u.i. martingale.*

*Proof* Apply the Burkholder-Davis-Gundy Inequality Theorem 2.4.1 with  $p = 1$ , we have

$$\mathbb{E}[M_T^*] \leq C \mathbb{E}\left[\left(\int_0^T |\sigma_t|^2 dt\right)^{\frac{1}{2}}\right] < \infty.$$

Then the local martingale  $M$  is a u.i. martingale. ■

## 2.5 The Martingale Representation Theorem

Given  $\sigma \in \mathbb{L}^2(\mathbb{F}, \mathbb{R}^d)$ , it is known that  $M_t := \int_0^t \sigma_s \cdot dB_s$  is a square integrable  $\mathbb{F}$ -martingale. The Martingale Representation Theorem deals with the opposite issue: given a square integrable  $\mathbb{F}$ -martingale  $M$ , does there exist  $\sigma \in \mathbb{L}^2(\mathbb{F}, \mathbb{R}^d)$  such that  $M_t = M_0 + \int_0^t \sigma_s \cdot dB_s$ ?

The answer to the above question is in general negative.

**Example 2.5.1** Let  $d = 1$  and  $B, \tilde{B}$  be independent  $\mathbb{F}$ -Brownian Motion. Then  $\tilde{B}$  is a square integrable  $\mathbb{F}$ -martingale, but there is no  $\sigma \in \mathbb{L}^2(\mathbb{F})$  such that  $\tilde{B}_t = \int_0^t \sigma_s dB_s$ .

*Proof* We prove by contradiction. Assume  $\tilde{B}_t = \int_0^t \sigma_s dB_s$  for some  $\sigma \in \mathbb{L}^2(\mathbb{F})$ . On one hand, for  $X_t^1 := \int_0^t \sigma_s dB_s$  and  $X_t^2 := \int_0^t 1 d\tilde{B}_s$ , applying Itô formula (2.3.17) we have

$$d|\tilde{B}_t|^2 = d(X_t^1 X_t^2) = X_t^1 d\tilde{B}_t + X_t^2 \sigma_t dB_t$$

and thus  $|\tilde{B}_t|^2$  is a local martingale. On the other hand, applying Itô formula (2.3.5) directly on  $|\tilde{B}|^2$  we obtain

$$d|\tilde{B}_t|^2 = 2\tilde{B}_t d\tilde{B}_t + dt$$

and thus it is not a local martingale. Contradiction. ■

The key issue here is that  $\tilde{B}$  is independent of  $B$  and thus is not  $\mathbb{F}^B$ -measurable. We have the following important result by using the filtration  $\mathbb{F}^B$ .

**Theorem 2.5.2** For any  $\xi \in \mathbb{L}^2(\mathcal{F}_T^B)$ , there exists unique  $\sigma \in \mathbb{L}^2(\mathbb{F}^B, \mathbb{R}^d)$  such that

$$\xi = \mathbb{E}[\xi] + \int_0^T \sigma_t \cdot dB_t. \quad (2.5.1)$$

Consequently, for any  $\mathbb{F}^B$ -martingale  $M$  such that  $\mathbb{E}[|M_T|^2] < \infty$ , there exists unique  $\sigma \in \mathbb{L}^2(\mathbb{F}^B, \mathbb{R}^d)$  such that

$$M_t = M_0 + \int_0^t \sigma_s \cdot dB_s. \quad (2.5.2)$$

*Proof* Again we assume  $d = 1$  for simplicity. First note that (2.5.2) is a direct consequence of (2.5.1). Indeed, for any  $\mathbb{F}^B$ -martingale  $M$  such that  $\mathbb{E}[|M_T|^2] < \infty$ , by (2.5.1) there exists unique  $\sigma \in \mathbb{L}^2(\mathbb{F}^B)$  such that

$$M_T = \mathbb{E}[M_T] + \int_0^T \sigma_t dB_t.$$

Denote

$$\tilde{M}_t := \mathbb{E}[M_T] + \int_0^t \sigma_s dB_s.$$

Then  $\tilde{M}$  is an  $\mathbb{F}^B$ -martingale and  $\tilde{M}_T = M_T$ . Thus

$$M_t = \mathbb{E}[M_T | \mathcal{F}_t^B] = \mathbb{E}[\tilde{M}_T | \mathcal{F}_t^B] = \tilde{M}_t.$$

In particular,

$$M_0 = \tilde{M}_0 = \mathbb{E}[M_T].$$

This implies (2.5.2) immediately.

We next prove the uniqueness of  $\sigma$  in (2.5.1). If there is another  $\tilde{\sigma} \in \mathbb{L}^2(\mathbb{F}^B)$  satisfying (2.5.1). Then

$$\int_0^T (\sigma_t - \tilde{\sigma}_t) dB_t = 0.$$

Square both sides and take expectations, we get

$$\mathbb{E}\left[\int_0^T |\sigma_t - \tilde{\sigma}_t|^2 dt\right] = 0.$$

That is,

$$\tilde{\sigma} = \sigma, \quad dt \times dP - \text{a.s.}$$

It remains to prove the existence in (2.5.1). We proceed in several steps.

*Step 1.* Assume  $\xi = g(B_T)$ , where  $g \in C_b^2(\mathbb{R})$ . Define

$$u(t, x) := \mathbb{E}\left[g(x + B_{T-t})\right] = \int_{\mathbb{R}} g(y) p(T-t, y-x) dy, \quad \text{where } p(t, x) := \frac{1}{\sqrt{2\pi t}} e^{-\frac{x^2}{2t}}. \quad (2.5.3)$$

Note that

$$\begin{aligned} \partial_t p(t, x) &= \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2t}} \left[ -\frac{1}{2} t^{-\frac{3}{2}} + \frac{x^2}{2} t^{-\frac{5}{2}} \right] \\ \partial_x p(t, x) &= \frac{1}{\sqrt{2\pi t}} e^{-\frac{x^2}{2t}} \left( -\frac{x}{t} \right), \quad \partial_{xx} p(t, x) = \frac{1}{\sqrt{2\pi t}} e^{-\frac{x^2}{2t}} \left[ \frac{x^2}{t^2} - \frac{1}{t} \right]. \end{aligned}$$

Then

$$\partial_t p(t, x) - \frac{1}{2} \partial_{xx} p(t, x) = 0.$$

One can easily check that  $u \in C_b^{1,2}([0, T] \times \mathbb{R})$  and

$$\partial_t u(t, x) + \frac{1}{2} \partial_{xx} u(t, x) = 0, \quad u(T, x) = g(x). \quad (2.5.4)$$

Now define

$$M_t := u(t, B_t), \quad \sigma_t := \partial_x u(t, B_t). \quad (2.5.5)$$

Apply Itô formula we have

$$du(t, B_t) = u_x(t, B_t)dB_t + [\partial_t u + \frac{1}{2}\partial_{xx}u](t, B_t)dt = \sigma_t dB_t.$$

Thus

$$g(B_T) = u(T, B_T) = u(0, 0) + \int_0^T \sigma_t dB_t = \mathbb{E}[g(B_T)] + \int_0^T \sigma_t dB_t.$$

Since  $\partial_x u$  is bounded, we see that  $\sigma \in \mathbb{L}^2(\mathbb{F}^B)$ , and therefore, (2.5.1) holds.

*Step 2.* Assume  $\xi = g(B_T)$  where  $g : \mathbb{R} \rightarrow \mathbb{R}$  is Borel measurable and bounded. Let  $g_n$  be a smooth mollifier of  $g$  as in Problem 1.4.14. Then  $g_n \in C_b^2(\mathbb{R})$  for each  $n$ ,  $|g_n| \leq C$  for all  $n$ , and  $g_n(x) \rightarrow g(x)$  for  $dx$ -a.e.  $x$ . Since  $B_T$  has density, the probability that  $B_T$  lies in a Lebesgue null set is 0. Then  $g_n(B_T) \rightarrow g(B_T)$  a.s. Applying the Dominated Convergence Theorem we get  $\lim_{n \rightarrow \infty} \mathbb{E}[|g_n(B_T) - g(B_T)|^2] = 0$ . Now for each  $n$ , by Step 1 there exists  $\sigma^n \in \mathbb{L}^2(\mathbb{F})$  such that  $g_n(B_T) = \mathbb{E}[g_n(B_T)] + \int_0^T \sigma_t^n dB_t$ . Then (2.5.1) follows from Problem 2.10.5.

