Stochastic Differential Equations and Applications

Guohuan Zhao

May 13, 2024

Contents

Notation:

- $\mathbb{N} = \{0, 1, 2, \cdots\}\}$, $\mathbb{R}_+ = [0, \infty)$.
- $\mathscr{L}^d$ is the Lebesgue measure on $\mathbb{R}^d$.
- $\mathbb{S}^d_+$ is the collection of all symmetric non-negative matrix in $\mathbb{R}^{d\times d}$. For any $\delta \in (0, 1]$, $\mathbb{S}^d_\delta := \left\{a \in \mathbb{S}^d_+ : a_{ij} = a_{ji}, \delta I_d \leqslant a \leqslant \delta^{-1} I_d\right\}$.
- We use $:=$ as a way of definition.
- 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 $a \leqslant Cb$ ($a \geqslant Cb$) for some constant C.
- W_t is a Brownian motion staring from 0 and W^x_t a Brownian motion staring from x.
- τ_D is the first exit time of a process from domain D; σ_Γ is the first hitting time of Γ.

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. $X : (\Omega, \mathcal{F}) \to (E, \mathcal{E})$ a measurable map, and $\mathcal{G}$ a σ-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 $A \in \mathcal{G}$, $\mathbf{E}(X; A) = \mathbf{E}(Y; A)$.

$Q_{\mathcal{G}} : \Omega \times \mathcal{E} \to [0, 1]$ 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*

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

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

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

(iii) (Jesen's inequality) If φ is a convex function, then

$$\mathbf{E}(\varphi(X)|\mathcal{G}) \leqslant \varphi(\mathbf{E}(X|\mathcal{G})). \tag{1.3}$$

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

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

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

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

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

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

$$\mathbf{E}f(X) = -\int_0^\infty f'(t)\mathbf{P}(X \leqslant t)\mathrm{d}t. \tag{1.5}$$

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 regard as a measurable map from $(\Omega, \mathcal{F}, \mathbf{P})$ to the space of functions $E^{\mathbf{T}}$. The Kolmogorov σ-field on $E^{\mathbf{T}}$ is the smallest σ-field making the projections $\pi_t : E^{\mathbf{T}} \ni f \mapsto f(t) \in E$ measurable. This definition ensures that a random map $\Omega \ni \omega \mapsto X_\cdot(\omega) \in E^{\mathbf{T}}$ is measurable if its component random variables $X_t : \Omega \to E$ are measurable for all $t \in \mathbf{T}$. Therefore, the mapping $\omega \mapsto X_\cdot(\omega)$ 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 $(X_{t_1}, X_{t_2},$
$..., X_{t_n})$ for every n and every finite subset $I = (t_1, t_2, ..., t_n)$ of $\mathbf{T}$. These can be specified as probability measures μ_I on $\mathbb{R}^n$. These μ_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 μ_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 (μ_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 $\mathbb{P}$ on $(E^{\mathbf{T}}, \mathcal{B}(E^{\mathbf{T}}))$ which is uniquely determined by the family (μ_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 $t_1, ..., t_n \in \mathbf{T}$ a probability measure $\mu_{t_1\cdots t_n}$ on E^n, and that these probability measures satisfy:*

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

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

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

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

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

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 $s \leqslant t, W_t - W_s$ *is a mean zero Gaussian random variable with variance* $t - s$*;*

(c) for all $s < t, W_t - W_s$ *is independent of* $\sigma\left(W_r; r \leqslant s\right)$*;*

(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 σ-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* $\mathbb{R}^n$*,* $p \geqslant 1$ *and* $s \in (n/p, 1)$*. Let* $f : D \to \mathbb{R}^d$ *be a measurable function. Assume*

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

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

$$\sup_{x,y\in D} \frac{|\tilde{f}(x) - \tilde{f}(y)|^p}{|x - y|^{sp-n}} \leqslant C \left(\iint_{D\times D} \frac{|f(x) - f(y)|^p_{\mathcal{B}}}{|x - y|^{n+sp}} \,\mathrm{d}x\mathrm{d}y\right)^{1/p}, \tag{1.6}$$

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*

$$\mathbf{E}|X_s - X_t|^p \leqslant c|t-s|^{1+\beta p}, \qquad \forall t, s \in I. \tag{1.7}$$

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

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

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 $Y : \Omega \times I \to \mathcal{B}$, such that for each $\omega \notin \mathcal{N}$,

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

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

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

By Fubini theorem, there exists a $\mathscr{L}$-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 $X_{t_n} \xrightarrow[I\setminus N \ni t_n \to t_0]{\mathbf{P}} X_{t_0}$. On the other hand, $X_{t_n} \doteq Y_{t_n} \to Y_{t_0}$, so we have $X_{t_0} \doteq Y_{t_0}$. Therefore, Y is a version of X. □

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

We already observed that as a consequence of Kolmogorov's continuity theorem, the Brownian paths are α-Hölder continuous for every $\alpha \in \left(0, \frac{1}{2}\right)$. 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 1.2.5 (law of iterated logarithm)**.** *Let $(B_t)_{t\geqslant 0}$ be a Brownian motion. For $s \geqslant 0$,*

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

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

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

Let us first prove that

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

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 $\left(e^{\alpha B_t-\frac{\alpha^2}{2}t}\right)_{t\geqslant 0}$, we have for $t\geqslant 0$:

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

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 $n\to+\infty$,

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

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

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

But,

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

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

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

We conclude:

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

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

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

Let us now prove that

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

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

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

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

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

implies

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

with

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

When $n \to +\infty$,

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

therefore,

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

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

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

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

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

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

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

We finally get

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

by letting $\theta \to 0$. □

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

Corollary 1.2.6. *Let $(B_t)_{t\geqslant 0}$ be a standard Brownian motion.*

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

1.3 Martingale

Let $(\Omega, \mathcal{F}, \mathbf{P})$ be a standard probability space. Let $\mathcal{F}_n$ $(n \in \mathbb{N})$ be an increasing sequence of σ-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 $\mathcal{F}_{t+} = \cap_{\varepsilon>0}\mathcal{F}_{t+\varepsilon}$) and each $\mathcal{F}_t$ is **complete** (i.e., $\mathcal{F}_t$ contains all $\mathbf{P}$-null sets).

