

# Stochastic Differential Equations and Applications

Guohuan Zhao

January 1, 2025

---

# Contents

<b>1</b>	<b>Brownian motion and Martingale</b>	<b>7</b>
1.1	Probabilistic terminology . . . . .	7
1.2	Brownian motion . . . . .	8
1.3	Martingale . . . . .	10
1.3.1	Discrete time . . . . .	10
1.3.2	Continuous time . . . . .	12
1.3.3	Some Applications of Martingale Theory . . . . .	17
1.4	Stochastic Integral . . . . .	21
1.5	Itô's formula and its applications . . . . .	24
1.5.1	Applications in martingale theory . . . . .	24
1.5.2	Applications in PDEs . . . . .	28
<b>2</b>	<b>Itô processes</b>	<b>31</b>
2.1	Support theorem . . . . .	31
2.2	ABP estimate and Generalized Itô's formula . . . . .	34
<b>3</b>	<b>Itô's stochastic differential equations</b>	<b>41</b>
3.1	Strong solutions . . . . .	41
3.1.1	Lipschitz conditions . . . . .	41
3.1.2	Definitions of solutions . . . . .	42
3.1.3	SDEs with Hölder drifts . . . . .	43
3.1.4	Stochastic Flow . . . . .	44
3.2	Weak solutions . . . . .	48
3.2.1	Uniqueness in law . . . . .	48
3.2.2	Markov properties . . . . .	49
<b>4</b>	<b>Applications to Elliptic PDEs</b>	<b>53</b>
4.1	Dirichlet Problems and Harmonic functions . . . . .	53
4.2	Once again on the hitting probability . . . . .	55
4.3	Harnack Inequality and Hölder estimate . . . . .	58
<b>5</b>	<b>Malliavin's proof of Hörmander's Theorem</b>	<b>63</b>
5.1	Integration by parts . . . . .	63
5.2	Malliavin Matrix . . . . .	69
5.3	Hörmander's Theorem . . . . .	71

<b>A</b>	<b>Useful facts</b>	<b>77</b>
A.1	Sobolev spaces . . . . .	77
<b>B</b>	<b>Harmonic analysis from the perspective of Probability</b>	<b>79</b>
B.1	Interpolation Theorems . . . . .	80
<b>C</b>	<b>Some basic results of PDEs</b>	<b>81</b>
C.1	Monge-Ampère Equation . . . . .	81
C.2	Schauder estimate . . . . .	84

**Notation:**

- We use  $:=$  as a way of definition.
- $\mathbb{N} = \{0, 1, 2, \dots\}$ ,  $\mathbb{R}_+ = [0, \infty)$ .
- $\mathcal{L}^d$  is the Lebesgue measure on  $\mathbb{R}^d$ .
- $B_r(x) = \{y \in \mathbb{R}^d : |y - x| < r\}$ ,  $B_r = B_r(0)$ .
- $Q_r(x) = \{y \in \mathbb{R}^d : |y_i - x_i| < r/2, i = 1, 2, \dots, d\}$ ,  $Q_r = Q_r(0)$ .
- $\mathbb{S}_+^d$  is the collection of all symmetric non-negative matrix in  $\mathbb{R}^{d \times d}$ . For any  $\delta \in (0, 1]$ ,  $\mathbb{S}_\delta^d := \{a \in \mathbb{S}_+^d : a_{ij} = a_{ji}, \delta I_d \leq a \leq \delta^{-1} I_d\}$ .
- The letter  $c$  or  $C$  with or without subscripts stands for an unimportant constant, whose value may change in different places. We use  $a \asymp b$  to denote that  $a$  and  $b$  are comparable up to a constant, and use  $a \lesssim b$  ( $a \gtrsim b$ ) to denote  $a \leq Cb$  ( $a \geq Cb$ ) for some constant  $C$ .
- $W_t$  is a Brownian motion starting from 0 and  $W_t^x$  a Brownian motion starting from  $x$ .
- $\tau_D$  is the first exit time of a process from domain  $D$ ;  $\sigma_\Gamma$  is the first hitting time of  $\Gamma$ .



# Chapter 1

## Brownian motion and Martingale

### 1.1 Probabilistic terminology

Let  $(\Omega, \mathcal{F}, \mathbf{P})$  be a probability space and  $(E, \mathcal{E})$  be a measurable space.  $X : (\Omega, \mathcal{F}) \rightarrow (E, \mathcal{E})$  a measurable map, and  $\mathcal{G}$  a  $\sigma$ -field  $\subseteq \mathcal{F}$ .

When  $E = \mathbb{R}$ , we define the **conditional expectation** of  $X$  given  $\mathcal{G}$ ,  $\mathbf{E}(X|\mathcal{G})$ , to be any random variable  $Y$  that satisfies

- (a)  $Y \in \mathcal{G}$ ;
- (b) for all  $A \in \mathcal{G}$ ,  $\mathbf{E}(X; A) = \mathbf{E}(Y; A)$ .

$Q_{\mathcal{G}} : \Omega \times \mathcal{E} \rightarrow [0, 1]$  is said to be a **regular conditional distribution** (RCD) for  $X$  given  $\mathcal{G}$  if

- (a) For each  $A \in \mathcal{E}$ ,  $\omega \mapsto Q_{\mathcal{G}}(\omega, A)$  is a version of  $\mathbf{E}(\mathbf{1}_A(X)|\mathcal{G})$ ;
- (b) For a.e.  $\omega \in \Omega$ ,  $A \mapsto Q_{\mathcal{G}}(\omega, A)$  is a probability measure.

If  $E = \Omega$ ,  $X(\omega) = \omega$ , then  $Q_{\mathcal{G}}$  is called a **regular conditional probability**.

The following results can be found in Durrett's book [Dur19].

**Proposition 1.1.1.** (i) If  $\mathcal{G}_1 \subseteq \mathcal{G}_2 \subseteq \mathcal{F}$ , then

$$\mathbf{E}[(X|\mathcal{G}_2)|\mathcal{G}_1] = \mathbf{E}(X|\mathcal{G}_1) \quad (1.1)$$

(ii) Assume that  $X \in \mathcal{F}$  and  $Y \in \mathcal{G} \subseteq \mathcal{F}$ , then

$$\mathbf{E}(XY|\mathcal{G}) = \mathbf{E}(X|\mathcal{G})Y. \quad (1.2)$$

(iii) (Jesen's inequality) If  $\varphi$  is a **convex** function, then

$$\mathbf{E}(\varphi(X)|\mathcal{G}) \leq \varphi(\mathbf{E}(X|\mathcal{G})). \quad (1.3)$$

**Proposition 1.1.2.** Let  $Q_{\mathcal{G}}$  be a RCD for  $X$  given  $\mathcal{G}$ . If  $f : E \rightarrow \mathbb{R}$  satisfying  $\mathbf{E}|f(X)| < \infty$ , then

$$\mathbf{E}(f(X)|\mathcal{G})(\omega) = \int_E f(x)Q_{\mathcal{G}}(\omega, dx) \quad \text{a.s..}$$

A measurable space  $(E, \mathcal{E})$  is said to be **standard** if  $(E, \mathcal{E})$  is isomorphic (as a measurable space) to  $(\mathbb{R}; \mathcal{B}(\mathbb{R}))$ .

**Theorem 1.1.3.** *RCD exists if  $(E, \mathcal{B}(E))$  is a standard measure space.*

**Proposition 1.1.4.** *Assume  $X \geq 0$ ,  $f : \mathbb{R}_+ \rightarrow \mathbb{R}_+$  such that  $f \in C^1(\mathbb{R}_+)$  and  $f(0) = 0$ . Then*

$$\mathbf{E}f(X) = \int_0^\infty f'(t)\mathbf{P}(X > t)dt. \quad (1.4)$$

**Exercise 1.1.1.** *If  $X \geq 0$ ,  $f : \mathbb{R}_+ \rightarrow \mathbb{R}_+$  such that  $f \in C^1(\mathbb{R}_+)$  and  $f(\infty) = 0$ . Then*

$$\mathbf{E}f(X) = - \int_0^\infty f'(t)\mathbf{P}(X \leq t)dt. \quad (1.5)$$

## 1.2 Brownian motion

A stochastic process defined on  $(\Omega, \mathcal{F}, \mathbf{P})$  taking value in  $E$  can be understood in various ways. It involves a collection of random variables  $X_t \in E$  indexed by a parameter set  $\mathbf{T}$  (usually,  $\mathbf{T} = \mathbb{N}$  or  $\mathbb{R}_+$ ), where  $X_t$  is a measurable map from  $(\Omega, \mathcal{F}, \mathbf{P})$  to  $(E, \mathcal{B}(E))$  for each  $t \in \mathbf{T}$ . The parameter set  $\mathbf{T}$  typically represents time and can be discrete or continuous. The process can also be regarded as a measurable map from  $(\Omega, \mathcal{F}, \mathbf{P})$  to the space of functions  $E^{\mathbf{T}}$ . The Kolmogorov  $\sigma$ -field on  $E^{\mathbf{T}}$  is the smallest  $\sigma$ -field making the projections  $\pi_t : E^{\mathbf{T}} \ni f \mapsto f(t) \in E$  measurable. This definition ensures that a random map  $\Omega \ni \omega \mapsto X(\omega) \in E^{\mathbf{T}}$  is measurable if its component random variables  $X_t : \Omega \rightarrow E$  are measurable for all  $t \in \mathbf{T}$ . Therefore, the mapping  $\omega \mapsto X(\omega)$  induces a measure on  $(E^{\mathbf{T}}, \mathcal{B}(E^{\mathbf{T}}))$  denoted by  $\mathbb{P}$ . The underlying probability model  $(\Omega, \mathcal{F}, \mathbf{P})$  is replaceable by the canonical model  $(\mathbb{P}, E^{\mathbf{T}}, \mathcal{B}(E^{\mathbf{T}}))$  with a specific choice of  $X_t(f) = \pi_t(f) = f(t)$ . In simpler terms, a stochastic process is just a probability measure  $\mathbb{P}$  on  $(E^{\mathbf{T}}, \mathcal{B}(E^{\mathbf{T}}))$ .

Another point of view is that the only relevant objects are the joint distributions of  $(X_{t_1}, X_{t_2}, \dots, X_{t_n})$  for every  $n$  and every finite subset  $I = (t_1, t_2, \dots, t_n)$  of  $\mathbf{T}$ . These can be specified as probability measures  $\mu_I$  on  $\mathbb{R}^n$ . These  $\mu_I$  cannot be totally arbitrary. If we allow different permutations of the same set, so that  $I$  and  $I'$  are permutations of each other then  $\mu_I$  and  $\mu_{I'}$  should be related by the same permutation. If  $I \subseteq I'$ , then we can obtain the joint distribution of  $(X_t)_{t \in I}$  by projecting the joint distribution of  $(X_t)_{t \in I'}$  from  $\mathbb{R}^{n'}$  to  $\mathbb{R}^n$  where  $n$  and  $n'$  are the cardinalities of  $I$  and  $I'$  respectively. A stochastic process can then be viewed as a family  $(\mu_I)$  of distributions on various finite dimensional spaces that satisfy the consistency conditions. A theorem of Kolmogorov says that this is not all that different. Any such consistent family arises from a  $\mathbf{P}$  on  $(E^{\mathbf{T}}, \mathcal{B}(E^{\mathbf{T}}))$  which is uniquely determined by the family  $(\mu_I)$ .

**Theorem 1.2.1** (Kolmogorov's consistency Theorem, cf. [Yan21]). *Let  $E$  be a standard measure space. Assume that we are given for every  $t_1, \dots, t_n \in \mathbf{T}$  a probability measure  $\mu_{t_1 \dots t_n}$  on  $E^n$ , and that these probability measures satisfy:*

(a). *for each  $\tau \in S_n$  and  $A_i \in \mathcal{B}(E)$ ,*

$$\mu_{t_1 \dots t_n}(A_1 \times \dots \times A_n) = \mu_{\tau(t_1) \dots \tau(t_n)}(A_{\tau(1)} \times \dots \times A_{\tau(n)});$$



(b). for each  $A_i \in \mathcal{B}(E)$ ,

$$\mu_{t_1 \dots t_n}(A_1 \times \dots \times A_{n-1} \times E) = \mu_{t_1 \dots t_{n-1}}(A_1 \times \dots \times A_{n-1}).$$

Then, there is a unique probability measure  $\mathbb{P}$  on  $(E^{\mathbf{T}}, \mathcal{B}(E^{\mathbf{T}}))$  such that for  $t_1, \dots, t_n \in \mathbf{T}$ ,  $A_1, \dots, A_n \in \mathcal{B}(E)$ :  $\mathbb{P}(f(t_1) \in A_1, \dots, f(t_n) \in A_n) = \mu_{t_1, \dots, t_n}(A_1 \times \dots \times A_n)$ .

**Definition 1.2.2.** A stochastic process  $(W_t)_{t \in \mathbb{R}_+}$  is a one-dimensional **Brownian motion** started at 0 if

- (a)  $W_0 = 0$ , a.s.;
- (b) for all  $s \leq t$ ,  $W_t - W_s$  is a mean zero Gaussian random variable with variance  $t - s$ ;
- (c) for all  $s < t$ ,  $W_t - W_s$  is independent of  $\sigma(W_r; r \leq s)$ ;
- (d) with probability 1 the map  $t \mapsto W_t(\omega)$  is continuous.

By Theorem 1.2.1, we can define a probability measure  $\mathbb{Q}$  on  $\mathbb{R}^{\mathbb{R}_+}(E = \mathbb{R}, \mathbf{T} = \mathbb{R}_+)$  such that the canonical process  $X_t(f) = f(t)$  satisfies (a), (b) and (c). However, whether the measure is concentrated on the space of continuous functions is not a simple question. In fact, since  $\mathbf{T} = \mathbb{R}_+$  is uncountable the space of bounded functions, continuous functions, etc., are **not** measurable sets of  $\mathbb{R}^{\mathbb{R}_+}$ . They do not belong to the natural  $\sigma$ -field. Essentially, in probability theory, the rules involve only a countable collection of sets at one time, and any information that involves the values of an uncountable number of measurable functions is beyond reach. There is an intrinsic reason for this. In probability theory, we can always change the values of a random variable on a set of measure 0, and we have not changed anything significant. Since we are allowed to mess up each function on a set of measure 0, we have to assume that each function has indeed been messed up on a set of measure 0. If we are dealing with a countable number of functions, the 'mess up' has occurred only on the countable union of these individual sets of measure 0, which, by the properties of a measure, is again a set of measure 0. On the other hand, if we are dealing with an uncountable set of functions, then these sets of measure 0 can possibly gang up on us.

Of course it would be foolish of us to mess things up unnecessarily. If we can clean things up and choose a nice version of our random variables we should do so. But we cannot really do this sensibly unless we decide first what nice means. We however face the risk of being too greedy and it may not be possible to have a version as nice as we seek. But then we can always change our mind.

**Lemma 1.2.3** (Fractional Sobolev inequality). *Let  $D$  be an open set in  $\mathbb{R}^n$ ,  $p \geq 1$  and  $s \in (n/p, 1)$ . Let  $f : D \rightarrow \mathbb{R}^d$  be a measurable function. Assume*

$$\iint_{D \times D} \frac{|f(x) - f(y)|^p}{|x - y|^{n+sp}} dx dy < \infty.$$

*Then there exists a version of  $f$ , say  $\tilde{f}$ , such that*

$$\sup_{x, y \in D} \frac{|\tilde{f}(x) - \tilde{f}(y)|^p}{|x - y|^{sp-n}} \leq C \left( \iint_{D \times D} \frac{|f(x) - f(y)|_{\mathcal{B}}^p}{|x - y|^{n+sp}} dx dy \right)^{1/p}, \quad (1.6)$$

*Here  $C$  only depends on  $n, s, p$  and  $D$ .*

**Theorem 1.2.4.** *Let  $I = [0, T]$ , and let  $p > 1$  and  $\beta \in (1/p, 1)$ . Assume  $(X_t)_{t \in I}$  satisfies*

$$\mathbf{E}|X_s - X_t|^p \leq c|t - s|^{1+\beta p}, \quad \forall t, s \in I. \quad (1.7)$$

*Then there exists a version of  $X, Y$  (for each  $t \in I$ ,  $\mathbf{P}(X_t = Y_t) = 1$ ) such that*

$$\mathbf{P} \left( \sup_{t \in I} \frac{|Y_t - Y_s|}{|t - s|^\alpha} \leq K \right) = 1,$$

*where  $\alpha \in (0, \beta - 1/p)$ ,  $K = K(\alpha, \beta, p, c, I, \omega)$  and  $\mathbf{E}K^p < \infty$ .*

*Proof.* Regard  $X$  as a measurable function from  $\Omega \times I$  to  $\mathbb{R}^d$ . By Lemma 1.2.3, there is a null set  $\mathcal{N} \subseteq \Omega$  and a measurable function  $Y : \Omega \times I \rightarrow \mathcal{B}$ , such that for each  $\omega \notin \mathcal{N}$ ,

$$\mathcal{L}^1(\{t \in I : Y_t(\omega) \neq X_t(\omega)\}) = 0,$$

and  $Y(\omega)$  is a continuous function. Moreover,

$$\|Y(\omega)\|_{C^\alpha(I)} \lesssim K(\omega) := \left( \iint_{I \times I} \frac{|X_t(\omega) - X_s(\omega)|^p}{|t - s|^{2+\alpha p}} \, ds dt \right)^{1/p} \in L^p(\mathbf{P}).$$

By Fubini theorem, there exists a  $\mathcal{L}$ -null set  $N \subseteq I$ , such that for each  $t \notin N$ ,  $\mathbf{P}(X_t \neq Y_t) = 0$ . For any  $t_0 \in N$ , by (1.7), one can see that  $X_{t_n} \xrightarrow[I \setminus N \ni t_n \rightarrow t_0]{\mathbf{P}} X_{t_0}$ . On the other hand,  $X_{t_n} \xrightarrow{a.s.} Y_{t_n} \rightarrow Y_{t_0}$ , so we have  $X_{t_0} \xrightarrow{a.s.} Y_{t_0}$ . Therefore,  $Y$  is a version of  $X$ .  $\square$

Thanks to Theorem 1.2.4 and the discussion after Definition 1.2.2, we get the existence of Brownian motion.

## 1.3 Martingale

### 1.3.1 Discrete time

Let  $(\Omega, \mathcal{F}, \mathbf{P})$  be a standard probability space. Let  $\mathcal{F}_n$  ( $n \in \mathbb{N}$ ) be an increasing sequence of  $\sigma$ -fields. A sequence of random variables  $X_n$  is **adapted** to  $\mathcal{F}_n$  if for each  $n$ ,  $X_n$  is  $\mathcal{F}_n$  measurable. Similarly a collection of random variables  $X_t$  ( $t \in \mathbb{R}_+$ ) is adapted to  $\mathcal{F}_t$  if each  $X_t$  is  $\mathcal{F}_t$  measurable. We say the filtration  $\mathcal{F}_t$  satisfies the usual conditions if  $\mathcal{F}_t$  is **right continuous** (i.e.,  $\mathcal{F}_t = \mathcal{F}_{t+}$  for all  $t$ , where  $\mathcal{F}_{t+} = \bigcap_{\varepsilon > 0} \mathcal{F}_{t+\varepsilon}$ ) and each  $\mathcal{F}_t$  is **complete** (i.e.,  $\mathcal{F}_t$  contains all  $\mathbf{P}$ -null sets). For the most part, in this and the following sections there will only be one probability measure and we assume the filtration  $\mathcal{F}_t$  satisfies the usual conditions.

We say  $\tau : \Omega \rightarrow \mathbb{N} \cup \{\infty\}$  is a **stopping time** if  $\tau$  satisfying  $\{\tau \leq n\} \in \mathcal{F}_n$  ( $\{\tau \leq t\} \in \mathcal{F}_t$ ), for each  $n \in \mathbb{N}$  ( $t \in \mathbb{R}_+$ ).

$\mathcal{F}_\tau$  is a  $\sigma$ -field containing all measurable sets  $A \in \mathcal{F}$  such that  $A \cap \{\tau \leq n\} \in \mathcal{F}_n$  ( $A \cap \{\tau \leq t\} \in \mathcal{F}_t$ ) for all  $n \in \mathbb{N}$  ( $t \in \mathbb{R}_+$ ).

Below we recall the results about discrete time martingales and submartingales that will be used. The proof of the subsequent statements can be found in Durrett's book [Dur19], and in many other books dealing with discrete time martingales.

**Definition 1.3.1.** Let  $X_n$  be a real-valued  $\mathcal{F}_n$ -adapted processes. If for each  $n$  and  $m < n$ ,  $X_n$  is integrable and  $\mathbf{E}(X_n|\mathcal{F}_m) \geq (\leq) X_m$  a.s., then we call  $X_n$  is a **submartingale** (**supermartingale**). We say  $X_n$  is a **martingale** if it is both a submartingale and a supermartingale.

Suppose  $M_n$  is a martingale, and  $\varphi$  is convex function. Assume that  $\mathbf{E}|\varphi(M_n)| < \infty$ , in the light of Jensen's inequality,  $\varphi(M_n)$  is a submartingale.

**Example 1.** Let  $\xi_1, \xi_2, \dots$  be a sequence of i.i.d random variable. Set  $X_n := \sum_{i=0}^n \xi_i$  and  $\mathcal{F}_n := \sigma(\xi_0, \dots, \xi_n)$ .

**Theorem 1.3.2** (Doob). If  $X_n \in \mathcal{F}_n$  is a submartingale then it can be uniquely decomposed as  $X_n = M_n + A_n$ , where  $M_n \in \mathcal{F}_n$  is martingale,  $A_0 = 0$ ,  $A_{n+1} \geq A_n$  almost surely and  $A_n$  is  $\mathcal{F}_{n-1}$ -measurable.

*Proof.* Let

$$A_0 = 0, \quad A_n := \sum_{k=1}^n \mathbf{E}(X_k - X_{k-1} | \mathcal{F}_{k-1})$$

and

$$M_n = X_n - A_n = X_0 + \sum_{k=1}^n [(X_k - X_{k-1}) - \mathbf{E}(X_k - X_{k-1} | \mathcal{F}_{k-1})].$$

Then  $M_n$  and  $A_n$  are the desired processes. The proof for the uniqueness result is easy.  $\square$

The following theorem lies at the basis of all other results for martingales.

**Theorem 1.3.3** (Doob's Optional stopping theorem). Assume that  $\sigma$  and  $\tau$  are two bounded stopping time, and  $X_t$  is a submartingale, then  $\mathbf{E}(X_\tau | \mathcal{F}_\sigma) \geq X_{\sigma \wedge \tau}$ .

**Lemma 1.3.4.** Let  $X_n$  be a submartingale, and  $\tau$  be a bounded stopping time and  $\tau \leq K$  (constant). Then

$$(i) \quad \mathbf{E}(X_K | \mathcal{F}_\tau) \geq X_\tau;$$

$$(ii) \quad X_{\tau \wedge n} \text{ is a } \mathcal{F}_n\text{-submartingale.}$$

*Proof.* (i). for each  $A \in \mathcal{F}_\tau$ , we will show that  $\mathbf{E}(X_K; A) \geq \mathbf{E}(X_\tau; A)$ . In fact,

$$\mathbf{E}(X_\tau; A) = \sum_{k=0}^K \mathbf{E}(X_k; \underbrace{A \cap \{\tau = k\}}_{\in \mathcal{F}_k}) \leq \sum_{k=0}^K \mathbf{E}(X_K; A \cap \{\tau = k\}) = \mathbf{E}(X_K; A).$$

(ii). For each  $A \in \mathcal{F}_{n-1}$ ,

$$\begin{aligned} \mathbf{E}(X_{\tau \wedge n}; A) &= \mathbf{E}(X_{\tau \wedge n}; A \cap \{\tau \leq n-1\}) + \mathbf{E}(X_{\tau \wedge n}; A \cap \{\tau > n-1\}) \\ &= \mathbf{E}(X_\tau; A \cap \{\tau \leq n-1\}) + \mathbf{E}(X_n; \underbrace{A \cap \{\tau > n-1\}}_{\in \mathcal{F}_{n-1}}) \\ &\geq \mathbf{E}(X_\tau; A \cap \{\tau \leq n-1\}) + \mathbf{E}(X_{n-1}; A \cap \{\tau > n-1\}) \\ &= \mathbf{E}(X_{\tau \wedge (n-1)}; A). \end{aligned}$$

$\square$

*Proof of Theorem 1.3.3.* By the above lemma, we have  $\mathbf{E}(X_\tau|\mathcal{F}_\sigma) = \mathbf{E}(X_{K \wedge \tau}|\mathcal{F}_\sigma) \geq X_{\sigma \wedge \tau}$ .  $\square$

**Theorem 1.3.5** (Doob's inequality). *Let  $M_n$  be a martingale. If  $M_n^* := \sup_{k \leq n} |M_k|$ , then*

$$\mathbf{P}(M_n^* > \lambda) \leq \lambda^{-1} \mathbf{E}(|M_n|; M_n^* > \lambda).$$

*Proof.* Let  $\tau = \inf\{k : |M_k| > \lambda\}$ . Noting that  $\{M_n^* > \lambda\} = \{\tau \leq n\}$ , we have

$$\begin{aligned} \lambda \mathbf{P}(M_n^* > \lambda) &= \lambda \mathbf{P}(\tau \leq n) \leq \mathbf{E}(|M_\tau|; \tau \leq n) \\ &\leq \mathbf{E}(|M_{\tau \wedge n}|; \tau \leq n) \leq \mathbf{E}(|M_n|; M_n^* > \lambda). \end{aligned}$$

$\square$

**Corollary 1.3.6.** *Let  $M_n$  be a martingale and  $\tau$  be a stopping time. For each  $p > 1$ ,  $\mathbf{E}|M_\tau^*|^p \leq C_p \mathbf{E}|M_{\tau \wedge n}|^p$ .*

Let  $a \leq b$ . Set  $\sigma_1 = \inf\{n \geq 0 : X_n \leq a\}$ ,  $\tau_1 = \inf\{n > \sigma_1 : X_n \geq b\}$ ,  $\sigma_2 = \inf\{n > \tau_1 : M_n \leq a\}$ ,  $\tau_2 = \inf\{n > \sigma_2 : X_n \geq b\}$ ,  $\dots$ , and  $U_N := \max\{k : \tau_k \leq N\}$ .

**Lemma 1.3.7** (Upcrossing inequality). *Suppose that  $X_N$  is a submartingale, then*

$$(b - a) \mathbf{E}U_N(a, b) \leq \mathbf{E}(X_N - a)^+.$$

*Proof.* We only prove the case that  $a = 0$  and  $X_k \geq 0$ .

$$X_N = \underbrace{X_{S_1 \wedge N}}_{\geq 0} + \underbrace{\sum_{i=1}^{\infty} X_{T_i \wedge N} - X_{S_i \wedge N}}_{\geq b U_N(0, b)} + \sum_{i=1}^{\infty} \underbrace{X_{S_{i+1} \wedge N} - X_{T_i \wedge N}}_{\text{positive expectation}}.$$

$\square$

Upcrossing inequality leads to

**Theorem 1.3.8.** *If  $X_n$  is a submartingale such that  $\sup_n \mathbf{E}X_n^+ < \infty$ , then  $X_n$  converges a.s. as  $n \rightarrow \infty$ .*

**Corollary 1.3.9.** *Suppose that  $X \in L^1(\mathbf{P}, \Omega)$ ,  $\mathcal{F}_n \uparrow \mathcal{F}_\infty$ , then*

$$\lim_{n \rightarrow \infty} \mathbf{E}(X|\mathcal{F}_n) = \mathbf{E}(X|\mathcal{F}_\infty), \quad \text{a.s. and in } L^1.$$

### 1.3.2 Continuous time

*All of the above results also hold for all right continuous martingale (submartingales) (see [Hua01]), except the Doob decomposition theorem.*

---

**Exercise 1.3.1.** Using the continuous-time version of Corollary 1.3.6 to show the following result:

Suppose that  $\{M_t^n\}_{t \geq 0}$  is a sequence of continuous square integrable martingales, and for each  $T > 0$ ,  $M_T^n - M_T^m \rightarrow 0$  in  $L^2$  as  $n, m \rightarrow \infty$ . Then there exists a continuous square integrable martingale  $M_t$  such that for each  $T > 0$ ,

$$(M^n - M)_T^* \xrightarrow{L^2} 0 \quad \text{as } n \rightarrow \infty.$$

The celebrated **Doob-Meyer decomposition** says that, under mild hypotheses, a submartingale can be decomposed into a martingale plus an increasing process. We are limiting ourselves to continuous submartingales, and the theorem we want to state is the following.

**Theorem 1.3.10.** Suppose that  $X_t$  is a continuous submartingale, then there exists a unique martingale  $M$  and a unique continuous adapted process  $A$  with finite bounded variation on each interval such that

$$A_0 = 0, \quad X_t = M_t + A_t.$$

We say that  $M_t$  is a **local martingale** if there exist stopping times  $\tau_n \uparrow \infty$  such that  $X_{\tau_n \wedge t}$  is a martingale for each  $n \in \mathbb{N}$ . A process is a **semimartingale** if it is the sum of a local martingale and a process that is locally of finite bounded variation (i.e., finite bounded variation on every interval  $[0, t]$ ). We will be dealing almost exclusively with continuous processes, so unless stated otherwise, all of our processes will have continuous paths.

If  $M$  is a continuous square integrable martingale, then  $M^2$  is a submartingale. In the light of Doob-Meyer's decomposition theorem, there is a (unique) increasing continuous process, denoted by  $\langle M \rangle_t$  such that  $M_t^2 - \langle M \rangle_t$  is a martingale. Particularly,  $\mathbf{E}M_t^2 - \mathbf{E}M_0^2 = \mathbf{E}\langle M \rangle_t$ . More general, if  $M$  is a local martingale, the **quadratic variation** of  $M$  is the unique increasing continuous process  $\langle M \rangle_t$  such that  $M_t^2 - \langle M \rangle_t$  is a local martingale.