*Step 3.* Assume  $\xi = g(B_{t_1}, \dots, B_{t_n})$ , where  $0 < t_1 < \dots < t_n \leq T$  and  $g : \mathbb{R}^n \rightarrow \mathbb{R}$  is Borel measurable and bounded. Denote  $g_n(x_1, \dots, x_n) := g(x_1, \dots, x_n)$ . Apply Step 2 on  $[t_{n-1}, t_n]$ , there exists  $\sigma^n \in \mathbb{L}^2(\mathbb{F}^B)$  such that

$$\begin{aligned} g_n(B_{t_1}, \dots, B_{t_n}) &= \mathbb{E}\left[g_n(B_{t_1}, \dots, B_{t_n}) | \mathcal{F}_{t_{n-1}}^B\right] + \int_{t_{n-1}}^{t_n} \sigma_t^n dB_t \\ &= g_{n-1}(B_{t_1}, \dots, B_{t_{n-1}}) + \int_{t_{n-1}}^{t_n} \sigma_t^n dB_t, \end{aligned}$$

where, since  $B$  has independent increments,

$$g_{n-1}(x_1, \dots, x_{n-1}) := \mathbb{E}\left[g_n(x_1, \dots, x_{n-1}, x_{n-1} + B_{t_{n-1}, t_n})\right]$$

is also Borel measurable and bounded. Repeating the arguments backwardly in time, we obtain

$$g_{i+1}(B_{t_1}, \dots, B_{t_{i+1}}) = g_i(B_{t_1}, \dots, B_{t_i}) + \int_{t_i}^{t_{i+1}} \sigma_t^{i+1} dB_t,$$

where

$$g_i(x_1, \dots, x_i) := \mathbb{E}\left[g_{i+1}(x_1, \dots, x_i, x_i + B_{t_i, t_{i+1}})\right].$$



Define

$$\sigma := \sum_{i=1}^n \sigma^i \mathbf{1}_{[t_{i-1}, t_i]}.$$

Then one can easily see that  $\sigma \in \mathbb{L}^2(\mathbb{F}^B)$  and satisfies the requirement.

*Step 4.* Assume  $\xi \in \mathbb{L}^\infty(\mathcal{F}_T^B)$ . For each  $n$ , denote  $t_i^n := \frac{iT}{2^n}$ ,  $i = 0, \dots, 2^n$ . Let  $\mathcal{F}_T^n$  be the  $\sigma$ -field generated by  $\{B_{t_i^n}, 0 \leq i \leq 2^n\}$  and define  $\xi_n := \mathbb{E}[\xi | \mathcal{F}_T^n]$ . By the Doob-Dynkin lemma we have

$$\xi_n = g_n(B_{t_1^n}, \dots, B_{t_{2^n}^n}) \quad \text{for some Borel measurable function } g_n.$$

Since  $\xi$  is bounded, then so is  $\xi_n$  and thus  $g_n$  is bounded. By Step 3 we get

$$\xi_n = E[\xi_n] + \int_0^T \sigma_t^n dB_t \quad \text{for some } \sigma^n \in \mathbb{L}^2(\mathbb{F}^B).$$

Since  $B$  is continuous, it is clear that  $\mathcal{F}_T^B := \vee_n \mathcal{F}_T^n$ . Note that  $\mathbb{E}[\xi | \mathcal{F}_T^B] = \xi$ . Then by Problem 1.4.2 (iii) and the Dominated Convergence Theorem we have

$$\lim_{n \rightarrow \infty} \mathbb{E}[|\xi_n - \xi|^2] = 0.$$

Now (2.5.1) again follows from Problem 2.10.5.

*Step 5.* In the general case, for each  $n$ , let  $\xi_n := (-n) \vee \xi \wedge n$ . Then  $|\xi_n| \leq n$  and thus by Step 4, there exists  $\sigma^n \in \mathbb{L}^2(\mathbb{F}^B)$  such that

$$\xi_n = \mathbb{E}[\xi_n] + \int_0^T \sigma_t^n dB_t.$$

Clearly  $\xi_n \rightarrow \xi$  for all  $\omega$ . Moreover,  $|\xi_n| \leq |\xi|$ . Then by the Dominated Convergence Theorem we have

$$\lim_{n \rightarrow \infty} \mathbb{E}[|\xi_n - \xi|^2] = 0,$$

and thus (2.5.1) follows from Problem 2.10.5 again. ■

**Remark 2.5.3** In the financial application in Section 2.8, the stochastic integrand  $\sigma$  is related to the hedging portfolio. In particular, from (2.5.5) we see that  $\sigma$  is the derivative of  $M$  with respect to  $B$ , and thus is closely related to the so-called delta hedging. In fact, this connection is true even in non-Markov case, by introducing the path derivatives in Section 9.4. ■

**Remark 2.5.4** The condition that  $\xi$  is  $\mathcal{F}_T^B$ -measurable is clearly crucial in Theorem 2.5.2. When  $\xi \in \mathbb{L}^2(\mathcal{F}_T)$  and  $\mathbb{F}$  is larger than  $\mathbb{F}^B$ , we may have the following

extended martingale representation theorem: there exists unique  $\sigma \in \mathbb{L}^2(\mathbb{F}, \mathbb{R}^d)$  such that

$$\xi = \mathbb{E}[\xi] + \int_0^T \sigma_t dB_t + N_T, \quad (2.5.6)$$

where  $N \in \mathbb{L}^2(\mathbb{F})$  is a martingale orthogonal to  $B$ , in the sense that the quadric covariation  $\langle N, B \rangle = 0$ , or equivalently that  $NB$  is also a martingale. See, e.g., Protter [196].  $\blacksquare$

## 2.6 The Girsanov Theorem

In this section we shall derive another probability measure from  $\mathbb{P}$ . To distinguish the two probability measures, we shall write  $\mathbb{P}$  explicitly. Recall that  $B$  is a  $d$ -dimensional  $(\mathbb{P}, \mathbb{F})$ -Brownian motion. Let  $\theta \in \mathbb{L}_{loc}^2(\mathbb{F}, \mathbb{P}, \mathbb{R}^d)$ , and define

$$M_t^\theta := \exp \left( \int_0^t \theta_s \cdot dB_s - \frac{1}{2} \int_0^t |\theta_s|^2 ds \right), \text{ which implies } M_t^\theta = 1 + \int_0^t M_s^\theta \theta_s \cdot dB_s. \quad (2.6.1)$$

Then  $M^\theta$  is a  $\mathbb{P}$ -local martingale. Moreover, we have

**Lemma 2.6.1** *Assume  $\theta \in \mathbb{L}^\infty(\mathbb{F}, \mathbb{P}; \mathbb{R}^d)$ . Then  $M^\theta \in \bigcap_{1 \leq p < \infty} \mathbb{L}^{\infty, p}(\mathbb{F}, \mathbb{P})$ . In particular,  $M^\theta$  is a u.i.  $(\mathbb{P}, \mathbb{F})$ -martingale.*

*Proof* For simplicity again we assume  $d = 1$ . Denote  $X_t := \int_0^t \theta_s dB_s$ . Since  $|\theta| \leq C_0$  for some constant  $C_0 > 0$ , by the Burkholder-Davis-Gundy Inequality we see that  $X \in \bigcap_{n \geq 1} \mathbb{L}^{\infty, n}(\mathbb{F}, \mathbb{P})$ . For  $n \geq 1$ , applying Itô formula we have

$$X_t^{2n} = 2n \int_0^t X_s^{2n-1} \theta_s dB_s + n(2n-1) \int_0^t X_s^{2n-2} |\theta_s|^2 ds.$$

Then

$$\mathbb{E}^\mathbb{P}[|X_t|^{2n}] = n(2n-1) \mathbb{E}^\mathbb{P} \left[ \int_0^t X_s^{2n-2} |\theta_s|^2 ds \right] \leq \frac{1}{2} C_0^2 (2n)(2n-1) \int_0^t \mathbb{E}^\mathbb{P}[|X_s|^{2n-2}] ds.$$

By induction one can easily check that

$$\mathbb{E}^\mathbb{P}[|X_t|^{2n}] \leq \frac{C_0^{2n}}{2^n} t^{2n} \leq \left( \frac{C_0 T}{\sqrt{2}} \right)^{2n}. \quad (2.6.2)$$

Then clearly  $\mathbb{E}[|X_t|^n] \leq C_1^n$ ,  $n \geq 1$ , for some constant  $C_1 > 0$ . Note that

$$|M_t^\theta|^p = \exp \left( pX_t - \frac{p}{2} \int_0^t |\theta_s|^2 ds \right) \leq \exp(pX_t) = \sum_{n=0}^{\infty} \frac{p^n X_t^n}{n!}.$$

Then

$$\mathbb{E}^{\mathbb{P}}[|M_t^\theta|^p] \leq \sum_{n=0}^{\infty} \frac{p^n}{n!} \mathbb{E}^{\mathbb{P}}[|X_t|^n] \leq \sum_{n=0}^{\infty} \frac{p^n C_1^n}{n!} = e^{pC_1} < \infty.$$

Now it follows from the Burkholder-Davis-Gundy inequality that  $M^\theta \in \bigcap_{1 \leq p < \infty} \mathbb{L}^{\infty,p}(\mathbb{F}, \mathbb{P})$ .  $\blacksquare$

Clearly  $M^\theta > 0$ , and the above lemma implies  $\mathbb{E}^{\mathbb{P}}[M_T^\theta] = M_0^\theta = 1$ . Then one can easily check that the following  $\mathbb{P}^\theta$  is a probability measure equivalent to  $\mathbb{P}$ :

$$\mathbb{P}^\theta(A) := \mathbb{E}^{\mathbb{P}}[M_T^\theta \mathbf{1}_A], \quad \forall A \in \mathcal{F}_T, \quad \text{or equivalently,} \quad d\mathbb{P}^\theta := M_T^\theta d\mathbb{P}. \quad (2.6.3)$$

We have the following lemma whose proof is left to the exercise.