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

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

Definition 1.3.1. *Let X_t be a real-valued $\mathcal{F}_t$-adapted processes. If for each t and $s < t$, X_t is integrable and $\mathbf{E}(X_t|\mathcal{F}_s) \geqslant (\leqslant) X_s$ a.s., then we call X_t is a* **submartingale (supermartingale)***. We say X_t is a* **martingale** *if it is both a submartingale and a supermartingale.*

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

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.

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, $A_n = 0$, $A_{n+1} \geqslant A_n$ almost surely and A_n is $\mathcal{F}_{n-1}$-measurable.*

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

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

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

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

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

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

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

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

$$\mathbf{E}(X_{\tau\wedge n}; A) = \mathbf{E}(X_{\tau\wedge n}; A \cap \{\tau \leqslant n-1\}) + \mathbf{E}(X_{\tau\wedge n}; A \cap \{\tau > n-1\})$$

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

□

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

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

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

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

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

□

Corollary 1.3.6. *Let M_n be a martingale and T be a stopping time. For each $p>1$, $\mathbf{E}|M_T^*|^p\leqslant C_p\mathbf{E}|M_T|^p$.*

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

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

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

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

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

□

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 $n\to\infty$.*

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

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

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 σ-field generated by the sets

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

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 μ, and instead of decomposing into equal subcubes, we could use any nested partition of sets we like, provided none of these sets had μ measure 0.

All of the above results also hold for all right continuous martingale (submartingales) (see [Hua01]).

Theorem 1.3.10. *Assume X is a continuous submartingale, then there exists a unique martingale M and a unique continuous increasing adapted process A such that*

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

If M is a continuous square integrable martingale, then M^2 is a submartingale. Thus, there exists a continuous increasing process, denoted by $\langle M \rangle$, the **quadratic variation** of M, such that $M^2 - \langle M \rangle$ is a martingale. Particularly, $\mathbf{E}M_t^2 - \mathbf{E}M_0^2 = \mathbf{E}\langle M \rangle_t$.

1.4 Stochastic Integral

From now on, unless stated otherwise, our processes have continuous paths.

Lemma 1.4.1. *Let M_t be a square integrable martingale (that is, $M_t \in L^2$ for every $t \geqslant 0$). Let $0 \leqslant s < t$ and let $s = t_0 < t_1 < \cdots < t_n = t$ be a division of the interval $[s,t]$. Then,*

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

Proof. For every $i = 1, \ldots, n$,

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

and the desired result follows by summing over i. □

We say that M_t if a **local martingale** if there exist stopping times $\tau_n \uparrow \infty$ such that $X_{\tau_n \wedge t}$ is a martingale for each $n \in \mathbb{N}$.

Theorem 1.4.2. *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 $\pi^n = \{(t_0^n, \cdots, t_{k_n}^n) : 0 = t_0^n < t_1^n < \cdots < t_{k_n}^n = t\}$ is an increasing sequence of subdivisions of $[0, t]$ with mesh going to 0 , then we have*

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

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

Lemma 1.4.3. *Let M_t be a continuous bounded martingale. Let $\pi^n = \{(t_0^n, \cdots, t_{k_n}^n) : 0 = t_0^n < t_1^n < \cdots < t_{k_n}^n = T\}$ be an increasing sequence of subdivisions of $[0, T]$ with mesh going to 0 , then for each n,*

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

is a martingale, and N_t^n convergent uniformly on compacts, with probability one to some square integrable martingale N_t.

Proof. It is easy to verify that N_t^n is a martingale. Let us fix $n \leqslant m$ and evaluate the product $\mathbf{E}\left(N_T^n N_T^m\right)$. 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 $\left(t_{j-1}^m, t_j^m\right]$ is contained in $\left(t_{i-1}^n, t_i^n\right]$. Indeed, suppose that $t_i^n \leqslant t_{j-1}^m$ (the symmetric case $t_j^m \leqslant t_{i-1}^n$ is treated in an analogous way).

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

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

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

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

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

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

As a special case of this relation, we have

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

Furthermore,

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

If we combine the last three displays, we get

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

Using the Cauchy-Schwarz inequality, we then have

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

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

$$\lim_{n,m\to\infty, n\leqslant m} \mathbf{E}\left[\sup_{1\leqslant j\leqslant p_m, i=i_{n,m}(j)} \left(M_{t^n_{i-1}} - M_{t^m_{j-1}}\right)^4\right] = 0.$$

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,

$$\mathbf{E}\left[\left(\sum_{1\leqslant j\leqslant p_m} \left(M_{t^m_j} - M_{t^m_{j-1}}\right)^2\right)^2\right] \leqslant C.$$

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

$$\begin{aligned}
&\mathbf{E}\left[\left(\sum_{1\leqslant j\leqslant p_m} \left(M_{t^m_j} - M_{t^m_{j-1}}\right)^2\right)^2\right] \\
&= \mathbf{E}\left[\sum_{1\leqslant j\leqslant p_m} \left(M_{t^m_j} - M_{t^m_{j-1}}\right)^4\right] + 2\mathbf{E}\left[\sum_{1\leqslant j<k\leqslant p_m} \left(M_{t^m_j} - M_{t^m_{j-1}}\right)^2 \left(M_{t^m_k} - M_{t^m_{k-1}}\right)^2\right] \\
&\leqslant 4A^2\mathbf{E}\left[\sum_{1\leqslant j\leqslant p_m} \left(M_{t^m_j} - M_{t^m_{j-1}}\right)^2\right] \\
&\quad + 2\sum_{j=1}^{p_m-1} \mathbf{E}\left[\left(M_{t^m_j} - M_{t^m_{j-1}}\right)^2 \mathbf{E}\left[\sum_{k=j+1}^{p_m} \left(M_{t^m_k} - M_{t^m_{k-1}}\right)^2 \mid \mathscr{F}_{t^m_j}\right]\right] \\
&= 4A^2\mathbf{E}\left[\sum_{1\leqslant j\leqslant p_m} \left(M_{t^m_j} - M_{t^m_{j-1}}\right)^2\right] \\
&\quad + 2\sum_{j=1}^{p_m-1} \mathbf{E}\left[\left(M_{t^m_j} - M_{t^m_{j-1}}\right)^2 \mathbf{E}\left[\left(M_T - M_{t^m_j}\right)^2 \mid \mathscr{F}_{t^m_j}\right]\right]
\end{aligned}$$