If  $X_t = M_t + A_t$ , where  $M_t$  is a local martingale and  $A_t$  has paths of locally finite bounded variation, then  $\langle X \rangle_t$  is defined to be  $\langle M \rangle_t$ . If  $X$  and  $Y$  are two semimartingales, we define

$$\langle X, Y \rangle_t = \frac{1}{2} (\langle X + Y \rangle_t - \langle X \rangle_t - \langle Y \rangle_t).$$

**Lemma 1.3.11.** Let  $M_t$  be a square integrable martingale (that is,  $M_t \in L^2$  for every  $t \geq 0$ ). Let  $0 \leq s < t$  and let  $s = t_0 < t_1 < \dots < t_n = t$  be a division of the interval  $[s, t]$ . Then,

$$\mathbf{E} \left[ \sum_{i=1}^n (M_{t_i} - M_{t_{i-1}})^2 \mid \mathcal{F}_s \right] = \mathbf{E} [M_t^2 - M_s^2 \mid \mathcal{F}_s] = \mathbf{E} [(M_t - M_s)^2 \mid \mathcal{F}_s].$$

*Proof.* For every  $i = 1, \dots, n$ ,

$$\begin{aligned} \mathbf{E} \left[ (M_{t_i} - M_{t_{i-1}})^2 \mid \mathcal{F}_s \right] &= \mathbf{E} \left[ \mathbf{E} \left[ (M_{t_i} - M_{t_{i-1}})^2 \mid \mathcal{F}_{t_{i-1}} \right] \mid \mathcal{F}_s \right] \\ &= \mathbf{E} \left[ \mathbf{E} \left[ M_{t_i}^2 \mid \mathcal{F}_{t_{i-1}} \right] - 2M_{t_{i-1}} \mathbf{E} \left[ M_{t_i} \mid \mathcal{F}_{t_{i-1}} \right] + M_{t_{i-1}}^2 \mid \mathcal{F}_s \right] \\ &= \mathbf{E} \left[ \mathbf{E} \left[ M_{t_i}^2 \mid \mathcal{F}_{t_{i-1}} \right] - M_{t_{i-1}}^2 \mid \mathcal{F}_s \right] \\ &= \mathbf{E} \left[ M_{t_i}^2 - M_{t_{i-1}}^2 \mid \mathcal{F}_s \right] \end{aligned}$$

and the desired result follows by summing over  $i$ .  $\square$

**Theorem 1.3.12.** *Let  $M_t$  be a continuous local martingale. There exists an increasing process denoted by  $\langle M \rangle_t$ , which is unique up to indistinguishability, such that  $M_t^2 - \langle M \rangle_t$  is a continuous local martingale. Furthermore, for every fixed  $t > 0$ , if  $\pi^n = \{(t_0^n, \dots, t_{k_n}^n) : 0 = t_0^n < t_1^n < \dots < t_{k_n}^n = t\}$  is an increasing sequence of subdivisions of  $[0, t]$  with mesh going to 0, then we have*

$$\langle M \rangle_t = \lim_{n \rightarrow \infty} \sum_{i=1}^{k_n} \left( M_{t_i^n} - M_{t_{i-1}^n} \right)^2 \quad (1.8)$$

*in probability. The process  $\langle M \rangle_t$  is called the quadratic variation of  $M_t$ .*

**Lemma 1.3.13.** *Let  $M_t$  be a continuous bounded martingale. Let  $\pi^n = \{(t_0^n, \dots, t_{k_n}^n) : 0 = t_0^n < t_1^n < \dots < t_{k_n}^n = T (\leq \infty)\}$  be an increasing sequence of subdivisions of  $[0, T]$  with mesh going to 0, then for each  $n$ ,*

$$N_t^n := \sum_{i=1}^{k_n} M_{t_{i-1}^n} (M_{t_i^n \wedge t} - M_{t_{i-1}^n \wedge t})$$

*is a martingale, and there exists a continuous square integrable martingale  $N_t$  such that*

$$\mathbf{E} \sup_{t \in [0, T]} |N_t^n - N_t|^2 \rightarrow 0.$$

*Proof.* It is easy to verify that  $N_t^n$  is a martingale. Let us fix  $n \leq m$  and evaluate the product  $\mathbf{E}(N_T^n N_T^m)$ . This product is equal to

$$\sum_{i=1}^{k_n} \sum_{j=1}^{k_m} \mathbf{E} \left[ M_{t_{i-1}^n} \left( M_{t_i^n} - M_{t_{i-1}^n} \right) M_{t_{j-1}^m} \left( M_{t_j^m} - M_{t_{j-1}^m} \right) \right].$$

In this double sum, the only terms that may be nonzero are those corresponding to indices  $i$  and  $j$  such that the interval  $(t_{j-1}^m, t_j^m]$  is contained in  $(t_{i-1}^n, t_i^n]$ . Indeed, suppose that  $t_i^n \leq t_{j-1}^m$  (the symmetric case  $t_j^m \leq t_{i-1}^n$  is treated in an analogous way).

Then, conditioning on the  $\sigma$ -field  $\mathcal{F}_{t_{j-1}^m}$ , we have

$$\begin{aligned} &\mathbf{E} \left[ M_{t_{i-1}^n} \left( M_{t_i^n} - M_{t_{i-1}^n} \right) M_{t_{j-1}^m} \left( M_{t_j^m} - M_{t_{j-1}^m} \right) \right] \\ &= \mathbf{E} \left[ M_{t_{i-1}^n} \left( M_{t_i^n} - M_{t_{i-1}^n} \right) M_{t_{j-1}^m} \mathbf{E} \left[ M_{t_j^m} - M_{t_{j-1}^m} \mid \mathcal{F}_{t_{j-1}^m} \right] \right] = 0. \end{aligned}$$

For every  $j = 1, \dots, k_m$ , write  $i_{n,m}(j)$  for the unique index  $i$  such that  $(t_{j-1}^m, t_j^m] \subset (t_{i-1}^n, t_i^n]$ . It follows from the previous considerations that

$$\mathbf{E}[N_T^n N_T^m] = \sum_{1 \leq j \leq k_m, i = i_{n,m}(j)} \mathbf{E} \left[ M_{t_{i-1}^n} \left( M_{t_i^n} - M_{t_{i-1}^n} \right) M_{t_{j-1}^m} \left( M_{t_j^m} - M_{t_{j-1}^m} \right) \right].$$

In each term  $\mathbf{E} \left[ M_{t_{i-1}^n} \left( M_{t_i^n} - M_{t_{i-1}^n} \right) M_{t_{j-1}^m} \left( M_{t_j^m} - M_{t_{j-1}^m} \right) \right]$ , we can now decompose

$$M_{t_i^n} - M_{t_{i-1}^n} = \sum_{k: i_{n,m}(k) = i} \left( M_{t_k^m} - M_{t_{k-1}^m} \right)$$

and we observe that, if  $k$  is such that  $i_{n,m}(k) = i$  but  $k \neq j$ ,

$$\mathbf{E} \left[ M_{t_{i-1}^n} \left( M_{t_k^m} - M_{t_{k-1}^m} \right) M_{t_{j-1}^m} \left( M_{t_j^m} - M_{t_{j-1}^m} \right) \right] = 0$$

(condition on  $\mathcal{F}_{t_{k-1}^m}$  if  $k > j$  and on  $\mathcal{F}_{t_j^m}$  if  $k < j$ ). The only case that remains is  $k = j$ , and we have thus obtained

$$\mathbf{E}[N_T^n N_T^m] = \sum_{1 \leq j \leq k_m, i = i_{n,m}(j)} \mathbf{E} \left[ M_{t_{i-1}^n} M_{t_{j-1}^m} \left( M_{t_j^m} - M_{t_{j-1}^m} \right)^2 \right].$$

As a special case of this relation, we have

$$\mathbf{E}[(N_T^m)^2] = \sum_{1 \leq j \leq k_m} \mathbf{E} \left[ M_{t_{j-1}^m}^2 \left( M_{t_j^m} - M_{t_{j-1}^m} \right)^2 \right].$$

Furthermore,

$$\begin{aligned} \mathbf{E}[(N_T^n)^2] &= \sum_{1 \leq i \leq k_n} \mathbf{E} \left[ M_{t_{i-1}^n}^2 \left( M_{t_i^n} - M_{t_{i-1}^n} \right)^2 \right] \\ &= \sum_{1 \leq i \leq k_n} \mathbf{E} \left[ M_{t_{i-1}^n}^2 \mathbf{E} \left[ \left( M_{t_i^n} - M_{t_{i-1}^n} \right)^2 \mid \mathcal{F}_{t_{i-1}^n}^n \right] \right] \\ &= \sum_{1 \leq i \leq k_n} \mathbf{E} \left[ M_{t_{i-1}^n}^2 \sum_{j: i_{n,m}(j) = i} \mathbf{E} \left[ \left( M_{t_j^m} - M_{t_{j-1}^m} \right)^2 \mid \mathcal{F}_{t_{i-1}^n}^n \right] \right] \\ &= \sum_{1 \leq j \leq k_m, i = i_{n,m}(j)} \mathbf{E} \left[ M_{t_{i-1}^n}^2 \left( M_{t_j^m} - M_{t_{j-1}^m} \right)^2 \right], \end{aligned}$$

If we combine the last three displays, we get

$$\mathbf{E}[(N_T^n - N_T^m)^2] = \mathbf{E} \left[ \sum_{1 \leq j \leq k_m, i = i_{n,m}(j)} \left( M_{t_{i-1}^n} - M_{t_{j-1}^m} \right)^2 \left( M_{t_j^m} - M_{t_{j-1}^m} \right)^2 \right].$$

Using the Cauchy-Schwarz inequality, we then have

$$\begin{aligned} \mathbf{E} [(N_T^n - N_T^m)^2] &\leq \mathbf{E} \left[ \sup_{1 \leq j \leq k_m, i=i_{n,m}(j)} \left( M_{t_{i-1}^n} - M_{t_{j-1}^m} \right)^4 \right]^{1/2} \\ &\quad \times \mathbf{E} \left[ \left( \sum_{1 \leq j \leq k_m} \left( M_{t_j^m} - M_{t_{j-1}^m} \right)^2 \right)^2 \right]^{1/2}. \end{aligned}$$

By the continuity of sample paths (together with the fact that the mesh of our subdivisions tends to 0) and dominated convergence, we have

$$\lim_{n, m \rightarrow \infty, n \leq m} \mathbf{E} \left[ \sup_{1 \leq j \leq k_m, i=i_{n,m}(j)} \left( M_{t_{i-1}^n} - M_{t_{j-1}^m} \right)^4 \right] = 0.$$

To complete the proof of the lemma, it is then enough to prove the existence of a finite constant  $C$  such that, for every  $m$ ,

$$\mathbf{E} \left[ \left( \sum_{1 \leq j \leq k_m} \left( M_{t_j^m} - M_{t_{j-1}^m} \right)^2 \right)^2 \right] \leq C.$$

Let  $A$  be a constant such that  $|M_t| \leq A$  for every  $t \geq 0$ . Expanding the square and using Proposition 3.14 twice, we have

$$\begin{aligned} &\mathbf{E} \left[ \left( \sum_{1 \leq j \leq k_m} \left( M_{t_j^m} - M_{t_{j-1}^m} \right)^2 \right)^2 \right] \\ &= \mathbf{E} \left[ \sum_{1 \leq j \leq k_m} \left( M_{t_j^m} - M_{t_{j-1}^m} \right)^4 \right] + 2\mathbf{E} \left[ \sum_{1 \leq j < k \leq k_m} \left( M_{t_j^m} - M_{t_{j-1}^m} \right)^2 \left( M_{t_k^m} - M_{t_{k-1}^m} \right)^2 \right] \\ &\leq 4A^2 \mathbf{E} \left[ \sum_{1 \leq j \leq k_m} \left( M_{t_j^m} - M_{t_{j-1}^m} \right)^2 \right] \\ &\quad + 2 \sum_{j=1}^{k_m-1} \mathbf{E} \left[ \left( M_{t_j^m} - M_{t_{j-1}^m} \right)^2 \mathbf{E} \left[ \sum_{k=j+1}^{k_m} \left( M_{t_k^m} - M_{t_{k-1}^m} \right)^2 \mid \mathcal{F}_{t_j^m} \right] \right] \\ &= 4A^2 \mathbf{E} \left[ \sum_{1 \leq j \leq k_m} \left( M_{t_j^m} - M_{t_{j-1}^m} \right)^2 \right] \\ &\quad + 2 \sum_{j=1}^{k_m-1} \mathbf{E} \left[ \left( M_{t_j^m} - M_{t_{j-1}^m} \right)^2 \mathbf{E} \left[ \left( M_T - M_{t_j^m} \right)^2 \mid \mathcal{F}_{t_j^m} \right] \right] \\ &\leq 12A^2 E \left[ \sum_{1 \leq j \leq k_m} \left( M_{t_j^m} - M_{t_{j-1}^m} \right)^2 \right] \\ &= 12A^2 E \left[ (M_T - M_0)^2 \right] \end{aligned}$$



---


$$\leq 48A^4$$

Thanks to Doob's inequality, we obtain our assertion.  $\square$

We now return to the proof of the theorem.

*Proof of Theorem 1.3.12.* Assume that  $M$  is a continuous bounded martingale with  $M_0 = 0$ . Note that

$$M_t^2 - 2N_t^n = \sum_{i=1}^{k_n} (M_{t_i \wedge t} - M_{t_{i-1} \wedge t})^2$$

are nondecreasing along the finite sequence  $(t_i^n, 0 \leq i \leq k_n)$ . By passing to the limit  $n \rightarrow \infty$  along the sequence  $(n_k)_{k \geq 1}$ , we get that the sample paths of  $M_t^2 - 2Y_t$  are continuous and nondecreasing on  $[0, T]$ , and it is a martingale. It follows that we can define an increasing process  $A_t^T$  such that  $A_t^T = M_t^2 - 2Y_t$  for every  $t \in [0, T]$  and clearly  $M_t^2 - A_t^T$  is a martingale. It is now standard to verify that  $A_t^T$  and  $A_t^{T'}$  coincide when  $t \leq T \wedge T'$ . It follows that we can define an increasing process  $\langle M \rangle$  such that  $\langle M \rangle_t = A_t^T$  for every  $t \in [0, T]$  and every  $T \geq 1$ , a.s., and clearly  $M_t^2 - \langle M \rangle_t$  is a martingale. By fact that  $N_T^n$  converges in  $L^2$  to  $N_T = \frac{1}{2} (M_T^2 - A_T^T)$ , we get that

$$\lim_{n \rightarrow \infty} \sum_{j=1}^{k_n} (M_{t_j^n} - M_{t_{j-1}^n})^2 = \langle M \rangle_T$$

in  $L^2$ . This completes the proof of the theorem in the case when  $M_0 = 0$  and  $M$  is bounded.

Let us consider the general case. Writing  $M_t = M_0 + m_t$ , so that  $M_t^2 = M_0^2 + 2M_0m_t + m_t^2$ , and noting that  $M_0m_t$  is a continuous local martingale, we see that we may assume that  $M_0 = 0$ . We then set

$$\tau_n = \inf \{t \geq 0 : |M_t| \geq n\}$$

and we can apply the bounded case to the stopped martingales  $M^{\tau_n}$ . Set  $A^{[n]} = \langle M^{\tau_n} \rangle$ . The uniqueness part of the theorem shows that the processes  $A_{t \wedge \tau_n}^{[n+1]}$  and  $A_t^{[n]}$  are indistinguishable. It follows that there exists an increasing process  $A$  such that, for every  $n$ , the processes  $A_{t \wedge \tau_n}$  and  $A_t^{[n]}$  are indistinguishable. By construction,  $M_{t \wedge \tau_n}^2 - A_{t \wedge \tau_n}$  is a martingale for every  $n$ , which precisely implies that  $M_t^2 - A_t$  is a continuous local martingale. We take  $\langle M \rangle_t = A_t$ , which completes the proof of the existence part of the theorem.

Finally, to get (1.8), it suffices to consider the case  $M_0 = 0$ . The bounded case then shows that (1.8) holds if  $M$  and  $\langle M \rangle_t$  are replaced respectively by  $M^{\tau_n}$  and  $\langle M \rangle_{t \wedge \tau_n}$  (even with convergence in  $L^2$ ). Then it is enough to observe that, for every  $t > 0$ ,  $P(t \leq \tau_n)$  converges to 1 when  $n \rightarrow \infty$ .  $\square$

### 1.3.3 Some Applications of Martingale Theory

For an example of a discrete martingale, let  $\Omega = [0, 1]$ ,  $\mathbf{P}$  Lebesgue measure, and  $f$  an integrable function on  $[0, 1]$ . Let  $\mathcal{F}_n$  be the  $\sigma$ -field generated by the sets

$$\{[k/2^n, (k+1)/2^n), k = 0, 1, \dots, 2^n - 1\}.$$

Let  $f_n = \mathbf{E}[f \mid \mathcal{F}_n]$ . If  $I$  is an interval in  $\mathcal{F}_n$ , shows that

$$f_n(x) = \frac{1}{|I|} \int_I f(y) dy \quad \text{if } x \in I.$$

$f_n$  is a particular example of what is known as a dyadic martingale. Of course,  $[0, 1]$  could be replaced by any interval as long as we normalize so that the total mass of the interval is 1. We could also divide cubes in  $\mathbb{R}^d$  into  $2^d$  subcubes at each step and define  $f_n$  analogously. Such martingales are called dyadic martingales. In fact, we could replace Lebesgue measure by any finite measure  $\mu$ , and instead of decomposing into equal subcubes, we could use any nested partition of sets we like, provided none of these sets had  $\mu$  measure 0.

**Exercise 1.3.2.** *Show that the martingale convergence theorem implies that  $f_n \rightarrow f$ , a.e.*

The martingale convergence theorem also provides the basis of the Calderón-Zygmund lemma. We give the standard proof, phrased in martingale language.

**Lemma 1.3.14** (Calderón-Zygmund). *Let  $f \geq 0$  and  $f \in L^1(Q_1)$ . Let  $\lambda > 0$ . There exists a closed set  $F$  and countably many pairwise disjoint open cubes  $Q^i$  such that  $|F^c \Delta \cup_i Q^i| = 0$ ,  $f \leq \lambda$  a.e. on  $F$ , and for each  $i$ ,*

$$\lambda \leq \frac{1}{|Q^i|} \int_{Q^i} f \leq 2^d \lambda. \quad (1.9)$$

Moreover,

$$|F^c| \leq \frac{1}{\lambda} \int f.$$

*Proof.* First of all, note that the last assertion is a consequence of (1.9), since (1.9) implies  $|Q^i| \leq \int_{Q^i} f / \lambda$ . Summing over  $i$  then gives

$$\left| \bigcup_i Q^i \right| \leq \int_{\cup_i Q^i} f / \lambda \leq \int f / \lambda.$$

Suppose next that  $Q$  is a cube with  $|Q|^{-1} \int_Q f < \lambda$  and look at the dyadic martingale  $f_n = \mathbf{E}[f \mid \mathcal{F}_n]$ , where  $\mathcal{F}_n$  is the partition of  $Q$  into  $2^{nd}$  equal cubes. Let  $\tau = \inf \{n : f_n > \lambda\}$ . Note by our assumption on  $Q$  that  $\tau > 0$ . For each  $n$ ,  $\{\tau = n\}$  is the union of cubes in  $\mathcal{F}_n$ , and the boundary of  $\{T = n\}$  has measure 0. Let

$$F^c = \bigcup_n \text{int}\{T = n\}.$$

Then  $F$  is closed. If  $n < \tau$ , then  $f_n \leq \lambda$ . By the martingale convergence theorem, on  $\{\tau = \infty\}$  we have  $f = \lim_n f_n \leq \lambda$  a.e.. Since  $F$  differs from the set

$$\{\tau = \infty\} = \{\tau < \infty\}^c = (\cup_n \{\tau = n\})^c$$

by a set of measure 0, then  $f \leq \lambda$  a.e. on  $F$ .

If  $Q$  is one of the cubes in  $\mathcal{F}_n$  contained in  $\{\tau = n\}$  and  $x \in \text{int}(Q)$ , then

$$\frac{1}{|Q|} \int_Q f = f_n(x) > \lambda.$$

By our definition of  $\tau$ ,  $f_{n-1}(x) \leq \lambda$ . Let  $\tilde{Q}$  be the element of  $\mathcal{F}_{n-1}$  containing  $Q$ . Then since  $|\tilde{Q}| = 2^d |Q|$ ,

$$\frac{1}{|Q|} \int_Q f \leq \frac{2^d}{|\tilde{Q}|} \int_{\tilde{Q}} f \leq 2^d \lambda.$$

Finally, take  $n_0$  large enough so that  $\|f\|_1/2^{n_0d} < \lambda/2$ . We will take each cube  $Q_j$  that has side length  $2^{n_0}$  and vertices at integer multiples of  $2^{n_0}$  and decompose  $Q_j$  into  $F_j$  and  $Q_j$  as above. If we then let  $F = \cup_j F_j$ , we have our result.  $\square$

We already observed that as a consequence of Kolmogorov's continuity theorem, the Brownian paths are  $\alpha$ -Hölder continuous for every  $\alpha \in (0, \frac{1}{2})$ . The next proposition, which is known as the law of iterated logarithm shows in particular that Brownian paths are not  $\frac{1}{2}$ -Hölder continuous.

**Theorem 1.3.15** (law of iterated logarithm). *Let  $(B_t)_{t \geq 0}$  be a Brownian motion. For  $s \geq 0$ ,*

$$\mathbf{P} \left( \liminf_{t \rightarrow 0} \frac{B_{t+s} - B_s}{\sqrt{2t \log \log \frac{1}{t}}} = -1, \limsup_{t \rightarrow 0} \frac{B_{t+s} - B_s}{\sqrt{2t \log \log \frac{1}{t}}} = 1 \right) = 1.$$

*Proof.* Thanks to the symmetry and invariance by translation of the Brownian motion, it suffices to show that:

$$\mathbf{P} \left( \limsup_{t \rightarrow 0} \frac{B_t}{\sqrt{2t \log \log \frac{1}{t}}} = 1 \right) = 1.$$

Let us first prove that

$$\mathbf{P} \left( \limsup_{t \rightarrow 0} \frac{B_t}{\sqrt{2t \log \log \frac{1}{t}}} \leq 1 \right) = 1.$$

Let us denote  $h(t) = \sqrt{2t \log \log \frac{1}{t}}$ . Let  $\alpha, \beta > 0$ , from Doob's maximal inequality applied to the martingale  $\left( e^{\alpha B_t - \frac{\alpha^2}{2} t} \right)_{t \geq 0}$ , we have for  $t \geq 0$ :

$$\mathbf{P} \left( \sup_{0 \leq s \leq t} \left( B_s - \frac{\alpha}{2} s \right) > \beta \right) = \mathbf{P} \left( \sup_{0 \leq s \leq t} e^{\alpha B_s - \frac{\alpha^2}{2} s} > e^{\alpha \beta} \right) \leq e^{-\alpha \beta}.$$

Let now  $\theta, \delta \in (0, 1)$ . Using the previous inequality for every  $n \in \mathbb{N}$  with  $t = \theta^n$ ,  $\alpha = \frac{(1+\delta)h(\theta^n)}{\theta^n}$ ,  $\beta = \frac{1}{2}h(\theta^n)$ , yields when  $n \rightarrow +\infty$ ,

$$\mathbf{P} \left( \sup_{0 \leq s \leq \theta^n} \left( B_s - \frac{(1+\delta)h(\theta^n)}{2\theta^n} s \right) > \frac{1}{2}h(\theta^n) \right) = O \left( \frac{1}{n^{1+\delta}} \right).$$

Therefore from Borel-Cantelli lemma, for almost every  $\omega \in \Omega$ , we may find  $N(\omega) \in \mathbb{N}$  such that for  $n \geq N(\omega)$ ,

$$\sup_{0 \leq s \leq \theta^n} \left( B_s(\omega) - \frac{(1+\delta)h(\theta^n)}{2\theta^n} s \right) \leq \frac{1}{2}h(\theta^n).$$

But,

$$\sup_{0 \leq s \leq \theta^n} \left( B_s(\omega) - \frac{(1+\delta)h(\theta^n)}{2\theta^n} s \right) \leq \frac{1}{2}h(\theta^n)$$

implies that for  $\theta^{n+1} \leq t \leq \theta^n$ ,

$$B_t(\omega) \leq \sup_{0 \leq s \leq \theta^n} B_s(\omega) \leq \frac{1}{2}(2+\delta)h(\theta^n) \leq \frac{(2+\delta)h(t)}{2\sqrt{\theta}}.$$

We conclude:

$$\mathbf{P} \left( \limsup_{t \rightarrow 0} \frac{B_t}{\sqrt{2t \log \log \frac{1}{t}}} \leq \frac{2+\delta}{2\sqrt{\theta}} \right) = 1.$$

Letting now  $\theta \rightarrow 1$  and  $\delta \rightarrow 0$  yields

$$\mathbf{P} \left( \limsup_{t \rightarrow 0} \frac{B_t}{\sqrt{2t \log \log \frac{1}{t}}} \leq 1 \right) = 1.$$

Let us now prove that

$$\mathbf{P} \left( \limsup_{t \rightarrow 0} \frac{B_t}{\sqrt{2t \log \log \frac{1}{t}}} \geq 1 \right) = 1.$$

Let  $\theta \in (0, 1)$ . For  $n \in \mathbb{N}$ , we denote

$$A_n = \left\{ \omega, B_{\theta^n}(\omega) - B_{\theta^{n+1}}(\omega) \geq (1 - \sqrt{\theta})h(\theta^n) \right\}.$$

Let us prove that  $\sum \mathbf{P}(A_n) = +\infty$ . The basic inequality

$$\int_a^{+\infty} e^{-\frac{u^2}{2}} du \geq \frac{a}{1+a^2} e^{-\frac{a^2}{2}},$$

implies

$$\mathbf{P}(A_n) = \frac{1}{\sqrt{2\pi}} \int_{a_n}^{+\infty} e^{-\frac{u^2}{2}} du \geq \frac{a_n}{1+a_n^2} e^{-\frac{a_n^2}{2}},$$

with

$$a_n = \frac{(1 - \sqrt{\theta})h(\theta^n)}{\theta^{n/2}\sqrt{1 - \theta}}.$$

When  $n \rightarrow +\infty$ ,

$$\frac{a_n}{1+a_n^2} e^{-\frac{a_n^2}{2}} = O\left(\frac{1}{n^{\frac{1+\theta-2\sqrt{\theta}}{1-\theta}}}\right),$$

therefore,

$$\sum \mathbf{P}(A_n) = +\infty.$$

As a consequence of the independence of the Brownian increments and of Borel-Cantelli lemma, the event

$$B_{\theta^n} - B_{\theta^{n+1}} \geq (1 - \sqrt{\theta})h(\theta^n)$$

will occur almost surely for infinitely many  $n$ 's. But, thanks to the first part of the proof, for almost every  $\omega$ , we may find  $N(\omega)$  such that for  $n \geq N(\omega)$ ,

$$B_{\theta^{n+1}} > -2h(\theta^{n+1}) \geq -2\sqrt{\theta}h(\theta^n).$$

Thus, almost surely, the event  $B_{\theta^n} > h(\theta^n)(1 - 3\sqrt{\theta})$  will occur for infinitely many  $n$ 's. This implies

$$\mathbf{P}\left(\limsup_{t \rightarrow 0} \frac{B_t}{\sqrt{2t \log \log \frac{1}{t}}} \geq 1 - 3\sqrt{\theta}\right) = 1.$$

We finally get

$$\mathbf{P}\left(\limsup_{t \rightarrow 0} \frac{B_t}{\sqrt{2t \log \log \frac{1}{t}}} \geq 1\right) = 1.$$

by letting  $\theta \rightarrow 0$ . □

As a straightforward consequence, we may observe that the time inversion invariance property of Brownian motion implies:

**Corollary 1.3.16.** *Let  $(B_t)_{t \geq 0}$  be a standard Brownian motion.*

$$\mathbf{P}\left(\liminf_{t \rightarrow +\infty} \frac{B_t}{\sqrt{2t \log \log t}} = -1, \limsup_{t \rightarrow +\infty} \frac{B_t}{\sqrt{2t \log \log t}} = 1\right) = 1.$$

## 1.4 Stochastic Integral

We write  $\mathbb{H}^2$  for the space of all continuous martingales  $M$  such that  $\mathbf{E}\langle M \rangle_\infty < \infty$  and  $M_0 = 0$ , with the usual convention that two indistinguishable processes are identified.

**Exercise 1.4.1.** *The space  $\mathbb{H}^2$  equipped with the scalar product  $\langle M, N \rangle_\infty$  is a Hilbert space.*

Let  $M \in \mathbb{H}^2$ ,  $0 = t_0 \leq t_1 \leq \dots \leq t_n < \infty$  and  $H_s(\omega) = \sum_{i=0}^{n-1} F_i(\omega) \mathbf{1}_{(t_i, t_{i+1}]}(s)$ , where  $F_i$  is bounded and  $\mathcal{F}_{t_i}$ -measurable. Denote the collection of all such  $H$  by  $\mathcal{S}_a$ .