**Lemma 2.6.2** *Let  $\xi \in \mathbb{L}^0(\mathcal{F}_T)$ . Then  $\mathbb{E}^{\mathbb{P}^\theta}[|\xi|] < \infty$  if and only if  $\mathbb{E}^{\mathbb{P}}[M_T^\theta |\xi|] < \infty$ . Moreover,*

$$\mathbb{E}^{\mathbb{P}^\theta}[\xi] = \mathbb{E}^{\mathbb{P}}[M_T^\theta \xi].$$

The next result is crucial.

**Lemma 2.6.3** *Let  $X \in \mathbb{L}^0(\mathbb{F})$  such that  $\mathbb{E}^{\mathbb{P}^\theta}[|X_t|] < \infty$  for each  $t$ . Then  $X$  is a  $\mathbb{P}^\theta$ -martingale if and only if  $M^\theta X$  is a  $\mathbb{P}$ -martingale. In particular,  $(M^\theta)^{-1}$  is a  $\mathbb{P}^\theta$ -martingale.*

*Proof* First, by Lemma 2.6.2 we see that  $\mathbb{E}^{\mathbb{P}^\theta}[|X_t|] < \infty$  implies

$$\mathbb{E}^{\mathbb{P}}[M_t^\theta |X_t|] = \mathbb{E}^{\mathbb{P}}[\mathbb{E}^{\mathbb{P}}[M_T^\theta | \mathcal{F}_t] |X_t|] = \mathbb{E}^{\mathbb{P}}[M_T^\theta |X_t|] = \mathbb{E}^{\mathbb{P}^\theta}[|X_t|] < \infty.$$

We claim that, for any  $\xi \in \mathbb{L}^1(\mathcal{F}_T, \mathbb{P}^\theta)$ ,

$$\mathbb{E}^{\mathbb{P}^\theta}[\xi | \mathcal{F}_t] = (M_t^\theta)^{-1} \mathbb{E}^{\mathbb{P}}[M_T^\theta \xi | \mathcal{F}_t], \quad (2.6.4)$$

Notice that  $X$  is a  $\mathbb{P}^\theta$ -martingale if and only if  $X_t = \mathbb{E}^{\mathbb{P}^\theta}[X_T | \mathcal{F}_t]$ . By (2.6.4), this is equivalent to  $M_t^\theta X_t = \mathbb{E}^{\mathbb{P}}[M_T^\theta \xi | \mathcal{F}_t]$ , which amounts to saying that  $M^\theta X$  is a  $\mathbb{P}$ -martingale.

We now prove (2.6.4). For any  $\eta \in \mathbb{L}^\infty(\mathcal{F}_t, \mathbb{P}) = \mathbb{L}^\infty(\mathcal{F}_t, \mathbb{P}^\theta)$ , applying Lemma 2.6.2 twice and noting that  $M^\theta$  is a  $\mathbb{P}$ -martingale we have

$$\begin{aligned} \mathbb{E}^{\mathbb{P}^\theta}[(M_t^\theta)^{-1} \mathbb{E}^{\mathbb{P}}[M_T^\theta \xi | \mathcal{F}_t] \eta] &= \mathbb{E}^{\mathbb{P}}[M_T^\theta (M_t^\theta)^{-1} \mathbb{E}^{\mathbb{P}}[M_T^\theta \xi | \mathcal{F}_t] \eta] \\ &= \mathbb{E}^{\mathbb{P}}[\mathbb{E}^{\mathbb{P}}[M_T^\theta \xi | \mathcal{F}_t] \eta] = \mathbb{E}^{\mathbb{P}}[M_T^\theta \xi \eta] = \mathbb{E}^{\mathbb{P}^\theta}[\xi \eta], \end{aligned}$$

which implies (2.6.4) immediately.  $\blacksquare$

We now prove the main result of this section.

**Theorem 2.6.4** Let  $\theta \in \mathbb{L}^\infty(\mathbb{F}, \mathbb{P}, \mathbb{R}^d)$ . The following  $B^\theta$  is a  $(\mathbb{P}^\theta, \mathbb{F})$ -Brownian motion:

$$B_t^\theta := B_t - \int_0^t \theta_s ds. \quad (2.6.5)$$

*Proof* For simplicity we assume  $d = 1$ . Apply Itô formula, we have

$$d(M_t^\theta B_t^\theta) = M_t^\theta [\theta_t B_t^\theta + 1] dB_t; \quad d\left(M_t^\theta [|B_t^\theta|^2 - t]\right) = M_t^\theta \left[2B_t^\theta + (|B_t^\theta|^2 - t)\theta_t\right] dB_t.$$

By Lemmas 2.6.1 and 2.6.3 we see that  $B^\theta$  and  $|B_t^\theta|^2 - t$  are  $\mathbb{P}^\theta$ -martingales.

To show that  $B^\theta$  is a  $(\mathbb{P}^\theta, \mathbb{F})$ -Brownian motion, we follow the arguments of the so-called Levy's characterization theorem. Fix  $0 \leq s < T$ . By the martingale properties we have

$$\mathbb{E}^{\mathbb{P}^\theta} [B_{s,t}^\theta | \mathcal{F}_s] = 0, \quad \mathbb{E}^{\mathbb{P}^\theta} [(B_{s,t}^\theta)^2 | \mathcal{F}_s] = t - s, \quad s \leq t \leq T. \quad (2.6.6)$$

Denote  $N_t^\theta := (M_s^\theta)^{-1} M_t^\theta$ . For each  $n \geq 2$ , applying Itô formula we have

$$d\left[N_t^\theta (B_{s,t}^\theta)^n\right] = [\cdots] dB_t + \frac{n(n-1)}{2} N_t^\theta (B_{s,t}^\theta)^{n-2} dt.$$

Then

$$\begin{aligned} \mathbb{E}^{\mathbb{P}^\theta} \left[ (B_{s,t}^\theta)^n | \mathcal{F}_s \right] &= \mathbb{E}^{\mathbb{P}} \left[ N_t^\theta (B_{s,t}^\theta)^n | \mathcal{F}_s \right] = \frac{n(n-1)}{2} \int_s^t \mathbb{E}^{\mathbb{P}} \left[ N_r^\theta (B_{s,r}^\theta)^{n-2} | \mathcal{F}_s \right] dr \\ &= \frac{n(n-1)}{2} \int_s^t \mathbb{E}^{\mathbb{P}^\theta} \left[ (B_{s,r}^\theta)^{n-2} | \mathcal{F}_s \right] dr. \end{aligned}$$

By induction one can easily derive from (2.6.6) that

$$\mathbb{E}^{\mathbb{P}^\theta} [(B_{s,t}^\theta)^{2n+1} | \mathcal{F}_s] = 0, \quad \mathbb{E}^{\mathbb{P}^\theta} [(B_{s,t}^\theta)^{2n} | \mathcal{F}_s] = \frac{(2n)!}{2^n n!} (t-s)^n. \quad (2.6.7)$$

Then, for any  $\alpha \in \mathbb{R}$ ,

$$\mathbb{E}^{\mathbb{P}^\theta} [e^{\alpha B_{s,t}^\theta} | \mathcal{F}_s] = \sum_{n=0}^{\infty} \frac{\alpha^n}{n!} \mathbb{E}^{\mathbb{P}^\theta} [(B_{s,t}^\theta)^n | \mathcal{F}_s] = \sum_{n=0}^{\infty} \frac{\alpha^n (t-s)^n}{2^n n!} = e^{\frac{\alpha(t-s)}{2}}. \quad (2.6.8)$$

This implies that, under  $\mathbb{P}^\theta$ ,  $B_{s,t}^\theta$  is independent of  $\mathcal{F}_s$  and has distribution  $N(0, t-s)$ . That is,  $B^\theta$  is a  $(\mathbb{P}^\theta, \mathbb{F})$ -Brownian motion. ■

**Remark 2.6.5** The above theorem is a special case of the Levy's martingale characterization of Brownian motion (see, e.g., Karatzas & Shreve [117]):

Let  $M$  be a continuous process with  $M_0 = 0$  and denote  $N_t := M_t^2 - t$ . Then  $M$  is a Brownian motion if and only if both  $M$  and  $N$  are martingales.