□

Proof of Theorem 1.4.2. □

Let M_t be a square integrable martingale, $0 = t_0 \leqslant t_1 \leqslant \cdots \leqslant t_n = T$ 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. Define

$$\int_0^t H_s \mathrm{d}M_s := \sum_{i=0}^{n-1} F_i(M_{t\wedge t_{i+1}} - M_{t\wedge t_i}).$$

Then

Lemma 1.4.4. *$t \mapsto \int_0^t H_s \mathrm{d}M_s$ is a L^2-martingale. Moreover, we have the following Itô isometry:*

$$\mathbf{E}\left(\int_0^t H_s \mathrm{d}M_s\right)^2 = \mathbf{E}\int_0^t H_s^2 \mathrm{d}\langle M\rangle_s. \tag{1.8}$$

Proof.

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

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

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

Therefore,

$$\mathbf{E}\left(\int_0^1 H_s \mathrm{d}M_s\right)^2 = \mathbf{E}\int_0^1 H_s^2 \mathrm{d}\langle M\rangle_s.$$

$t \mapsto \int_0^t H_s \mathrm{d}W_s = \sum_{i=0}^{n-1} H_{t_i}(W_{t\wedge t_{i+1}} - W_{t\wedge t_i})$ is a continuous martingale. □

We then can use this to extend the above construction to more general H_s satisfying $\int_0^t H_s^2 \mathrm{d}\langle M\rangle_s < \infty$ by taking limits in L^2. 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, $\int_0^t H_s \mathrm{d}X_s$ is given by

$$\int_0^t H_s \mathrm{d}X_s = \int_0^t H_s \mathrm{d}M_s + \int_0^t H_s \mathrm{d}A_s,$$

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

Proposition 1.4.5.

$$\left\langle \int_0^\cdot H_s \mathrm{d}M_s \right\rangle_t = \int_0^t H_s^2 \mathrm{d}\langle M\rangle_s.$$

Let $N_t = \int_0^t H_s \mathrm{d}M_s$

$$\int_0^t K_s \mathrm{d}N_s = \int_0^t K_s H_s \mathrm{d}M_s$$

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 $i \in 1, \cdots d\}$) is a continuous semimartingale and $f \in C^2(\mathbb{R}^d)$, then*

$$\begin{aligned} & f\left(X_t\right)-f\left(X_0\right) \\ = & \int_0^t \sum_{i=1}^d \partial_i f\left(X_s\right) \mathrm{d} X_s^i+\frac{1}{2} \int_0^t \sum_{i, j=1}^d \partial_{i j} f\left(X_s\right) \mathrm{d}\left\langle X^i, X^j\right\rangle_s \end{aligned} \tag{1.9}$$

(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 $\mathbf{P}\left(\int_0^T H_s \mathrm{d}\langle M\rangle_s<\infty\right)=1$, then by definition the Stratonovitch integral of H_t with respect to M_t is defined as

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

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*

$$f(M_t)=f(M_0)+\int_0^t \partial_i f(X_s) \circ \mathrm{d} M_s^i, \quad t \geqslant 0 .$$

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

$$\mathbf{E} \sup_{t \in[0, \tau]}|M_t|^p \asymp_p \mathbf{E}\langle M\rangle_\tau^{p / 2}, \quad p \in(0, \infty) \tag{1.10}$$

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

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

By Doob's inequality and Hölder's inequality,

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

Step 2: using Lenglart'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

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

By using the previous argument, we now have

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

As a conclusion, we obtained that d □

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

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

We shall use this lemma to prove the following Another approach to proving (1.10) is utilizing "good-λ" inequality (cf. [RY13]).

Theorem 1.5.5 (Lévy's theorem)**.** *If X_t is a d-dimensional $\mathscr{F}$-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 $\mathscr{F}$-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, $\exp(\mathrm{i}\xi \cdot X_t + \frac{1}{2}|\xi|^2 t)$ 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 $s < t$,

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

Thus,

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

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 $\mathscr{F}$ so that X is a s-dimensional $\mathscr{F}$-Brownian motion. □

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

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

$$\mathrm{d}X_t = X_t \mathrm{d}M_t, \quad X_0 = 1.$$

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 $\mathscr{E}(M)_t$ is a martingale, we define a new probability measure $\mathbb{Q}$ by setting the restriction of $\mathrm{d}\mathbb{Q}/\mathrm{d}\mathbb{P}$ to $\mathcal{F}_t$ to be $\mathscr{E}(M)_t$, 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 $Y_t\mathscr{E}(M)_t$ is a martingale under $\mathbb{P}$. By Itô's formula,

$$\begin{aligned}\mathrm{d}Y_t\mathscr{E}(M)_t =& \mathscr{E}(M)_t\mathrm{d}X_t - \mathscr{E}(M)_t\mathrm{d}\langle X, M\rangle_t + Y_t\mathscr{E}(M)_t\mathrm{d}M_t + \mathrm{d}\langle X, \mathscr{E}(M)\rangle_t \\ =& \mathscr{E}(M)_t\mathrm{d}X_t + Y_t\mathscr{E}(M)_t\mathrm{d}M_t.\end{aligned}$$

Therefore, $Y_t\mathscr{E}(M)_t$ 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.