For any  $H \in \mathcal{S}_a$ , define

$$(H \cdot M)_t = \int_0^t H_s dM_s := \sum_{i=0}^{n-1} F_i (M_{t \wedge t_{i+1}} - M_{t \wedge t_i}).$$

We have

**Lemma 1.4.1.** *Let  $M \in \mathbb{H}^2$  and  $H \in \mathcal{S}_a$ . Then  $(H \cdot M) \in \mathbb{H}^2$ , and*

$$\langle H \cdot M \rangle_t = \int_0^t H_s^2 d\langle M \rangle_s. \quad (1.10)$$

*In particular, we have the following Itô isometry:*

$$\mathbf{E} \left( \int_0^t H_s dM_s \right)^2 = \mathbf{E} \int_0^t H_s^2 d\langle M \rangle_s. \quad (1.11)$$

*Proof.*

$$\begin{aligned} \mathbf{E} \left( \int_0^t H_s dM_s \right)^2 &= \mathbf{E} \sum_i H_{t_i}^2 (M_{t \wedge t_{i+1}} - M_{t \wedge t_i})^2 \\ &\quad + 2\mathbf{E} \sum_{i < j} H_{t_i} H_{t_j} (M_{t \wedge t_{i+1}} - M_{t \wedge t_i})(M_{t \wedge t_{j+1}} - M_{t \wedge t_j}) =: I_1 + I_2. \end{aligned}$$

$$\begin{aligned} I_1(t) &= \sum_i \mathbf{E} \mathbf{E} (H_{t_i}^2 (M_{t \wedge t_{i+1}} - M_{t \wedge t_i})^2 | \mathcal{F}_{t \wedge t_i}) = \sum_i \mathbf{E} [H_{t_i}^2 \mathbf{E} ((M_{t \wedge t_{i+1}} - M_{t \wedge t_i})^2 | \mathcal{F}_{t \wedge t_i})] \\ &= \sum_i \mathbf{E} H_{t_i}^2 (\langle M \rangle_{t \wedge t_{i+1}} - \langle M \rangle_{t \wedge t_i}) = \mathbf{E} \int_0^t H_s^2 d\langle M \rangle_s, \end{aligned}$$

$$I_2 = 2 \sum_{i < j} \mathbf{E} [H_{t_i} H_{t_j} (M_{t \wedge t_{i+1}} - M_{t \wedge t_i}) \mathbf{E} ((M_{t \wedge t_{j+1}} - M_{t \wedge t_j}) | \mathcal{F}_{t \wedge t_j})] = 0.$$

Therefore,

$$\mathbf{E} \left( \int_0^t H_s dM_s \right)^2 = \mathbf{E} \int_0^t H_s^2 d\langle M \rangle_s.$$

Similarly, one can also prove that

$$\mathbf{E} \left\{ \left[ \left( \int_{t_0}^t H_s dM_s \right)^2 - \int_{t_0}^t H_s^2 d\langle M \rangle_s \right] \middle| \mathcal{F}_{t_0} \right\} = 0, \quad t_0 < t,$$

which yields (1.10). It is also easy to verify that  $t \mapsto \int_0^t H_s dW_s = \sum_{i=0}^{n-1} H_{t_i} (W_{t \wedge t_{i+1}} - W_{t \wedge t_i})$  is a continuous martingale, we let it to the reader.  $\square$

We then can use the above lemma and Exercise 1.3.1 to extend the above construction to more general  $H_s$  satisfying the following condition: there exists a sequence  $H^n \in \mathcal{S}_a$  such that

$$\mathbf{E} \int_0^\infty (H_s^n - H_s)^2 d\langle M \rangle_s < \infty.$$

We denote the progressive  $\sigma$ -field on  $\mathbb{R}_+ \times \Omega$  by  $\mathcal{P}$ , and if  $M \in \mathbb{H}^2$ , we let  $L_{\mathcal{P}}^2(M)$  be the set of all progressive processes  $H$  such that

$$\mathbf{E} \left[ \int_0^\infty H_s^2 d\langle M \rangle_s \right] < \infty,$$

with the convention that two progressive processes  $H$  and  $H'$  satisfying this integrability condition are identified if  $H_s = H'_s = 0$ ,  $d\langle M \rangle_s$  a.e., a.s. We can view  $L_{\mathcal{P}}^2(M)$  as an ordinary  $L^2$  space, namely

$$L_{\mathcal{P}}^2(M) = L^2(\mathbb{R}_+ \times \Omega, \mathcal{P}, d\langle M \rangle_s dP)$$

where  $d\langle M \rangle_s dP$  refers to the finite measure on  $(\mathbb{R}_+ \times \Omega, \mathcal{P})$  that assigns the mass

$$\mathbf{E} \left[ \int_0^\infty 1_A(\omega, s) d\langle M \rangle_s \right]$$

to a set  $A \in \mathcal{P}$  (the total mass of this measure is  $\mathbf{E}\langle M \rangle_\infty = \|M\|_{\mathbb{H}^2}^2$ ). Just like any  $L^2$  space, the space  $L_{\mathcal{P}}^2(M)$  is a Hilbert space for the scalar product. Just like any  $L^2$  space, the space  $L_{\mathcal{P}}^2(M)$  is a Hilbert space for the scalar product

$$(H, K)_{L_{\mathcal{P}}^2(M)} = \mathbf{E} \left[ \int_0^\infty H_s K_s d\langle M \rangle_s \right].$$

**Proposition 1.4.2.** *For every  $M \in \mathbb{H}^2$ ,  $\mathcal{S}_a$  is dense in  $L_{\mathcal{P}}^2(M)$ .*

Therefore, by the above result and the discussion after Lemma 1.4.1, for each  $M \in \mathbb{H}^2$  and  $H \in L_{\mathcal{P}}^2(M)$ , we can define

$$(H \cdot M)_t = \int_0^t H_s dM_s := \lim_{n \rightarrow \infty} \int_0^t H_s^n dM_s,$$

where  $H^n \in \mathcal{S}_a$  such that  $H^n \rightarrow H$  in  $L_{\mathcal{P}}^2(M)$ . Moreover,  $(H \cdot M) \in \mathbb{H}^2$  and

$$\langle H \cdot M \rangle_t = \int_0^t H_s^2 d\langle M \rangle_s.$$

Let  $M, N \in \mathbb{H}^2$ ,  $H \in L_{\mathcal{P}}^2(M)$  and  $K \in L_{\mathcal{P}}^2(N)$ . Since  $H \cdot M$  and  $K \cdot N$  are martingales in  $\mathbb{H}^2$ , we have, for every  $t \in [0, \infty]$ ,

$$\mathbf{E} \left[ \left( \int_0^t H_s dM_s \right) \left( \int_0^t K_s dN_s \right) \right] = \mathbf{E} \left[ \int_0^t H_s K_s d\langle M, N \rangle_s \right],$$

In particular,

$$\mathbf{E} \left[ \left( \int_0^t H_s dM_s \right)^2 \right] = \mathbf{E} \left[ \int_0^t H_s^2 d\langle M \rangle_s \right].$$

**Proposition 1.4.3.** *Let  $M, N \in \mathbb{H}^2$ , and let  $H \in L^2(M)$ . If  $K$  is a progressive process, we have  $KH \in L^2(M)$  if and only if  $K \in L^2(H \cdot M)$ . If the latter properties hold,*

$$(KH) \cdot M = K \cdot (H \cdot M).$$

For general continuous local martingale, we can employ standard localization argument to define the above integral. For  $X_t = M_t + A_t$ , a semimartingale,  $\int_0^t H_s dX_s$  is given by

$$\int_0^t H_s dX_s = \int_0^t H_s dM_s + \int_0^t H_s dA_s,$$

where the first integral on the right is a stochastic integral and the second integral on the right is a Riemann-Stieltjes integral.

## 1.5 Itô's formula and its applications

### 1.5.1 Applications in martingale theory

We list some important results in stochastic calculus.

**Theorem 1.5.1** (Itô's formula). *If each  $X_t^i$  (for each  $i \in 1, \dots, d$ ) is a continuous semimartingale and  $f \in C^2(\mathbb{R}^d)$ , then*

$$\begin{aligned} & f(X_t) - f(X_0) \\ &= \int_0^t \sum_{i=1}^d \partial_i f(X_s) dX_s^i + \frac{1}{2} \int_0^t \sum_{i,j=1}^d \partial_{ij} f(X_s) d\langle X^i, X^j \rangle_s \end{aligned} \quad (1.12)$$

(see [Hua01, Theorem 13.5]).

It is often useful to use the language of Stratonovitch's integration to study stochastic differential equations because the Itô's formula takes a much nicer form. If  $M_t$  is an  $\mathcal{F}_t$ -adapted real valued local martingale and if  $H_t$  is an  $\mathcal{F}_t$ -adapted continuous semimartingale satisfying  $\mathbf{P}\left(\int_0^T H_s d\langle M \rangle_s < \infty\right) = 1$ , then by definition the Stratonovitch integral of  $H_t$  with respect to  $M_t$  is defined as

$$\int_0^T H_t \circ dM_t = \int_0^T H_t dM_t + \frac{1}{2} \langle H, M \rangle_T.$$

By using Stratonovitch integral instead of Itô's, the Itô formula reduces to the classical change of variable formula.

**Theorem 1.5.2.** *Let  $M_t$  be a  $d$ -dimensional continuous semimartingale. Let now  $f$  be a  $C^2$  function. We have*

$$f(M_t) = f(M_0) + \int_0^t \partial_i f(X_s) \circ dM_s^i, \quad t \geq 0.$$



**Theorem 1.5.3** (Burkholder-Davis-Gundy inequalities). *If  $M_t$  is a continuous martingale with  $M_0 = 0$ , and  $\tau$  is a stopping time, then*

$$\mathbf{E} \sup_{t \in [0, \tau]} |M_t|^p \asymp_p \mathbf{E} \langle M \rangle_\tau^{p/2}, \quad p \in (0, \infty) \quad (1.13)$$

*Proof. Step 1:* for any  $p \geq 2$ , by Itô's formula

$$|M_\tau|^p = p \int_0^\tau \operatorname{sgn}(M_t) |M_t|^{p-2} M_t dM_t + \frac{p(p-1)}{2} \int_0^\tau |M_t|^{p-2} d\langle M \rangle_t;$$

By Doob's inequality and Hölder's inequality,

$$\begin{aligned} \mathbf{E}(M_\tau^*)^p &\lesssim_p \mathbf{E}|M_\tau|^p \lesssim_p \mathbf{E}((M_\tau^*)^{p-2} \langle M \rangle_\tau) \\ &\leq (\mathbf{E}(M_\tau^*)^p)^{1-\frac{2}{p}} (\mathbf{E} \langle M \rangle_\tau^{\frac{p}{2}})^{\frac{2}{p}}; \end{aligned}$$

*Step 2:* using Lengart's domination inequality, we can get the proof for the case  $p \in (0, 2)$ .

We proceed now to the proof of the left hand side inequality. We have,

$$M_t^2 = \langle M \rangle_t + 2 \int_0^t M_s dM_s.$$

Therefore, we get

$$\mathbf{E} \left( \langle M \rangle_T^{\frac{p}{2}} \right) \lesssim \mathbf{E}(M_T^*)^p + \mathbf{E} \left( \sup_{0 \leq t \leq T} \left| \int_0^t M_s dM_s \right|^{p/2} \right).$$

By using the previous argument, we now have

$$\begin{aligned} 2^{\frac{p}{2}} \mathbf{E} \left( \sup_{0 \leq t \leq T} \left| \int_0^t M_s dM_s \right|^{p/2} \right) &\leq C \mathbf{E} \left( \left( \int_0^T M_s^2 d\langle M \rangle_s \right)^{p/4} \right) \\ &\leq C \mathbf{E} \left( (M_T^*)^{p/2} \langle M \rangle_T^{p/4} \right) \leq C (\mathbf{E}(M_T^*)^p)^{1/2} (\mathbf{E} \langle M \rangle_T^{p/2})^{1/2} \\ &\leq \varepsilon' \mathbf{E}(M_T^*)^p + C_{\varepsilon'} \mathbf{E} \langle M \rangle_T^{p/2} \leq \varepsilon. \end{aligned}$$

As a conclusion, we obtained that  $d$  □

**Proposition 1.5.4** (Lengart). *Let  $X_t$  be a positive adapted right-continuous process and  $A_t$  be an increasing process. Assume that for every bounded stopping time  $\tau$ ,  $\mathbf{E}(X_\tau | \mathcal{F}_0) \leq \mathbf{E}(A_\tau | \mathcal{F}_0)$ . Then, for every  $\kappa \in (0, 1)$ ,*

$$\mathbf{E}(X_T^*)^\kappa \leq \frac{2-\kappa}{1-\kappa} \mathbf{E}(A_T^\kappa).$$

We shall use this lemma to prove the following Another approach to proving (1.13) is utilizing "good- $\lambda$ " inequality (cf. [RY13]).

**Theorem 1.5.5** (Lévy's theorem). *If  $X_t$  is a  $d$ -dimensional  $\mathcal{F}$ -adapted process, each of whose coordinates is a continuous local martingale, and  $\langle X^i, X^j \rangle_t = \delta_{ij}t$ , then  $X_t$  is a  $d$ -dimensional  $\mathcal{F}$ -Brownian motion.*

*Proof.* Let  $\xi \in \mathbb{R}^d$ . Then  $\xi \cdot X_t$  is a continuous local martingale with quadratic variation  $\langle \xi \cdot X \rangle_t = |\xi|^2 t$ . By Itô's formula,  $\exp(i\xi \cdot X_t + \frac{1}{2}|\xi|^2 t)$  is a continuous local martingale. This complex continuous local martingale is bounded on every finite interval and is therefore a (true) martingale, in the sense that its real and imaginary parts are both martingales. Hence, for every  $s < t$ ,

$$\mathbf{E} \left[ \exp \left( i\xi \cdot X_t + \frac{1}{2}|\xi|^2 t \right) \middle| \mathcal{F}_s \right] = \exp \left( i\xi \cdot X_s + \frac{1}{2}|\xi|^2 s \right)$$

Thus,

$$\mathbf{E} [\exp (i\xi \cdot (X_t - X_s)) | \mathcal{F}_s] = \exp \left( \frac{1}{2}|\xi|^2 (t - s) \right).$$

This implies  $X_t - X_s$  is independent with  $\mathcal{F}_s$  and  $X_t - X_s \sim \mathcal{N}(0, t - s)$ .

Finally,  $X$  is adapted and has independent increments with respect to the filtration  $\mathcal{F}$  so that  $X$  is a  $s$ -dimensional  $\mathcal{F}$ -Brownian motion.  $\square$

Let  $M_t$  be a continuous local martingale with  $M_0 = 0$ . Set  $\mathcal{E}(M)_t := \exp(M_t - \langle M \rangle_t / 2)$ .

**Lemma 1.5.6.**  $\mathcal{E}(M)_t$  is a continuous local martingale, and is the unique solution to

$$dX_t = X_t dM_t, \quad X_0 = 1.$$

**Theorem 1.5.7** (Girsanov theorem). *Let  $X_t$  and  $M_t$  be two continuous local martingales under  $\mathbb{P}$  with  $M_0 = 0$   $\mathbb{P}$ -a.s.. Assume that  $\mathcal{E}(M)_t$  is a martingale, we define a new probability measure  $\mathbb{Q}$  by setting the restriction of  $d\mathbb{Q}/d\mathbb{P}$  to  $\mathcal{F}_t$  to be  $\mathcal{E}(M)_t$ , then  $X_t - \langle X, M \rangle_t$  is a martingale under  $\mathbb{Q}$  and the quadratic variation of  $X_t$  is the same under  $\mathbb{P}$  and  $\mathbb{Q}$ .*

*Proof.* By localization, we can assume  $X$  is a martingale. Set  $Y_t = X_t - \langle X, M \rangle_t$ . We only need to verify that  $Y_t \mathcal{E}(M)_t$  is a martingale under  $\mathbb{P}$ . By Itô's formula,

$$\begin{aligned} dY_t \mathcal{E}(M)_t &= \mathcal{E}(M)_t dX_t - \mathcal{E}(M)_t d\langle X, M \rangle_t + Y_t \mathcal{E}(M)_t dM_t + d\langle X, \mathcal{E}(M) \rangle_t \\ &= \mathcal{E}(M)_t dX_t + Y_t \mathcal{E}(M)_t dM_t. \end{aligned}$$

Therefore,  $Y_t \mathcal{E}(M)_t$  is a martingale, which implies

$$\mathbf{E}_{\mathbb{Q}}(Y_t; A) = \mathbf{E}_{\mathbb{Q}}(Y_s; A), \quad \forall A \in \mathcal{F}_s,$$

i.e.

$$\mathbf{E}_{\mathbb{Q}}(Y_t | \mathcal{F}_s) = Y_s.$$

$\square$

**Exercise 1.5.1** (Onsager–Machlup function). *Let  $h \in C^2([0, 1])$  with  $h(0) = 0$ . Prove that*

$$\lim_{\varepsilon \rightarrow 0} \frac{\mathbf{P}(\|\omega - h\| < \varepsilon)}{\mathbf{P}(\|\omega\| < \varepsilon)} = \exp \left( -\frac{1}{2} \int_0^1 |\dot{h}_s|^2 ds \right).$$

---

**Theorem 1.5.8** (Dambis–Dubins–Schwarz). *Let  $M$  be a continuous local martingale such that  $\langle M \rangle_\infty = \infty$ . There exists a Brownian motion  $B$  such that*

$$M_t = B_{\langle M \rangle_t}.$$

The proof of Dambis–Dubins–Schwarz’s Theorem can also be found in [Hua01].

**Theorem 1.5.9** (Exponential martingale inequality). *Let  $M_t$  be a continuous martingale,  $\tau$  a bounded stopping time, then*

$$\mathbf{P} \left( \sup_{t \leq \tau} |M_t| > \lambda \ \& \ \langle M \rangle_\tau \leq \mu \right) \leq 2e^{-\frac{\lambda^2}{2\mu}}.$$

*We need*

**Lemma 1.5.10.** *Let  $W$  be a 1-dimensional Brownian motion. Then for any  $\lambda, t > 0$*

$$\mathbf{P} \left( \sup_{s \in [0, t]} |W_s| > \lambda \right) \leq 2e^{-\frac{\lambda^2}{2t}}$$

*Proof.* Let  $X_t = e^{a|W_t|}$  with  $a > 0$ . Since  $x \mapsto e^{a|x|}$  is a convex function,  $X_t$  is a submartingale. By Doob’s inequality (see Theorem 1.3.5), we have

$$\mathbf{P}(W_t^* > \lambda) = \mathbf{P}(X_t^* > e^{a\lambda}) \leq e^{-a\lambda} \mathbf{E}X_t = \frac{e^{-a\lambda}}{\pi t} \int_0^\infty e^{ax - \frac{x^2}{2t}} dx = 2e^{\frac{a^2 t}{2} - a\lambda}.$$

Taking  $a = \lambda/t$ , we obtain

$$\mathbf{P}(W_t^* > \lambda) \leq 2e^{-\frac{\lambda^2}{2t}}.$$

□

*Proof of Theorem 1.5.9.* By Theorem 1.5.8,  $M_t$  is a time change of a Brownian motion  $W_t$ . So the desired probability is bounded by

$$\mathbf{P} \left( \sup_{t \leq T} |W_t| > \lambda \ \& \ \langle W \rangle_T < \mu \right),$$

where  $T$  is a stopping time. Since  $\langle W \rangle_T = T$ , the probability above is in turn bounded by

$$\mathbf{P} \left( \sup_{t \leq \mu} |W_t| > \lambda \right) \leq 2e^{-\frac{\lambda^2}{2\mu}}.$$

□

### 1.5.2 Applications in PDEs

In chapter 3, we will consider the following stochastic differential equation (SDE):

$$X_t^x = x + \int_0^t \sigma(X_s^x) dW_s + \int_0^t b(X_s^x) ds, \quad x \in \mathbb{R}^d.$$

Here  $W$  is a  $n$ -dimensional BM, and  $\sigma : \mathbb{R}^d \rightarrow \mathbb{R}^{d \times n}$  and  $b : \mathbb{R}^d \rightarrow \mathbb{R}^d$ .

In chapter 4, we will study the following Poisson equation:

$$\lambda u - Lu = f,$$

where  $L = \frac{1}{2} \sigma_{ik} \sigma_{jk} \partial_{ij} + b_i \partial_i$ . The relationship between these two subjects can be easily established by Itô's formula.

Here we consider the case that  $L = \frac{1}{2} \Delta$ .

**Theorem 1.5.11.** *Suppose  $u$  is a  $C_b^2$  function satisfying the above Poisson equation with  $L = \frac{1}{2} \Delta$ . Then*

$$u(x) = \mathbf{E} \int_0^\infty e^{-\lambda t} f(W_t^x) dt$$

*Proof.* Applying Itô's formula, we have  $du(W_t^x) = dM_t + \frac{1}{2} \Delta u(W_t^x) dt$ . So

$$\begin{aligned} e^{-\lambda t} u(W_t^x) - u(x) &= \int_0^t e^{-\lambda s} dM_s + \int_0^t e^{-\lambda s} Lu(X_s^x) ds \\ &\quad - \lambda \int_0^t e^{-\lambda s} u(X_s^x) ds. \end{aligned}$$

Taking expectation, we get what we claimed.  $\square$

Let us now let  $D$  be a nice bounded domain, e.g., a ball. Poisson's equation in  $D$  requires one to find a function  $u$  such that

$$\begin{cases} \lambda u - Lu = f & \text{in } D \\ u = 0 & \text{on } \partial D, \end{cases}$$

where  $\lambda \geq 0$ . Here we can allow  $\lambda$  to be equal to 0. Recall that if  $L = \Delta$  ( $X_t$  is a Brownian motion), then the time to exit  $D$ , namely,  $\tau_D := \inf\{t : X_t \notin D\}$ , is finite almost surely.

**Theorem 1.5.12.** *Suppose  $u$  is a solution to Poisson's equation in a bounded domain  $D$  that is  $C^2$  in  $D$  and continuous on  $\bar{D}$ . Assume also that*

$$\mathbf{E}(\tau_D < \infty) = 1,$$

where  $\tau_D = \inf\{t > 0 : X_t^x \notin D\}$ . Then

$$u(x) = \mathbf{E} \int_0^{\tau_D} e^{-\lambda s} f(X_s^x) ds.$$

---

**Exercise 1.5.2.** *Prove Theorem 1.5.12.*

In this exercise,  $d \geq 3$ . Let  $K$  be a compact subset of the open unit ball of  $\mathbb{R}^d$ , and  $T_K := \inf \{t \geq 0 : B_t \in K\}$ . We assume that  $D := \mathbb{R}^d \setminus K$  is connected. We also consider a function  $g$  defined and continuous on  $K$ . The goal of the exercise is to determine all functions  $u : \bar{D} \rightarrow \mathbb{R}$  that satisfy: (P)  $u$  is bounded and continuous on  $\bar{D}$ , harmonic on  $D$ , and  $u(y) = g(y)$  if  $y \in \partial D$ . (This is the Dirichlet problem in  $D$ , but in contrast with Sect. 7.3 above,  $D$  is unbounded here.) We fix an increasing sequence  $(R_n)_{n \geq 1}$  of reals, with  $R_1 \geq 1$  and  $R_n \uparrow \infty$  as  $n \rightarrow \infty$ . For every  $n \geq 1$ , we set  $T_{(n)} := \inf \{t \geq 0 : |B_t| \geq R_n\}$ .

*Proof.* 1. Suppose that  $u$  satisfies (P). Prove that, for every  $n \geq 1$  and every  $x \in D$  such that  $|x| < R_n$ ,

$$u(x) = E_x \left[ g(B_{T_K}) \mathbf{1}_{\{T_K \leq T_{(n)}\}} \right] + E_x \left[ u(B_{T_{(n)}}) \mathbf{1}_{\{T_{(n)} \leq T_K\}} \right].$$

2. Show that, by replacing the sequence  $(R_n)_{n \geq 1}$  with a subsequence if necessary, we may assume that there exists a constant  $\alpha \in \mathbb{R}$  such that, for every  $x \in D$ ,

$$\lim_{n \rightarrow \infty} E_x \left[ u(B_{T_{(n)}}) \right] = \alpha,$$

and that we then have

$$\lim_{|x| \rightarrow \infty} u(x) = \alpha.$$

3. Show that, for every  $x \in D$ ,

$$u(x) = E_x \left[ g(B_{T_K}) \mathbf{1}_{\{T_K < \infty\}} \right] + \alpha P_x(T_K = \infty).$$

4. Assume that  $D$  satisfies the exterior cone condition at every  $y \in \partial D$  (this is defined in the same way as when  $D$  is bounded). Show that, for any choice of  $\alpha \in \mathbb{R}$ , the formula of question 3. gives a solution of the problem (P).  $\square$



# Chapter 2

## Itô processes

Let  $a_t = \frac{1}{2}\sigma_t\sigma_t^T$  and

$$x_t = \int_0^t \sigma_s \cdot dW_s + \int_0^t b_s ds. \quad (2.1)$$

For simplicity, we always assume that  $a \in \mathbb{S}_\delta^d$ .

### 2.1 Support theorem

The following result is a simplify version of Stroock-Varadhan's support theorem, which is taken from [Bas98].

**Theorem 2.1.1** (Support theorem). *Suppose  $\sigma, \sigma^{-1}$  and  $b$  are bounded,  $x_t$  satisfies (2.1). Suppose  $\varphi : [0, 1] \rightarrow \mathbb{R}^d$  is continuous with  $\varphi(0) = 0$ . Then for each  $\varepsilon > 0$ , there exists a constant  $c > 0$  depending only on  $\varepsilon$ , the modulus of continuity of  $\varphi$ , and the bounds on  $b, \sigma$  and  $\sigma^{-1}$  such that*

$$\mathbf{P} \left( \sup_{t \in [0, 1]} |x_t - \varphi(t)| \leq \varepsilon \right) \geq c. \quad (2.2)$$

This can be phrased as saying the graph of  $x_t$  stays inside an  $\varepsilon$ -tube about  $\varphi$ .

To prove Theorem 2.1, we need some auxiliary lemmas.

By Lemma 1.5.10, there is a constant  $\delta_0 > 0$  such that

$$\inf_{|x| \leq 1/3} \mathbf{P} \left( \sup_{t \in [0, \delta_0]} |W_t^x| \leq 1 \right) \geq 5/6. \quad (2.3)$$

**Lemma 2.1.2.** *Let  $W$  be a 1-dimensional Brownian motion. For any  $\varepsilon > 0$  and  $T > 0$ , there is a constant  $c(\varepsilon, T) > 0$  such that*

$$\mathbf{P} \left( \sup_{s \in [0, T]} |W_s| \leq \varepsilon \right) \geq c(\varepsilon, T).$$

## 2.1 Support theorem

*Proof.* We assume  $T = 1$ . By the scaling property of Brownian motion ( $\varepsilon^{-1}W_t \stackrel{d}{=} W_{\varepsilon^{-2}t}$ ), we only need to show

$$\mathbf{P} \left( \sup_{t \in [0, \varepsilon^{-2}]} |W_t| \leq 1 \right) \geq c(\varepsilon) > 0.$$

It is easy to see that  $\inf_{|x| \leq 1/3} \mathbf{P}(|W_{\delta_1}^x| \leq 1/3) \geq 1/3$  (for some  $0 < \delta_1 \ll 1$ ), which together with (2.3) implies that

$$\inf_{|x| \leq 1/3} \mathbf{P} \left( \sup_{t \in [0, \delta_2]} |W_t^x| \leq 1, |W_{\delta_2}^x| \leq 1/3 \right) \geq \frac{1}{6}, \quad \delta_2 = \delta_0 \wedge \delta_1 > 0.$$

By the Markov property of  $W$ ,

$$\inf_{|x| \leq 1/3} \mathbf{P} \left( \sup_{t \in [0, k\delta_2]} |W_t^x| \leq 1, |W_{\delta_2}^x| \leq 1/3 \right) \geq 6^{-k}.$$

Letting  $k = \lceil \varepsilon^{-2}\delta_2^{-1} \rceil + 1$ , we get

$$\mathbf{P} \left( \sup_{t \in [0, \varepsilon^{-2}]} |W_t| \leq 1 \right) \geq \inf_{|x| \leq 1/3} \mathbf{P} \left( \sup_{t \in [0, \varepsilon^{-2}]} |W_t^x| \leq 1, |W_{\delta_2}^x| \leq 1/3 \right) \geq 6^{-\varepsilon^{-2}\delta_2^{-1}-1} =: c(\varepsilon) > 0.$$

□

**Lemma 2.1.3.** *Suppose  $X_0 = 0$ ,  $X_t = M_t + A_t$  is a continuous semimartingale with  $dA_t/dt$  and  $d\langle M \rangle_t/dt$  bounded above by  $N_1$  and  $d\langle M \rangle_t/dt$  bounded below by  $N_2 > 0$ . If  $\varepsilon > 0$  and  $T > 0$ , then*

$$\mathbf{P} \left( \sup_{t \in [0, T]} |X_t| < \varepsilon \right) \geq c(\varepsilon, T, N_1, N_2) > 0.$$

*Proof.* Let  $\tau(t) = \inf\{u > 0 : \langle M \rangle_u > t\}$ . Then  $\tau(t) \asymp t$ , and  $B_t = M_{\tau(t)}$  is a Brownian motion due to Lemma 2.1.2. Then  $Y_t := X_{\tau(t)} = B_t + \int_0^t b_s ds$  with  $|b_s| \leq C(N_1, N_2)$ . Our assertion will follow if we can show

$$\mathbf{P} \left( \sup_{t \in [0, T]} |Y_t| \leq \varepsilon \right) \geq c > 0.$$

We now use Girsanov's theorem. Define a probability measure  $\mathbf{Q}$  by

$$d\mathbf{Q}/d\mathbf{P} = \mathcal{E}_T(-b) := \exp \left( - \int_0^T b_s dB_s - \frac{1}{2} \int_0^T |b_s|^2 ds \right) \quad \text{on } \mathcal{F}_T.$$

By Girsanov's theorem, under  $\mathbf{Q}$ ,  $Y_t$  is a Brownian motion. Therefore,

$$\mathbf{Q}(A) \geq c > 0, \quad A = \left\{ \sup_{t \in [0, T]} |Y_t| \leq \varepsilon \right\}.$$

By Hölder's inequality,

$$c \leq \mathbf{Q}(A) \leq E^{\mathbf{P}}(\mathcal{E}_T(-b)\mathbf{1}_A) \leq [E^{\mathbf{P}}\mathcal{E}_T^2(-b)]^{\frac{1}{2}} [\mathbf{P}(A)]^{\frac{1}{2}}.$$

Since  $b$  is bounded, it is easy to verify that  $E^{\mathbf{P}}\mathcal{E}_T^2(-b) < \infty$ . This yields  $\mathbf{P}(A) \geq c > 0$ . □



Now we are on the point to give

*Proof of Theorem 2.1. Step 1:* We first consider the case and  $\varphi = 0$ . Let  $z \in \partial B_{\varepsilon/4}$ . Applying Itô's formula with  $f(x) = |x - z|^2$  and setting  $y_t = |x_t - z|^2$ , then

$$y_t = z^2 + \int_0^t (x_s - z) \cdot dx_s + \int_0^t \text{tr} a_s ds, \quad \langle y \rangle_t = \int_0^t (x_s - z)^T a_s (x_s - z) ds \asymp y_t$$

$\langle y \rangle_t \geq c\varepsilon^2$  before  $\tau := \inf\{s > 0 : |y_s - y_0| \geq (\varepsilon/8)^2\}$ . If we set  $z_t$  equal to  $y_t$  for  $t \leq \tau$  and equal to some Brownian motion for  $t$  larger than this stopping time, then Lemma 2.1.3 applies (for  $z_t$ ) and

$$\mathbf{P} \left( \sup_{t \in [0, T]} |x_t| \leq \varepsilon \right) \geq \mathbf{P} \left( \sup_{t \in [0, T]} |y_t - y_0| \leq (\varepsilon/8)^2 \right) = \mathbf{P} \left( \sup_{t \in [0, T]} |z_t - z_0| \leq (\varepsilon/8)^2 \right) > 0.$$

*Step 2:* Without loss of generality, we may assume  $\varphi$  is differentiable with a derivative bounded by a constant. Define a new probability measure  $\mathbf{Q}$  by

$$d\mathbf{Q}/d\mathbf{P} = \exp \left( - \int_0^T \varphi'(s) \sigma_s^{-1} dW_s - \frac{1}{2} \int_0^T |\varphi'(s) \sigma_s^{-1}|^2 ds \right) \quad \text{on } \mathcal{F}_T.$$

Noting that

$$\left\langle - \int_0^\cdot \varphi'(s) \sigma_s^{-1} dW_s, x \right\rangle_t = \int_0^t \varphi'(s) ds = -\varphi(t).$$

So by the Girsanov theorem, under  $\mathbf{Q}$  each component of  $x_t$  is a semimartingale and  $n_t^i := x_t^i - \int_0^t b_s^i ds - \varphi^i(t)$  is a martingale for each  $i = 1, \dots, d$ . Therefore,

$$B_t := \int_0^t \sigma_s^{-1} dn_s$$

is a continuous local martingale with  $\langle B^i, B^j \rangle_t = \delta_{ij} t$  under  $\mathbf{Q}$ . Therefore  $B_t$  is a  $d$ -dimensional Brownian motion under  $\mathbf{Q}$ . Since

$$x_t - \varphi(t) = \int_0^t \sigma_s dB_s + \int_0^t b_s ds,$$

by *Step 1*,  $\mathbf{Q}(\sup_{t \in [0, T]} |x_t - \varphi(t)| < \varepsilon) \geq c > 0$ . similarly to the last paragraph of the proof for Lemma 2.1.3, we conclude

$$\mathbf{P} \left( \sup_{t \in [0, T]} |x_t - \varphi(t)| < \varepsilon \right) \geq c > 0.$$

□

## 2.2 ABP estimate and Generalized Itô's formula

Below we will use the an analytic result due to Aleksandroff to study the Itô process given by (2.1). For simplicity, *in this section, we assume that  $b = 0$ .*

**Proposition 2.2.1** (Aleksandroff). *Let  $f$  be a nonnegative function on  $B_2$  such that  $f^d$  has finite integral over  $B_2$  and  $f = 0$  outside  $B_2$ . Then there exists a **nonpositive convex** function  $u$  on  $B_2$  such that*

(i) *for any  $x \in B_2$ ,*

$$|u(x)| \leq C \left( \int_{B_2} f^d dx \right)^{\frac{1}{d}}; \quad (2.4)$$

(ii) *for any constant  $a \in \mathbb{S}_+^d$ ,  $\varepsilon > 0$  and  $x \in B_2$ ,*

$$a_{ij} \partial_{ij} u_\varepsilon(x) \geq d \sqrt[d]{\det a} f_\varepsilon(x), \quad (2.5)$$

where  $v_\varepsilon = v * \zeta_\varepsilon$  and  $\zeta_\varepsilon$  is a standard mollifier.