(2.6.9)

The result follows similar arguments, but involves the general martingale theory, and we omit it. ■

We conclude the section with the martingale representation theorem for  $(\mathbb{P}^\theta, B^\theta)$ . For  $\xi \in \mathbb{L}^2(\mathcal{F}_T^\theta, \mathbb{P}^\theta)$ , the result follows from the standard martingale representation Theorem 2.5.2. For  $\xi \in \mathbb{L}^2(\mathcal{F}_T, \mathbb{P}^\theta)$ , as seen in Example 2.5.1, the result is in general not true. The nontrivial interesting case is  $\xi \in \mathbb{L}^2(\mathcal{F}_T^B, \mathbb{P}^\theta)$ . We note that

for  $\theta \in \mathbb{L}^\infty(\mathbb{F}^B, \mathbb{P})$ , we have  $\mathbb{F}^{B^\theta} \subset \mathbb{F}^B$ , but in general  $\mathbb{F}^{B^\theta} \neq \mathbb{F}^B$ . (2.6.10)

A counterexample for  $\mathbb{F}^{B^\theta} \neq \mathbb{F}^B$  is provided by Tsirelson [229]. Nevertheless, we still have

**Theorem 2.6.6** Assume  $\theta \in \mathbb{L}^\infty(\mathbb{F}^B, \mathbb{P}, \mathbb{R}^d)$ . Then for any  $\xi \in \mathbb{L}^2(\mathcal{F}_T^B, \mathbb{P}^\theta)$ , there exists ( $\mathbb{P}^\theta$ -a.s.) unique  $\sigma \in \mathbb{L}^2(\mathbb{F}_T^B, \mathbb{P}^\theta, \mathbb{R}^d)$  such that

$$\xi = \mathbb{E}^{\mathbb{P}^\theta}[\xi] + \int_0^T \sigma_t \cdot dB_t^\theta.$$

We remark that in general we cannot expect  $\sigma$  to be  $\mathbb{F}^{B^\theta}$ -measurable.

*Proof* Assume for simplicity that  $d = 1$ . By the truncation arguments in Step 5 of Theorem 2.5.2, we may assume without loss of generality that  $\xi$  is bounded. Denote  $X_t := \mathbb{E}^{\mathbb{P}^\theta}[\xi | \mathcal{F}_t^B]$ . Then  $X$  is a bounded  $(\mathbb{P}^\theta, \mathbb{F}^B)$ -martingale. By Lemmas 2.6.1 and 2.6.3,  $M^\theta X$  is a  $(\mathbb{P}, \mathbb{F}^B)$ -square integrable martingale. By Theorem 2.5.2, there exists  $\tilde{\sigma} \in \mathbb{L}^2(\mathbb{F}^B, \mathbb{P})$  such that

$$d(M_t^\theta X_t) = \tilde{\sigma}_t dB_t.$$

Apply Itô formula, we have

$$\begin{aligned} d(M_t^\theta)^{-1} &= -(M_t^\theta)^{-2} M_t^\theta \theta_t dB_t + (M_t^\theta)^{-3} |M_t^\theta \theta_t|^2 dt = (M_t^\theta)^{-1} \left[ -\theta_t dB_t + |\theta_t|^2 dt \right]; \\ dX_t &= d \left[ (M_t^\theta)^{-1} (M_t^\theta X_t) \right] \\ &= (M_t^\theta)^{-1} \tilde{\sigma}_t dB_t + M_t^\theta X_t (M_t^\theta)^{-1} \left[ -\theta_t dB_t + |\theta_t|^2 dt \right] - \tilde{\sigma}_t (M_t^\theta)^{-1} \theta_t dt \\ &= \left[ (M_t^\theta)^{-1} \tilde{\sigma}_t - X_t \theta_t \right] dB_t. \end{aligned}$$

This proves the result with  $\sigma_t := (M_t^\theta)^{-1} \tilde{\sigma}_t - X_t \theta_t$ . ■

**Remark 2.6.7**

- (i) In the option pricing theory in Section 2.8, Girsanov theorem is a convenient tool to find the so-called risk neutral probability measure.
- (ii) In stochastic control theory, see Section 4.5.2, Girsanov theorem is a powerful tool to stochastic optimization problem with drift control in weak formulation.
- (iii) Note that  $\mathbb{P}^\theta$  is equivalent to  $\mathbb{P}$ . For stochastic optimization problem with diffusion control in weak formulation, the involved probability measures are typically mutually singular. Then Girsanov theorem is not enough. We shall introduce new tools in Part III to address these problems. ■

**Remark 2.6.8** The Girsanov theorem holds true under weaker assumptions on  $\theta$ , see Theorem 7.2.3 and Problem 7.5.2 below. ■

**2.7 The Doob-Meyer Decomposition**

The result in this section actually holds for general setting and under much weaker conditions, see, e.g., Karatzas & Shreve [117]. However, for simplicity we shall only present a special case.

**Theorem 2.7.1** Assume  $\mathbb{F} = \mathbb{F}^B$  and let  $X \in \mathbb{S}^2(\mathbb{F})$  be a continuous submartingale. Then there exists unique decomposition  $X_t = X_0 + \int_0^t Z_s \cdot dB_s + K_t$ , where  $Z \in \mathbb{L}^2(\mathbb{F}, \mathbb{R}^d)$ ,  $K \in \mathbb{I}^2(\mathbb{F})$  with  $K_0 = 0$ . Moreover, there exists a constant  $C > 0$ , depending only on  $d$ , such that

$$\mathbb{E} \left[ \int_0^T |Z_t|^2 dt + |K_T|^2 \right] \leq C \mathbb{E}[|X_T^*|^2]. \quad (2.7.1)$$

*Proof* For simplicity we assume  $d = 1$ . We first prove the uniqueness. Assume  $Z' \in \mathbb{L}^2(\mathbb{F})$  and  $K' \in \mathbb{I}^2(\mathbb{F})$  with  $K'_0 = 0$  provide another decomposition. Then, denoting  $\Delta Z := Z - Z'$ ,  $\Delta K := K - K'$ ,

$$\int_0^t \Delta Z_s dB_s = -\Delta K_t, \quad 0 \leq t \leq T.$$

For each  $n \geq 1$ , denote  $t_i := t_i^n := \frac{i}{n}T$ ,  $i = 0, \dots, n$ . Then, noting that  $K, K'$  are increasing,

$$\begin{aligned} \mathbb{E} \left[ \int_0^T |\Delta Z_t|^2 dt \right] &= \sum_{i=0}^{n-1} \mathbb{E} \left[ \left( \int_{t_i}^{t_{i+1}} \Delta Z_t dB_t \right)^2 \right] = \sum_{i=0}^{n-1} \mathbb{E} \left[ |\Delta K_{t_{i+1}} - \Delta K_{t_i}|^2 \right] \\ &= \sum_{i=0}^{n-1} \mathbb{E} \left[ |K_{t_i, t_{i+1}} - K'_{t_i, t_{i+1}}|^2 \right] \leq \sum_{i=0}^{n-1} \mathbb{E} \left[ |K_{t_i, t_{i+1}} + K'_{t_i, t_{i+1}}|^2 \right] \\ &\leq \mathbb{E} \left[ \sup_{0 \leq i \leq n-1} [K_{t_i, t_{i+1}} + K'_{t_i, t_{i+1}}] [K_T + K'_T] \right]. \end{aligned}$$

Since  $K, K'$  are continuous, send  $n \rightarrow \infty$  and apply the Dominated Convergence Theorem, we obtain  $\mathbb{E}\left[\int_0^T |\Delta Z_t|^2 dt\right] = 0$ . Then  $Z = Z'$ , which implies further that  $K = K'$ .

We now prove the existence. Let  $t_i$  be as above,  $M_{t_0}^n := K_{t_0}^n := 0$ , and for  $i = 0, \dots, n-1$ ,

$$M_{t_{i+1}}^n := M_{t_i}^n + X_{t_{i+1}} - \mathbb{E}_{t_i}[X_{t_{i+1}}], \quad K_{t_{i+1}}^n := K_{t_i}^n + \mathbb{E}_{t_i}[X_{t_{i+1}}] - X_{t_i}. \quad (2.7.2)$$

Then clearly  $M_{t_i}^n$  is an  $\{\mathcal{F}_{t_i}\}_{0 \leq i \leq n}$ -martingale and, since  $X$  is a submartingale,  $K_{t_i}^n \in \mathbb{L}^0(\mathcal{F}_{t_{i-1}})$  is increasing in  $i$ . Note that

$$\begin{aligned} \mathbb{E}\left[|X_{t_{i+1}}|^2 - |X_{t_i}|^2\right] &= \mathbb{E}\left[|M_{t_i, t_{i+1}}^n + K_{t_i, t_{i+1}}^n + X_{t_i}|^2 - |X_{t_i}|^2\right] \\ &= \mathbb{E}\left[|M_{t_i, t_{i+1}}^n|^2 + |K_{t_i, t_{i+1}}^n|^2 + 2X_{t_i}K_{t_i, t_{i+1}}^n\right] \\ &\geq \mathbb{E}\left[|M_{t_i, t_{i+1}}^n|^2 - 2X_T^*K_{t_i, t_{i+1}}^n\right]. \end{aligned}$$