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

□

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, τ a bounded stopping time, then*

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

We need

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

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

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

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

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

$$\mathbf{P}\left(W_t^* > \lambda\right) \leqslant 2\mathrm{e}^{-\frac{\lambda^2}{2t}}.$$

□

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

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

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

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

□

1.5.2 Applications in PDEs

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

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

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

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:

Theorem 1.5.11. *Suppose u is a C_b^2 function satisfying the above Poisson equation. Then*

$$u(x) = \mathbf{E}\int_0^\infty e^{-\lambda t} f\left(X_t^x\right)\mathrm{d}t$$

Proof. Applying Itô's formula, we have $\mathrm{d}u(X_t^x) = \mathrm{d}M_t + Lu(X_t^x)\mathrm{d}t$. So

$$\begin{aligned} e^{-\lambda t}u\left(X_t^x\right) - u\left(x\right) = &\int_0^t e^{-\lambda s}\mathrm{d}M_s + \int_0^t e^{-\lambda s}Lu\left(X_s^x\right)\mathrm{d}s \\ &- \lambda\int_0^t e^{-\lambda s}u\left(X_s^x\right)\mathrm{d}s. \end{aligned}$$

Taking expection, we get what we claimed. □

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 $\lambda \geqslant 0$. Here we can allow λ 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 $\tau_D = \inf\{t > 0 : x \notin D\}$. Then

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

Exercise 1.5.1. *Prove Theorem 1.5.12.*

Chapter 2

Itô processes

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

$$x_t = \int_0^t \sigma_s \cdot \mathrm{d}W_s + \int_0^t b_s \mathrm{d}s. \tag{2.1}$$

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

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

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

This can be phrased as saying the graph of x_t stays inside an ε-tube about φ.

To prove Theorem 2.1, we need some auxiliary lemmas.

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

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

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*

$$\mathbf{P}\left(\sup_{s\in[0,T]} |W_t| \leqslant \varepsilon\right) \geqslant c(\varepsilon, T).$$

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

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

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

$$\inf_{|x|\leqslant 1/3}\mathbf{P}\left(\sup_{t\in[0,\delta_2]}|W^x_t| \leqslant 1,\ |W^x_{\delta_2}| \leqslant 1/3\right) \geqslant \frac{1}{6}, \quad \delta_2 = \delta_0 \wedge \delta_1 > 0.$$

By the Markov property of W,

$$\inf_{|x|\leqslant 1/3}\mathbf{P}\left(\sup_{t\in[0,k\delta_2]}|W^x_t| \leqslant 1,\ |W^x_{\delta_2}| \leqslant 1/3\right) \geqslant 6^{-k}.$$

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

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

□

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

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

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 $Y_t := X_{\tau(t)} = B_t + \int_0^t b_s\mathrm{d}s$ with $|b_s| \leqslant C(N_1, N_2)$. Our assertion will follow if we can show

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

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

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

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

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

By Hölder's inequqlity,

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

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

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

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

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

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

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

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

Noting that

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

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

$$B_t := \int_0^t \sigma_s^{-1}\mathrm{d}n_s$$

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 udner $\mathbf{Q}$. Since

$$x_t - \varphi(t) = \int_0^t \sigma_s \mathrm{d}B_s + \int_0^t b_s \mathrm{d}s,$$

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

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

□

2.2 ABP estimate and Generalized Itô's formula

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

Proposition 2.2.1 (Alexsandroff). *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 $x \in B_2$,

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

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

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

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

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

In Appendix B.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:

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

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*

$$\mathbf{E} \int_0^{\tau_R(x)} f(x + x_t) \sqrt[d]{\det a_t}\, \mathrm{d}t \leqslant C(d) R \|f\|_{L^d(B_R)}. \tag{2.7}$$

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,

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

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

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

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

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

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

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

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

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 γ be a finite stopping time, then

$$\mathbf{E}\left[\int_\gamma^{\tau_R(x)} f\left(x + x_t\right) \sqrt[d]{\det a_t}\, \mathbf{1}_{\{\gamma \leqslant \tau_R(x)\}} \,\mathrm{d}t \middle| \mathcal{F}_\gamma\right] \leqslant C(d) R \|f\|_{L_d(B_R)}. \tag{2.8}$$

Theorem 2.2.4. *There are constants C, μ depending only on d, such that*

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

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

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

Lemma 2.2.5.

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

Proof. We can assume $x = 0$.

We claim that

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

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

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

So we get what we desired. $\square$

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

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

Using this to give another proof for (2.9).

Corollary 2.2.4 basically says that τ_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 τ_R.

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

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

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

$$\mathrm{d}\phi(x_t)=\nabla\phi(x_t)\cdot\sigma_t\mathrm{d}W_t+a_t^{ij}\partial_{ij}\phi(x_t)\mathrm{d}t,$$

which yields that

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

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

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

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

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

$$\mathbf{E}F(X) = -\int_0^\infty F'(t)\mathbf{P}(X \leqslant t)\mathrm{d}t,$$

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

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

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

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

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

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

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

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 τ^k, $k = 1,\cdots,N$, as the first exit time of x_t from $B_{k/N}$. Also set γ^k be the first exit times of x_t from $B_{N^{-1}}(x_{\tau^{k-1}})$ after τ^{k-1} ($\tau^{k-1} \leqslant \gamma^k \leqslant \tau^k$). obviously,

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

By the conditional version of (2.13),

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

Therefore,

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

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

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

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

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, δ, κ and naturally, also on γ, such that for any $R \in (0,\infty), x \in B_{\kappa R}$, and closed $\Gamma \subset B_R$ satisfying $|\Gamma| \geqslant \gamma |B_R|$ we have*

$$\mathbf{P}\left(\sigma_\Gamma(x) \leqslant \tau_R(x)\right) \geqslant q(\gamma),$$

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

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

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

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

where the constants C depend only on d, δ, κ. By denoting

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

□

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

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*

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

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

Hint: Using support theorem.