(2.4) is called Alexandroff–Bakelman–Pucci estimate in PDE literature.

In Appendix C.1, we provide the proof for Proposition 2.2.1 based on the very initial knowledge of the solvability of the following Monge–Ampère equations and estimates of its solutions:

$$\det \nabla^2 u(x) = f \quad \text{in } D, \quad (2.6)$$

which, actually, after a long development became also one of the cornerstones of the theory of fully nonlinear elliptic partial differential equations.

Set

$$\tau_R(x) = \inf \{t > 0 : x + x_t \notin B_R\}.$$

Proposition 2.2.1 implies

**Theorem 2.2.2** (Krylov [Kry09]). *There is a constant  $C(d)$  such that for any  $R > 0$ , and nonnegative Borel  $f$  given on  $\mathbb{R}^d$ , we have*

$$\mathbf{E} \int_0^{\tau_R(x)} f(x + x_t) \sqrt[d]{\det a_t} dt \leq C(d) R \|f\|_{L^d(B_R)}. \quad (2.7)$$

*Proof.* By scaling, we only need to consider the case  $R = 1$ . We can assume  $f \in C_c^\infty(B_1)$ .

By Itô's formula,

$$u_\varepsilon(x + x_{t \wedge \tau_1(x)}) - u_\varepsilon(x) = \int_0^{t \wedge \tau_1(x)} a_s^{ij} \partial_{ij} u_\varepsilon(x + x_s) ds + m_{t \wedge \tau_1(x)}$$

Taking expectation, letting  $t \rightarrow \infty$  and using Proposition 2.2.1, we get

$$\int_0^{\tau_1(x)} \sqrt[d]{\det a_t} f_\varepsilon(x + x_t) dt \leq d^{-1} \int_0^{\tau_1(x)} a_t^{ij} \partial_{ij} u_\varepsilon(x + x_t) dt$$

$$\leq \frac{2}{d} \sup_{x \in B_1} |u(x)| \leq C(d) \|f\|_{L^d(B_2)} = C(d) \|f\|_{L^d(B_1)}.$$

Letting  $\varepsilon \rightarrow 0$ , we obtain our assertion.  $\square$

We should point out that here we do not need to assume  $a \in \mathbb{S}_\delta^d$ .

**Remark 2.2.3.** (i) (2.7) implies that if  $x_t$  is a Itô's process given by (2.1) with  $\sigma$  non-degenerate, then the process  $t \mapsto \int_0^t f(x_s) ds$  is well-defined.

(ii) Suppose  $x_t$  is a Itô process given by (2.1),  $a \in \mathbb{S}_\delta^d$  and  $b$  satisfying  $|b_t| \leq \mathbf{b}(x_t)$  with some  $\mathbf{b} \in L^d$ . In this case, Krylov [Kry21a] also proved (2.7) with  $\|f\|_{L^d(D)}$  replaced by  $\|f\|_{L^{d-\varepsilon}(D)}$  for some  $\varepsilon = \varepsilon(d, \delta, \|\mathbf{b}\|) > 0$ .

Theorem 2.2.2 as many results below admits a natural generalization with conditional expectations. This generalization is obtained by tedious and not informative repeating the proof with obvious changes. We mean the following which we call the conditional version of Theorem 2.2.2. Let  $\gamma$  be a finite stopping time, then

$$\mathbf{E} \left[ \int_\gamma^{\tau_R(x)} f(x + x_t) \sqrt[d]{\det a_t} \mathbf{1}_{\{\gamma \leq \tau_R(x)\}} dt \middle| \mathcal{F}_\gamma \right] \leq C(d) R \|f\|_{L^d(B_R)}. \quad (2.8)$$

**Theorem 2.2.4.** There are constants  $C, \mu$  depending only on  $d$ , such that

$$\mathbf{E} \exp \left( \frac{\mu \tau_R(x)}{\delta R^2} \right) \leq C, \quad \forall R \in (0, \infty) \text{ and } x \in B_R. \quad (2.9)$$

In particular, for each  $\lambda > 0$ ,

$$\mathbf{P}(\tau_R(x) \geq \lambda) \leq C \exp \left( -\frac{\mu \lambda}{\delta R^2} \right). \quad (2.10)$$

**Lemma 2.2.5.**

$$\mathbf{E} \tau_R(x)^n \leq n! (CR^2/\delta)^n.$$

*Proof.* We can assume  $x = 0$ .

We claim that

$$I_n(t) := \mathbf{E}([\tau_R - t]_+^n | \mathcal{F}_t) \leq n! (CR^2/\delta)^n. \quad (2.11)$$

Of course, (2.11) implies our desired result.

When  $n = 1$ , (2.8) implies (2.11). If our assertion is true for a given  $n$ , then

$$\begin{aligned}
 I_{n+1}(t) &= (n+1)! \mathbf{E} \left( \int \mathbf{1}_{t < t_1 < \dots < t_{n+1} < \tau_R} dt_1 \cdots dt_{n+1} \middle| \mathcal{F}_t \right) \\
 &= (n+1)! \int dt_1 \cdots dt_{n+1} \mathbf{E} \left( \mathbf{1}_{t < t_1 < \dots < t_n < \tau_R} \mathbf{1}_{t_n < t_{n+1} < \tau_R} \middle| \mathcal{F}_t \right) \\
 &= (n+1)! \int dt_1 \cdots dt_{n+1} \mathbf{E} \left[ \mathbf{1}_{t < t_1 < \dots < t_n < \tau_R} \mathbf{E} \left( \mathbf{1}_{t_n < t_{n+1} < \tau_R} \middle| \mathcal{F}_{t_n} \right) \middle| \mathcal{F}_t \right] \\
 &= (n+1) \mathbf{E} \left[ n! \int \mathbf{1}_{t < t_1 < \dots < t_n < \tau_R} dt_1 \cdots dt_n \int \mathbf{E} \left( \mathbf{1}_{t_n < t_{n+1} < \tau_R} \middle| \mathcal{F}_{t_n} \right) dt_{n+1} \middle| \mathcal{F}_t \right] \\
 &= (n+1) \mathbf{E} \left\{ [\tau_R - t]_+^n \mathbf{E} \left[ \int_{t_n}^{\tau_R} \mathbf{1}_{B_R}(x_{t_{n+1}}) dt_{n+1} \middle| \mathcal{F}_{t_n} \right] \middle| \mathcal{F}_t \right\} \\
 &\stackrel{(2.8)}{\leq} (n+1) C \delta^{-1} R^2 I_n(t) \stackrel{(2.11)}{\leq} (n+1)! (CR^2/\delta)^{n+1}.
 \end{aligned}$$

So we get what we desired.  $\square$

**Exercise 2.2.1.** Let  $B$  be a one-dimensional BM. Let  $I = (-1, 1)$ . Prove that

$$\mathbf{E} \tau_I^n \leq C^n n!.$$

Using this to give another proof for (2.9).

Corollary 2.2.4 basically says that  $\tau_R$  is smaller than a constant times  $R^2$ . We want to show that in a sense the converse is also true:  $R^2$  is basically smaller than a constant times  $\tau_R$ .

**Lemma 2.2.6.** *There exists  $C$  depending only on  $d$  such that*

$$\mathbf{P}(\tau_R/R^2 \leq t) \leq C \delta^{-1} t, \quad \forall t, R > 0. \quad (2.12)$$

*Proof.* We only need to prove the case  $R = 1$ . Let  $\phi$  be a  $C^2$  function that is zero at 0, one on  $\partial B_1$ , with  $\partial_{ij}\phi$  bounded by a constant. By Itô's formula

$$d\phi(x_t) = \nabla \phi(x_t) \cdot \sigma_t dW_t + a_t^{ij} \partial_{ij} \phi(x_t) dt,$$

which yields that

$$\phi(x_{t \wedge \tau_1}) = \mathbf{E} \int_0^{t \wedge \tau_1} a_s^{ij} \partial_{ij} \phi(x_s) ds \leq C \delta^{-1} t.$$

Since  $\phi(x_{t \wedge \tau_1}) \geq \mathbf{1}_{\{\tau_1 \leq t\}}$ , we get  $\mathbf{P}(\tau_1 \leq t) \leq C \delta^{-1} t$ .  $\square$

**Lemma 2.2.7.** *There is a constant  $\mathfrak{R} = R(d, \delta)$  such that*

$$\mathbf{E} \exp(-\tau_{\mathfrak{R}}) \leq 1/2.$$

*Proof.* Let  $X$  be a non-negative random variable, and let  $F : \mathbb{R}_+ \rightarrow \mathbb{R}$  be a decreasing function with  $F(\infty) = 0$ . Then

$$\mathbf{E}F(X) = - \int_0^\infty F'(t) \mathbf{P}(X \leq t) dt,$$

due to Fubini's theorem. Set  $X = \tau_R$  and  $F(t) = e^{-t}$ . Then

$$\mathbf{E}e^{-\tau_R} = \int_0^\infty e^{-t} \mathbf{P}(\tau_R \leq t) dt \leq \int_0^\infty e^{-t} [1 \wedge (C\delta^{-1}R^{-2}t)] dt \leq C\delta^{-1}R^{-2}.$$

We set  $\mathfrak{R} = \sqrt{2C/\delta}$ . □

**Exercise 2.2.2.** For any  $R \in (0, \infty)$

$$\mathbf{E} \exp \left( - (\mathfrak{R}/R)^2 \tau_R \right) \leq 1/2. \quad (2.13)$$

**Theorem 2.2.8.** For any  $\kappa \in (0, 1)$ ,  $R \in (0, \infty)$ ,  $x \in B_{\kappa R}$ , and  $\lambda \geq 0$ ,

$$\mathbf{E} \exp \left( -\lambda \tau_R(x) \right) \leq 2e^{-\sqrt{\lambda}(1-\kappa)R/K}, \quad (2.14)$$

where  $K = \mathfrak{R}/\log 2$ .

*Proof.* Recall that  $\tau_R(x)$  is the first exit time of  $x + x_t$  from  $B_R$ . Let  $\tau'_R(x)$  be the first exit time of  $x + x_t$  from  $B_{(1-\kappa)R}(x)$ . It follows that in the proof of , we may assume that  $\kappa = 0$  and  $x = 0$ . Then, as usual we may assume that  $R$ . In that case take  $N$ , to be specified later, and introduce  $\tau^k$ ,  $k = 1, \dots, N$ , as the first exit time of  $x_t$  from  $B_{k/N}$ . Also set  $\gamma^k$  be the first exit times of  $x_t$  from  $B_{N-1}(x_{\tau^{k-1}})$  after  $\tau^{k-1}$  ( $\tau^{k-1} \leq \gamma^k \leq \tau^k$ ). obviously,

$$\tau_1 \geq (\gamma^1 - \tau_0) + (\gamma^2 - \tau^1) + \dots + (\gamma^N - \tau^{N-1}).$$

By the conditional version of (2.13),

$$\mathbf{E} \left\{ \exp \left[ -\mathfrak{R}^2 N^2 (\gamma^k - \tau^{k-1}) \right] \middle| \mathcal{F}_{\tau^{k-1}} \right\} \leq 1/2.$$

Therefore,

$$\begin{aligned} & \mathbf{E} \left[ \exp \left( -\mathfrak{R}^2 N^2 \tau_1 \right) \right] \\ & \leq \mathbf{E} \left[ \prod_{k=1}^N \exp \left( -\mathfrak{R}^2 N^2 (\gamma^k - \tau^{k-1}) \right) \right] \\ & \leq \mathbf{E} \left\{ \prod_{k=1}^{N-1} \exp \left( -\mathfrak{R}^2 N^2 (\gamma^k - \tau^{k-1}) \right) \mathbf{E} \left[ \exp \left( -\mathfrak{R}^2 N^2 (\gamma^N - \tau^{N-1}) \right) \middle| \mathcal{F}_{\tau^{N-1}} \right] \right\} \\ & \leq \frac{1}{2} \mathbf{E} \left[ \prod_{k=1}^{N-1} \exp \left( -\mathfrak{R}^2 N^2 (\gamma^k - \tau^{k-1}) \right) \right] \leq \dots \leq (1/2)^N. \end{aligned} \quad (2.15)$$

Choosing  $N = \lceil \sqrt{\lambda}/\mathfrak{R} \rceil$ , we get (2.14). □

**Exercise 2.2.3.** For any  $R, t > 0$ ,

$$\mathbf{P}(\tau_R(x) \leq tR^2) \leq 2 \exp\left(-\frac{\beta(1-\kappa)^2}{t}\right), \quad (2.16)$$

where  $\beta = \beta(\mathfrak{R}) \in (0, 1)$ .

The above estimates for first exit times have many important applications.

**Proposition 2.2.9.** For any  $\kappa \in (0, 1)$  there is a function  $q(\gamma), \gamma \in (0, 1)$ , depending only on  $d, \delta, \kappa$  and naturally, also on  $\gamma$ , such that for any  $R \in (0, \infty), x \in B_{\kappa R}$ , and closed  $\Gamma \subset B_R$  satisfying  $|\Gamma| \geq \gamma |B_R|$  we have

$$\mathbf{P}(\sigma_\Gamma(x) \leq \tau_R(x)) \geq q(\gamma),$$

where  $\sigma_\Gamma(x)$  is the first time the process  $x + x_t$  hits  $\Gamma$ . Furthermore,  $q(\gamma) \rightarrow 1$  as  $\gamma \uparrow 1$ .

*Proof.* By using scaling we reduce the general case to the one in which  $R = 1$ . In that case for any  $\varepsilon > 0$  we have

$$\begin{aligned} \mathbf{P}(\sigma_\Gamma(x) > \tau_1(x)) &\leq \mathbf{P}\left(\tau_1(x) = \int_0^{\tau_1(x)} \mathbf{1}_{B_1 \setminus \Gamma}(x + x_t) dt\right) \\ &\leq \mathbf{P}(\tau_1(x) \leq \varepsilon) + \varepsilon^{-1} \mathbf{E} \int_0^{\tau_1(x)} I_{B_1 \setminus \Gamma}(x + x_t) dt. \end{aligned}$$

In light of Theorem 2.2.2, we can estimate the right-hand side and then obtain

$$\begin{aligned} \mathbf{P}(\sigma_\Gamma(x) > \tau_1(x)) &\leq 2e^{-C/\varepsilon} + C\varepsilon^{-1} |B_1 \setminus \Gamma|^{1/d} \\ &\leq 2e^{-C/\varepsilon} + C\varepsilon^{-1} (1 - \gamma)^{1/d} \end{aligned}$$

where the constants  $C$  depend only on  $d, \delta, \kappa$ . By denoting

$$q(\gamma) = 1 - \inf_{\varepsilon > 0} (2e^{-C/\varepsilon} + C\varepsilon^{-1} (1 - \gamma)^{1/d}).$$

□

Note that in the above result, we have no assumption on the shape of the set  $\Gamma$ .

**Exercise 2.2.4.** For any  $\kappa \in (0, 1), R \in (0, \infty)$ . For any  $x \in B_1$  and  $B_{\kappa R}(y) \subseteq B_R$ , we have

$$\mathbf{P}(\sigma_{B_{\kappa R}(y)}(x) < \tau_R(x)) \geq \zeta(\kappa) > 0,$$

where  $\zeta(\kappa) > 0$  depends only on  $d, \delta$ , and naturally, also on  $\kappa$ .

**Hint:** Using support theorem.

**Theorem 2.2.10.** Let  $p \geq d$ . Then there exists constants  $C$  depending only on  $d, \delta$ , such that for any  $\lambda > 0$  and Borel nonnegative  $f$  given on  $\mathbb{R}^d$  we have

$$\mathbf{E} \int_0^\infty e^{-\lambda t} f(X_t) dt \leq C \lambda^{\frac{d}{2p}-1} \|f\|_p. \quad (2.17)$$

*Proof.* Let  $\gamma$  be a stopping time and  $\gamma'$  be the first exit time of  $x_t$  from  $B_R(x_\gamma)$  after  $\gamma$ . By the conditional version of (2.14),

$$\mathbf{E} \left[ \exp(-\lambda(\gamma' - \gamma)) \middle| \mathcal{F}_\gamma \right] \leq 2e^{-\sqrt{\lambda}R/K}.$$

Choosing  $R = K/\sqrt{\lambda}$ , then

$$\mathbf{E} \left[ \exp(-\lambda(\gamma' - \gamma)) \middle| \mathcal{F}_\gamma \right] \leq 2/e < 1.$$

Let  $\tau^0 = 0$  and  $\tau^k$  be the first exit time of  $x_t$  from  $B_R(x_{\tau^{k-1}})$  after  $\tau^{k-1}$ . As the proof for (2.15), we have

$$\mathbf{E} e^{-\lambda\tau^k} = \mathbf{E} \prod_{i=1}^k e^{-\lambda(\tau^i - \tau^{i-1})} \leq (2/e)^k. \quad (2.18)$$

If (2.18) holds, then

$$\begin{aligned} \mathbf{E} \int_0^\infty e^{-\lambda t} f(x_t) dt &\leq \sum_{k=1}^\infty \mathbf{E} \left[ e^{-\lambda\tau^{k-1}} \mathbf{E} \left( \int_{\tau^{k-1}}^{\tau^k} f(x_t) dt \middle| \mathcal{F}_{\tau^{k-1}} \right) \right] \\ &\stackrel{(2.7)}{\leq} \sum_{k=1}^\infty \mathbf{E} \left( C\delta^{-1} R \|f\|_{L^d(B_R(x_{\tau^{k-1}}))} e^{-\lambda\tau^{k-1}} \right) \\ &\leq C\delta^{-1} R^{2-\frac{d}{p}} \|f\|_p \sum_{k=0}^\infty \mathbf{E} e^{-\lambda\tau^k} \leq C\delta^{-1} \|f\|_p \sum_{k=0}^\infty (2/e)^k \\ &\leq C \|f\|_p / \lambda^{1-(d/2p)}. \end{aligned}$$

□

**Theorem 2.2.11** (Generalized Itô's formula, see Krylov-[Kry09]). *Let  $x_t$  be a Itô process given by (2.1). Suppose that  $a \in \mathbb{S}_\delta^d$  and  $b$  are bounded, then for any  $u \in W_{loc}^{2,p}$  with  $p \geq d$ , we have*

$$u(x_t) - u(x_0) = \int_0^t \nabla u(x_s) \sigma_s dW_s + \int_0^t a_s^{ij} \partial_{ij} u(x_s) ds \quad (2.19)$$

*Proof.* We only need to consider the case  $u \in W^{2,d}$ . Let  $\eta \in C_c^\infty(B_1)$  with  $\int \eta = 1$ . Set  $\eta_\varepsilon(x) = \varepsilon^{-d} \eta(x/\varepsilon)$  and  $u_\varepsilon = u * \eta_\varepsilon$ . By Itô's formula,

$$u_\varepsilon(x_t) - u_\varepsilon(x_0) = \int_0^t \nabla u_\varepsilon(x_s) \sigma_s dW_s + \int_0^t a_s^{ij} \partial_{ij} u_\varepsilon(x_s) ds. \quad (2.20)$$

**Fact:** by Sobolev embedding theorem, we have

$$W^{2,d} \hookrightarrow C_b; \quad \|\nabla u\|_{2d} \leq C(\|\nabla^2 u\|_{L^d} + \|\nabla u\|_{L^d}). \quad (2.21)$$

Since  $u \in C_b$ , by letting  $\varepsilon \rightarrow 0$ , one sees the left-hand side of (2.20) goes to  $u(x_t) - u(x_0)$  as  $\varepsilon \rightarrow 0$ . For the right-hand side of (2.20). By Doob's maximal inequality

$$\begin{aligned}
& \mathbf{E} \sup_{t \in [0, T]} \left| \int_0^t \nabla u_\varepsilon(x_s) \sigma_s dW_s - \int_0^t \nabla u_{\varepsilon'}(x_s) \sigma_s dW_s \right|^2 \\
& \leq C \mathbf{E} \int_0^T |\nabla u_\varepsilon - \nabla u_{\varepsilon'}|^2(x_s) ds \leq C \|\nabla u_\varepsilon - \nabla u_{\varepsilon'}\|_{L^{2d}}^2 \\
& \stackrel{(2.21)}{\leq} C \|u_\varepsilon - u_{\varepsilon'}\|_{W^{2,d}} \rightarrow 0, \quad \varepsilon, \varepsilon' \rightarrow 0.
\end{aligned}$$

Similarly, we can also show that the second integral on the right-hand side of (2.20) also converges to  $\int_0^t a_s^{ij} \partial_{ij} u(x_s) ds$   $\square$

**Remark 2.2.12.** *The above generalized Itô's formula also holds for Itô process given by (2.1), where  $a \in \mathbb{S}_\delta^d$ , and  $b$  satisfying  $|b_t| \leq \mathbf{b}(x_t)$  with  $\mathbf{b} \in L^d$ .*



# Chapter 3

## Itô's stochastic differential equations

### 3.1 Strong solutions

One of the main object in this course is the following SDE:

$$dX_t^i = \sigma_k^i(X_t) dW_t^k + b^i(X_t) dt, \quad X_0 = \xi \in \mathcal{F}_0. \quad (3.1)$$

Given  $(\Omega, \mathcal{F}, \mathbf{P}, \mathcal{F}, W_t)$ , we say (3.1) has a **pathwise solution** if there exists a continuous  $\mathcal{F}_t$ -adapted process  $X_t$  satisfying (3.1). We say that we have **pathwise uniqueness** for (3.1) if whenever  $X_t$  and  $Y_t$  are two solutions, then there exists a set  $\mathcal{N}$  such that  $\mathbf{P}(\mathcal{N}) = 0$  and for all  $\omega \notin \mathcal{N}$ , we have  $X_t = Y_t$  for all  $t$ .

#### 3.1.1 Lipschitz conditions

**Theorem 3.1.1** (Itô). *Suppose  $\sigma$  and  $b$  are Lipschitz. Then there exists a unique pathwise solution to the SDE (3.1) for any  $X_0 \in L^2(\Omega, \mathcal{F}_0, \mathbf{P})$ .*

*Proof.* Let  $B$  denote the set of all continuous processes  $\xi$  that are adapted to the filtration  $\mathcal{F}_t$  and satisfy

$$\|\xi\|_B := \left( \mathbf{E} \sup_{t \in [0, T]} |\xi_t|^2 \right)^{1/2} < \infty.$$

Here  $T$  is a positive number which will be determined later. It is not hard to verify that  $B$  is a Banach space. Define a map  $\mathcal{T}$  on  $B$  by

$$(\mathcal{T}(\xi))_t := X_0 + \int_0^t \sigma(\xi_s) \cdot dW_s + \int_0^t b(\xi_s) ds, \quad t \in [0, T].$$

By (1.13) (or Doob's inequality) and Lipschitz condition on the coefficients,

$$\begin{aligned}
\|\mathcal{T}(\xi) - \mathcal{T}(\eta)\|_B^2 &= \mathbf{E} \sup_{t \in [0, T]} |\mathcal{T}(\xi)_t - \mathcal{T}(\eta)_t|^2 \\
&\leq 2\mathbf{E} \sup_{t \in [0, T]} \left| \int_0^t (\sigma(\xi_s) - \sigma(\eta_s)) dW_s \right|^2 + 2\mathbf{E} \sup_{t \in [0, T]} \left| \int_0^t (b(\xi_s) - b(\eta_s)) ds \right|^2 \\
&\stackrel{(1.13)}{\leq} C\mathbf{E} \int_0^T |\sigma(\xi_s) - \sigma(\eta_s)|^2 ds + C\mathbf{E} \left( \int_0^T |b(\xi_s) - b(\eta_s)| ds \right)^2 \\
&\leq C(T + T^2) \mathbf{E} \sup_{t \in [0, T]} |\xi_t - \eta_t|^2 = C_1(T + T^2) \|\xi - \eta\|_B^2.
\end{aligned}$$

Choosing  $T > 0$  sufficiently small such that  $C_1(T + T^2) \leq 1/2$ , then  $\mathcal{T}$  is a Contraction mapping on  $B$ . Banach fixed-point theorem yields that  $\mathcal{T}$  has a unique fixed point, which is the unique pathwise solution to (3.1). We can extend the same result to arbitrarily time intervals.  $\square$

### 3.1.2 Definitions of solutions

1. **strong solution exists** to (3.1): if given the Brownian motion  $W_t$  there exists a process  $X_t$  satisfying (3.1) such that  $X_t$  is adapted to the filtration generated by  $W_t$ .
2. **weak solution exists** to (3.1): if there exists  $(\Omega, \mathcal{F}, \mathbf{P}, \mathcal{F}; X_t, W_t)$  such that  $W_t$  is a  $\mathcal{F}$ -Brownian motion and the equation (3.1) holds.
3. **weak uniqueness**: if whenever  $(\Omega, \mathcal{F}, \mathbf{P}, \mathcal{F}; X_t, W_t)$  and  $(X, \mathcal{G}, \mathbf{Q}, \mathcal{G}; Y_t, B_t)$  are two weak solutions, then the laws of the processes  $X$  and  $Y$  are equal; **Joint uniqueness in law** means the joint law of  $(X, W)$  and  $(Y, B)$  are equal.

A fundamental result is

**Theorem 3.1.2** (Yamada-Watanabe-Engelbert [Eng91]). *The following two conditions are equivalent.*

- (i) *For every initial distribution, there exists a weak strong solution to (3.1) and the solution to (3.1) is pathwise unique.*
- (ii) *For every initial distribution, there exists a strong strong solution to (3.1) and the solution to (3.1) is jointly unique in law.*

*If one (and therefore both) of these conditions is satisfied then every solution to (3.1) is a strong solution.*

### 3.1.3 SDEs with Hölder drifts

For strong well-posedness, if the diffusion coefficient  $\sigma$  is non-degenerate, then the condition of  $b$  can be weakened.

**Theorem 3.1.3** (Krylov [Kry21b]). *Suppose that  $a \in \mathbb{S}_\delta^d$  and  $\nabla\sigma, b \in L^d(\mathbb{R}^d)$ , then equation (3.1) admits a unique strong solution.*

Of course, we will not to prove such a strong result here, but a simpler one below.

**Theorem 3.1.4** (Zvonkin). *Equation (3.1) admits a unique strong solution, provided that  $a \in \mathbb{S}_\delta^d$  and  $\sigma$  is Lipschitz, and  $b \in C^\alpha(\mathbb{R}^d)$  ( $\forall \alpha > 0$ ).*

Let

$$Lu = a_{ij}\partial_{ij}u + b_i\partial_iu.$$

Consider

$$\lambda u - Lu = f. \quad (3.2)$$

We need the following analytic result.

**Lemma 3.1.5.** *Suppose  $a \in \mathbb{S}_\delta^d$  and  $a, b \in C^\alpha$ . There exists a constant  $\lambda_0 > 0$  such that for any  $\lambda \geq \lambda_0$  and  $f \in C^\alpha$ , equation (3.2) admits a unique solution in  $C^{2,\alpha}$ . Moreover,*

$$\lambda\|u\|_\alpha + \|\nabla^2 u\|_\alpha \leq C\|f\|_\alpha, \quad (3.3)$$

where  $C$  only depends on  $d, \delta$ , and  $\|a\|_\alpha$  and  $\|b\|_\alpha$ .

The proof for the above lemma can be founded in Appendix C.2. Here we give the *Sketch of the proof for Lemma 3.1.5*:

- (i) If  $L = \Delta$  and  $f \in \mathcal{S}(\mathbb{R}^d)$ , then for each  $\lambda > 0$ , one can use Fourier transformation to solve (3.2), i.e.  $u = \mathcal{F}^{-1}[\mathcal{F}(f) \cdot (\lambda + 4\pi^2|\cdot|^2)] \in \cap_{s>0} H^s \subseteq C_b^\infty$ . Moreover, (3.3) can also be proved by Fourier analysis method (see Appendix C.2);
- (ii) For any  $L$  satisfying the conditons in Lemma 3.1.5, and any  $u \in C^{2,\alpha}$ , one can prove that (3.3) holds true for any  $\lambda$  sufficiently large via frozen coefficient method;
- (iii) Let  $\chi$  be a cutoff function and  $\zeta$  be a mollifier. For any  $f \in C^\alpha$ , we set  $f_\varepsilon = \chi_\varepsilon(f * \zeta_\varepsilon)$ . Here  $\chi_\varepsilon(x) = \chi(x/\varepsilon)$  and  $\zeta_\varepsilon(x) = \varepsilon^{-d}\zeta(x/\varepsilon)$ . Using (i), for each  $\varepsilon > 0$ , there is a smooth soluiton, say  $u_\varepsilon$ , to (3.2) with  $L$  and  $f$  replaced by  $\Delta$  and  $f_\varepsilon$ . The limit of  $(u_\varepsilon)$ ,  $u$ , satisfies  $\lambda u - \Delta u = f$ , and  $u$  also satisfies (3.3);
- (iv) In the light of (3.3) and the method of continuity (see lemma below), one can obtain the solvability of (3.2) in  $C^{2,\alpha}$ .

**Lemma 3.1.6** (Method of continuity). *Let  $B$  be a Banach space,  $V$  a normed vector space, and  $T_t$  a norm continuous family of bounded linear operators from  $B$  into  $V$ . Assume that there exists a positive constant  $C$  such that for every  $t \in [0, 1]$  and every  $x \in B$ ,*

$$\|x\|_B \leq C\|T_t x\|_V.$$

*Then  $T_0$  is a surjective if and only if  $T_1$  is surjective as well.*

*Proof of Theorem 3.1.4.* Since  $\sigma$  and  $b$  are bounded continuous, weak solution exists to (3.1) (see [Hua01]). Thanks to Theorem 3.1.2, we only need to prove the pathwise uniqueness.

Let  $\lambda \geq \lambda_0$ . Here  $\lambda_0$  is the same number in Lemma 3.1.5. Consider the following equation

$$\lambda \mathbf{u}_\lambda - L\mathbf{u}_\lambda = b.$$

By Lemma 3.1.5 and interpolation theorem

$$\|\nabla \mathbf{u}\|_\alpha \leq \|\mathbf{u}\|_\alpha^{\frac{1}{2}} \|\nabla^2 \mathbf{u}\|_\alpha^{\frac{1}{2}} \leq C\lambda^{-\frac{1}{2}} \|b\|_\alpha.$$

Choosing  $\lambda$  sufficiently large so that  $C\lambda^{-\frac{1}{2}} < 1/2$ . Set  $\phi(x) = x + \mathbf{u}(x)$ , then  $\phi : \mathbb{R}^d \rightarrow \mathbb{R}^d$  is a  $C^{1,\alpha}$ -homeomorphism.