This implies, noting that  $M^n$  is a martingale,

$$\begin{aligned} \mathbb{E}[|M_T^n|^2] &= \sum_{i=0}^{n-1} \mathbb{E}\left[|M_{t_i, t_{i+1}}^n|^2\right] \leq \mathbb{E}\left[|X_T|^2 - |X_0|^2 + 2X_T^*K_T^n\right] \\ &= \mathbb{E}\left[|X_T|^2 - |X_0|^2 + 2X_T^*[X_T - X_0 - M_T^n]\right] \leq \mathbb{E}\left[C|X_T^*|^2 + \frac{1}{2}|M_T^n|^2\right]. \end{aligned}$$

Then

$$\mathbb{E}[|M_T^n|^2] \leq C\mathbb{E}[|X_T^*|^2], \text{ which implies further that } \mathbb{E}[|K_T^n|^2] \leq C\mathbb{E}[|X_T^*|^2]. \quad (2.7.3)$$

Now by the martingale representation Theorem 2.5.2, for each  $n$  there exists  $Z^n \in \mathbb{L}^2(\mathbb{F})$  such that  $M_T^n = \int_0^T Z_t^n dB_t$ . Denote  $K_t^n := \sum_{i \geq 0} K_{t_i, t_{i+1}}^n \mathbf{1}_{[t_i, t_{i+1})}$ . By (2.7.3) and applying Theorem 1.3.7, we may assume without loss of generality that  $(Z^n, K^n)$  converges weakly to certain  $(Z, K) \in \mathbb{L}^2(\mathbb{F})$ . Applying Problem 2.10.11 (ii) and (iii) we see that  $M^n$  converges weakly to  $M := \int_0^T Z_s dB_s$  and  $\mathbb{E}[\int_0^T |Z_t|^2 dt] \leq C\mathbb{E}[|X_T^*|^2]$ . Moreover, since  $X_{t_i} = X_0 + M_{t_i}^n + K_{t_i}^n$  and  $X$  is continuous. By Problem 2.10.11 (i) it is clear that  $X_t = X_0 + M_t + K_t$ . In particular, this implies that  $K$  is continuous and  $\mathbb{E}[|K_T|^2] \leq C\mathbb{E}[|X_T^*|^2]$ .

It remains to show that  $K$  is increasing. Note that each  $K^n$  is increasing. Let  $\hat{K}^n$  be the convex combination of  $K^n$  as in Theorem 1.3.8, then  $\hat{K}^n$  is also increasing and  $\lim_{n \rightarrow \infty} \mathbb{E}\left[\int_0^T |\hat{K}_t^n - K_t|^2 dt\right] = 0$ . By otherwise choosing a further subsequence, we have  $\int_0^T |\hat{K}_t^n - K_t|^2 dt \rightarrow 0$ , a.s. This clearly implies that  $K$  is increasing, a.s. ■

## 2.8 A Financial Application

Consider the Black-Scholes model on a filtered probability space  $(\Omega, \mathcal{F}, \mathbb{P})$  with a one-dimensional  $(\mathbb{P}, \mathbb{F})$ -Brownian motion  $B$ . The financial market consists of two assets: a bank account (or bond) with constant interest rate  $r$  (continuously compounded), and a stock with price  $S_t$ :

$$dS_t = S_t \left[ \mu dt + \sigma dB_t \right], \text{ or equivalently } S_t = S_0 \exp \left( \sigma B_t + \left( \mu - \frac{1}{2} \sigma^2 \right) t \right), \quad (2.8.1)$$

where the constants  $\mu$  and  $\sigma > 0$  stand for the appreciation and volatility of the stock, respectively. Let  $\xi \in \mathbb{L}^2(\mathcal{F}_T)$  be a European option with maturity time  $T$ , namely at time  $T$  the option is worth  $\xi$ . Now our goal is to find the *fair* price  $Y_0$  of  $\xi$  at time 0, or more generally the fair price  $Y_t$  at time  $t \in [0, T]$ . Clearly  $Y_T = \xi$ .

### 2.8.1 Pricing via Risk Neutral Measure

We first note that, due to the presence of the interest, we should consider the discounted prices:

$$\bar{S}_t = e^{-rt} S_t, \quad \bar{Y}_t = e^{-rt} Y_t. \quad (2.8.2)$$

One natural guess for the option price is that

$$Y_0 = \bar{Y}_0 = \mathbb{E}^{\mathbb{P}}[\bar{Y}_T] = \mathbb{E}^{\mathbb{P}}[e^{-rT} \xi]. \quad (2.8.3)$$

However, the above guess cannot be true in general. Indeed, if we set  $\xi = S_T$ , then following (2.8.3) we should have  $\bar{Y}_0 = \mathbb{E}^{\mathbb{P}}[\bar{S}_T]$ , or more generally  $\bar{Y}_t = \mathbb{E}^{\mathbb{P}}[\bar{S}_T | \mathcal{F}_t]$ . That is,  $\bar{Y}$  should be a  $\mathbb{P}$ -martingale. However, obviously in this case we should have  $Y_t = S_t$  and thus  $\bar{Y}_t = \bar{S}_t$ . Applying Itô formula we have

$$d\bar{S}_t = \bar{S}_t \left[ (\mu - r) dt + \sigma dB_t \right]. \quad (2.8.4)$$

Then  $\bar{S}$  is not a  $\mathbb{P}$ -martingale unless  $\mu = r$ .

If we want to use price formula in the form of (2.8.3), from the above discussion it seems necessary that  $\bar{S}$  needs to be martingale. We thus introduce the following concept.

**Definition 2.8.1** A probability measure  $\bar{\mathbb{P}}$  on  $\Omega$  is called a *risk neutral measure*, also called *martingale measure*, if

- (i)  $\bar{\mathbb{P}}$  is equivalent to  $\mathbb{P}$ ;
- (ii)  $\bar{S}$  is a  $\bar{\mathbb{P}}$ -martingale.

In contrast to  $\bar{\mathbb{P}}$ , we call the original  $\mathbb{P}$  the market measure.



To construct  $\bar{\mathbb{P}}$ , our main tool is the Girsanov theorem. By (2.8.4), it is clear that

$$d\bar{S}_t = \bar{S}_t \sigma dB_t^{-\theta}, \quad (2.8.5)$$

where  $\theta := \frac{\mu-r}{\sigma}$  is the *Sharpe ratio* of the stock, and  $dB_t^{-\theta} := dB_t + \theta dt$ . Consider the  $\mathbb{P}^{-\theta}$  in Section 2.6. Then  $\mathbb{P}^{-\theta} \sim \mathbb{P}$  and  $B^{-\theta}$  is a  $\mathbb{P}^{-\theta}$ -Brownian motion. Now it follows from (2.8.5) that  $\bar{S}$  is a  $\mathbb{P}^{-\theta}$ -martingale, and thus  $\bar{\mathbb{P}} = \mathbb{P}^{-\theta}$  is a risk neutral measure.

We will justify in the next subsection that  $\bar{Y}$  should also be a  $\bar{\mathbb{P}}$ -martingale. Then we obtain the following pricing formula, in the spirit of (2.8.3) but under the risk neutral measure  $\bar{\mathbb{P}}$  instead of the market measure  $\mathbb{P}$ :

$$\bar{Y}_t = \mathbb{E}^{\bar{\mathbb{P}}}[e^{-rT}\xi | \mathcal{F}_t], \quad \text{or equivalently,} \quad Y_t = \mathbb{E}^{\bar{\mathbb{P}}}[e^{-r(T-t)}\xi | \mathcal{F}_t]. \quad (2.8.6)$$

## 2.8.2 Hedging the Option

Assume an investor invests in the market with portfolio  $(\lambda_t, h_t)_{0 \leq t \leq T}$ . The corresponding portfolio value is:

$$V_t := \lambda_t e^{rt} + h_t S_t. \quad (2.8.7)$$

Note that  $\mathbb{F}$  stands for the information flow, thus it is natural to require  $(\lambda, h)$  to be  $\mathbb{F}$ -measurable. Moreover, we shall assume the investor invests only in this market, which induces the following concept:

**Definition 2.8.2** *An  $\mathbb{F}$ -measurable portfolio  $(\lambda, h)$  is called self-financing if, in addition to certain integrability conditions which we do not discuss in detail,*

$$dV_t = \lambda_t de^{rt} + h_t dS_t. \quad (2.8.8)$$

The fairness of the price is based on the following arbitrage free principle.

### Definition 2.8.3

(i) *We say a self-financing portfolio  $(\lambda, h)$  has arbitrage opportunity if*

$$V_0 = 0, \quad V_T \geq 0, \quad \mathbb{P}\text{-a.s.}, \quad \text{and} \quad \mathbb{P}(V_T > 0) > 0. \quad (2.8.9)$$

(ii) *We say the market consisting of the bond and stock is arbitrage free if there is no self-financing portfolio  $(\lambda, h)$  admitting arbitrage opportunity.*

The following theorem is called the first fundamental theorem of mathematical finance, which holds true in much more general models.