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

$$\mathbf{E}\int_0^\infty e^{-\lambda t} f\left(X_t\right) \mathrm{d}t \leqslant C\lambda^{\frac{d}{2p}-1}\left\|f\right\|_p. \tag{2.17}$$

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

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

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

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

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

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

If (2.18) holds, then

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

□

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

$$u(x_t) - u(x_0) = \int_0^t \nabla u(x_s)\sigma_s \mathrm{d}W_s + \int_0^t a_s^{ij}\partial_{ij}u(x_s)\mathrm{d}s \tag{2.19}$$

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,

$$u_\varepsilon(x_t) - u_\varepsilon(x_0) = \int_0^t \nabla u_\varepsilon(x_s)\sigma_s \mathrm{d}W_s + \int_0^t a_s^{ij}\partial_{ij}u_\varepsilon(x_s))\mathrm{d}s. \tag{2.20}$$

Fact: by Sobolev embedding theorem, we have

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

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

$$\begin{aligned}
&\mathbf{E} \sup_{t\in[0,T]} \left| \int_0^t \nabla u_\varepsilon(x_s)\sigma_s \mathrm{d}W_s - \int_0^t \nabla u_{\varepsilon'}(x_s)\sigma_s \mathrm{d}W_s \right|^2 \\
\leqslant& C\mathbf{E} \int_0^T |\nabla u_\varepsilon - \nabla u_{\varepsilon'}|^2(x_s)\mathrm{d}s \leqslant C\|\nabla u_\varepsilon - \nabla u_{\varepsilon'}\|_{L^{2d}}^2 \\
\overset{(2.21)}{\leqslant}& C\|u_\varepsilon - u_{\varepsilon'}\|_{W^{2,d}} \to 0, \quad \varepsilon, \varepsilon' \to 0.
\end{aligned}$$

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

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 $|b_t| \leqslant \mathfrak{b}(x_t)$ with $\mathfrak{b} \in L^d$.*

Chapter 3

Itô's stocahstic differential equations

3.1 Strong solutions

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

$$\mathrm{d}X_t^i = \sigma_k^i(X_t)\mathrm{d}W_t^k + b^i(X_t)\mathrm{d}t, \quad X_0 = \xi \in \mathcal{F}_0. \tag{3.1}$$

Given $(\Omega, \mathcal{F}, \mathbf{P}, \mathscr{F}, W_t)$, 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 σ 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 ξ 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

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

By (1.10) (or Doob's inequality) and Lipschitz condition on the coefficients,

$$\begin{aligned}
&\|\mathcal{T}(\xi)-\mathcal{T}(\eta)\|_B^2=\mathbf{E}\sup_{t\in[0,T]}|\mathcal{T}(\xi)_t-\mathcal{T}(\eta)_t|^2\\
\leqslant&2\mathbf{E}\sup_{t\in[0,T]}\left|\int_0^t(\sigma(\xi_s)-\sigma(\eta_s))\mathrm{d}W_s\right|^2+2\mathbf{E}\sup_{t\in[0,T]}\left|\int_0^t(b(\xi_s)-b(\eta_s))\mathrm{d}s\right|^2\\
\overset{(1.10)}{\leqslant}&C\mathbf{E}\int_0^T|\sigma(\xi_s)-\sigma(\eta_s)|^2\mathrm{d}s+C\mathbf{E}\left(\int_0^T|b(\xi_s)-b(\eta_s)|\mathrm{d}s\right)^2\\
\leqslant&C(T+T^2)\mathbf{E}\sup_{t\in[0,T]}|\xi_t-\eta_t|^2=C_1(T+T^2)\|\xi-\eta\|_B^2.
\end{aligned}$$

Choosing $T>0$ sufficiently small such that $C_1(T+T^2)\leqslant 1/2$, 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. □

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 $(\Omega,\mathcal{F},\mathbf{P},\mathscr{F};X_t,W_t)$ such that W_t is a $\mathscr{F}$-Brownian motion and the equation (3.1) holds.

3. **weak uniqueness**: if whenever $(\Omega,\mathcal{F},\mathbf{P},\mathscr{F};X_t,W_t)$ and $(X,\mathcal{G},\mathbf{Q},\mathscr{G};Y_t,B_t)$ 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 σ is non-degenerate, then the condition of b can be weakened.

Theorem 3.1.3 (Krylov [Kry21b]). *Suppose that $a \in \mathbb{S}^d_\delta$ 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 simper one below.

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

Let

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

Consider

$$\lambda u - Lu = f. \tag{3.2}$$

We need the following analytic result.

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

$$\lambda\|u\|_\alpha + \|\nabla^2 u\|_\alpha \leqslant C\|f\|_\alpha, \tag{3.3}$$

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

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

(i) If $L = \Delta$ and $f \in \mathscr{S}(\mathbb{R}^d)$, then for each $\lambda > 0$, one can use Fourier transformation to solve (3.2), i.e. $u = \mathcal{F}^{-1}\left[\mathcal{F}(f) \cdot (\lambda + 4\pi^2|\cdot|^2)\right] \in \cap_{s>0} H^s \subseteq C_b^\infty$. Moreover, (3.3) can also be proved by Fourier analysis method (see Appendix B.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 λ sufficiently large via frozen coefficient method;

(iii) Let χ be a cutoff function and ζ 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_ε, to (3.2) with L and f replaced by Δ and f_ε. The limit of (u_ε), 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 $x \in B$,*

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

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

Proof of Theorem 3.1.4. Since σ 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 $\lambda \geqslant \lambda_0$. Here λ_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

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

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

$$\mathrm{d}Y_t^i = (\delta_j^i + \partial_j \mathbf{u}^i)(X_t)\sigma_{jk}(X_t)\mathrm{d}W_t^k + \left[a_{jk}(X_t)\partial_{jk}\mathbf{u}^i(X_t) + (\delta_j^i + \partial_j \mathbf{u}^i)(X_t)b^j(X_t)\right]\mathrm{d}t$$

i.e.

$$\begin{aligned}
\mathrm{d}Y_t =& [(I + \nabla \mathbf{u})\sigma] \circ \phi^{-1}(Y_t)\mathrm{d}W_t + \left[a : \nabla^2 \mathbf{u} + (I + \nabla \mathbf{u})b\right] \circ \phi^{-1}(Y_t)\mathrm{d}t \\
=& \underbrace{[(I + \nabla \mathbf{u})\sigma] \circ \phi^{-1}}_{=:\widetilde{\sigma}}(Y_t)\mathrm{d}W_t + \underbrace{\lambda u \circ \phi^{-1}}_{=:\widetilde{b}}(Y_t)\mathrm{d}t.
\end{aligned}$$

Similarly, $\mathrm{d}Y'_t = \widetilde{\sigma}(Y'_t)\mathrm{d}W_t + \widetilde{b}(Y'_t)\mathrm{d}t$. Since $\widetilde{\sigma}$ and $\widetilde{b}$ are both $C^{1,\alpha}$ functions, as in the proof for Theorem 3.1.1, we have

$$\mathbf{E}|Y_t - Y'_t|^2 \leqslant C\int_0^t \mathbf{E}|Y_s - Y'_s|^2 \mathrm{d}s.$$

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

3.1.4 Stochastic Flow

Consider (3.1) .