Assume that  $X$  and  $X'$  are two solutions to (3.1). Set  $Y_t = \phi(X_t)$  and  $Y'_t = \phi(X'_t)$ . Then by Itô's formula,

$$dY_t^i = (\delta_j^i + \partial_j \mathbf{u}^i)(X_t) \sigma_{jk}(X_t) dW_t^k + [a_{jk}(X_t) \partial_{jk} \mathbf{u}^i(X_t) + (\delta_j^i + \partial_j \mathbf{u}^i)(X_t) b^j(X_t)] dt$$

i.e.

$$\begin{aligned} dY_t &= [(I + \nabla \mathbf{u})\sigma] \circ \phi^{-1}(Y_t) dW_t + [a : \nabla^2 \mathbf{u} + (I + \nabla \mathbf{u})b] \circ \phi^{-1}(Y_t) dt \\ &= \underbrace{[(I + \nabla \mathbf{u})\sigma] \circ \phi^{-1}(Y_t)}_{=: \tilde{\sigma}} dW_t + \underbrace{[a : \nabla^2 \mathbf{u} + (I + \nabla \mathbf{u})b] \circ \phi^{-1}(Y_t)}_{=: \tilde{b}} dt. \end{aligned}$$

Similarly,  $dY'_t = \tilde{\sigma}(Y'_t) dW_t + \tilde{b}(Y'_t) dt$ . Since  $\tilde{\sigma}$  and  $\tilde{b}$  are both  $C^{1,\alpha}$  functions, as in the proof for Theorem 3.1.1, we have

$$\mathbf{E}|Y_t - Y'_t|^2 \leq C \int_0^t \mathbf{E}|Y_s - Y'_s|^2 ds.$$

This yields  $Y_t = Y'_t$ , due to Gronwall's inequality. Since  $\phi$  is one-to-one,  $X_t = X'_t$ .  $\square$

### 3.1.4 Stochastic Flow

Consider (3.1) .

**Theorem 3.1.7.** *If  $\sigma$  and  $b$  are Lipschitz, then there exist versions of  $X(t, x)$  that are jointly continuous in  $t$  and  $x$  a.s.*

*Proof.*

$$X(t, x) - X(t, y) = x - y + \int_0^t [\sigma(X(s, x)) - \sigma(X(s, y))] dW_s + \int_0^t [b(X(s, x)) - b(X(s, y))] ds$$

By the Burkholder-Davis-Gundy inequalities, for any  $t \in [0, 1]$ ,

$$\mathbf{E} \sup_{s \in [0, t]} \left| \int_0^s [\sigma(X_r(x)) - \sigma(X_r(y))] dW_r \right|^p$$

$$\begin{aligned} &\leq C \mathbf{E} \left( \int_0^t |X_s(x) - X(s, y)|^2 ds \right)^{p/2} \\ &\leq C \mathbf{E} \int_0^t |X_s(x) - X(s, y)|^p ds. \end{aligned}$$

Set  $g(t) = \mathbf{E} \sup_{s \in [0, t]} |X(s, x) - X(s, y)|^p$ . Then

$$g(t) \leq C|x - y|^p + \int_0^t g(s) ds, \quad t \in [0, 1].$$

Gronwall's inequality yields

$$\mathbf{E} \sup_{t \in [0, 1]} |X(t, x) - X(t, y)|^p \leq C|x - y|^p, \quad \forall p \geq 2.$$

Moreover,

$$\mathbf{E} |X(t, x) - X(s, y)|^p \leq C \left( |x - y| + |t - s|^{\frac{1}{2}} \right)^p, \quad x, y \in \mathbb{R}^d, \quad t, s \in [0, 1], \quad p \geq 2.$$

This together with Lemma 1.2.3 implies that there is a version of continuous version of  $(t, x) \mapsto X(t, x)$  such that

$$\|X(\omega)\|_{C^\alpha([0, 1]; \dot{C}^\beta(B_R))} \leq K(\omega)$$

with  $\alpha \in (0, 1/2)$  and  $\beta \in (0, 1)$ , and  $K \in L^p$  for all  $p \geq 1$ .  $\square$

**Remark 3.1.8.** *The above result also holds if  $\sigma$  and  $b$  are  $\omega$ -dependent and  $\|\sigma\|_{C^1} + \|b\|_{C^1} \leq L$  a.s., for some constant  $L$ .*

The collection of processes  $X(t, x)$  is called a flow. If  $\sigma$  and  $b$  are smoother functions, then  $X(t, x)$  will be smoother in  $x$ . If in we take derivative, and use the chain rule, formally we obtain

$$\begin{aligned} \partial_j X_t^i(x) &= \delta_j^i + \int_0^t \partial_l \sigma_k^i(X(s, x)) \partial_j X^l(s, x) dW_s^k \\ &\quad + \int_0^t \partial_l b^i(X(s, x)) \partial_j X^l(s, x) ds. \end{aligned}$$

To make this more precise, suppose  $\sigma$  and  $b$  are in  $C^2$  and are bounded with bounded first and second derivatives and consider the SDE

$$dJ(t, x) = \partial_l \sigma(X(t, x)) J^l(t, x) dW_t + \nabla b(X(t, x)) J(t, x) dt, \quad Y_0 = I \quad (3.4)$$

Follow the proof of Theorem 3.1.7, we have

**Proposition 3.1.9.** *Assume  $\sigma, b \in C_b^2$ . A pathwise solution to (3.4) exists and is unique. The solution has moments of all orders. If  $J(t, x)$  denotes the solution, versions of  $J(t, x)$  exist that are jointly Hölder continuous in  $t$  and  $x$ , and*

$$\mathbf{E} \sup_{t \in [0, 1]} |J(t, x) - J(t, y)|^p \leq C|x - y|^p, \quad x, y \in \mathbb{R}^d, \quad p \geq 1.$$

**Exercise 3.1.1.** *Prove the above Lemma.*

We now prove the differentiability of  $X(t, x)$ .

**Theorem 3.1.10.** *Suppose  $\sigma, b \in C_b^{k-1, \alpha}$ . Then  $x \mapsto X(t, x)$  is  $C^{k-1, \alpha}$  a.s., and  $\nabla X(t, x) = J(t, x)$ .*

*Proof.* For simplicity we take  $b = 0$  and  $k = 2$ . Set

$$\nabla_j^h X_t^i(x) := \frac{1}{|h|} [X_t^i(x + e_j h) - X_t^i(x)], \quad h \in (-1, 1)$$

and

$$Z^h(t, x) := \nabla^h X(t, x) - J(t, x).$$

Noting that

$$\nabla_j^h X^i(t, x) = \delta_j^i + \underbrace{\int_0^t \left[ \int_0^1 \partial_l \sigma_k^i(\tau X(s, x + e_j h) + (1 - \tau)X(s, x)) d\tau \right]}_{\leq \|\nabla \sigma\|_\infty} \nabla_j^h X^l(s, x) dW_s^k,$$

as we done in the proof of Theorem 3.1.7, it is not hard to prove

$$\sup_{x \in \mathbb{R}^d; h \in (-1, 1)} \mathbf{E} \sup_{t \in [0, 1]} |\nabla^h X(t, x)|^p < \infty, \quad p \geq 1. \quad (3.5)$$

By Taylor expansion,

$$\begin{aligned} & \nabla_j^h X^i(t, x) - \nabla_j^h X^i(t, y) \\ &= \int_0^t \underbrace{\left[ \int_0^1 \partial_l \sigma_k^i(\tau X(s, x + e_j h) + (1 - \tau)X(s, x)) d\tau \right]}_{\leq \|\nabla \sigma\|_\infty} [\nabla_j^h X^l(t, x) - \nabla_j^h X^l(t, y)] dW_s^k \\ &+ \int_0^t \underbrace{\left[ \int_0^1 \partial_l \sigma_k^i(\tau X(s, x + e_j h) + (1 - \tau)X(s, x)) - \partial_l \sigma_k^i(\tau X(s, y + e_j h) + (1 - \tau)X(s, y)) d\tau \right]}_{\leq \|\nabla^2 \sigma\|_\infty (|X(s, x + e_j h) - X(s, y + e_j h)| + |X(s, x) - X(s, y)|)} \nabla_j^h X^l(s, y) dW_s^i. \end{aligned}$$

Let

$$g(t) = \mathbf{E} \sup_{s \in [0, t]} |\nabla^h X(s, x) - \nabla^h X(s, y)|^p.$$

Then BDG and Gronwall's inequality yield

$$\begin{aligned} g(t) &\leq C \int_0^t \mathbf{E} [(|X(s, x + h) - X(s, y + h)| + |X(s, x) - X(s, y)|) |\nabla^h X(s, y)|]^p ds \\ &\leq C \left[ \mathbf{E} \int_0^t (|X(s, x + h) - X(s, y + h)| + |X(s, x) - X(s, y)|)^{2p} ds \right]^{1/2} \\ &\quad \left[ \mathbf{E} \int_0^t |\nabla^h X(s, y)|^{2p} ds \right]^{1/2} \stackrel{(3.5)}{\leq} C |x - y|^p, \quad p \geq 2. \end{aligned}$$

This together with Proposition 3.1.9 implies

$$\mathbf{E} \sup_{t \in [0,1]} |Z^h(t, x) - Z^h(t, y)|^p \leq C|x - y|^p, \quad p \geq 1. \quad (3.6)$$

On the other hand,

$$\begin{aligned} Z_j^{h,i}(t, x) &= \int_0^t \left\{ \nabla_j^h [\sigma_k^i(X(s, x))] - \partial_l \sigma_k^i(X(s, x)) J_j^l(s, x) \right\} dW_s^k \\ &= \int_0^t \left\{ \partial_l \sigma_k^i(X(s, x)) [\nabla_j^h X^l(s, x) - J_j^l(s, x)] + R_{jk}^{h,i}(s, x) \right\} dW_s^k \\ &= \int_0^t \left[ \partial_l \sigma_k^i(X(s, x)) Z_j^{h,l}(s, x) + R_{jk}^{h,i}(s, x) \right] dW_s^k, \end{aligned}$$

where

$$R_{jk}^{h,i}(s, x) = \nabla_j^h [\sigma_k^i(X(s, x))] - \partial_l \sigma_k^i(X(s, x)) \nabla_j^h X^l(s, x),$$

and

$$|R^h(s, x)| \leq |h| \|\nabla^2 \sigma\|_\infty |\nabla^h X(s, x)|^2.$$

Again utilizing BDG inequality and Gronwall's inequality, we see that for each  $p \geq 2$ ,

$$\begin{aligned} \mathbf{E} \sup_{\tau \in [0,t]} |Z^h(\tau, x)|^p &\leq C \mathbf{E} \left\{ \int_0^t \left[ |Z^h(s, x)|^2 + (R^h(s, x))^2 \right] ds \right\}^{p/2} \\ &\leq C \int_0^t \mathbf{E} \sup_{\tau \in [0,s]} |Z^h(\tau, x)|^p ds + C|h|^p \int_0^t \mathbf{E} |\nabla^h X(s, x)|^{2p} ds, \end{aligned}$$

which together with (3.5) implies

$$\mathbf{E} \sup_{t \in [0,1]} |Z^h(t, x)|^p \leq C|h|^p, \quad |h| < 1. \quad (3.7)$$

Combining (3.6) and (3.7), we obtain

$$\mathbf{E} \sup_{t \in [0,1]} |Z^h(t, x) - Z^h(t, y)|^p \leq C|x - y|^{\theta p} |h|^{(1-\theta)p}, \quad \theta \in [0, 1], \quad p \geq 1, \quad |h| < 1.$$

Thanks to Lemma (1.2.3), for any  $\alpha \in (0, 1)$ ,

$$\lim_{|h| \rightarrow 0} \mathbf{E} \sup_{t \in [0,1]} \|Z^h(t, \cdot)\|_{C^\alpha}^p = 0, \quad p \geq 1.$$

Therefore,  $X(t, \cdot) \in C^{1,\alpha}$  a.s. and  $\nabla X(t, x) = J(t, x)$ . □

One can also show (see Ikeda and Watanabe [IW14]) that the map  $x \mapsto X(t, x)$  is one-to-one and onto  $\mathbb{R}^d$ .

## 3.2 Weak solutions

In this section, we study the weak well-posedness of (3.1). The core is to study the regularity of following resolvent equation:

$$\lambda u - Lu = f, \quad (3.8)$$

where  $L = a_{ij}\partial_{ij}$  and  $a \in \mathbb{S}_\delta^d$  and uniformly continuous.

### 3.2.1 Uniqueness in law

**Theorem 3.2.1** (Stroock-Varadhan). *Under the assumptions that  $\sigma, b$  are bounded,  $\sigma$  is continuous and  $\sigma(x)\sigma^t(x) > 0$  for each  $x \in \mathbb{R}^d$ . Then SDE (3.1) has a weak solution, and the distribution of such solution is unique.*

Our strategy is

- (a) Utilizing Girsanov transformation to simplify the problem to the case without drift term;
- (b) Using generalized Itô's formal and  $L^p$ -estimate for the resolvent equation to show the uniqueness of law( $X_t$ ).
- (c) Proving the law of  $(X_t)_{t \geq 0}$  is unique by induction.

**Lemma 3.2.2.** *Let  $L = \Delta$  and  $p \in (1, \infty)$ . For any  $f \in L^p$ , there exists a unique solution  $u \in W^{2,p}$  solving (3.8). Moreover,  $u$  satisfies*

$$\lambda \|u\|_p + \|\nabla u^2\|_p \leq C \|f\|_p, \quad (3.9)$$

where  $C$  only depends on  $d$  and  $p$ .

**Theorem 3.2.3.** *Let  $p \in (1, \infty)$ . There exists a constant  $\lambda_0 = \lambda_0(d, p, \omega_a) > 0$  such that for any  $\lambda \geq \lambda_0$  and  $f \in L^p$ , equation (3.8) admits a unique solution  $u \in W^{2,p}$ .*

*Proof.* Assume that  $u \in W^{2,p}$ . We want to show that for sufficiently large  $\lambda$ , it holds that

$$\lambda \|u\|_p + \|u\|_{W^{2,p}} \leq C \|\lambda u - Lu\|_p. \quad (3.10)$$

Suppose we have (3.10). Let  $T_0 = \lambda - \Delta$  and  $T_1 = \lambda - L$ , and  $B = W^{2,p}$  and  $V = L^p$ . Utilizing Lemma 3.1.6 and Lemma 3.2.2, we can see that (3.8) has a solution in  $W^{2,p}$ .

Now let us prove (3.10). Let  $f := \lambda u - Lu$ . Assume  $\zeta \in C_c^\infty(B_2)$  such that  $\zeta \geq 0$ ,  $\zeta \equiv 1$  in  $B_1$ . Set  $\zeta_\varepsilon^z = \zeta((x - z)/\varepsilon)$ . Then

$$\lambda(u\zeta_\varepsilon^z) - a_{ij}(z)\partial_{ij}(u\zeta_\varepsilon^z) = f\zeta_\varepsilon^z - 2a_{ij}\partial_i u \partial_j \zeta_\varepsilon^z - a_{ij}\partial_{ij}\zeta_\varepsilon^z u + (a_{ij} - a_{ij}(z))\partial_{ij}(u\zeta_\varepsilon^z).$$

By Lemma 3.2.2, we get

$$\begin{aligned} \lambda \|u\zeta_\varepsilon^z\|_p + \|\nabla^2(u\zeta_\varepsilon^z)\|_p &\leq C\omega_a(2\varepsilon)\|\nabla^2(u\zeta_\varepsilon^z)\|_p + C\|f\|_{L^p(B_{2\varepsilon}(z))} \\ &\quad + C\varepsilon^{-1}\|\nabla u\|_{L^p(B_{2\varepsilon}(z))} + C\varepsilon^{-2}\|u\|_{L^p(B_{2\varepsilon}(z))}. \end{aligned}$$



Choosing  $\varepsilon_0 > 0$  sufficiently small such that  $C\omega_a(2\varepsilon_0) \leq 1/2$ , then

$$\begin{aligned} & \lambda \|u\|_{L^p(B_{\varepsilon_0}(z))} + \|\nabla^2 u\|_{L^p(B_{\varepsilon_0}(z))} \\ & \leq C \|f\|_{L^p(B_{2\varepsilon_0}(z))} + C\varepsilon_0^{-1} \|\nabla u\|_{L^p(B_{2\varepsilon_0}(z))} + C\varepsilon_0^{-2} \|u\|_{L^p(B_{2\varepsilon_0}(z))}. \end{aligned} \quad (3.11)$$

**Fact:** There exist constants  $c = c(d, p, \varepsilon) > 0$  and  $C = C(d, p, \varepsilon) > 0$ , and a sequence  $\{z_i\}_{i \in \mathbb{N}} \subseteq \mathbb{R}^d$  such that

$$c \sum_i \int |h \zeta_\varepsilon^{z_i}|^p \leq \int |h|^p \leq C \sum_i \int |h \zeta_\varepsilon^{z_i}|^p. \quad (3.12)$$

By (3.11) and (3.12), we obtain

$$\lambda \|u\|_p^p + \|\nabla^2 u\|_p^p \leq C \|f\|_p^p + C \|\nabla u\|_p^p + C \|u\|_p^p,$$

where  $C$  only depends on  $d, p$  and  $\omega_a$ . Using interpolation theorem, one can see that

$$\lambda \|u\|_p + \|\nabla u\|_p + \|\nabla^2 u\|_p^p \leq \frac{1}{2} \|\nabla^2 u\|_p + \frac{\lambda_0}{2} \|u\|_p + C \|f\|_p,$$

where  $\lambda_0 \geq 1$  is a constant only depends on  $d, p$  and  $\omega_a$ . Therefore, for any  $\lambda \geq \lambda_0 \geq 1$ , we have

$$\lambda \|u\|_p + \|u\|_{W^{2,p}} \leq C \|f\|_p.$$

□

Now let  $f \in C_c^\infty(\mathbb{R}^d)$ . Assume that  $u \in W^{2,d}$  is a solution to (3.8) for some  $\lambda \geq \lambda_0$ . Applying Generalized Itô's formula, one can see that

$$d(e^{-\lambda t} u(X_{s+t})) = e^{-\lambda t} [-\lambda u(X_{s+t}) + Lu(X_{s+t})] + e^{-\lambda t} \nabla u(X_{s+t}) \sigma(X_{s+t}) dW_{s+t}.$$

Taking expectation conditional on  $\mathcal{F}_s$ , we get

$$u(X_s) = \mathbf{E}(u(X_s) | \mathcal{F}_s) = \int_0^\infty e^{-\lambda t} \mathbf{E}(f(X_{s+t}) | \mathcal{F}_s) dt, \quad \forall \lambda \gg 1.$$

This implies that  $\mathbf{P}(X_{s+t} \in \cdot | \mathcal{F}_s)$  is unique and  $\mathbf{P}(X_{s+t} \in \cdot | \mathcal{F}_s) = \mathbf{P}(X_{s+t} \in \cdot | X_s)$ . Using this fact, then the uniqueness in law of  $X_t$  can be obtained by induction.

### 3.2.2 Markov properties

Define  $\mathcal{W}$  to be the set of all continuous functions from  $\mathbb{R}_+$  to  $\mathbb{R}^d$ . Suppose that for each starting point  $x$  the SDE (3.1) has a solution that is unique in law. Let us denote the solution by  $X(x, t, \omega)$ . For each  $x$  define a probability measure  $\mathbb{P}_x$  on  $\mathcal{W}$  so that

$$\begin{aligned} & \mathbf{P}(X(x, t_1) \in A_1, \dots, X(x, t_n) \in A_n) \\ & = \mathbb{P}_x(\omega(t_1) \in A_1, \dots, \omega(t_n) \in A_n). \end{aligned}$$

Let  $\mathcal{G}_t^0$  be the  $\sigma$ -algebra generated by  $\{\omega_s : s \leq t\}$ . We complete these  $\sigma$ -fields by considering all sets that are in the  $\mathbb{P}_x$  completion of  $\mathcal{G}_t^0$  for all  $x$ . Finally, we obtain a right continuous filtration by letting  $\mathcal{G}_t := \cap_{\varepsilon > 0} \mathcal{G}_{t+\varepsilon}^0$ . We then extend  $\mathbb{P}_x$  to  $\mathcal{G}_\infty$ .

Shift operators  $\theta_t : \mathcal{W} \ni \omega \mapsto \omega(t + \cdot) \in \mathcal{W}$ .

The **strong Markov property** is the assertion that

$$\mathbb{E}_x(Y \circ \theta_\tau | \mathcal{G}_\tau)(\omega) = \mathbb{E}_{Z_\tau}(Y),$$

whenever  $x \in \mathbb{R}^d$ ,  $Y \in \mathcal{G}_\infty$  is bounded, and  $\tau$  a finite stopping time.

To prove the strong Markov property it suffices to show

$$\mathbb{E}_x(f(Z_{\tau+t}) | \mathcal{G}_\tau) = \mathbb{E}_{Z_\tau} f(Z_t), \quad (3.13)$$

for all  $x \in \mathbb{R}^d$ ,  $f \in C_c(\mathbb{R}^d)$  and  $\tau$  a bounded stopping time.

**Theorem 3.2.4.** *Suppose the solution to (3.1) is weakly unique for each  $x$ . Then  $(\mathbb{P}_x, X_t)$  is a strong Markov process.*

*Proof.* We have the equation

$$Z_t = Z_0 + \int_0^t \sigma(Z_s) dB_s + \int_0^t b(Z_s) ds,$$

where  $B$  is a Brownian motion, not necessarily the same as the one in (3.1). Set  $Z'_t = Z_{\tau+t}$  and  $B'_t = B_{\tau+t} - B_\tau$ . Then

$$Z'_t = Z'_0 + \int_0^t \sigma(Z'_s) dB'_s + \int_0^t b(Z'_s) ds. \quad (3.14)$$

Let  $Q_\tau(\omega, \omega')$  be the regular conditional distribution for  $\mathbf{E}(\cdot | \mathcal{F}_\tau)$ .

**Claim:**

- $W'$  is a Brownian motion with respect to the measure  $Q_\tau(\omega, \cdot)$  for almost every  $\omega$ ;
- $Z'_0(\omega') = Z_{\tau(\omega)}(\omega)$ ,  $Q_\tau(\omega, \cdot)$ -a.s..

The uniqueness in law tells us that

$$\mathbf{E}^{Q_\tau} f(Z'_t) = \mathbb{E}_{Z_\tau} f(Z_t), \quad \mathbb{P}_x - a.s..$$

On the other hand, by definition

$$\mathbf{E}^{Q_\tau} f(Z'_t) = \mathbf{E}^{Q_\tau} f(Z_{\tau+t}) = \mathbb{E}_x(f(Z_{\tau+t}) | \mathcal{G}_\tau)$$

Thus, we get (3.13).

Our task now is to prove the claim. We need the following

**Fact:** If  $B_t$  is a  $\mathcal{G}_t$ -Brownian motion and  $\tau$  is a finite stopping time, then  $B_{\tau+t} - B_\tau$  is a  $\mathcal{G}_{\tau+t}$ -Brownian motion.

Using the above fact, we have

$$\begin{aligned}
& \mathbf{E}^{Q_\tau} \exp \left( i \sum_{k=1}^{n-1} \lambda_k \cdot (B_{\tau+t_{k+1}} - B_{\tau+t_k}) \right) \\
&= \mathbb{E}_x \left[ \exp \left( i \sum_{k=1}^{n-1} \lambda_k \cdot (B_{\tau+t_{k+1}} - B_{\tau+t_k}) \right) \middle| \mathcal{G}_\tau \right] \\
&= \exp \left( \sum_{k=1}^{n-1} \lambda_k^2 (t_{k+1} - t_k) / 2 \right),
\end{aligned}$$

we get what we claimed. □

From now on, by a slight abuse of notation, we will say  $(\mathbb{P}_x, X_t)$  is a strong Markov family when  $(\mathbb{P}_x, Z_t)$  is a strong Markov family.



# Chapter 4

## Applications to Elliptic PDEs

Let  $X_t$  be the solution to (3.1) with  $X_0 = x$ . We will write  $(\mathbb{P}_x, X_t)$  for the strong Markov process corresponding to  $\sigma$  and  $b$  (This can be ensured by assuming  $\sigma, b \in C_b^1$ , or  $a \in \mathbb{S}_\delta$ ,  $a$  is continuous and  $b$  is bounded).

We have considered Poisson's equation in Chapter 1. Let  $u$  be a  $C_b^2$  solution to (3.2). Then by Theorem 1.5.11,

$$u(x) = \mathbb{E}_x \int_0^\infty e^{-\lambda t} f(X_t) dt.$$

We have also studied the Poisson's equation in a nice bounded domain:

$$\begin{cases} \lambda u - Lu = f & \text{in } D \\ u = 0 & \text{on } \partial D, \end{cases} \quad (4.1)$$

and showed that

$$u(x) = \mathbb{E}_x \int_0^{\tau_D} e^{-\lambda s} f(X_s) ds,$$

if  $\mathbb{P}_x(\tau_D < \infty) = 1$ .

### 4.1 Dirichlet Problems and Harmonic functions

Let  $D$  be a ball (or other nice bounded domain) and let us consider the solution to the Dirichlet problem: given  $g$  a continuous function on  $\partial D$ , find  $u \in C(\bar{D})$  such that  $u$  is  $C^2$  in  $D$  and

$$\begin{cases} Lu = 0 & \text{in } D \\ u = g & \text{on } \partial D. \end{cases} \quad (4.2)$$

**Theorem 4.1.1.** *The solution to (4.2) satisfies*

$$u(x) = \mathbb{E}_x g(X_{\tau_D}),$$

*provided that  $\mathbb{P}_x(\tau_D < \infty) = 1$ .*

*Proof.* Let  $\tau_n = \inf \{t : \text{dist}(X_t, \partial D) < 1/n\}$ . By Itô's formula,

$$u(X_{t \wedge \tau_n}) = u(X_0) + \text{martingale} + \int_0^{t \wedge \tau_n} Lu(X_s) ds.$$

Since  $Lu = 0$  inside  $D$ , taking expectations shows

$$u(x) = \mathbb{E}_x u(X_{t \wedge \tau_n}).$$

We let  $t \rightarrow \infty$  and then  $n \rightarrow \infty$ . By dominated convergence, we obtain  $u(x) = \mathbb{E}_x u(X_{\tau_D})$ . This is what we want since  $u = g$  on  $\partial D$ .  $\square$

**Exercise 4.1.1.** Theorem 4.1.1 implies the weak maximum principle:  $\max_D u \leq \max_{\partial D} u$ .

**Exercise 4.1.2.** Theorem 2.1 implies the strong maximum principle: if  $u$  is not a constant function, then for each  $x \in D$ ,  $u(x) < \max_{\partial D} u$ .

**Theorem 4.1.2.** Let  $g$  be continuous on  $\partial D$ . Suppose that  $\mathbb{P}_x(\tau_D < \infty) = 1$  and  $u(x) = \mathbb{E}_x g(X_{\tau_D})$  is continuous on  $\bar{D}$  and  $C^2$  on  $D$ . Suppose the coefficients of  $L$  are continuous. Then  $Lu = 0$  in  $D$ .

*Proof.* Let  $B_r(x) \subseteq D$ . By the strong Markov property, we have

$$\begin{aligned} u(x) &= \mathbb{E}_x g(X_{\tau_D}) = \mathbb{E}_x g(X_{\tau_D} \circ \theta_{\tau_{B_r(x)}}) = \mathbb{E}_x \left\{ \mathbb{E}_x \left[ g(X_{\tau_D} \circ \theta_{\tau_{B_r(x)}}) \middle| \mathcal{F}_{\tau_{B_r(x)}} \right] \right\} \\ &= \mathbb{E}_x \left[ \mathbb{E}_{X_{\tau_{B_r(x)}}} g(X_{\tau_D}) \right] = \mathbb{E}_x u(X_{\tau_{B_r(x)}}). \end{aligned}$$

Noting that  $u \in C^2(D)$ , by Itô's formula,

$$u(X_{t \wedge \tau_{B_r(x)}}) - u(x) = \int_0^{t \wedge \tau_{B_r(x)}} Lu(X_s) ds + M_{t \wedge \tau_{B_r(x)}},$$

where  $M$  is a martingale. Taking expectations and letting  $t \rightarrow \infty$ ,

$$0 = \frac{1}{\mathbb{E}_x \tau_{B_r(x)}} \mathbb{E}_x \int_0^{\tau_{B_r(x)}} Lu(X_s) ds.$$

By the continuity of  $Lu$  and letting  $r \rightarrow 0$ , we get  $Lu(x) = 0$ .  $\square$

If  $Lu = 0$  in  $D$ , we say  $u$  is  $L$ -harmonic in  $D$ .

One can also the following Schrödinger type operator:

$$L_q u = Lu + qu.$$

Equation involving the above operator are considerably simpler than the quantum mechanics Schrödinger equation because here all terms are real-valued.

---

**Theorem 4.1.3.** *Let  $D$  be a nice bounded domain, and  $q \in C(\bar{D})$  and  $g \in C(\partial D)$ . Let  $u \in C^2(D) \cap C(\bar{D})$ . that agrees with  $g$  on  $\partial D$  and satisfies  $L_q u = 0$  in  $D$ . If*

$$\mathbb{E}_x \exp \left( \int_0^{\tau_D} q^+(X_s) ds \right) < \infty,$$

then

$$u(x) = \mathbb{E}_x \left[ g(X_{\tau_D}) \exp \left( \int_0^{\tau_D} q(X_s) ds \right) \right].$$

**Exercise 4.1.3.** *Prove Theorem 4.1.3.*

**Exercise 4.1.4.** *Using (2.7) to show: there exists  $\varepsilon > 0$  such that if  $B \subseteq Q_1, x \in Q_{1/2}$ , and  $|Q_1 - B| < \varepsilon$ , then*

$$\mathbb{E}_x \int_0^{\tau_{Q_1}} 1_B(X_s) ds \geq c > 0,$$

where  $c$  is a constant only depends on  $d, \delta$  and  $\varepsilon$ .

## 4.2 Once again on the hitting probability

Recall that

$$x_t = \int_0^t \sigma_s dW_s, \quad a = \frac{1}{2} \sigma \sigma^t \in \mathbb{S}_\delta^d.$$

In this section, we want to prove following important hitting probability estimate, which is a refined version of Proposition 2.2.9. This was first found by Krylov-Safonov [KS79].

Recall that

$$\sigma_\Gamma(x) = \inf \{t > 0 : x + x_t \in \Gamma\} \quad \text{and} \quad \tau_Q = \inf \{t > 0 : x + x_t \notin Q\}.$$

**Theorem 4.2.1.** *There is a increasing function  $p : (0, 1) \rightarrow (0, 1)$ , which only depends on  $d$  and  $\delta$ , such that for any  $\Gamma \subset Q_1$  and  $x \in Q_{1/2}$ ,*

$$\mathbf{P}(\sigma_\Gamma(x) < \tau_{Q_1}(x)) \geq p(|\Gamma|). \quad (4.3)$$

Before prove Theorem 4.2.1, we need some preparation.

One tool is a corollary of the Calderón-Zygmund cube decomposition. Let  $Q_1$  be the unit cube. We split it into  $2^n$  cubes of half side. We do the same splitting with each one of these  $2^n$  cubes and we iterate this process. The cubes obtained in this way are called dyadic cubes.

If  $Q$  is a dyadic cube different from  $Q_1$ , we say that  $\tilde{Q}$  is the predecessor of  $Q$  if  $Q$  is one of the  $2^n$  cubes obtained from dividing  $\tilde{Q}$ .