**Theorem 2.8.4** *The market is arbitrage free if and only if there exists a risk neutral measure  $\bar{\mathbb{P}}$ .*

By the previous subsection, the Black-Scholes market is arbitrage free. We remark that, since  $\bar{\mathbb{P}}$  is equivalent to  $\mathbb{P}$ , so (2.8.9) holds under  $\bar{\mathbb{P}}$  as well.

Now given an option  $\xi$ , let  $Y_t$  denote its market price. We may consider an extended market  $(e^{rt}, S_t, Y_t)$ , and we can easily extend the concept of arbitrage free to this market.

**Definition 2.8.5** *We say  $Y$  is a fair price, also called arbitrage free price, if the market  $(e^{rt}, S_t, Y_t)$  is arbitrage free.*

**Definition 2.8.6** *Given  $\xi$ , we say a self-financing portfolio  $(\lambda, h)$  is a hedging portfolio of  $\xi$  if  $V_T = \xi$ ,  $\mathbb{P}$ -a.s.*

**Proposition 2.8.7** *If  $(\lambda, h)$  is a hedging portfolio of  $\xi$ , then  $Y_t := V_t$  is the unique fair price.*

*Proof* The fairness of  $V$  involves the martingale properties and we leave the proof to interested readers. To illustrate the main idea, we prove only that, if  $Y_0 > V_0$ , then there will be arbitrage opportunity in the extended market  $(e^{rt}, S_t, Y_t)$ . Indeed, in this case, consider the portfolio:  $(\lambda_t + Y_0 - V_0, h_t, -1)$ , with value

$$\tilde{V}_t := [\lambda_t + Y_0 - V_0]e^{rt} + h_t S_t - Y_t = V_t - Y_t + [Y_0 - V_0]e^{rt}.$$

Note that

$$\begin{aligned} d\tilde{V}_t &= dV_t - dY_t + [Y_0 - V_0]de^{rt} = \lambda_t de^{rt} + h_t dS_t - dY_t + [Y_0 - V_0]de^{rt} \\ &= [\lambda_t + Y_0 - V_0]de^{rt} + h_t dS_t + (-1)dY_t. \end{aligned}$$

That is, the portfolio is self-financing. Note that

$$\tilde{V}_0 = V_0 - Y_0 + [Y_0 - V_0]e^{r0} = 0;$$

$$\tilde{V}_T = V_T - Y_T + [Y_0 - V_0]e^{rT} = \xi - \xi + [Y_0 - V_0]e^{rT} = [Y_0 - V_0]e^{rT} > 0, \quad \mathbb{P}\text{-a.s.}$$

Then the portfolio  $(\lambda_t + Y_0 - V_0, h_t, -1)$  has arbitrage opportunity. ■

We next find the hedging portfolio in the Black-Scholes model. Our main tool is the martingale representation theorem. Consider the discounted portfolio value  $\bar{V}_t := e^{-rt} V_t$ . By (2.8.8) and (2.8.5) we have

$$d\bar{V}_t = h_t d\bar{S}_t = h_t \bar{S}_t \sigma dB_t^{-\theta}. \quad (2.8.10)$$

That is,  $\bar{V}$  is a  $\bar{\mathbb{P}}$ -martingale, where, again,  $\bar{\mathbb{P}} := \mathbb{P}^{-\theta}$ . Note that  $\bar{V}_T = e^{-rT} \xi$ . Assume

$$\xi \in \mathbb{L}^2(\mathcal{F}_T^B, \bar{\mathbb{P}}). \quad (2.8.11)$$

Then by the generalized martingale representation Theorem 2.6.6, there exists  $Z \in \mathbb{L}^2(\mathbb{F}^B, \bar{\mathbb{P}})$  such that

$$e^{-rT}\xi = \mathbb{E}^{\bar{\mathbb{P}}}[e^{-rT}\xi] + \int_0^T Z_t dB_t^{-\theta}. \quad (2.8.12)$$

This induces the hedging portfolio (and the price) immediately:

$$\bar{V}_t := \mathbb{E}^{\bar{\mathbb{P}}}[e^{-rT}\xi | \mathcal{F}_t^B], \quad h_t := \frac{Z_t}{\bar{S}_t \sigma}, \quad \lambda_t := \bar{V}_t - h_t \bar{S}_t = \bar{V}_t - \frac{Z_t}{\sigma}. \quad (2.8.13)$$

The hedging portfolio is closely related to the important notion of completeness of the market.

**Definition 2.8.8** *The market is called complete if all option  $\xi \in \mathbb{L}^0(\mathcal{F}_T)$  satisfying appropriate integrability condition can be hedged.*

From the above analysis we see that the Black-Scholes market is complete if

$$\mathbb{F} = \mathbb{F}^B. \quad (2.8.14)$$

We conclude this subsection with the second fundamental theorem of mathematical finance, which also holds true in much more general models.

**Theorem 2.8.9** *Assume the market is arbitrage free. Then the market is complete if and only if the risk neutral measure  $\bar{\mathbb{P}}$  is unique.*

### 2.8.3 Some Further Discussion

We first note that one rationale of using Brownian motion to model the stock price lies in the central limit Theorem 1.1.2. As a basic principle in finance, the supply and demand have great impact on the price. That is, the buy orders will push the stock price up, while the sell orders will push the stock price down. Assume there are many small investors in the market and they place their order independently. Then by the central limit Theorem 1.1.2, the accumulative price impact of their trading induces the normal distribution. In the rest of this subsection we discuss two subtle issues.

First, as we see in (2.8.14), even for Black-Scholes model, the completeness relies on the information setting. In a more general model,  $\mathbb{F}$ ,  $\mathbb{F}^B$ , and  $\mathbb{F}^S$  can be all different. The investor's portfolio  $(\lambda, h)$  has to be measurable with respect to the filtration the investor actually observes. While in different situation the real information can be different, typically the investor indeed observes  $S$  and thus  $\mathbb{F}^S$  is accessible to the investor. As discussed in the previous paragraph, observing  $B$  essentially means the investor observes numerous other (small) investors (and possibly other random factors). This is not that natural in practice. Moreover, note

that in Theorem 1.1.2, the convergence is in distribution sense, not in pointwise sense. Then even one observes a path of the portfolios of all small investors, one typically does not know a corresponding path of  $B$ . So in this sense, at least in some applications, it makes more sense to use  $\mathbb{F}^S$  than to use  $\mathbb{F}^B$ . This implies that in these applications one should use weak formulation, as we will do in Part III. In Parts I and II, however, we will nevertheless use strong formulation, namely use  $\mathbb{F}^B$ . This could be reasonable in some other applications, and still makes perfect sense in this particular application when  $\mathbb{F}^S = \mathbb{F}^B$ , which is true in, e.g., Black-Scholes model.

The next is the pathwise stochastic integration. Recall that for an elementary process  $\sigma \in \mathbb{L}_0^2(\mathbb{F})$ , the Itô integral  $(\int_0^T \sigma_t dB_t)(\omega) = \int_0^T \sigma_t(\omega) dB_t(\omega)$  is defined in pathwise manner. For general  $\sigma \in \mathbb{L}^2(\mathbb{F})$ , however,  $\int_0^T \sigma_t dB_t$  is defined as the  $\mathbb{L}^2$ -limit of  $\int_0^T \sigma_t^n dB_t$ , where  $\sigma^n \in \mathbb{L}_0^2(\mathbb{F})$  is an approximation of  $\sigma$ . As a consequence,  $\int_0^T \sigma_t dB_t$  is defined only in a.s. sense, with the null set arbitrary and up to the particular version we want to choose. In particular, for any given  $\omega$ , since  $\mathbb{P}(\{\omega\}) = 0$ , the value  $(\int_0^T \sigma_t dB_t)(\omega)$  is arbitrary. In other words, in our application, assume we have observed a path  $S_t(\omega)$  and decided a path  $h_t(\omega)$ , the value of  $\int_0^T h_t dS_t$  at this particular observed  $\omega$  is actually arbitrary. This is of course not desirable. We shall mention that in real practice, the portfolio  $h$  should be discrete, and thus the issue does not exist. But nevertheless, theoretically this is a subtle issue we face in such applications.

One way to get around of this difficulty is to use pathwise integration. Assume, under certain conditions,  $\lim_{n \rightarrow \infty} \int_0^T \sigma_t^n dB_t = \int_0^T \sigma_t dB_t$  in a.s. sense, with a common exceptional null set  $E_0$  independent of our choice of the approximation  $\sigma^n$ . Then we may fix a version:  $(\int_0^T \sigma_t dB_t)(\omega) := \lim_{n \rightarrow \infty} (\int_0^T \sigma_t^n dB_t)(\omega) \mathbf{1}_{E_0^c}(\omega)$ . If we are lucky that the observed path  $\omega$  is not in  $E_0$ , then we may use the limit of  $(\int_0^T \sigma_t^n dB_t)(\omega)$  as the value of  $(\int_0^T \sigma_t dB_t)(\omega)$ . Another powerful tool to study pathwise analysis is the rough path theory, which approximates  $B(\omega)$  by smooth paths. We have some discussion along this line in Problem 2.10.14.