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

Proof.

$$X(t,x) - X(t,y) = x - y + \int_0^t [\sigma(X(s,x)) - \sigma(X(s,y))]\,\mathrm{d}W_s + \int_0^t [b(X(s,x)) - b(X(s,y))]\,\mathrm{d}s$$

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

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

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

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

$$
g(t)\leqslant C|x-y|^p+\int_0^t g(s)\mathrm{d}s,\quad t\in[0,1].
$$

Gronwall's inequality yields

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

Moreover,

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

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

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

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

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

The collection of processes $X(t,x)$ is called a flow. If σ 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

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

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

$$
\mathrm{d}J(t,x)=\partial_l\sigma(X(t,x))J^l(t,x)\mathrm{d}W_t+\nabla b(X(t,x))J(t,x)\mathrm{d}t,\quad Y_0=I \tag{3.4}
$$

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*

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

Exercise 3.1.1. *Prove the above Lemma.*

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

Theorem 3.1.10. *Suppose $\sigma, b \in C_b^k$. 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

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

and

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

Noting that

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

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

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

By Taylor expansion,

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

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

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

This together with Proposition 3.1.9 implies

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

On the other hand,

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

where

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

and

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

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

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

which together with (3.5) implies

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

Combining (3.6) and (3.7), we obtain

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

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

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

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:

$$\lambda u - Lu = f, \tag{3.8}$$

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

3.2.1 Uniqueness in law

Theorem 3.2.1 (Stroock-Varadhan)**.** *Under the assumptions that σ, b are bounded, σ 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 formual and L^p-estimate for the resolvent equation to show the uniqueness of law(X_t).

(c) Proving the law of $(X_t)_{t\geqslant 0}$ 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*

$$\lambda \|u\|_p + \|\nabla u^2\|_p \leqslant C\|f\|_p, \tag{3.9}$$

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 $\lambda \geqslant \lambda_0$ 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 sufficinetly large λ, it holds that

$$\lambda \|u\|_p + \|u\|_{W^{2,p}} \leqslant C\|\lambda u - Lu\|_p. \tag{3.10}$$

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 $\zeta \geqslant 0$, $\zeta \equiv 1$ in B_1. Set $\zeta^z_\varepsilon = \zeta((x-z)/\varepsilon)$. Then

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

By Lemma 3.2.2, we get

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

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

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

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

$$c\sum_i \int |h\zeta_\varepsilon^{z_i}|^p \leqslant \int |h|^p \leqslant C\sum_i \int |h\zeta_\varepsilon^{z_i}|^p. \tag{3.12}$$

By (3.11) and (3.12), we obtain

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

where C only depends on d, p and ω_a. Using intepolation theorem, one can see that

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

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

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

□

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 $\lambda \geqslant \lambda_0$. Applying Generalized Itô's formula, one can see that

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

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

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

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

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

Let $\mathcal{G}_t^0$ be the σ-algebra generated by $\{\omega_s : s \leqslant t\}$. We complete these σ-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 τ a finite stopping time.

To prove the strong Markov property it suffices to show

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

for all $x \in \mathbb{R}^d$, $f \in C_c(\mathbb{R}^d)$ and τ 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

$$Z_t = Z_0 + \int_0^t \sigma(Z_s)\,\mathrm{d}B_s + \int_0^t b(Z_s)\,\mathrm{d}s,$$

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

$$Z'_t = Z'_0 + \int_0^t \sigma(Z'_s)\,\mathrm{d}B'_s + \int_0^t b(Z'_s)\,\mathrm{d}s. \tag{3.14}$$

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 ω;
- $Z'_0(\omega') = Z_{\tau(\omega)}(\omega)$, $Q_\tau(\omega, \cdot)$-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 τ 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

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

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 σ 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,

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

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

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

and showed that

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

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 ∂D, find $u \in C(\bar{D})$ such that u is C^2 in D and

$$\begin{cases} Lu = 0 \text{ in } D \\ u = g \text{ on } \partial D. \end{cases} \tag{4.2}$$

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

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

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

Proof. Let $\tau_n = \inf\{t : \operatorname{dist}(X_t, \partial D) < 1/n\}$. 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 $t \to \infty$ and then $n \to \infty$. By dominated convergence, we obtain $u(x) = \mathbb{E}_x u(X_{\tau_D})$. This is what we want since $u = g$ on ∂D. □

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

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* ∂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,

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

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

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

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

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

One can also the following Schrödinger type operator:

$$L_q u = Lu + cu.$$

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 ∂D and satisfies $L_q u = 0$ in D. If*

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

then

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

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*

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

where c is a constant only depends on d, δ and ε.

4.2 Once again on the hitting probability

Recall that

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

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 $p : (0,1) \to (0,1)$, which only depends on d and δ, such that for any $\Gamma \subset Q_1$ and $x \in Q_{1/2}$,*

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

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 $\widetilde{Q}$ is the predecessor of Q if Q is one of the 2^n cubes obtained from dividing $\widetilde{Q}$.

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

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

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

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

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

We subdivide Q_1 into 2^n dyadic cubes. If Q is one of these 2^n subcubes of Q_1 and satisfies $|Q \cap \Gamma|/|Q| \leqslant \gamma$, we then split Q into 2^n dyadic cubes. We iterate this process. In this way, we pick a family $Q^1 \cdot Q^2 \ldots$ 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 $|Q \cap \Gamma|/|Q| \leqslant \gamma < 1$. Applying the Lebesgue differentiation theorem to $\mathbf{1}_A$, we get that $\mathbf{1}_A(x) \leqslant \gamma < 1$ 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 $\{\widetilde{Q}^i\}_{i \in \widetilde{\mathcal{I}}}$ are pairwise disjoint. We clearly have that

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

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

$$\frac{\left|\widetilde{Q}^i \cap \Gamma\right|}{\left|\widetilde{Q}^i\right|} \leqslant \gamma, \quad \forall i \in \widetilde{\mathcal{I}}.$$

We conclude that

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

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* $\widetilde{Q}$ *is the predecessor of* Q, *then for each* $x \in \widetilde{Q}(\kappa)$,

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

where $p'(\kappa)$ *only depends on* d, δ *and* κ.