We also let  $Q(\kappa)$  denote the cube with the same center as  $Q$  but side length  $\kappa$  times as long.

**Lemma 4.2.2** (Krylov-Safonov [KS79]). *Let  $\gamma \in (0, 1)$ . If  $\Gamma \subset Q_1$  and  $|\Gamma| \leq \gamma$ , then there exists a sequence of dyadic cubes, say  $\{Q^i\}_{i \in \mathbb{I}}$  such that*

1. the interiors of the  $Q^i$  are pairwise disjoint;
2.  $|\Gamma \cap Q^i| > \gamma|Q^i|$  and  $|\Gamma \cap \tilde{Q}^i| \leq \gamma|\tilde{Q}^i|$ , for each  $i \in \mathcal{I}$ ;
3.  $|\Gamma| \leq \gamma|E|$  and  $|\Gamma \setminus E| = 0$ , where  $E = \cup_{i \in \mathcal{I}} \tilde{Q}^i$ .

*Proof.* We use the Calderón-Zygmund decomposition. We have that

$$\frac{|Q_1 \cap \Gamma|}{|Q_1|} = |\Gamma| \leq \gamma.$$

We subdivide  $Q_1$  into  $2^n$  dyadic cubes. If  $Q$  is one of these  $2^n$  subcubes of  $Q_1$  and satisfies  $|Q \cap \Gamma|/|Q| \leq \gamma$ , we then split  $Q$  into  $2^n$  dyadic cubes. We iterate this process. In this way, we pick a family  $Q^1 \cdot Q^2 \dots$  of dyadic cubes (different from  $Q_1$ ) such that

$$\frac{|Q^i \cap \Gamma|}{|Q^i|} > \gamma, \quad \forall i \in \mathcal{I}$$

If  $x \notin \cup_{i \in \mathcal{I}} Q^i$  then  $x$  belongs to an infinite number of closed dyadic cubes  $Q$  with diameters tending to zero, such that  $|Q \cap \Gamma|/|Q| \leq \gamma < 1$ . Applying the Lebesgue differentiation theorem to  $\mathbf{1}_A$ , we get that  $\mathbf{1}_A(x) \leq \gamma < 1$  for a.e.  $x \notin \cup_{i \in \mathcal{I}} Q^i$ . Hence  $A \subset \cup_{i \in \mathcal{I}} Q^i$ , except for a set of measure zero.

Consider the family of predecessors of the cubes  $Q^i$ , and relabel them so that  $\{\tilde{Q}^i\}_{i \in \tilde{\mathcal{I}}}$  are pairwise disjoint. We clearly have that

$$\Gamma \subseteq \cup_{i \in \mathcal{I}} Q^i \subseteq \cup_{i \in \tilde{\mathcal{I}}} \tilde{Q}^i =: E,$$

except for a set of measure zero. From the way we chose the cubes  $Q^i$ ,

$$\frac{|\tilde{Q}^i \cap \Gamma|}{|\tilde{Q}^i|} \leq \gamma, \quad \forall i \in \tilde{\mathcal{I}}.$$

We conclude that

$$|\Gamma| \leq \sum_{i \in \tilde{\mathcal{I}}} |\tilde{Q}^i \cap \Gamma| \leq \gamma \sum_{i \in \tilde{\mathcal{I}}} |\tilde{Q}^i| = \gamma \left| \bigcup_{i \in \tilde{\mathcal{I}}} \tilde{Q}^i \right| \leq \gamma|E|,$$

that finishes the proof of Lemma 4.2.2. □

The second tool is support theorem, which implies

**Lemma 4.2.3.** *Let  $\kappa \in (3/4, 1)$ . Suppose that  $\tilde{Q}$  is the predecessor of  $Q$ , then for each  $x \in \tilde{Q}(\kappa)$ ,*

$$\mathbf{P} \left( \sigma_{Q(\frac{1}{2})}(x) < \tau_{\tilde{Q}}(x) \right) \geq p'(\kappa) > 0,$$

where  $p'(\kappa)$  only depends on  $d, \delta$  and  $\kappa$ .



---

*Proof of Theorem 4.2.1.* Define

$$p(\gamma) = \inf \left\{ \mathbf{P}(\sigma_\Gamma(x) < \tau_{Q_1}(x)) : a \in \mathbb{S}_\delta^d, x \in Q_{1/2}, \Gamma \subset Q_1, |\Gamma| \geq \gamma \right\}.$$

By Proposition 2.2.9, we know that there exists a constant  $b \in (0, 1)$  such that  $p(b) > 0$ .

We want to prove that for each  $\gamma \in (0, b]$ ,

$$p(\gamma) > 0 \text{ implies } p(\theta\gamma) > 0, \text{ where } \theta = \frac{1+b}{2} < 1.$$

Assume that  $p(\gamma) > 0$  for some  $\gamma \in (0, b]$ , and  $\Gamma \subseteq Q_1$  with  $|\Gamma| \geq \theta\gamma$ . Let  $Q_i$  and  $E = \cup_{i \in \mathcal{I}} \tilde{Q}^i$  be the sets in Lemma 4.2.2. Then

$$|E| \geq |\Gamma|/\gamma \geq \theta = \frac{1+b}{2}.$$

Therefore, we can find a finite subset of  $\mathcal{I}$ , say  $\mathcal{I}_0$ , and  $\kappa \in (3/4, 1)$  such that

$$A := \bigcup_{i \in \mathcal{I}_0} \tilde{Q}^i(\kappa) \text{ with } |A| \geq b.$$

Since  $|A| \geq b|Q_1|$ , by Proposition 2.2.9,

$$\mathbf{P}(\sigma_A(x) < \tau_{Q_1}(x)) \geq p(b) > 0, \quad \forall x \in Q_{1/2}. \quad (4.4)$$

Suppose that  $y \in \partial A = \cup_{i \in \mathcal{I}_0} \partial \tilde{Q}^i(\kappa)$ , then  $y \in \partial \tilde{Q}^i(\kappa)$  for some  $i \in \mathcal{I}_0$ . In this case,

$$\mathbf{P}(\sigma_{Q^i(1/2)}(y) < \tau_{Q_1}(y)) \geq \mathbf{P}(\sigma_{Q^i(1/2)}(y) < \tau_{\tilde{Q}^i}(y)) \geq p'(\kappa) > 0,$$

due to Lemma 4.2.3. Set

$$B = \bigcup_{i \in \mathcal{I}_0} Q^i(1/2).$$

Then

$$\begin{aligned} \mathbf{P}(\sigma_B(y) < \tau_{Q_1}(y)) &\geq \inf_{y \in \partial A} \mathbf{P}(\sigma_{Q^i(1/2)}(y) < \tau_{\tilde{Q}^i}(y)) \\ &\geq p'(\kappa) > 0, \quad \forall y \in \partial A. \end{aligned}$$

The conditional version of above estimate we need below is

$$\mathbf{P}(\sigma'_B < \tau'_{Q_1} | \mathcal{F}_{\sigma_A(x)}) \geq p'(\kappa) > 0. \quad (4.5)$$

where

$$\sigma'_B := \inf \{t > \sigma_A(x) : x + x_t \in B\} \quad \text{and} \quad \tau'_{Q_1} := \inf \{t > \sigma_A(x) : x + x_t \notin Q_1\}.$$

Suppose that  $z \in \partial B$ , then  $z \in \partial Q^i(1/2)$  for some  $i \in \mathcal{I}_0$ . Since  $|\Gamma \cap Q_i| > \gamma|Q_i|$ , by our assumption

$$\mathbf{P}(\sigma_{\Gamma \cap Q^i}(z) < \tau_{Q_1}(y)) \geq \mathbf{P}(\sigma_{\Gamma \cap Q^i}(z) < \tau_{Q^i}(z)) \geq p(\gamma) > 0, \quad \forall z \in \partial B.$$

Set

$$D = \bigcup_{i \in \mathcal{I}_0} Q_i.$$

Then

$$\begin{aligned} \mathbf{P}(\sigma_\Gamma(z) < \tau_{Q_1}(z)) &\geq \mathbf{P}(\sigma_{\Gamma \cap D}(z) < \tau_{Q_1}(z)) \\ &\geq \inf_{i \in \mathcal{I}_0} \mathbf{P}(\sigma_{\Gamma \cap Q_i}(z) < \tau_{Q_i}(z)) \geq p(\gamma) > 0, \quad \forall z \in \partial B. \end{aligned}$$

The conditional version of above estimate we need below is

$$\mathbf{P}(\sigma''_\Gamma < \tau''_{Q_1} | \mathcal{F}_{\sigma'_B}) \geq p(\gamma) > 0. \quad (4.6)$$

where

$$\sigma''_\Gamma := \inf \{t > \sigma'_B : x + x_t \in \Gamma\} \quad \text{and} \quad \tau''_{Q_1} := \inf \{t > \sigma'_B : x + x_t \notin Q_1\}$$

Therefore, for each  $x \in Q_{1/2}$ ,

$$\begin{aligned} &\mathbf{P}(\sigma_\Gamma(x) < \tau_{Q_1}(x)) \\ &\geq \mathbf{P}(\sigma_A(x) < \tau_{Q_1}(x); \sigma'_\Gamma < \tau'_{Q_1} | \mathcal{F}_{\sigma_A(x)}) \\ &= \mathbf{E} \left[ \mathbf{1}_{\sigma_A(x) < \tau_{Q_1}(x)} \mathbf{P}(\sigma'_\Gamma < \tau'_{Q_1} | \mathcal{F}_{\sigma_A(x)}) \right] \\ &= \mathbf{E} \left[ \mathbf{1}_{\sigma_A(x) < \tau_{Q_1}(x)} \mathbf{P}(\sigma'_B < \tau'_{Q_1}; \sigma''_\Gamma < \tau''_{Q_1} | \mathcal{F}_{\sigma_A(x)}) \right] \\ &= \mathbf{E} \left\{ \mathbf{1}_{\sigma_A(x) < \tau_{Q_1}(x)} \mathbf{E} \left[ \mathbf{1}_{\sigma'_B < \tau'_{Q_1}} \mathbf{P}(\sigma''_\Gamma < \tau''_{Q_1} | \mathcal{F}_{\sigma'_B}) | \mathcal{F}_{\sigma_A(x)} \right] \right\} \\ &\stackrel{(4.6)}{\geq} p(\gamma) \mathbf{E} \left[ \mathbf{1}_{\sigma_A(x) < \tau_{Q_1}(x)} \mathbf{P}(\sigma'_B < \tau'_{Q_1} | \mathcal{F}_{\sigma_A(x)}) \right] \\ &\stackrel{(4.5)}{\geq} p(\gamma) p'(\kappa) \mathbf{P}(\sigma_A(x) < \tau_{Q_1}(x)) \\ &\stackrel{(4.4)}{\geq} p(b) p(\gamma) p'(\kappa) \geq p'(\kappa) p^2(\gamma) > 0. \end{aligned}$$

Since the above estimate holds for any  $\Gamma \subseteq Q_1$  with  $|\Gamma| \geq \theta\gamma$ , we get  $p(\theta\gamma) > 0$ , provided that  $p(\gamma) > 0$ . Noting that  $\theta < 1$ , we obtain that  $p(\gamma) > 0$  for all  $\gamma \in (0, 1)$ .  $\square$

## 4.3 Harnack Inequality and Hölder estimate

In this section, we prove some theorems of Krylov and Safonov [KS81] concerning (positive)  $L$ -harmonic functions. Let  $\delta \in (0, 1)$ . Set

$$\begin{aligned} \mathcal{P}(\delta) := \left\{ \{ \mathbb{P}_x \}_{x \in \mathbb{R}^n} : (\mathbb{P}_x, X) \text{ is the strong Markov process} \right. \\ \left. \text{associate with some } a(\cdot) \in \mathbb{S}_\delta^d \right\}. \end{aligned}$$

Let

$$[u]_{\alpha; D} := \sup_{x, y \in D} \frac{|u(x) - u(y)|}{|x - y|^\alpha} \quad \text{and} \quad \text{Osc}_D u := \sup_{x \in D} u(x) - \inf_{x \in D} u(x).$$

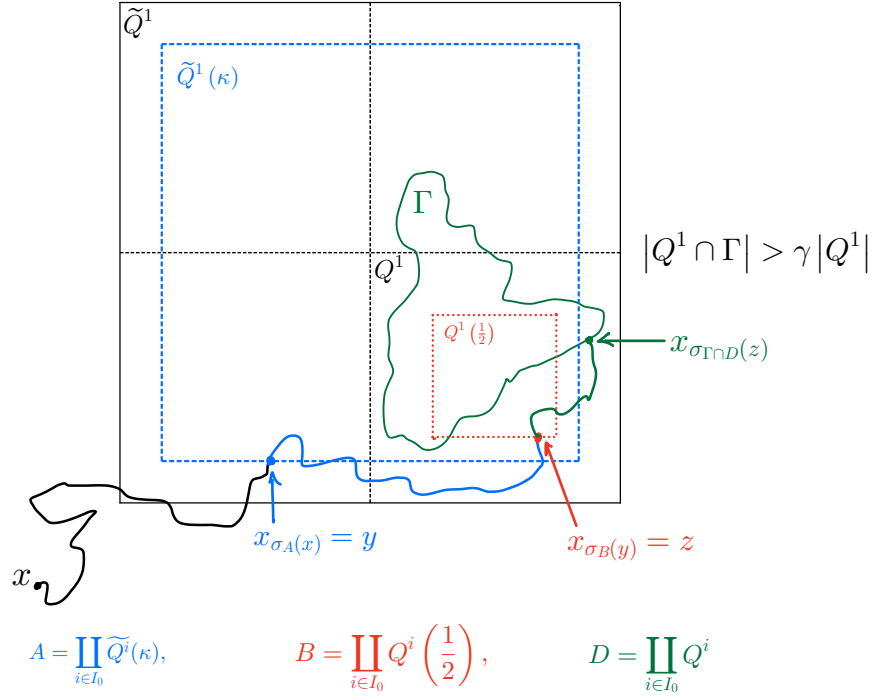


Figure 4.1: Hitting Prob.

**Theorem 4.3.1** (Hölder estimate). *Suppose  $u$  is bounded in  $Q_1$  and  $Lu = 0$  in  $Q_1$ . Then there exist  $\alpha$  and  $C$  only depending on  $d$  and  $\delta$  such that*

$$[u]_{\alpha; Q_{1/2}} \leq C \|u\|_{L^\infty(Q_1)}. \quad (4.7)$$

*Proof. Claim:* there exists a constant  $\rho \in (0, 1)$  such that for any  $z \in Q_{1/2}$ ,  $r \leq 1/2$ ,

$$\text{Osc}_{Q_{r/2}(z)} u \leq \rho \text{Osc}_{Q_r(z)} u. \quad (4.8)$$

Assume the claim is true. Suppose that  $x, y \in Q_{1/2}$  and  $|x - y| \ll 1$ , let  $k \in \mathbb{N}$  such that  $2^{-k-1} \leq |x - y| < 2^{-k}$ .

$$\begin{aligned} |u(x) - u(y)| &\leq \text{Osc}_{Q_{2^{-k}}(x)} u \leq \rho \text{Osc}_{Q_{2^{-k+1}}(x)} u \leq \dots \leq C \rho^k \|u\|_{L^\infty(Q_1)} \\ &\leq C \rho^{-\log_2 |x-y|} \|u\|_{L^\infty(Q_1)} \leq C |x - y|^{-\log_2 \rho} \|u\|_{L^\infty(Q_1)}. \end{aligned}$$

Therefore, the above claim implies (4.7) with  $\alpha = \log_2 \rho^{-1}$ .

To prove (4.8). Without loss of generality, we can assume  $\inf_{x \in Q_r(z)} u = 0$  and  $\sup_{x \in Q_r(z)} u = 1$ . In this case,  $\text{Osc}_{Q_r(z)} u = 1$ . Let  $B := \{x \in Q_{r/2} : u(x) \geq 1/2\}$ , we may assume  $|B| \geq \frac{1}{2}|Q_{r/2}|$ , if not, we replace  $u$  by  $1 - u$ . For any  $x \in Q_{r/2}$ , by Itô's

formula, Theorem 4.2.1 and scaling,

$$u(x) = \mathbb{E}_x u(X_{\tau_{Q_r} \wedge \sigma_B}) \geq \frac{1}{2} \mathbb{P}_x(\sigma_B < \tau_{Q_r}) \geq \frac{1}{2} p(2^{-d-1}).$$

Hence we get

$$\text{Osc}_{Q_{r/2}(z)} u \leq 1 - \frac{1}{2} p(2^{-d-1}) =: \rho = \rho \text{Osc}_{Q_r(z)} u.$$

□

**Theorem 4.3.2** (Harnack inequality). *Suppose  $a \in \mathbb{S}_\delta^d$  and  $L = a_{ij} \partial_{ij}$ . There exists  $C$  depending only on  $\delta$  such that if  $u$  is nonnegative, bounded in  $Q_4$ , and  $u(X_{t \wedge \tau_{Q_4}})$  is a martingale, then  $u(x) \leq Cu(y)$  if  $x, y \in Q_1$ .*

*Proof.* If we look at  $u + \delta$  and let  $\delta \rightarrow 0$ , we may assume  $u > 0$ . By looking at  $Cu$ , we may assume  $\inf_{Q_{1/2}} u = 1$ . By Theorem , we know that  $u$  is Hölder continuous in  $Q_1$ , so there exists

$$y \in Q_{1/2} \text{ such that } u(y) = 1.$$

We want to show that  $u$  is bounded above by a constant in  $Q_1$ , where the constant depends only on  $\delta$ .

By the support theorem and scaling, if  $x \in Q_{1/2}$ , there exists  $\delta$  such that

$$\mathbb{P}_y \left( \sigma_{Q_{1/2}(x)} < \tau_{Q_2} \right) \geq \delta.$$

By scaling, if  $z \in Q_{1/2}(x)$ , then  $\mathbb{P}_z \left( \sigma_{Q_{1/4}(x)} < \tau_{Q_2} \right) \geq \delta$ . So by the strong Markov property,

$$\mathbb{P}_z \left( \sigma_{Q_{1/4}(x)} < \tau_{Q_2} \right) \geq \delta^2.$$

Repeating and using induction,

$$\mathbb{P}_y \left( \sigma_{Q_{2^{-k}}(x)} < \tau_{Q_2} \right) \geq \delta^k.$$

Then

$$\begin{aligned} 1 = u(y) &\geq \mathbb{E}_y \left[ u \left( X_{\sigma_{Q_{2^{-k}}(x)}} \right); \sigma_{Q_{2^{-k}}(x)} < \tau_{Q_2} \right] \\ &\geq \delta^k \left( \inf_{Q_{2^{-k}}(x)} u \right), \end{aligned}$$

or

$$\inf_{Q_{2^{-k}}(x)} u \leq \delta^{-k}, \quad \forall k \geq 1. \tag{4.9}$$

By the proof of Theorem 4.3.1, there exists  $\rho < 1$  such that

$$\text{Osc}_{Q_{2^{-k-1}}(x)} u \leq \rho \text{Osc}_{Q_{2^{-k}}(x)} u.$$

Take  $N$  large so that  $\rho^{-N} \geq 1/(\delta - \delta^2)$ . Then

$$\text{Osc}_{Q_{2^{-N-k}}(x)} u \geq \rho^{-N} \text{Osc}_{Q_{2^{-k}}(x)} u \geq \frac{1}{\delta - \delta^2} \text{Osc}_{Q_{2^{-k}}(x)} u.$$

---

Take  $K$  large so that  $\sqrt{d}2^{N-K} < 1/8$ . Suppose there exists  $x_0 \in Q_1(y)$  such that  $u(x_0) \geq \delta^{-K-1}$ .

We will construct a sequence  $x_1, x_2, \dots$  by induction such that  $u(x_j) \geq \delta^{-K-j-1}$ .

Suppose we have  $x_j \in Q_{2^{N+1-K-j}}(x_{j-1})$  with  $u(x_j) \geq \delta^{-K-j-1}$ ,  $j \leq n$ . Since  $|x_j - x_{j-1}| < \sqrt{d}2^{N+1-K-j}$ ,  $1 \leq j \leq n$ , and  $|x_0 - y| \leq 1$ , then  $|x_n - y| < 2$ . Since  $u(x_n) \geq \delta^{-K-n-1}$  and by (4.9),  $\inf_{Q_{2^{-K-n}}(x_n)} u \leq \delta^{-K-n}$ ,

$$\operatorname{Osc}_{Q_{2^{-K-n}}(x_n)} u \geq \delta^{-K-n} (\delta^{-1} - 1).$$

So  $\operatorname{Osc}_{Q_{2^{N-K-n}}(x_n)} u \geq \delta^{-K-n-2}$ , which implies that there exists  $x_{n+1} \in Q_{2^{N-K-n}}(x_n)$  with  $u(x_{n+1}) \geq \delta^{-K-n-2}$  because  $u$  is nonnegative. By induction we obtain a sequence  $x_n$  with  $x_n \in Q_3(y)$  and  $u(x_n) \rightarrow \infty$ . This contradicts the boundedness of  $u$  on  $Q_4$ . Therefore  $u$  is bounded on  $Q_1$  by  $\delta^{-K-1}$ .  $\square$



# Chapter 5

## Malliavin's proof of Hörmander's Theorem

Let  $V_0, V_1, \dots, V_n : \mathbb{R}^d \rightarrow \mathbb{R}^d$  be vector fields satisfying  $C^\infty$ -boundedness conditions. Consider

$$\begin{cases} dX_t(x) &= \sum_{l=1}^n V_l(X_t) \circ dW_s^l + V_0(X_t)dt = V(X_t) \circ dW_s + V_0(X_t)dt, \\ X_0(x) &= x. \end{cases} \quad (5.1)$$

The Malliavin calculus is a method originally developed for proving smoothness of  $p_t(x, y)$  in the variable  $y$ , where  $p_t(x, y)$  is the transition density of a process associated to an operator with smooth coefficients. The basic idea involves an integration by parts formula in an infinite-dimensional space.

There are two main approaches, one using the Girsanov transformation and the other using the Ornstein-Uhlenbeck operator. We follow the Girsanov approach pioneered by Bismut [Bis81] to obtain the integration by parts formula.

### 5.1 Integration by parts

Let  $d \geq 1$ , and  $\xi$  be a  $d$ -dimensional standard Gaussian random variable. Then

$$\mathbf{P} \circ \xi^{-1}(dx) = \mu(dx) = \frac{1}{(2\pi)^{\frac{d}{2}}} e^{-\frac{|x|^2}{2}} dx.$$

Suppose  $\varphi \in C_c^\infty(\mathbb{R}^d)$  and  $h \in \mathbb{R}^d$ , integration by parts formula yields that

$$\begin{aligned} \mathbf{E}[\nabla_h \varphi(\xi)] &= \frac{1}{(2\pi)^{\frac{d}{2}}} \int_{\mathbb{R}^d} \nabla_h \varphi(x) e^{-\frac{|x|^2}{2}} dx \\ &= \frac{1}{(2\pi)^{\frac{d}{2}}} \int_{\mathbb{R}^d} \varphi(x) \langle x, h \rangle e^{-\frac{|x|^2}{2}} dx = \mathbf{E}[\varphi(\xi) \langle \xi, h \rangle]. \end{aligned} \quad (5.2)$$

One can also verify that

$$\mathbf{E}[\partial^\alpha \varphi(\xi)] = \mathbf{E}[\varphi(\xi) P_\alpha(\xi)],$$

where  $\alpha \in \mathbb{N}^d$  and  $P_\alpha$  is a polynomial.

**Remark 5.1.1.** (i) Charles Stein also showed that if (5.2) holds for all bounded, continuous and piecewise continuously differentiable functions  $\varphi$  with  $\mathbf{E}|\varphi'(\xi)| < \infty$ , then  $\xi$  has a standard normal distribution.

(ii) If  $h$  is a smooth vector field on  $\mathbb{R}^d$ . Then,

$$\begin{aligned} \mathbf{E}\langle \nabla \varphi(\xi), h \rangle &= \int \langle \nabla \varphi, h \rangle d\mu = - \int \varphi \underbrace{(\operatorname{div} h - \langle h, \cdot \rangle)}_{\operatorname{div}_\mu h} d\mu \\ &= \mathbf{E}\{\varphi(\xi)[\langle h, \xi \rangle - \operatorname{div} h(\xi)]\}. \end{aligned} \quad (5.3)$$

The operator  $L := -\operatorname{div}_\mu \nabla = \Delta - x \cdot \nabla$  is called the Ornstein-Uhlenbeck operator.

The following lemma is a criterion for smooth densities

**Lemma 5.1.2.** Suppose  $\xi : \Omega \rightarrow \mathbb{R}^d$ . Suppose for each  $k$  there exists  $C_k$  such that

$$|\mathbf{E} \nabla^k \varphi(\xi)| \leq C_k \|\varphi\|_\infty$$

whenever  $\varphi \in C_c^\infty$ . Then there exists  $\rho$  smooth such that

$$\mathbf{P}(\xi \in A) = \int_A \rho(x) dx$$

for all Borel sets  $A$ .

*Proof.* Let  $\mu = \mathbf{P} \circ \xi^{-1} \in \mathcal{P}(\mathbb{R}^d) \subseteq \mathcal{D}'(\mathbb{R}^d)$ , and  $\mu_\varepsilon(x) = \int_{\mathbb{R}^d} \varrho_\varepsilon(x-y) \mu(dy)$ . By our assumption, for any  $\alpha = (\alpha_1, \dots, \alpha_d) \in \mathbb{N}^d$  with  $\alpha_1 + \dots + \alpha_d = k$ ,

$$\begin{aligned} (-1)^{|\alpha|} \langle \varphi, \partial^\alpha \mu_\varepsilon \rangle &= \langle \partial^\alpha \varphi, \mu_\varepsilon \rangle = \langle \partial^\alpha \varphi * \varrho_\varepsilon, \mu \rangle \\ &= \int_{\mathbb{R}^d} \partial^\alpha \varphi * \varrho_\varepsilon(x) \mu(dx) \\ &\leq C_k \|\varphi\|_\infty, \quad \forall \varphi \in C_c^\infty, \quad k \in \mathbb{N}. \end{aligned}$$

This implies that  $\sup_\varepsilon \|\nabla^k \mu_\varepsilon\|_1 \leq C_k$ . In the light of Sobolev embedding, one can see that  $\sup_\varepsilon \|\nabla^k \mu_\varepsilon\|_\infty \leq C'_k$ . Therefore,  $\mu_\varepsilon \rightarrow \rho \in C_b^\infty$ .  $\square$

The main tool in the proof is the Malliavin calculus with its integration by part formula in Wiener space (infinite-dimension space), which was developed precisely in order to provide a probabilistic proof of Hörmander's Theorem. It essentially relies on the fact that the image of a Gaussian measure under a “smooth” submersion that is sufficiently integrable possesses a smooth density with respect to Lebesgue measure.

Below we set  $\Omega = C([0, 1]; \mathbb{R}^n)$  and  $(\Omega, \mathcal{F}, \mathcal{F}_t, \mathbf{P})$  be the Wiener space. Then the canonical process  $\omega_t$  is a  $n$ -dimensional Brownian motion under  $\mathbf{P}$ .

**Lemma 5.1.3.** Suppose  $F, G \in C_b^1(\Omega)$  (both  $F$  and  $G$  are bounded and have bounded Fréchet derivative). Suppose  $h_s$  is adapted and bounded and let  $H_t = \int_0^t h_s ds$ . Then

$$\mathbf{E}(\mathcal{D}_H F G) = \mathbf{E} \left[ F \left( -\mathcal{D}_H G + G \int_0^1 \dot{H}_t d\omega_t \right) \right]. \quad (5.4)$$

The left-hand side represents the expectation of the Fréchet derivative at  $\omega$ . in the direction  $H.(\omega)$ .



---

*Proof.* We first prove the case that  $G = 1$ . Let

$$X_t^\varepsilon = \omega_t + \varepsilon \int_0^t h_s ds$$

Let

$$M_t^\varepsilon = \exp \left( -\varepsilon \int_0^t h_s d\omega_s - \frac{\varepsilon^2}{2} \int_0^t |h_s|^2 ds \right).$$

Let  $\mathbf{P}_\varepsilon$  be defined by  $d\mathbf{P}_\varepsilon/d\mathbf{P} = M_t^\varepsilon$  on  $\mathcal{F}_t$ . By Girsanov's theorem, under  $\mathbf{P}_\varepsilon$  the process

$$\omega_t - \left\langle \omega, -\varepsilon \int_0^t h_s d\omega_s \right\rangle_t = \omega_t + \varepsilon \int_0^t h_s ds = X_t^\varepsilon$$

is a martingale with the same quadratic variation as  $\omega_t$ , namely  $t$ . By Theorem 1.5.5, under  $\mathbf{P}_\varepsilon$  the process  $X_t$  is a Brownian motion. Therefore

$$\mathbf{E}_\varepsilon F(X^\varepsilon) = \mathbf{E} F.$$

On the other hand,

$$\begin{aligned} & \mathbf{E}_\varepsilon F(X^\varepsilon) \\ &= \mathbf{E} \left[ F \left( \omega + \varepsilon \int_0^\cdot h_s ds \right) \exp \left( -\varepsilon \int_0^1 h_s d\omega_s + \frac{\varepsilon^2}{2} \int_0^1 |h_s|^2 ds \right) \right]. \end{aligned}$$

By, the right-hand side of is independent of  $\varepsilon$ . We differentiate with respect to  $\varepsilon$  and set  $\varepsilon = 0$ . The assumptions on  $h$  and  $F$  allow us to interchange the operations of differentiation and expectation by use of the dominated convergence theorem and we obtain

$$\mathbf{E} \left[ F \int_0^1 \dot{H}_s d\omega_s \right] = \mathbf{E}[\mathcal{D}_H F].$$

Now replacing  $F$  in the above identity with  $FG$ , and using chain rule for  $\mathcal{D}_H$ , we obtain our assertion.  $\square$

Let

$$W_0^{1,2}([0, 1]; \mathbb{R}^d) = \left\{ H \in C([0, 1]; \mathbb{R}^n) : H(0) = 0, \dot{H} \in L^2([0, 1]; \mathbb{R}^n) \right\}$$

and

$$\langle H, H' \rangle_{W_0^{1,2}} := \int_0^1 \dot{H}(s) \cdot \dot{H}'(s) ds.$$

**Exercise 5.1.1.** *Let*

$$A_\varepsilon = \left\{ \omega : \bigcup_{h \in W_0^{1,2}} \|\omega - h\| < \varepsilon \right\}$$

*Prove that*

$$\mathbf{P}(W_0^{1,2}) = 0, \quad \mathbf{P}(A_\varepsilon) = 1.$$

Suppose that  $F$  and  $H$  satisfy the same conditions in Lemma 5.1.3. Then by (5.4),

$$\mathbf{E}[\mathcal{D}_H F] = \mathbf{E}\left[F \int_0^1 \dot{H}(t) d\omega_t\right] = \mathbf{E}[F(\dot{H} \cdot \omega)], \quad (5.5)$$

where  $(\dot{H} \cdot \omega)$  is the Itô integral  $\int_0^1 \dot{H}_t d\omega_t$ .