## 2.9 Bibliographical Notes

The materials in this section are very standard in the literature. We refer to the classical reference Karatzas & Shreve [117] for a comprehensive presentation of properties of Brownian motions, some of which are more general or deeper than the results here. We also refer to Revuz & Yor [206] for a more general continuous martingale theory, and Protter [196] for a general semimartingale theory, including semimartingales with jumps.

For the financial application in Section 2.8, Shreve [209, 210] provides an excellent exposition. For the pathwise stochastic integration, we refer to Wong & Zakai [236, 237], Bichteler [17], Follmer [91], Willinger & Taqqu [235], Karandikar [119], and Nutz [160]. The rough path theory was initiated by Lyons [140]. We refer

interested readers to the book Friz & Hairer [94]. We also note that the pathwise stochastic integration is closely related to the quasi-sure stochastic integration in Section 12.1.1.

## 2.10 Exercises

**Problem 2.10.1** Prove Propositions 2.1.2 and 2.1.3. ■

**Problem 2.10.2**

- (i) Let  $X_t := \int_0^t b_s ds$  for some  $b \in \mathbb{L}_{loc}^1(\mathbb{F})$  and  $0 \leq s < t \leq T$ . Show that  $\bigvee_s^t(X) = \int_s^t |b_r| dr$ , a.s.
- (ii) Let  $X$  be as in Definition 2.1.5. Show that  $\langle X \rangle$  is increasing in  $t$ , a.s. That is,  $\langle X \rangle_t - \langle X \rangle_s \in \mathbb{S}^d$  is nonnegatively definite for all  $0 \leq s < t \leq T$ .
- (iii) Let  $x_i \in \mathbb{R}$ ,  $b^i \in \mathbb{L}^{1,2}(\mathbb{F}, \mathbb{R})$ ,  $\sigma^i \in \mathbb{L}^2(\mathbb{F}, \mathbb{R}^d)$ , and  $X_t^i := x_i + \int_0^t b_s^i ds + \int_0^t \sigma_s^i \cdot dB_s$ ,  $i = 1, 2$ . For any  $\pi : 0 = t_0 < \dots < t_n = T$ , denote

$$\langle X^1, X^2 \rangle_T^\pi := \sum_{i=0}^{n-1} X_{t_i, t_{i+1}}^1 X_{t_i, t_{i+1}}^2.$$

Show that  $\langle X^1, X^2 \rangle_T^\pi \rightarrow \int_0^T \sigma_t^1 \cdot \sigma_t^2 dt$  in  $\mathbb{L}^1(\mathcal{F}_T)$ , as  $|\pi| \rightarrow 0$ . ■

**Problem 2.10.3** This problem concerns the general Doob's maximum inequality, extending Lemma 2.2.4. Let  $X \in \mathbb{L}^1(\mathbb{F})$  be a right continuous nonnegative submartingale. Then

$$\mathbb{P}(X_T^* \geq \lambda) \leq \frac{1}{\lambda^p} \mathbb{E}[|X_T|^p \mathbf{1}_{\{X_T^* \geq \lambda\}}], \quad \text{for all } \lambda > 0, p \geq 1;$$

$$\mathbb{E}[|X_T^*|^p] \leq \left(\frac{p}{p-1}\right)^p \mathbb{E}[|X_T|^p], \quad \text{for all } p > 1; \quad \text{and} \quad \mathbb{E}[X_T^*] \leq \frac{e}{e-1} \mathbb{E}[1 + X_T(\ln(X_T))^+].$$

We remark that the  $|M|$  in Lemma 2.2.4 is a nonnegative submartingale, thanks to Jensen's inequality. Thus Lemma 2.2.4 is indeed a special case here. ■

**Problem 2.10.4** Prove the extended Itô formula Theorem 2.3.4. ■

**Problem 2.10.5** Let  $\eta_n \in \mathbb{L}^2(\mathcal{F}_0)$ ,  $\sigma^n \in \mathbb{L}^2(\mathbb{F}, \mathbb{R}^d)$ , and denote  $\xi_n := \eta_n + \int_0^T \sigma_t^n \cdot dB_t$ ,  $n \geq 1$ . Assume  $\lim_{n \rightarrow \infty} \mathbb{E}[|\xi_n - \xi|^2] = 0$  for some  $\xi \in \mathbb{L}^2(\mathcal{F}_T)$ . Then there exists unique  $\sigma \in \mathbb{L}^2(\mathbb{F}, \mathbb{R}^d)$  such that  $\xi = \mathbb{E}[\xi | \mathcal{F}_0] + \int_0^T \sigma_t \cdot dB_t$ , and  $\lim_{n \rightarrow \infty} \mathbb{E}\left[|\eta_n - \mathbb{E}[\xi | \mathcal{F}_0]|^2 + \int_0^T |\sigma_t^n - \sigma_t|^2 dt\right] = 0$ . ■

**Problem 2.10.6** Let  $\sigma \in \mathbb{L}^2(\mathbb{F}^B, \mathbb{S}^d)$  such that  $\sigma > 0$ , and  $X_t := \int_0^t \sigma_s dB_s$ . Show that the augmented filtrations of  $X$  and  $B$  are equal:  $\overline{\mathbb{F}^X}^{\mathbb{P}} = \overline{\mathbb{F}^B}^{\mathbb{P}}$ . ■

**Problem 2.10.7** Let  $p, q \in [1, \infty]$  be conjugates.

- (i) Assume  $X \in \mathbb{L}^{\infty,p}(\mathbb{F})$ ,  $Y \in \mathbb{L}^{2,q}(\mathbb{F})$  with appropriate dimensions so that  $XY$  takes values in  $\mathbb{R}^d$ . Show that  $M_t := \int_0^t (X_s Y_s) \cdot dB_s$  is a u.i.  $\mathbb{F}$ -martingale.
- (ii) Find a counterexample such that  $X \in \mathbb{L}^{2,p}(\mathbb{F})$ ,  $Y \in \mathbb{L}^{2,q}(\mathbb{F})$ , but  $M_t := \int_0^t (X_s Y_s) \cdot dB_s$  is not uniformly integrable.
- (iii) Find a counterexample such that  $M$  is a local martingale, but not a martingale.
- (iv) Find a counterexample such that  $M$  is a martingale, but not uniformly integrable.

Note that the  $M$  in (ii) is a local martingale, so it serves as a counterexample either for (iii) or for (iv). ■

**Problem 2.10.8** Let  $d = 1$  (for simplicity). Prove the following stochastic Fubini theorem:

$$\int_0^T \left[ u_t \int_0^t v_s ds \right] dB_t = \int_0^T \left[ v_s \int_s^T u_t dB_t \right] ds, \quad \forall u, v \in \mathbb{L}^\infty(\mathbb{F}, \mathbb{R}).$$

We remark that, unless  $u$  is deterministic, the following result is not true:

$$\int_0^T \left[ u_t \int_0^t v_s dB_s \right] dt = \int_0^T \left[ v_s \int_s^T u_t dt \right] dB_s.$$

In fact, the stochastic integrand in the right side above is in general not  $\mathbb{F}$ -adapted. ■

**Problem 2.10.9** This problem concerns general martingale theory. Let  $d = 1$ ,  $M$  a continuous  $\mathbb{F}$ -martingale with  $\mathbb{E}[|M_T|^2] < \infty$ ,  $K \in \mathbb{I}^1(\mathbb{F})$  with  $K_0 = 0$ , and  $M^2 - K$  is also a martingale.

- (i) For any bounded  $\sigma = \sum_{i=0}^{n-1} \sigma_{t_i} \mathbf{1}_{[t_i, t_{i+1})} \in \mathbb{L}^0(\mathbb{F})$ , denote  $\int_0^T \sigma_s dM_s := \sum_{i=0}^{n-1} \sigma_{t_i} M_{t_{i+1}} - M_{t_i}$ . Show that

$$\mathbb{E} \left[ \left| \int_0^T \sigma_s dM_s \right|^2 \right] = \mathbb{E} \left[ \int_0^T |\sigma_s|^2 dK_s \right].$$

- (ii) For any  $\sigma \in \mathbb{L}^0(\mathbb{F})$  such that  $\mathbb{E} \left[ \int_0^T |\sigma_s|^2 dK_s \right] < \infty$ , show that there exist bounded elementary processes  $\sigma^n \in \mathbb{L}^0(\mathbb{F})$  such that  $\lim_{n \rightarrow \infty} \mathbb{E} \left[ \int_0^T |\sigma_s^n - \sigma_s|^2 dK_s \right] = 0$ .
- (iii) For  $\sigma$  and  $\sigma^n$  as in (ii), show that  $\int_0^T \sigma^n dM_s$  converges in  $\mathbb{L}^2$ , and the limit is independent of the choices of  $\sigma^n$ . Thus we may define  $\int_0^T \sigma_s dM_s := \lim_{n \rightarrow \infty} \int_0^T \sigma_s^n dM_s$ .
- (iv) For  $\sigma$  as in (iii), define  $Y_t := \int_0^t \sigma_s dM_s$  similarly. Show that  $Y$  is still an  $\mathbb{F}$ -martingale.