Proof of Theorem 4.2.1. Define

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

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 $\gamma\in(0,b]$,

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

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

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

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

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

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

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

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

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

due to Lemma 4.2.3. Set

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

Then

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

The conditional version of above estimate we need below is

$$\mathbf{P}\left(\sigma'_B<\tau'_{Q_1}\middle|\mathcal{F}_{\sigma_A(x)}\right)\geqslant p'(\kappa)>0.\tag{4.5}$$

where

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

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

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

Set

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

Then

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

The conditional version of above estimate we need below is

$$\mathbf{P}\left(\sigma''_\Gamma < \tau''_{Q_1} \middle| \mathcal{F}_{\sigma'_B}\right) \geqslant p(\gamma) > 0. \tag{4.6}$$

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 $x \in Q_{1/2}$,

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

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

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

$$\begin{aligned}\mathscr{P}(\delta) := \Big\{\{\mathbb{P}_x\}_{x\in\mathbb{R}^n} : (\mathbb{P}_x, X) &\textit{ is the strong Markov process}\\ &\textit{associate with some } a(\cdot) \in \mathbb{S}^d_\delta\Big\}.\end{aligned}$$

Let

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

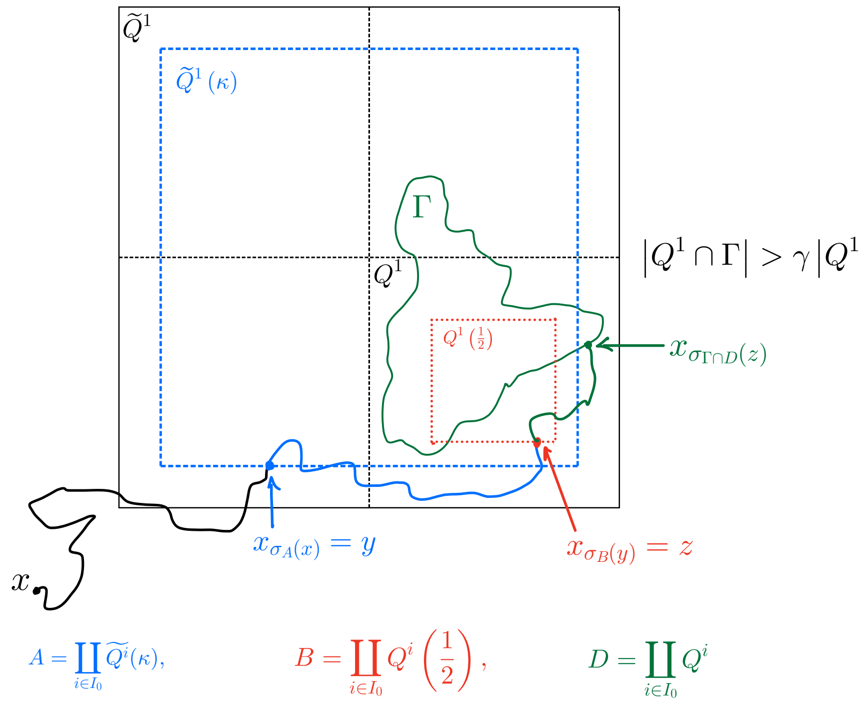

Figure 4.1: Hitting Prob.

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

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

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

$$\operatorname*{Osc}_{Q_{r/2}(z)} u \leqslant \rho \operatorname*{Osc}_{Q_r(z)} u. \tag{4.8}$$

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 $2^{-k-1} \leqslant |x-y| < 2^{-k}$.

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

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, $\operatorname{Osc}_{Q_r(z)} u = 1$. Let $B := \{x \in Q_{r/2} : u(x) \geqslant 1/2\}$, we may assume $|B| \geqslant \frac{1}{2}|Q_{r/2}|$, 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,

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

Hence we get

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

□

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

Proof. If we look at $u + \delta$ and let $\delta \to 0$, 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

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

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

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

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

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

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

Repeating and using induction,

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

Then

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

or

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

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

$$\operatorname*{Osc}_{Q_{2^{-k-1}}(x)} u \leqslant \rho \operatorname*{Osc}_{Q_{2^{-k}}(x)} u.$$

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

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

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

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

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

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

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

Chapter 5

Malliavin's proof of Hörmander's Theorem

Let $V_0, V_1, \ldots, V_n : \mathbb{R}^d \to \mathbb{R}^d$ be vector fields satisfying C^∞-boundedness conditions.

Consider

$$\begin{cases} \mathrm{d}X_t(x) & = \sum_{=1}^n V_l(X_t) \circ \mathrm{d}W_s^l + V_0(X_t)\mathrm{d}t = V(X_t) \circ \mathrm{d}W_s + V_0(X_t)\mathrm{d}t, \\ X_0(x) & = x. \end{cases} \tag{5.1}$$

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 $d \geqslant 1$, and ξ be a d-dimensional standard Gaussian random variable. Then

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

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

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

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_α 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 φ with $\mathbf{E}|\varphi'(\xi)| < \infty$, then ξ has a standard normal distribution.*

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

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

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

The following lemma is a criterion for smooth densities

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

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

whenever $\varphi \in C_c^k$. Then there exists ρ smooth such that

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

for all Borel sets A.

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

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

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

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 ω_t is a n-dimensional Brownian motion under $\mathbf{P}$.

Lemma 5.1.3. *Suppose F and G map Ω to $\mathbb{R}$ are bounded, and both F and G have bounded Fréchet derivative at each point of Ω. Suppose h_s is adapted and bounded and let $H_t = \int_0^t h_s\mathrm{d}s$. Then*

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

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

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

$$X_t^\varepsilon = \omega_t + \varepsilon \int_0^t h_s \mathrm{d}s$$

Let

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

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

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

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

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

On the other hand,

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

By, the right-hand side of is independent of ε. We differentiate with respect to ε 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

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

Now replacing F in the above identity with FG, and using chain rule for D, we obtain our assertion. ☐

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

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

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

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

where $(u\cdot\omega)$ is the Itô integral $\int_0^1 u(t)\mathrm{d}\omega_t$.