Compare (5.3) with (5.5). The adapted process  $H$  can be regarded as “divergence-free” vector field on Wiener space. In this case,

$$\langle h, x \rangle \rightsquigarrow (\dot{H} \cdot \omega) \text{ “} = \text{” } \langle H, \omega \rangle_{W_0^{1,2}} \quad \text{and} \quad \text{“div} H \text{”} = 0.$$

Note that

$$(W_0^{1,2}, \|\cdot\|_{W_0^{1,2}}) \simeq (L^2([0, 1]; \mathbb{R}^n); \|\cdot\|_2) =: (\mathcal{H}, |\cdot|_{\mathcal{H}}).$$

Let  $\{h_k\}_{k \in \mathbb{N}}$  be a normal orthogonal basis of  $\mathcal{H}$ , and  $W_0^{1,2} \ni H_k(t) = \int_0^t h_k(s) ds$ . Set

$$D_k F := \mathcal{D}_{H_k} F, \quad F \in C_b^1(\Omega).$$

Since  $\sum_k (D_k F)^2$  is bounded (since  $F \in C_b^1(\Omega)$ ), we can define

$$DF := \sum_k D_k F h_k \in L^\infty(\Omega; \mathcal{H}), \quad \text{and} \quad |DF|_{\mathcal{H}} = \sqrt{\sum_k (D_k F)^2}.$$

Next we going to identify “div  $H$ ”, when  $H$  is not adapted. Suppose that

$$u_k \in C_b^1(\Omega) \quad \text{and} \quad u = \sum_k u_k h_k,$$

where the summation contains only finite non-zero terms. Then by (5.4),

$$\begin{aligned} \mathbf{E} \langle DF, u \rangle_{\mathcal{H}} &= \mathbf{E} \left\langle DF, \sum_k u_k h_k \right\rangle_{\mathcal{H}} = \mathbf{E} \left( \sum_k D_k F u_k \right) \\ &= \mathbf{E} \left( F \sum_k u_k (h_k \cdot \omega) \right) - \mathbf{E} \left( F \sum_k D_k u_k \right) \\ &= \mathbf{E} \left[ F \left( \overbrace{\sum_k u_k (h_k \cdot \omega)}^{\text{“}=\text{”} \langle u, \omega \rangle_{\mathcal{H}}} - \overbrace{\sum_k D_k u_k}^{\text{“}=\text{”} \text{div} u} \right) \right] \end{aligned}$$

The operator  $\delta$  is called the divergence operator (cf. [Nua06]). In particular, for any  $h \in \mathcal{H}$  and  $F, G \in C_b^1(\Omega)$ , we have

$$\mathbf{E}[\langle DF, h \rangle G] = \mathbf{E}[FG(h \cdot \omega)] - \mathbf{E}[F \langle DG, h \rangle]. \quad (5.6)$$

The main examples of  $F$  we will consider later is  $\varphi(X_t)$ , where  $\varphi$  is smooth and  $X_t$  solves an SDE. However, the Itô map  $F : \omega \mapsto X_t(\omega)$  is even not continuous from  $C([0, 1]; \mathbb{R}^n)$  to  $C([0, 1]; \mathbb{R}^d)$ . So we need some extension. In fact, as a consequence of the integration by parts formula, we can show that the operator  $D$  is closable from  $L^2(\Omega, \mathbf{P}; \mathbb{R})$  to  $L^2(\Omega, \mathbf{P}; \mathcal{H})$ , which guarantees that it is “well-behaved” from a functional-analytic point of view.

**Proposition 5.1.4.** *The operator  $D$  is closable. In other words if, for some sequence  $F_n$ , one has  $F_n \rightarrow 0$  in  $L^2(\Omega, \mathbf{P}; \mathbb{R})$  and  $DF_n \rightarrow u$  in  $L^2(\Omega, \mathbf{P}; \mathcal{H})$ , then  $u = 0$ .*

**Exercise 5.1.2.** *Prove the above proposition. (Hint: show that  $C_b^1(\Omega)$  is dense in  $L^2(\Omega, \mathbf{P})$ .)*

We henceforth denote by  $\mathcal{W}^{1,2}$  the domain of the closure of  $D$  and we do not distinguish between  $D$  and its closure. One can of course apply the Malliavin differentiation operator repeatedly, thus yielding an unbounded closed operator  $D^k$  from  $L^2(\Omega, \mathbf{P})$  to  $L^2(\Omega, \mathbf{P}; \mathcal{H}^{\otimes k})$ . We denote the domain of this operator by  $\mathcal{W}^{k,2}$ . Actually, a similar proof shows that powers of  $D$  are closable as unbounded operators from  $L^p(\Omega, \mathbf{P})$  to  $L^p(\Omega, \mathbf{P}; \mathcal{H}^{\otimes k})$  for every  $p \geq 1$ . We denote the domain of these operators by  $\mathcal{W}^{k,p}$ . Furthermore, for any Hilbert space  $\mathcal{K}$ , we denote by  $\mathcal{W}^{k,p}(\mathcal{K})$  the domain of  $D^k$  viewed as an operator from  $L^p(\Omega, \mathbf{P}; \mathcal{K})$  to  $L^p(\Omega, \mathbf{P}; \mathcal{H}^{\otimes k} \otimes \mathcal{K})$ . We call a random variable belonging to  $\mathcal{W}^{k,p}$  for every  $k, p \geq 1$  “Malliavin smooth” and we write  $\mathcal{S} = \bigcap_{k,p} \mathcal{W}^{k,p}$  as well as  $\mathcal{S}(\mathcal{K}) = \bigcap_{k,p} \mathcal{W}^{k,p}(\mathcal{K})$ . The Malliavin smooth random variables play a role analogous to that of Schwartz test functions in finite-dimensional analysis.

As mentioned above the main examples of  $\mathcal{S}$  functionals will be considered is  $\varphi(X_t)$ .

**Proposition 5.1.5.** *Let  $X_t$  be the  $d$ -dimensional process that is a solution to (5.1). Assume that  $\sigma$  and  $b$  are  $C_b^\infty$ . Then  $X_t \in \mathcal{S}$ .*

**Lemma 5.1.6.**

$$\mathbf{E}|\delta u|^2 = \mathbf{E}|u|^2 + \sum_k \mathbf{E}\langle Du_k, Du_k \rangle_{\mathcal{H}}$$

*Proof.* Suppose  $u = \sum_{k=1}^N u_k h_k$  with  $u_k \in C_b^2(\Omega)$ . Recall that

$$\delta u = u_k(h_k \cdot \omega) - D_k u_k.$$

Then

$$\delta D_{k'} u = \delta(D_{k'} u_k h_k) = D_{k'} u_k(h_k \cdot \omega) - D_k D_{k'} u_k$$

and

$$D_{k'} \delta u = D_{k'}(u_k(h_k \cdot \omega) - D_k u_k) = D_{k'} u_k(h_k \cdot \omega) + u_{k'} - D_{k'} D_k u_k$$

Therefore,  $D\delta u - \delta Du = u$ .

$$\mathbf{E}|\delta u|^2 = \mathbf{E}\langle u, D\delta u \rangle = \mathbf{E}|u|_{\mathcal{H}}^2 + \mathbf{E}\langle u, \delta Du \rangle_{\mathcal{H}} = \mathbf{E}|u|_{\mathcal{H}}^2 + \mathbf{E}|Du|_{\mathcal{H} \otimes \mathcal{H}}^2 = \|u\|_{\mathcal{W}^{1,2}}^2.$$

□

**Proposition 5.1.7.** *For every  $p \geq 1$  there exist constants  $K$  and  $C$  such that, for every separable Hilbert space  $\mathcal{K}$  and every  $u \in \mathcal{S}(\mathcal{H} \otimes \mathcal{K})$ , one has the bound*

$$\mathbf{E}|\delta u|^p \leq C \sum_{0 \leq k \leq K} (\mathbf{E}|D^k u|^{2p})^{1/2}$$

*Proof.* By the above Lemma, we only prove the case that  $p > 2$ . Using the definition of  $\delta$  combined with the chain rule for  $D$ , Proposition 3.8, and Young's inequality, we obtain the bound

$$\begin{aligned} \mathbf{E}|\delta u|^p &= (p-1)\mathbf{E}(|\delta u|^{p-2}\langle u, D\delta u \rangle) = (p-1)\mathbf{E}|\delta u|^{p-2}(|u|^2 + \langle u, D\delta u \rangle_{\mathcal{H}}) \\ &\leq \frac{1}{2}\mathbf{E}|\delta u|^p + C\mathbf{E}(|u|^p + |u|^{p/2}|D\delta u|^{p/2}), \end{aligned}$$

for some constant  $C$ . We now use Hölder's inequality which yields

$$\mathbf{E}(|u|^{p/2}|\delta Du|^{p/2}) \leq (\mathbf{E}|u|^{2p})^{1/4} (\mathbf{E}|D\delta u|^{2p/3})^{3/4}.$$

Combining this with the above, we conclude that there exists a constant  $C$  such that

$$\mathbf{E}|\delta u|^p \leq C (\mathbf{E}|Du|^{2p})^{1/2} + (\mathbf{E}|D\delta u|^{2p/3})^{3/2}.$$

The proof is concluded by a simple inductive argument. □

Suppose that  $F \in \mathcal{S}$ . Define

$$LF = \delta(DF) = \sum_k \left( D_k D_k F - D_k F \int_0^1 h_k(t) d\omega_t \right)$$

Then  $L$  is the operator corresponding to the Ornstein-Uhlenbeck operator  $\Delta - x \cdot \nabla$  in finite dimensional case, and

**Exercise 5.1.3.** *Prove that*

$$\mathbf{E}[(LF)G] = \mathbf{E}[F(LG)].$$

Noting that

$$L(FG) = (LF)G + F(LG) + 2D_k F D_k G,$$

we have

$$\mathbf{E}[(LF)G] = \mathbf{E}[F(LG)] = -\mathbf{E}[D_k F D_k G].$$

**Theorem 5.1.8.** *Let  $p > 1$ ,  $F \in \mathcal{W}^{1,p}$  and . Assume that  $h_t$  is adapted,  $H_t = \int_0^t h_s ds$  and  $|H|_{\mathcal{H}}$  is bounded. Then*

(i)  $D_H F_n$  converges to  $\langle DF, H \rangle =: D_H F$  in  $L^p$  and

$$\mathbf{E}[D_H F] = \mathbf{E} \left[ F \int_0^1 h_s d\omega_s \right].$$

(ii) it holds that

$$\varphi(F) \in \mathcal{W}^{1,p} \quad \text{and} \quad D\varphi(F) = \nabla\varphi(F)DF, \quad \forall \varphi \in C_b^1.$$

(iii) if  $F, G \in \mathcal{W}^{2,p}$  with  $p \geq 2$ ,

$$\mathbf{E}[(LF)G] = \mathbf{E}[F(LG)] = -\mathbf{E}[D_k F D_k G]. \quad (5.7)$$

*Proof.* We apply Lemma 5.1.3 discussed above to  $F_n$  and let  $n \rightarrow \infty$ . The convergence of  $\mathbf{E}[D_H F_n(W)]$  can be seen by

$$|D_H F_n - D_H F_m| = \langle H, DF_n - DF_m \rangle_{\mathcal{H}} \leq |H|_{\mathcal{H}} |D(F_n - F_m)|_{\mathcal{H}} \rightarrow 0 \text{ in } L^p, \quad p > 1.$$

Since

$$\mathbf{E} \left( \int_0^1 h_s d\omega_s \right)^{p'} \leq C_p \mathbf{E} \left( \int_0^1 |h_s|^2 ds \right)^{p'/2} \leq C \mathbf{E} |H|_{\mathcal{H}}^{p'} < \infty, \quad 1 < p' < \infty,$$

the convergence of  $\mathbf{E} \left[ F_n \int_0^1 h_s d\omega_s \right]$  follows from the  $L^p$  convergence of  $F_n$  to  $F$  in  $L^p$  with  $p > 1$  and the Hölder inequality. The second assertion can be obtained by similar way.  $\square$

## 5.2 Malliavin Matrix

Now assume that  $F = (F_1, \dots, F_d)$ , each  $F_i$  is real-valued and  $F_i \in \cap_{k,p \geq 1} \mathcal{W}^{k,p}$ , and that  $\varphi \in C_b^\infty(\mathbb{R}^d)$ . Let

$$\gamma_F^{ij} := \langle DF^i, DF^j \rangle_{\mathcal{H}}.$$

Then

$$D(\varphi(F)) = \partial_i \varphi(F) DF^i$$

and

$$\langle D(\varphi(F)), DF^j \rangle_{\mathcal{H}} = \partial_i \varphi(F) \gamma_F^{ij}.$$

This yields that

$$\partial_i \varphi(F) = \langle D(\varphi(F)), DF^j \rangle_{\mathcal{H}} (\gamma_F^{-1})^{ji}, \quad (5.8)$$

provided that  $\gamma_F$  is invertible.

**Exercise 5.2.1.** Assume  $F \in \cap_{k,p \geq 1} \mathcal{W}^{k,p}$  and  $\gamma_F^{-1} \in \cap_{p \geq 1} L^p$ . Then

$$\gamma_F^{-1} \in \cap_{k,p \geq 1} \mathcal{W}^{k,p} \quad \text{and} \quad D\gamma_F^{-1} = -\gamma_F^{-1}(D\gamma_F)\gamma_F^{-1}.$$

**Theorem 5.2.1.** Suppose that  $F \in \cap_{k,p \geq 1} \mathcal{W}^{k,p}$  and  $\gamma_F^{-1} \in \cap_{p \geq 1} L^p$ . Then

$$|\mathbf{E} \nabla \varphi(F)| \leq C \|\varphi\|_\infty.$$

*Proof.* Thanks to (5.8) and Theorem 5.1.8,

$$\begin{aligned}\mathbf{E}\nabla^T\varphi(F) &= \mathbf{E} [\langle D(\varphi(F)), \gamma_F^{-1}DF \rangle_{\mathcal{H}}] \\ &= \mathbf{E} [\varphi(F)\delta(\gamma_F^{-1}DF)]\end{aligned}$$

This yields that

$$|\mathbf{E}\nabla^T\varphi(F)| \leq \|\varphi\|_{\infty}\mathbf{E}[\delta(\gamma_F^{-1}DF)] \leq C(p, \|F\|_{\mathcal{W}^{k,p}}, \|\gamma_F^{-1}\|_p)\|\varphi\|_{\infty}, \quad p \gg 1.$$

□

We need to calculate  $D_kX_t$ , where  $X_t$  solves (5.1). By definition,

$$X_t(\omega) = x + \int_0^t V(X_s) \circ d\omega_s + \int_0^t V_0(X_s)ds$$

and

$$\begin{aligned}X_t(\omega + \varepsilon H_k) &= x + \int_0^t V(X_s(\omega + \varepsilon H_k)) \circ d(\omega_s + \varepsilon H_k(s)) + \int_0^t V_0(X_s(\omega + \varepsilon H_k))ds \\ &= x + \int_0^t V(X_s(\omega + \varepsilon H_k)) \circ d\omega_s + \varepsilon \int_0^t V(X_t(\omega + \varepsilon H_k))h_k(s)ds \\ &\quad + \int_0^t V_0(X_s(\omega + \varepsilon H_k))ds\end{aligned}$$

Taking the difference, dividing by  $\varepsilon$ , and letting  $\varepsilon \rightarrow 0$ , we obtain that

$$D_kX_t = \int_0^t \partial_j V_l(X_s) D_kX_s^j \circ d\omega_s^l + \int_0^t \partial_j V_0(X_s) D_kX_s^j ds + \int_0^t V_l(X_s) h_k^l(s) ds. \quad (5.9)$$

Recall that  $J(t) = \nabla X_t$ , then

$$dJ_j^i(t) = \partial_l V_k^i(X_t) J_j^l(t) \circ d\omega_t^k + \partial_l V_0^i(X_t) J_j^l(t) dt, \quad J(0) = I. \quad (5.10)$$

Let  $Z(t) : \Omega \rightarrow \mathbb{R}^{d \times d}$  be the solution to

$$dZ_j^i(t) = -Z_l^i(t) \partial_j V_k^l(X_t) \circ d\omega_t^k - Z_l^i(t) V_0^l(X_t) dt, \quad Z(0) = I, \quad (5.11)$$

By Itô's formula, one can verify that

$$d(Z(t)J(t)) = Z(t) \circ dJ(t) + dZ(t) \circ J(t) = 0,$$

which yields that  $Z(t) = J^{-1}(t)$ .

**Proposition 5.2.2.**

$$D_kX_t = J(t) \int_0^t J^{-1}(s) V(X_s) h_k(s) ds \quad (5.12)$$

Consequently,

$$\gamma_{X_t} = \langle DX_t, DX_t^T \rangle = J(t) \int_0^t J^{-1}(s) V(X_s) V^T(X_s) [J^{-1}(s)]^T ds J^T(t). \quad (5.13)$$

*Proof.* By Itô's formula and (5.10),

$$\begin{aligned}
& d \left[ J(t) \int_0^t J^{-1}(s) V(X_s) h_k(s) ds \right] \\
&= dJ(t) \int_0^t J^{-1}(s) V(X_s) h_k(s) ds + V(X_t) h_k(t) dt \\
&\stackrel{(5.10)}{=} \nabla V(X_t) \left[ J(t) \int_0^t J^{-1}(s) V(X_s) h_k(s) ds \right] \circ d\omega_t \\
&\quad + \nabla V_0(X_t) \left[ J(t) \int_0^t J^{-1}(s) V(X_s) h_k(s) ds \right] dt + V(X_t) h_k(t) dt
\end{aligned}$$

Therefore,  $t \mapsto J(t) \int_0^t J^{-1}(s) V(X_s) h_k(s) ds$  satisfies the same equation as  $D_k X_t$ , which yields (5.12).  $\square$

Set

$$C(t) := \int_0^t J^{-1}(s) V(X_s) V^T(X_s) [J^{-1}(s)]^T ds.$$

### 5.3 Hörmander's Theorem

Let  $U(x) = \sum_{i=1}^n U^i(x) \partial_i = U^i(x) \partial_i$ ,  $V(x) = V^i(x) \partial_i$ . Define the Lie Bracket  $[U, V]$  as:

$$[U, V](x) := U^i(x) \partial_i (V^j(x)) \partial_j - V^i(x) \partial_i (U^j(x)) \partial_j = [U^i \partial_i V^j - V^i \partial_i U^j](x) \partial_j$$

Define

$$S_0 = \{V_i : i > 0\}, S_{k+1} = S_k \cup \{[U, V_j] : U \in S_k, j \geq 0\},$$

and

$$\mathcal{V}^k = \text{span}\{V : V \in S_k\}.$$

We say that the vector fields  $V_0, V_1, \dots, V_n$  satisfy the parabolic Hörmander condition

$$\bigcup_{k \geq 0} \mathcal{V}^k(x) = \mathbb{R}^d, \quad \forall x \in \mathbb{R}^d \quad (\text{H})$$

Why we consider this kind of condition?

**Theorem 5.3.1** (Stroock-Varadhan's support theorem). *The law of the solution to (5.1) on path space is supported by the closure of those smooth curves that, at every point  $(t, x)$ , are tangent to the hyperplane spanned by  $\{\hat{V}_0(x, t), \dots, \hat{V}_N(x, t)\}$ , where we set*

$$\hat{V}_0(x, t) = \begin{pmatrix} V_0(x) \\ 1 \end{pmatrix}, \quad \hat{V}_j(x, t) = \begin{pmatrix} V_j(x) \\ 0 \end{pmatrix}, \quad j = 1, 2, \dots, N.$$

For a smooth manifold  $\mathcal{M}$ , recall that  $E \subset T\mathcal{M}$  is a smooth subbundle of dimension  $d$  if  $E_x \subset T_x \mathcal{M}$  is a vector space of dimension  $d$  at every  $x \in \mathcal{M}$  and if the dependency  $x \rightarrow E_x$  is smooth. (Locally,  $E_x$  is the linear span of finitely many smooth vector fields on  $\mathcal{M}$ .) A subbundle is called integrable if, whenever  $U, V$  are vector fields on  $\mathcal{M}$  taking values in  $E$ , their Lie bracket  $[U, V]$  also takes values in  $E$ . With these definitions at hand, recall the well-known Frobenius integrability theorem from differential geometry:

**Theorem 5.3.2.** *Let  $\mathcal{M}$  be a smooth  $n$ -dimensional manifold and let  $E \subset T\mathcal{M}$  be a smooth vector bundle of dimension  $d < n$ . Then  $E$  is integrable if and only if there (locally) exists a smooth foliation of  $M$  into leaves of dimension  $d$  such that, for every  $x \in \mathcal{M}$ , the tangent space of the leaf passing through  $x$  is given by  $E_x$ .*

In view of this result, Hörmander's condition is not surprising. Indeed, if we define  $E(x, t) = \bigcup_{k \geq 0} \mathcal{V}^k(x, t)$ , then this gives us a subbundle of  $\mathbb{R}^{N+1}$  which is integrable. Note that the dimension of  $E(x, t)$  could in principle depend on  $(t, x)$ , but since the dimension is a lower semicontinuous function, it will take its maximal value on an open set. If, on some open set, this maximal value is less than  $n + 1$ , then support theorem tells us that, there exists a submanifold (with boundary)  $\mathcal{M}^- \subset \mathcal{M}$  of dimension strictly less than  $n$  such that  $T_{(y,s)}\mathcal{M}^- = E_{(y,s)}$  for every  $(y, s) \in \mathcal{M}^-$ . In particular, all the curves appearing in the Stroock-Varadhan support theorem and supporting the law of the solution to (1.1) must lie in  $\mathcal{M}^-$  until they reach its boundary. As a consequence, since  $\mathcal{M}^-$  is always transverse to the sections with constant  $t$ , the solutions at time  $t$  will, with positive probability, lie in a submanifold of  $\mathcal{M}$  of strictly positive codimension. This immediately implies that the transition probabilities cannot be continuous with respect to Lebesgue measure.

**Theorem 5.3.3.** *Consider (5.1) and assume that all vector fields have bounded derivatives of all orders. If it satisfies (H), then its solutions admit a smooth density with respect to Lebesgue measure and the corresponding Markov semigroup maps bounded functions into smooth functions.*

We only need to prove  $\det C_t \in L^{\infty-}$ .

We need the following useful lemma.

**Lemma 5.3.4.** *Let  $M$  be a random, symmetric, positive semidefinite matrix with entries in  $L^{\infty-}$ . Assume that for  $p$  sufficient large, there exists a constant  $C_p$  and an  $\varepsilon_p > 0$  such that for  $0 < \varepsilon < \varepsilon_p$  we have*

$$\sup_{|\xi|=1} \mathbf{P}(\xi^T M \xi \leq \varepsilon) \leq C_p \varepsilon^p.$$

Then  $(\det M)^{-1} \in L^{\infty-}$ .

*Proof.*  $\forall t > 1$ , Choose  $\{\xi_1, \dots, \xi_m\} \subset S^N$ , such that  $\sup_{|\xi|=1} \min_{k \leq m} |\xi - \xi_k| \leq t^{-2}$  and  $m \leq Ct^{2N}$ .

$\forall \xi \in S^N$ , we can find a vector  $\xi_k$  such that,

$$\xi^T M \xi = \xi_k^T M \xi_k + \xi_k^T M (\xi - \xi_k) + (\xi - \xi_k)^T M \xi \geq \xi_k^T M \xi_k - 2\|M\|t^{-2}$$

So we get

$$\left\{ \inf_{|\xi|=1} \xi^T M \xi < t^{-1} \right\} \setminus \bigcup_{k=1}^m \{ \xi_k^T M \xi_k < 3t^{-1} \} \subset \{ \|M\| > t \}$$



Now,

$$\begin{aligned}
\mathbf{P}(\|M^{-1}\| > t) &= \mathbf{P}\left(\inf_{|\xi|=1} \xi^T M \xi < t^{-1}\right) \\
&\leq \mathbf{P}\left(\cup_{k=1}^m \{\xi_k^T M \xi_k < 3t^{-1}\}\right) + \mathbf{P}(\|M\| > t) \\
&\leq C C_{2N+p} t^{2N} t^{-2N-p} + \mathbf{E}(\|M\|^p) t^{-p} \\
&\leq C_p t^{-p}
\end{aligned}$$

$\forall q \geq 1$ ,

$$\begin{aligned}
\mathbf{E}(\det M^{-1})^q &\leq \mathbf{E}\|M^{-1}\|^q \leq C \int_0^\infty t^{q-1} \mathbf{P}(\|M^{-1}\| > t) dt \\
&\leq \int_0^\infty 1 \wedge t^{q-1} C_q t^{-q-1} dt < \infty.
\end{aligned}$$

□

Fixed  $\xi \in S^N$ , define  $Z_U(t) = \xi^T J_t^{-1} U(X_t)$ . Using Ito's formula,

$$\begin{aligned}
dZ_U(t) &= -\xi^T J_t^{-1} V'_0(X_t) U(X_t) dt - \xi^T J_t^{-1} V'_j(X_t) U(X_t) \circ dW_t^j \\
&\quad + \xi^T J_t^{-1}(X_t) U'(X_t) V_0(X_t) dt + \xi^T J_t^{-1}(X_t) U'(X_t) V_j(X_t) \circ dW_t^j \\
&= \xi^T J_t^{-1} [V_0, U](X_t) dt + \xi^T J_t^{-1} [V_j, U](X_t) \circ dW_t^j \\
&= Z_{[V_0, U]}(t) + Z_{[V_j, U]}(X_t) \circ dW_t^j \\
&= \left[ Z_{[V_0, U]}(t) + \frac{1}{2} Z_{[V_j, [V_j, U]]} \right] dt + Z_{[V_j, U]}(X_t) \cdot dW_t^j.
\end{aligned} \tag{5.14}$$

Before going to prove the main theorem, we need some technical lemmas.

**Lemma 5.3.5.** *Suppose  $f \in C^{1+\alpha}([0, 1])$ , then*

$$\|f\|_{C^1} \leq C \|f\|_{C^0}^{\frac{\alpha}{1+\alpha}} \cdot \|f\|_{C^{1+\alpha}}^{\frac{1}{1+\alpha}}.$$

**Lemma 5.3.6.** *Suppose  $Y_t = \int_0^t \sigma_s dB_s$ ,  $\mathbf{E}\|\sigma\|_\infty^p \leq K_p < \infty$ , then  $\forall \alpha \in (0, \frac{1}{2})$ ,  $p \geq 1$ ,  $\mathbf{E}\|Y\|_{C^\alpha}^p \leq C(K, \alpha, p)$ .*

*Proof.* Choose  $\beta, p$  such that

$$\alpha + \frac{1}{p} < \beta < \frac{1}{2},$$

by Lemma 1.2.3,

$$\mathbf{E}\|Y\|_{C^\alpha}^p \leq C \mathbf{E} \left( \int_0^1 \int_0^1 \frac{|Y_s - Y_t|^p}{|s - t|^{1+\beta p}} ds dt \right) = C \int_0^1 \int_0^1 \frac{\mathbf{E}|Y_s - Y_t|^p}{|s - t|^{1+\beta p}} ds dt.$$

Thanks to BDG inequality,

$$\mathbf{E}|Y_s - Y_t|^p \leq C \mathbf{E} \left( \int_s^t \|\sigma_r\|_\infty^2 dr \right)^{p/2} \leq C_p |s - t|^{p/2}.$$

Combining the above inequalities,

$$\mathbf{E}\|Y\|_{C^\alpha}^p \leq C \int_0^1 \int_0^1 |s-t|^{-1-\beta p+p/2} ds dt < \infty.$$

□

Suppose  $U$  is a smooth vector field with bounded derivatives of all orders. By the above lemma, it's easy to see

$$\|Z_U\|_{C^\alpha} \in L^{\infty-}.$$

**Definition 5.3.7.** Let  $\{A\}_{\varepsilon \in (0,1)}$ ,  $\{B\}_{\varepsilon \in (0,1)}$  be two family of random events.

$A_\varepsilon \rightarrow_\varepsilon B_\varepsilon$  means  $\forall p > 1$ , there exists a constant  $C_p$ , such that,

$$\mathbf{P}(A_\varepsilon \setminus B_\varepsilon) \leq C_p \varepsilon^p.$$

**Lemma 5.3.8** (Quantitative version of Doob-Meyer's decomposition). *Let  $W$  be a  $d$ -dimensional Wiener process,  $a$  and  $b$  be  $\mathbb{R}$  respectively  $\mathbb{R}^d$ -valued adapted processes such that, for  $\alpha < 1/2$ , we have  $\|a\|_\alpha, \|b\|_\alpha \in L^{\infty-}$ . Moreover, let  $Z$  be defined by*

$$Z_t = Z_0 + \int_0^t a(s) ds + \int_0^t b_j(s) dW_s^j.$$

*Then there exists a constant  $r \in (0, 1)$  such that*

$$\{\|Z\|_\infty < \varepsilon\} \rightarrow_\varepsilon \{\|a\|_\infty < \varepsilon^r\} \cap \{\|b\|_\infty < \varepsilon^r\}.$$

*Proof.*

$$Z_t^2 = Z_0^2 + \int_0^t (2a_s Z_s + |b(s)|^2) ds + \int_0^t 2b_j(s) Z_s dW_s^j. \quad (5.15)$$

$\|a\|_\alpha \in L^{\infty-} \rightarrow \{\|a\|_\infty \leq \varepsilon^{-1/4}\} \rightarrow_\varepsilon \emptyset$ . Hence,  $\{\|Z\|_\infty \leq \varepsilon\} \rightarrow_\varepsilon \{\|\int_0^\cdot 2a_s Z_s ds\|_\infty \leq \varepsilon^{3/4}\}$ . Similarly  $\{\|Z\|_\infty \leq \varepsilon\} \rightarrow_\varepsilon \{\|\int_0^1 |b_j(s) Z_s|^2 ds\|_\infty \leq \varepsilon^{3/2}\}$ . Using exponential martingale inequality,

$$\left\{ \left\| \int_0^1 |b_j(s) Z_s|^2 ds \right\|_\infty \leq \varepsilon^{3/2} \right\} \rightarrow_\varepsilon \left\{ \left\| \int_0^\cdot 2b_j(s) dW_s^j \right\|_\infty \leq \varepsilon^{2/3} \right\}.$$

From (5.15), we get

$$\{\|Z\|_\infty \leq \varepsilon\} \rightarrow_\varepsilon \left\{ \int_0^1 |b(s)|^2 ds \leq \varepsilon^{2/3} \right\} \rightarrow_\varepsilon \left\{ \int_0^1 |b(s)| ds \leq \varepsilon^{1/3} \right\}.$$

Combining the above relation, (5.3.5) and  $\{\|b\|_{1/3} \leq \varepsilon^{-1/4}\} \rightarrow_\varepsilon \emptyset$ , we get

$$\{\|Z\|_\infty \leq \varepsilon\} \rightarrow_\varepsilon \{\|b\|_\infty \leq \varepsilon^{1/16}\}.$$

Using the same argument, we can prove

$$\{\|Z\|_\infty \leq \varepsilon\} \rightarrow_\varepsilon \{\|a\|_\infty \leq \varepsilon^{1/80}\}.$$

□

---

*Proof of Theorem.* Notice

$$\xi^T C_t \xi = \sum_{j=1}^d \int_0^t |Z_{V_j}(s)|^2 ds. \quad (5.16)$$

Using Lemma 5.3.5, we get

$$\{\xi^T C_t \xi \leq \varepsilon\} \rightarrow_\varepsilon \{\|Z_{H_k}\|_\infty \leq \varepsilon^q\}$$

By Lemma 5.3.8,

$$\{\xi^T C_t \xi \leq \varepsilon\} \rightarrow_\varepsilon \bigcap_{V \in \mathcal{H}_k} \{\|Z_V\|_\infty \leq \varepsilon^{q_k}\}$$

for suitable  $q_k > 0$ . Now observe that  $Z_V(0) = \langle x, U(x_0) \rangle$ . By Hörmander's condition,  $\mathcal{V}^{k'}(x_0) = \mathbb{R}^N$  for  $k$  large enough. However, if  $\mathcal{V}^{k'}(x_0) = \mathbb{R}^N$ , we can pick  $V \in \mathcal{V}^{k'}(x_0)$  such that  $|Z_V(0)| = |\langle x, V(x_0) \rangle| \geq \varepsilon_0$ , so that the right-hand-side in the above equation is the empty set. We have thus proved  $\{\xi^T C_t \xi \leq \varepsilon\} \rightarrow_\varepsilon \emptyset$ . Now, using Lemma 5.3.4, we complete the proof.  $\square$



# Appendix A

## Useful facts

**Lemma A.0.1** (Area formula). *Consider a locally Lipschitz function  $f : \mathbb{R}^d \rightarrow \mathbb{R}^d$  and a Borel set  $A \subseteq \mathbb{R}^d$ . Then the function  $y \mapsto N_A(y) := \text{card}\{f^{-1}(y) \cap A\}$  is measurable and*

$$\int_A |\det(\nabla f(x))| dx = \int_{\mathbb{R}^n} N_A(y) dy \geq \mathcal{L}^d(f(A)).$$

Consequently, for any  $g \geq 0$ ,

$$\int_{f(A)} g(y) dy \leq \int_A g(f(x)) |\det \nabla f(x)| dx. \quad (\text{A.1})$$

### A.1 Sobolev spaces

Let  $W^{k,p}$  denote the Sobolev space consisting of all real-valued functions on  $\mathbb{R}^d$  whose weak derivatives up to order  $k$  are functions in  $L^p$ . Here  $k$  is a non-negative integer and  $1 \leq p < \infty$ . The first part of the Sobolev embedding theorem states that

**Theorem A.1.1** (Sobolev). *If  $k > l$ , and  $1 \leq p < q < \infty$  are two real numbers such that*

$$\frac{1}{p} - \frac{k-l}{d} = \frac{1}{q},$$

*then*

$$W^{k,p} \hookrightarrow W^{l,q}.$$

The second part of the Sobolev embedding theorem applies to embeddings in Hölder spaces  $C^{r,\alpha}$ .