We remark that the above process  $K$  is called the quadratic variation of  $M$ , and is also denoted as  $\langle M \rangle$ . Its existence can actually be proved. ■

**Problem 2.10.10** Prove Lemma 2.6.2. ■

**Problem 2.10.11** Assume  $X_n \rightarrow X$ ,  $Y_n \rightarrow Y$  weakly in  $\mathbb{L}^2(\mathbb{F})$  and have appropriate dimensions.

- (i)  $X_n + Y_n \rightarrow X + Y$  weakly in  $\mathbb{L}^2(\mathbb{F})$ .
- (ii)  $\int_0^\cdot X_s^n \cdot dB_s \rightarrow \int_0^\cdot X_s \cdot dB_s$  weakly in  $\mathbb{L}^2(\mathbb{F})$ .
- (iii)  $\mathbb{E}[\int_0^T |X_t|^2 dt] \leq \liminf_{n \rightarrow \infty} \mathbb{E}[\int_0^T |X_t^n|^2 dt]$ . ■

**Problem 2.10.12** We note that Theorem 2.7.1 does not hold true for semimartingales in the following sense. Let  $d = 1$ . For any  $n$ , find a counterexample  $X_t = \int_0^t b_s ds + \int_0^t \sigma_s dB_s$ , where  $b \in \mathbb{L}^{1,2}(\mathbb{F})$  and  $\sigma \in \mathbb{L}^2(\mathbb{F})$  such that

$$\mathbb{E}\left[\left(\int_0^T |b_t| dt\right)^2 + \int_0^T |\sigma_t|^2 dt\right] > n\mathbb{E}[|X_T^*|^2].$$

■

**Problem 2.10.13** This problem concerns the Stratonovich integral  $\int_0^T X_t \circ dB_t$ , for which the integrand  $X$  requires some regularity. To be specific, let  $X_t := x + \int_0^t b_s ds + \int_0^t \sigma_s dB_s$ , where  $x \in \mathbb{R}^d$ ,  $b \in \mathbb{L}^{1,2}(\mathbb{F}, \mathbb{R}^d)$ , and  $\sigma \in \mathbb{L}^2(\mathbb{F}, \mathbb{R}^{d \times d})$ .

- (i) For any  $\pi : 0 = t_0 < \dots < t_n = T$ , denote

$$S(\pi) := \sum_{i=0}^{n-1} X_{\frac{t_i+t_{i+1}}{2}} \cdot B_{t_i, t_{i+1}}.$$

Show that  $S_M(\pi) \rightarrow \int_0^T X_t \cdot dB_t + \frac{1}{2} \int_0^T \text{tr}(\sigma_t) dt$  in  $\mathbb{L}^2(\mathcal{F}_T)$ , as  $|\pi| \rightarrow 0$ . We thus define the Stratonovich integral as

$$\int_0^T X_t \circ dB_t := \lim_{|\pi| \rightarrow 0} S_M(\pi) = \int_0^T X_t \cdot dB_t + \frac{1}{2} \int_0^T \text{tr}(\sigma_t) dt. \quad (2.10.1)$$

- (ii) The Stratonovich integral can be approximated in a different way. For each  $\pi$ , let  $B^\pi$  denote the linear interpolation of  $(t_i, B_{t_i})_{0 \leq i \leq n}$ , namely

$$B_t^\pi := \sum_{i=0}^{n-1} \left[ B_{t_i} \frac{t_{i+1} - t}{t_{i+1} - t_i} + B_{t_{i+1}} \frac{t - t_i}{t_{i+1} - t_i} \right] \mathbf{1}_{(t_i, t_{i+1}]}. \quad (2.10.2)$$

Then  $B^\pi$  is absolutely continuous in  $t$  and thus the following integration is well defined:

$$\tilde{S}(\pi) := \int_0^T X_t \cdot dB_t^\pi.$$

Show that  $\lim_{|\pi| \rightarrow 0} \tilde{S}(\pi) = \int_0^T X_t \circ dB_t$  in  $\mathbb{L}^2$ -sense.

- (iii) Similarly we may define  $Y_t := \int_0^t X_s \circ dB_s = \int_0^t X_s \cdot dB_s + \frac{1}{2} \int_0^t \text{tr}(\sigma_s) ds$ . We shall note that  $Y$  is in general not a martingale. Prove the following chain rule for Stratonovich integral:

$$df(t, Y_t) = \partial_t f(t, Y_t) dt + [\partial_x f(t, Y_t) X_t] \circ dB_t,$$

for any  $f : [0, T] \times \mathbb{R} \rightarrow \mathbb{R}$  smooth enough. ■

**Problem 2.10.14** This problem concerns a.s. convergence of stochastic integration. Given  $X \in \mathbb{L}^2(\mathbb{F}, \mathbb{R}^d)$ , denote  $Y_T := \int_0^T X_t \cdot dB_t$ , and, for a partition  $\pi : 0 = t_0 < \dots < t_n = T$ ,

$$Y_T^\pi := \sum_{i=0}^{n-1} X_{t_i} \cdot B_{t_i, t_{i+1}}. \quad (2.10.2)$$

Let  $\alpha \in (0, 1]$  and  $\beta > 0$  be two constants, and  $\{\pi_m\}_{m \geq 1}$  a sequence of partitions such that  $|\pi_m| \leq m^{-\beta}$ . At below, all limits are in the sense of a.s. convergence.

- (i) Assume  $X$  is uniformly Hölder- $\alpha$  continuous and  $\beta > \frac{1}{2\alpha}$ . Show that  $\lim_{m \rightarrow \infty} Y_T^{\pi_m} = Y_T$ , a.s. (Hint: show that  $\mathbb{E} \left[ \sum_{m=1}^{\infty} |Y_T^{\pi_m} - Y_T|^2 \right] < \infty$ .)
- (ii) Assume  $dX_t = \sigma_t dB_t$ ,  $\sigma \in \mathbb{L}^2(\mathbb{F}, \mathbb{R}^{d \times d})$ , and  $\beta > 1$ . Show that  $\lim_{m \rightarrow \infty} Y_T^{\pi_m} = Y_T$ , a.s.
- (iii) Assume  $dX_t = \sigma_t dB_t$ ,  $\sigma \in \mathbb{L}^\infty(\mathbb{F}, \mathbb{R}^{d \times d})$ , and  $\beta > \frac{1}{2}$ . Show that  $\lim_{m \rightarrow \infty} Y_T^{\pi_m} = Y_T$ , a.s.
- (iv) Assume  $d = 1$ ,  $X$  is as in (ii),  $\sigma$  is uniformly Hölder- $\alpha$  continuous, and  $\beta > \frac{1}{1+2\alpha}$ . Denote

$$Y_T^{2,\pi} := \sum_{i=0}^{n-1} \left[ X_{t_i} B_{t_i, t_{i+1}} + \sigma_{t_i} \frac{|B_{t_i, t_{i+1}}|^2 - (t_{i+1} - t_i)}{2} \right], \quad (2.10.3)$$

which we call the second order approximation. Show that  $\lim_{m \rightarrow \infty} Y_T^{2,\pi_m} = Y_T$ , a.s.

- (v) Consider the same setting as in (iv). Assume further that  $d\sigma_t = \theta_t dB_t$ ,  $\theta$  is uniformly Hölder- $\alpha$  continuous, and  $\beta > \frac{1}{1+3\alpha}$ . Denote

$$Y_T^{3,\pi} := \sum_{i=0}^{n-1} \left[ X_{t_i} B_{t_i, t_{i+1}} + \sigma_{t_i} \frac{|B_{t_i, t_{i+1}}|^2 - (t_{i+1} - t_i)}{2} + \theta_{t_i} \frac{(B_{t_i, t_{i+1}})^3 - 3B_{t_i, t_{i+1}}(t_{i+1} - t_i)}{6} \right], \quad (2.10.4)$$

which we call the third order approximation. Show that  $\lim_{m \rightarrow \infty} Y_T^{3,\pi_m} = Y_T$ , a.s.



We remark that in all the above cases,  $|\pi_m|$  converges to 0 with a rate  $\beta$ , and the exceptional null set of the a.s. convergence depends on  $\{\pi_m\}_{m \geq 1}$ . In the setting of (iv), by rough path theory one can show that there is a common null set  $E_0$  such that  $\lim_{|\pi| \rightarrow 0} Y_T^{2,\pi}(\omega)$  exists for all  $\omega \notin E_0$ . ■

Backward Stochastic Differential Equations  
From Linear to Fully Nonlinear Theory

Zhang, J.

2017, XVI, 388 p., Hardcover

ISBN: 978-1-4939-7254-8