**Theorem A.1.2** (Morrey). *If  $d < pk$  and*

$$r + \alpha = k - \frac{d}{p}$$

*with  $\alpha \in (0, 1)$ , then one has the embedding*

$$W^{k,p} \hookrightarrow C^{r,\alpha}.$$



# Appendix B

## Harmonic analysis from the perspective of Probability

Singular integrals are central to harmonic analysis and are intimately connected with the study of partial differential equations. Singular integral is an integral operator

$$T(f)(x) = \int K(x, y) f(y) \, dy,$$

whose kernel function  $K : \mathbb{R}^d \times \mathbb{R}^d \rightarrow \mathbb{R}$  is singular along the diagonal  $x = y$ .

Typical examples of integral operators are the Riesz transforms, which are a family of generalizations of the Hilbert transform to Euclidean spaces of dimension  $d \geq 2$ . Specifically, the Riesz transforms of a complex-valued function  $f$  are defined by

$$R_i f(x) = c_d \lim_{\epsilon \rightarrow 0} \int_{\mathbb{R}^d \setminus B_\epsilon(x)} \frac{(x_i - y_i) f(y)}{|x - y|^{d+1}} \, dy, \quad i = 1, \dots, d.$$

The constant  $c_d$  is a dimensional normalization given by  $c_d = \frac{1}{\pi \omega_{d-1}} = \frac{\Gamma((d+1)/2)}{\pi^{(d+1)/2}}$ . The limit is written in various ways, often as a principal value, or as a convolution with the tempered distribution

$$K(x) = \frac{1}{\pi \omega_{d-1}} p.v. \frac{x_j}{|x|^{d+1}}.$$

The Riesz transforms are given by a Fourier multiplier. Indeed, the Fourier transform of  $R_i f$  is given by

$$\mathcal{F}(R_i f)(\xi) = -i \frac{\xi_i}{|\xi|} (\mathcal{F} f)(\xi)$$

A particular consequence of this last observation is that the Riesz transform defines a bounded linear operator in  $L^2$ .

**Theorem B.0.1.** *For each  $i \in \{1, \dots, d\}$ ,  $R_i$  is bounded on  $L^p$  with  $p \in (1, \infty)$  and satisfy weak-type  $(1,1)$  estimates:*

$$|\{x \in \mathbb{R}^d : |R_i f(x)| > \lambda\}| \leq C_d \|f\|_1 / \lambda. \quad (\text{A.1})$$

## B.1 Interpolation Theorems

The following simple Interpolation theorem is useful.

**Lemma B.1.1.**    1. *Let  $p \in [1, \infty)$ . There exists constant  $C$  such that*

$$\|\nabla u\|_p \leq C \|\nabla^2 u\|_p^{\frac{1}{2}} \|u\|_p^{\frac{1}{2}}. \quad (\text{A.2})$$

*Let  $\alpha \in (0, \alpha)$ . There exists constant  $C$  such that*

2.

$$[\nabla u]_\alpha \leq C [\nabla^2 u]_\alpha^{\frac{1}{2}} [u]_\alpha^{\frac{1}{2}}. \quad (\text{A.3})$$



# Appendix C

## Some basic results of PDEs

### C.1 Monge-Ampère Equation

To motivate the definition of weak solutions to (2.6), given an open set  $D \subset \mathbb{R}^n$ , consider  $u : D \rightarrow \mathbb{R}$  a convex function of class  $C^2$  satisfying (2.6) for some  $f : D \rightarrow \mathbb{R}^+$ . Then given any Borel set  $E \subset D$ , it follows by the area formula that

$$\int_E f \, dx = \int_E \det D^2 u \, dx = |\nabla u(E)|.$$

Notice that while the above argument needs  $u$  to be of class  $C^2$ , the identity

$$\int_E f = |\nabla u(E)|$$

makes sense if  $u$  is only of class  $C^1$ . To find a definition when  $u$  is merely convex one could try to replace the gradient  $\nabla u(x)$  with the subdifferential  $\partial u(x)$  and ask for the above equality to hold for any Borel set  $E$ . Here  $\partial u(x)$  is given by

$$\partial u(x) := \{p \in \mathbb{R}^d : u(y) \geq u(x) + \langle p, y - x \rangle \quad \forall y \in D\}.$$

This motivates the following definition:

**Definition C.1.1.** *Given an open set  $D \subset \mathbb{R}^n$  and a convex function  $u : D \rightarrow \mathbb{R}$ , we define the Monge-Ampère measure associated to  $u$  by*

$$\mu_u(E) := \left| \bigcup_{x \in E} \partial u(x) \right|$$

The basic idea of Alexandrov was to say that  $u$  is a weak solution of (2.6) if  $\mu_u|_D = \nu|_D$ .

**Lemma C.1.2.** *Let  $u, v : D \rightarrow \mathbb{R}$  be convex functions. Then*

$$\mu_{u+v} \geq \mu_u + \mu_v \quad \text{and} \quad \mu_{\lambda u} = \lambda^n \mu_u \quad \forall \lambda > 0.$$

The following result is the celebrated Alexandrov maximum principle.

**Theorem C.1.3.** *Let  $D$  be an open bounded convex set, and let  $u : D \rightarrow \mathbb{R}$  be a convex function such that  $u|_{\partial D} = 0$ . Then there exists a dimensional constant  $C = C(d)$  such that*

$$|u(x)| \leq C(d) \operatorname{diam}(D)^{\frac{d-1}{d}} \operatorname{dist}(x, \partial D)^{\frac{1}{d}} |\partial u(D)|^{\frac{1}{d}}, \quad \forall x \in D. \quad (\text{B.1})$$

*Proof.* Let  $(x, u(x))$  be a point on the graph of  $u$ , and consider the convex “conical” function  $y \mapsto \widehat{C}_x(y)$  with vertex at  $(x, u(x))$  that vanishes on  $\partial D$ . Since  $u \leq \widehat{C}_x$  in  $D$  (by the convexity of  $u$ ), Lemma 2.7 implies that

$$\left| \partial \widehat{C}_x(x) \right| \leq \left| \partial \widehat{C}_x(D) \right| \leq |\partial u(D)|;$$

so, to conclude the proof, it suffices to bound  $|\partial \widehat{C}_x(x)|$  from below. It is not hard to see

- $\partial \widehat{C}_x(x)$  contains the ball  $B_\rho$  with  $\rho = |u(x)|/\operatorname{diam}(D)$
- $\partial \widehat{C}_x(x)$  contains a vector of norm  $|u(x)|/\operatorname{dist}(x, \partial D)$

Thus,

$$\partial \widehat{C}_x(x) \supset B_\rho(0) \cup \{q\}, \quad |q| = |u(x)|/\operatorname{dist}(x, \partial D).$$

Since  $\partial \widehat{C}_x(x)$  is convex, it follows that  $\partial \widehat{C}_x(x)$  contains the cone  $\mathcal{C}$  generated by  $q$  and  $\Sigma_q := \{p \in B_\rho : \langle p, q \rangle = 0\}$ . Therefore

$$c(d)\rho^{d-1}|q| = |\mathcal{C}| \leq |\partial u(D)|.$$

□

**Theorem C.1.4.** *Let  $D$  be an open bounded convex set, and let  $\nu$  be a Borel measure on  $D$  with  $\nu(D) < \infty$ . Then there exists a unique convex function  $u : D \rightarrow \mathbb{R}$  solving the Dirichlet problem*

$$\begin{cases} \mu_u = \nu & \text{in } D \\ u = 0 & \text{on } \partial D \end{cases}$$

*Proof.* By the stability result proved in Lemma below, since any finite measure can be approximated in the weak\* topology by a finite sum of Dirac deltas, we only need to solve the Dirichlet problem when  $\nu = \sum_{i=1}^N \alpha_i \delta_{x_i}$  with  $x_i \in D$  and  $\alpha_i > 0$ . To prove existence of a solution, we use the so-called Perron method: we define

$$\mathcal{S}[\nu] := \{v : \Omega \rightarrow \mathbb{R} \text{ convex} : v|_{\partial \Omega} = 0, \mu_v \geq \nu \text{ in } \Omega\}$$

and we show that the largest element in  $\mathcal{S}[\nu]$  is the desired solution. We split the argument into several steps.

*Step 1:*  $\mathcal{S}[\nu] \neq \emptyset$ . To construct an element of  $\mathcal{S}[\nu]$ , we consider the “conical” function  $C_{x_i}$ , that is 0 on  $\partial \Omega$  and takes the value  $-1$  at its vertex  $x_i$ . The Monge–Ampère measure of this function is concentrated at  $x_i$  and has mass equal to some positive number  $\beta_i$  corresponding to the measure of the set of supporting hyperplanes at  $x_i$ . Now, consider

the convex function  $\bar{v} = \sum_{i=1}^N \lambda C_{x_i}$ , where  $\lambda$  has to be chosen. We notice that  $\bar{v}|_{\partial\Omega} = 0$ . In addition, provided  $\lambda$  is sufficiently large, Lemma below implies that

$$\mu_{\bar{v}} \geq \sum_{i=1}^N \mu_{\lambda \hat{C}_{x_i}} = \sum_{i=1}^N \lambda^d \mu_{\hat{C}_{x_i}} = \sum_{i=1}^N \lambda^d \beta_i \delta_{x_i} \geq \sum_{i=1}^N \alpha_i \delta_{x_i} = \nu.$$

This yields  $\bar{v} \in \mathcal{S}[\nu]$ .

*Step 2:*  $v_1, v_2 \in \mathcal{S}[\nu] \Rightarrow w := \max\{v_1, v_2\} \in \mathcal{S}[\nu]$ . Set

$$\Omega_0 := \{v_1 = v_2\}, \quad \Omega_1 := \{v_1 > v_2\}, \quad \text{and} \quad \Omega_2 := \{v_1 < v_2\}$$

Also, given a Borel set  $E \subseteq \Omega$ , consider  $E_i = E \cap \Omega_i$ .

Since  $\Omega_1$  and  $\Omega_2$  are open sets,  $w|_{\Omega_1} = v_1$  and  $w|_{\Omega_2} = v_2$ ,

$$\partial w(E_1) = \partial v_1(E_1), \quad \partial w(E_2) = \partial v_2(E_2).$$

In addition, since  $w = v_1$  on  $\Omega_0$  and  $w \geq v_1$  everywhere else, we have

$$\partial v_1(E) \subseteq \partial w(E_0).$$

Therefore,

$$\mu_w(E) \geq \mu_{v_1}(E_0 \cup E_1) + \mu_{v_2}(E_2) \geq \nu(E).$$

*Step 3:*  $u := \sup_{v \in \mathcal{S}[\nu]} v$  belongs to  $\mathcal{S}[\nu]$ . Let  $w_m \uparrow u$  locally uniformly. Then  $\mu_{w_m} \rightharpoonup * \mu_u$ . Also, we deduce immediately that  $u|_{\partial\Omega} = 0$  by construction; hence,  $u \in \mathcal{S}[\nu]$ .

*Step 4:* The measure  $\mu_u$  is supported at the points  $\{x_1, \dots, x_N\}$ . Otherwise, there exists a set  $E \subseteq D$  such that

$$E \cap \{x_1, \dots, x_N\} = \emptyset \quad \text{and} \quad |\partial u(E)| = \mu_u(E) > 0$$

Therefore,

$$|\partial u(E) \setminus [\cup_{i=1}^N \partial u(x_i) \cup \partial u(\partial D)]| = |\partial u(E)| > 0$$

Let  $x_0 \in E$  and  $p \in \partial u(x_0) \setminus [\cup_{i=1}^N \partial u(x_i) \cup \partial u(\partial D)]$ . Then there exists  $\delta > 0$  such that

$$u \geq \ell_{x_0, p} + 2\delta \quad \text{on} \quad \{x_1, \dots, x_N\} \cup \partial\Omega, \tag{B.2}$$

where  $\ell_{x_0, p}(x) = u(x_0) + p \cdot (x - x_0)$ . Set  $\bar{u} := \max\{\ell_{x_0, p} + \delta, u\} \geq u$ . Notice that  $\bar{u}$  is convex,  $\bar{u} \geq u$ , and it follows by (B.2) that  $\bar{u} = u$  in a neighborhood of  $\{x_1, \dots, x_N\} \cup \partial\Omega$ . In particular,  $\bar{u}|_{\partial\Omega} = 0$  and  $\partial \bar{u}(x_i) = \partial u(x_i)$ , which implies that  $u \log eqq \bar{u} \in \mathcal{S}[\nu]$ . This is a contradiction.

*Step 5:*  $\mu_u = \nu$ . By *Step 3* and *Step 4*, we know that  $\mu_u = \sum_{i=1}^N \beta_i \delta_{x_i}$  with  $\beta_i \geq \alpha_i$ . Assume that  $\beta_1 = \mu_u(x_1) > \nu(x_1) = \alpha_1$ .

Since  $\partial u(x_j)$  is a convex set of positive measure, pick a vector  $p \in \mathbb{R}^n$  that belongs to the interior of  $\partial u(x_j)$ , define  $\ell_{x_j, p}(z) := u(z) + \langle p, z - x_j \rangle$ , and consider the function  $U := u - \ell_{x_j, p}$ . Notice that  $\partial U(z) = \partial u(z) - p$  for all  $z$ , which implies, in particular, that  $|\partial U(x_j)| = \beta_j$  and that 0 belongs to the interior of  $\partial U(x_j)$ . Choose  $\delta > 0$  small enough so that

$$\{U \leq U(x_j) + \delta\} \cap (\{x_1, \dots, x_N\} \cup \partial\Omega) = \emptyset,$$

and define the function

$$\tilde{U}(z) := \begin{cases} U(z) & \text{if } U > U(x_j) + \delta \\ (1 - \delta)U(z) + \delta[U(x_j) + \delta] & \text{if } U \leq U(x_j) + \delta. \end{cases}$$

Taking  $\delta > 0$  even smaller if necessary, we observe that

$$|\partial \tilde{U}(x_j)| = (1 - \delta)^n |\partial U(x_j)| = (1 - \delta)^n \beta_j > \alpha_j.$$

Notice that  $\tilde{U} \geq U$  and  $\tilde{U} = U$  in a neighborhood of  $\{x_1, \dots, x_N\} \cup \partial\Omega$ . Hence, considering  $\tilde{u} := \tilde{U} + \ell_{x_j, p}$  (see Figure 2.7), we see that  $\tilde{u} \geq u$ ,

$$\tilde{u}|_{\partial\Omega} = u|_{\partial\Omega} = 0, \quad \partial \tilde{u}(x_i) = \partial u(x_i) \quad \forall i \neq j, \quad \text{and} \quad |\partial \tilde{u}(x_j)| > \alpha_j.$$

Thus,  $\tilde{u} \in \mathcal{S}[v]$ , but this contradicts the maximality of  $u$  and concludes the proof.  $\square$

## C.2 Schauder estimate

Let  $\mathcal{S}$  be the Schwartz space of all rapidly decreasing functions, and  $\mathcal{S}'$  the dual space of  $\mathcal{S}$  called Schwartz generalized function (or tempered distribution) space. Given  $f \in \mathcal{S}$ , let  $\mathcal{F}f = \hat{f}$  be the Fourier transform defined by

$$\hat{f}(\xi) := \int_{\mathbb{R}^d} e^{-i2\pi\xi \cdot x} f(x) dx.$$

Let  $\chi : \mathbb{R}^d \rightarrow [0, 1]$  be a smooth radial function with

$$\chi(\xi) = 1, \quad |\xi| \leq 1, \quad \chi(\xi) = 0, \quad |\xi| \geq 3/2.$$

Define

$$\varphi(\xi) := \chi(\xi) - \chi(2\xi).$$

It is easy to see that  $\varphi \geq 0$  and  $\text{supp } \varphi \subset B_{3/2} \setminus B_{1/2}$  and formally

$$\sum_{j=-k}^k \varphi(2^{-j}\xi) = \chi(2^{-k}\xi) - \chi(2^{k+1}\xi) \xrightarrow{k \rightarrow \infty} 1. \quad (\text{B.3})$$

In particular, if  $|j - j'| \geq 2$ , then

$$\text{supp } \varphi(2^{-j}\cdot) \cap \text{supp } \varphi(2^{-j'}\cdot) = \emptyset.$$

From now on we shall fix such  $\chi$  and  $\varphi$  and define

$$\Delta_j f := \mathcal{F}^{-1}(\varphi(2^{-j}\cdot)\mathcal{F}f), \quad j \in \mathbb{Z}.$$

We first recall the following useful lemmas.

---

**Lemma C.2.1.** *Let  $\alpha \in (0, 1)$ . For any  $u \in C^\alpha$ , it holds that*

$$\frac{1}{C} \sup_{j \in \mathbb{Z}} 2^{j\alpha} \|\Delta_j u\|_\infty \leq [u]_\alpha \leq C \sup_{j \in \mathbb{Z}} 2^{j\alpha} \|\Delta_j u\|_\infty, \quad (\text{B.4})$$

where

$$[u]_\alpha := \sup_{x \neq y} \frac{|u(x) - u(y)|}{|x - y|^\alpha},$$

and  $C$  only depends on  $d$  and  $\alpha$ .

*Proof.* □

**Lemma C.2.2.** *There is a constant  $C = C(d, \alpha)$ , such that for any  $u \in C^{2,\alpha}$ ,*

$$[\nabla^2 u]_\alpha \leq C[\Delta u]_\alpha. \quad (\text{B.5})$$

*Proof.* Define

$$\varphi^{kl}(\xi) := \frac{\xi_k \xi_l}{|\xi|^2} \varphi(\xi), \quad h^{kl}(x) := \mathcal{F}^{-1}(\varphi^{kl})(x); \quad \varphi_j^{kl}(\xi) := \varphi^{kl}(2^{-j}\xi), \quad h_j^{kl}(x) := 2^{jd} h^{kl}(2^j x).$$

It is easy to see

$$\partial_{kl} u = \sum_{j \in \mathbb{Z}} u_j^{kl} := \sum_{j \in \mathbb{Z}} \varphi_j^{kl}(D) f = \sum_{j \in \mathbb{Z}} h_j^{kl} * f,$$

For any  $k, l \in \{1, 2, \dots, d\}$ , there is a constant  $C$  only depending on  $\alpha, d$  such that

$$\|u_j^{kl}\|_\infty \leq C 2^{-j\alpha} [f]_\alpha, \quad \forall j \in \mathbb{Z}. \quad (\text{B.6})$$

For any  $x \in \mathbb{R}^d$ , noticing  $h^{kl} \in \mathcal{S}(\mathbb{R}^d)$  and  $\int h^{kl} = \varphi(0) = 0$ , we get

$$\begin{aligned} |u_j^{kl}(x)| &= \left| \int_{\mathbb{R}^d} h_j^{kl}(y) f(x - y) dy \right| = \left| \int_{\mathbb{R}^d} h^{kl}(z) (f(x - 2^{-j}z) - f(x)) dz \right| \\ &\leq \int_{\mathbb{R}^d} |h^{kl}(z)| \cdot [f]_\alpha |2^{-j}z|^\alpha dz \leq C[f]_\alpha 2^{-j\alpha}. \end{aligned}$$

By this,

$$\begin{aligned} |\partial_{kl} u(x) - \partial_{kl} u(y)| &\leq \sum_{j \leq K} |u_j^{kl}(x) - u_j^{kl}(y)| + \sum_{j > K} |u_j^{kl}(x) - u_j^{kl}(y)| \\ &\leq |x - y| \cdot \sum_{j \leq K} \|\nabla u_j^{kl}\|_\infty + 2 \sum_{j > K} \|u_j^{kl}\|_\infty \\ &\leq C[f]_\alpha |x - y| \sum_{j \leq K} 2^{(1-\alpha)j} + C[f]_\alpha \sum_{j > K} 2^{-j\alpha} \\ &\leq C[f]_\alpha (|x - y| 2^{(1-\alpha)K} + 2^{-\alpha K}) \end{aligned}$$

Choosing  $K \in \mathbb{Z}$  such that  $2^{-K} \leq |x - y| < 2^{-K+1}$ , we obtain

$$|\partial_{kl} u(x) - \partial_{kl} u(y)| \leq C_1 [f]_\alpha \cdot |x - y|^\alpha.$$

So we complete our proof. □

**Lemma C.2.3.** Suppose  $f : \mathbb{R}_+ \rightarrow \mathbb{R}_+$ , and for any  $0 \leq s < t \leq 1$ ,

$$f(s) \leq \theta f(t) + A(t-s)^{-\gamma} + B,$$

for some  $\gamma > 0$ , then

$$f(s) \leq C(\gamma, \theta) (A(t-s)^{-\gamma} + B) \quad (\text{B.7})$$

*Proof.* Let  $t_0 = t$ ,  $t_i = s + (1-\tau)\tau^i(t-s)$ , where  $\tau \in (\theta^{1/\gamma}, 1)$ . By iteration,

$$\begin{aligned} f(s) &= f(t_0) \leq \theta f(t_1) + A(t_i - t_{i+1})^{-\gamma} + B \\ &\leq \theta^2 f(t_2) + \theta A(t_2 - t_1)^{-\gamma} + \theta B + A(t_1 - t_0)^{-\gamma} + B \\ &\leq \dots \leq \theta^k f(t_k) + A(1-\tau)^{-\gamma}(t-s)^{-\gamma} \sum_{i=0}^{k-1} (\theta\tau^{-\gamma})^i + B \sum_{i=0}^{k-1} \theta^i \\ &\leq C \left( \frac{A}{(t-s)^\gamma} + B \right). \end{aligned}$$

□

**Theorem C.2.4.** Suppose

$$Lu = a^{ij} \partial_{ij} u + b^i \partial_i u + cu = f, \quad \text{in } B_R, \quad (\text{B.8})$$

and

$$\delta |\xi|^2 \leq a^{ij} \xi_i \xi_j, \quad (\text{B.9})$$

$$\begin{aligned} &(\|a\|_{L^\infty(B_R)} + R^\alpha [a]_{\alpha; B_R}) + (R\|b\|_{L^\infty(B_R)} + R^{1+\alpha} [b]_{\alpha; B_R}) \\ &+ (R^2\|c\|_{L^\infty(B_R)} + R^{2+\alpha} [c]_{\alpha; B_R}) \leq \Lambda. \end{aligned} \quad (\text{B.10})$$

Then,

$$[\nabla^2 u]_{\alpha; B_{R/2}} \leq C \left\{ [f]_{\alpha; B_R} + R^{-\alpha} \|f\|_{L^\infty(B_R)} + R^{-2-\alpha} \|u\|_{L^\infty(B_R)} \right\}.$$

*Proof.* Suppose  $\eta \in C_c^\infty$ ,  $\eta(x) = 1$  if  $x \in B_\rho$ ,  $\eta(x) = 0$  if  $x \in B_r^c$  and

$$(r-\rho)^k \|\nabla^k \eta\|_\infty + (r-\rho)^{k+\alpha} [\nabla^k \eta]_\alpha \leq C(d, k).$$

Let  $v := u\eta$ , then

$$a_o^{ij} \partial_{ij} v = (a_o^{ij} - a^{ij}) \eta \cdot \partial_{ij} u + 2a_o^{ij} \partial_i u \partial_j \eta + a_o^{ij} u \partial_{ij} \eta + (b^i \partial_i u) \cdot \eta + (cu) \cdot \eta + f\eta.$$

$$\begin{aligned} [\nabla^2 u]_{\alpha; B_\rho} &\leq [\nabla^2 v]_\alpha \leq C_1 \left\{ r^\alpha [\nabla^2 u]_{\alpha; B_r} + (r-\rho)^{-\alpha} \|\nabla^2 u\|_{L^\infty(B_r)} \right. \\ &\quad + (r-\rho)^{-1} [\nabla u]_{\alpha; B_r} + (r-\rho)^{-1-\alpha} \|\nabla u\|_{L^\infty(B_r)} + (r-\rho)^{-2} [u]_{\alpha; B_r} \\ &\quad \left. + (r-\rho)^{-2-\alpha} \|u\|_{L^\infty(B_r)} + [f]_{\alpha; B_r} + (r-\rho)^{-\alpha} \|f\|_{L^\infty(B_r)} \right\} \end{aligned}$$

There is a constant  $r_0 \in (0, 1)$  such that  $C_1 r_0^\alpha \leq 1/4$ . By interpolation, there is a constant  $C$  such that

$$C_1 \left\{ (r-\rho)^{-\alpha} \|\nabla^2 u\|_{L^\infty(B_r)} + (r-\rho)^{-1} [\nabla u]_{\alpha; B_r} + (r-\rho)^{-1-\alpha} \|\nabla u\|_{L^\infty(B_r)} + (r-\rho)^{-2} [u]_{\alpha; B_r} \right.$$

---


$$+ (r - \rho)^{-2-\alpha} \|u\|_{L^\infty(B_r)} \Big\} \leq \frac{1}{4} [\nabla^2 u]_{\alpha; B_r} + C(r - \rho)^{-2-\alpha} \|u\|_{L^\infty(B_r)}$$

Combing the above inequalities and [C.2.3](#), we obtain that for any  $0 \leq \rho < r \leq r_0$ ,

$$[\nabla^2 u]_{\alpha; B_\rho} \leq C \left\{ (r - \rho)^{-2-\alpha} \|u\|_{L^\infty(B_r)} + [f]_{\alpha; B_r} + (r - \rho)^{-\alpha} \|f\|_{L^\infty(B_r)} \right\}.$$

Hence, by choosing  $\rho = \frac{r_0}{2}$ ,  $r = r_0$  and using finite cover technique, we find if  $u$  satisfies  $Lu = f$  in  $B_1$  and

$$\lambda |\xi|^2 \leq a^{ij} \xi_i \xi_j; \quad \|a\|_{C^\alpha(B_1)} + \|b\|_{C^\alpha(B_1)} + \|c\|_{C^\alpha(B_1)} \leq \Lambda,$$

then

$$[\nabla^2 u]_{\alpha; B_{1/2}} \leq C \left\{ [f]_{\alpha; B_1} + \|f\|_{L^\infty(B_1)} + \|u\|_{L^\infty(B_1)} \right\}.$$

By rescaling, it is easy to obtain our result. □





# Bibliography

- [Bas98] Richard F Bass. *Diffusions and elliptic operators*. Springer Science & Business Media, 1998.
- [Bis81] Jean-Michel Bismut. Martingales, the malliavin calculus and hypoellipticity under general hörmander’s conditions. *Zeitschrift für Wahrscheinlichkeitstheorie und verwandte Gebiete*, 56(4):469–505, 1981.
- [Dur19] Rick Durrett. *Probability: theory and examples*, volume 49. Cambridge university press, 2019.
- [Eng91] Hans-Jürgen Engelbert. On the theorem of t. yamada and s. watanabe. *Stochastics: An International Journal of Probability and Stochastic Processes*, 36(3-4):205–216, 1991.
- [Hua01] Zhiyuan Huang. *Fundamentals of stochastic analysis, version 2 (in Chinese)*. Science Press, 2001.
- [IW14] Nobuyuki Ikeda and Shinzo Watanabe. *Stochastic differential equations and diffusion processes*, volume 24. Elsevier, 2014.
- [Kry09] N. V. Krylov. *Controlled diffusion processes*, volume 14 of *Stochastic Modelling and Applied Probability*. Springer-Verlag, Berlin, 2009. Translated from the 1977 Russian original by A. B. Aries, Reprint of the 1980 edition.
- [Kry21a] NV Krylov. On diffusion processes with drift in  $L_d$ . *Probability Theory and Related Fields*, 179(1):165–199, 2021.
- [Kry21b] NV Krylov. On strong solutions of Itô’s equations with  $\sigma \in W_d^1$  and  $b \in L_d$ . *The Annals of Probability*, 49(6):3142–3167, 2021.
- [KS79] N. V. Krylov and M. V. Safonov. An estimate for the probability of a diffusion process hitting a set of positive measure. *Dokl. Akad. Nauk SSSR*, 245(1):18–20, 1979.
- [KS81] N V Krylov and M V Safonov. A certain property of solutions of parabolic equations with measurable coefficients. *Mathematics of the USSR-Izvestiya*, 16(1):151, feb 1981.

## BIBLIOGRAPHY

---

- [Nua06] David Nualart. *The Malliavin calculus and related topics*, volume 1995. Springer, 2006.
- [RY13] Daniel Revuz and Marc Yor. *Continuous martingales and Brownian motion*, volume 293. Springer Science & Business Media, 2013.
- [Yan21] Jiaan Yan. *Lecture note on Measure Theory*. Science Press, China, third edition, 2021.