Compare (5.3) with (5.5). The adapted process H can be regard 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 $W_0^{1,2}\ni H_k(t)=\int_0^t h_k(s)\mathrm{d}s$. Set

$$D_kF := \mathcal{D}_{H_k}F.$$

If $\sum_k (D_kF)^2$ is bounded, we can define

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

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

$$u = \sum_k u_k h_k \in L^\infty(\Omega;\mathcal{H}), \quad u_k\in\mathcal{F}.$$

Then by (5.4),

$$\begin{aligned}\mathbf{E}\,\langle DF, u\rangle_{\mathcal{H}} =& \mathbf{E}\Big\langle DF, \sum_k u_k h_k\Big\rangle_{\mathcal{H}} = \mathbf{E}\left(\sum_k D_kF\,u_k\right)\\ =&\mathbf{E}\Big(F\sum_k u_k(h_k\cdot\omega)\Big) - \mathbf{E}\Big(F\sum_k D_ku_k\Big)\\ =&\mathbf{E}\Big[F\Big(\underbrace{\underbrace{\sum_k u_k(h_k\cdot\omega)}_{\text{``=''}\langle u,\omega\rangle_{\mathcal{H}}} - \underbrace{\sum_k D_ku_k}_{\text{``=''div}u}}_{-\mathrm{div}_{\mathbf{P}}u=:\delta(u)}\Big)\Big]\end{aligned}$$

The operator δ is called the divergence operator.

The main examples of F we will consider later is $\varphi(X_t)$, where φ is smooth and X_t solves an SDE. However, the Itô map $F:\omega.\mapsto X.(\omega)$ is even not continuous from $C([0,1];\mathbb{R}^n)$ to $C([0,1];\mathbb{R}^d)$. So we need some extension.

Let $p\geqslant 1$, and F be a function on Ω. We say F, G is in $\mathscr{W}^{1,p}$ if there exist functionals F_n on Ω that are bounded, continuous, Fréchet differentiable with bounded and continuous Fréchet derivatives, and with $F_n\to F$ in L^p and $|D(F_m-F_n)|_{\mathcal{H}}\to 0$ in L^p. In this case, $F\in L^p$ and there exists an $\mathcal{H}$-valued random map, denoted by DF such that $|DF_n - DF|_{\mathcal{H}}\to 0$ in L^p, and we do not distinguish between D and its closure. We also follow [Nua06] in denoting the adjoint of D by δ. One can of course apply the Malliavin differentiation operator repeatedly, thus yielding an unbounded closed operator D^k from $L^p(\Omega;\mathbb{R})$ to $L^p\left(\Omega;\mathcal{H}^{\otimes k}\right)$. We denote the domain of this operator by $\mathscr{W}^{k,p}$.

Furthermore, for any Hilbert space $\mathcal{K}$, we denote by $\mathscr{W}^{k,p}(\mathcal{K})$ the domain of D^k viewed as an operator from $L^p(\Omega;\mathcal{K})$ to $L^p\left(\Omega;\mathcal{H}^{\otimes k}\otimes\mathcal{K}\right)$. We call a random variable belonging to $\mathscr{W}^{k,p}$ for every $k,p\geqslant 1$ "Malliavin smooth" and we write $\mathscr{S}=\bigcap_{k,p}\mathscr{W}^{k,p}$ as well as $\mathscr{S}(\mathcal{K})=\bigcap_{k,p}\mathscr{W}^{k,p}(\mathcal{K})$. 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 $\mathscr{W}^{k,p}$ functionals will be considered is $\varphi(X_t)$.

Proposition 5.1.4. *Let X_t be the d-dimensional process that is a solution to (5.1). Assume that σ and b are C_b^∞. Then $X_t\in\mathscr{S}$.*

Proposition 5.1.5. *For every p there exist constants K and C such that, for every separable Hilbert space $\mathcal{K}$ and every $u\in\mathscr{S}(\mathcal{H}\otimes\mathcal{K})$, one has the bound*

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

Proof. Take now $p>2$. Using the definition of δ combined with the chain rule for D, Proposition 3.8 , and Young's inequality, we obtain the bound

$$\begin{aligned}\mathbf{E}|\delta u|^p&=(p-1)\mathbf{E}\left(|\delta u|^{p-2}\langle u,D\delta u\rangle\right)=(p-1)\mathbf{E}|\delta u|^{p-2}\left(|u|^2+\langle u,\delta Du\rangle\right)\\&\leqslant\frac{1}{2}\mathbf{E}|\delta u|^p+c\mathbf{E}\left(|u|^p+|u|^{p/2}|\delta Du|^{p/2}\right),\end{aligned}$$

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

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

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

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

The proof is concluded by a simple inductive argument. □

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

$$LF=\delta(DF)=\sum_k\left(D_kD_kF-D_kF\int_0^1 h_k(t)\mathrm{d}\omega_t\right)$$

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

Exercise 5.1.1. *Prove that*

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

Noting that

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

we have

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

Theorem 5.1.6. *Let $p > 1$, $F \in \mathscr{W}^{1,p}$ and . Assume that h_t is adapted, $H_t = \int_0^t h_s \mathrm{d}s$ and $|H|_{\mathcal{H}}$ is bounded. Then*

(i) D_HF_n convergets to $\langle DF, H\rangle =: D_HF$ in L^p and

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

(ii) it holds that

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

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

$$\mathbf{E}[(LF)G] = \mathbf{E}[F(LG)] = -\mathbf{E}[D_kFD_kG]. \tag{5.6}$$

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

$$|D_HF_n - D_HF_m| = \langle H, DF_n - DF_m\rangle_{\mathcal{H}} \leqslant |H|_{\mathcal{H}}|D(F_n - F_m)|_{\mathcal{H}} \to 0 \text{ in } L^p,\ p > 1.$$

Since

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

the convergence of $\mathbf{E}\left[F_n\int_0^1 h_s \mathrm{d}\omega_s\right]$ 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 $F = (F_1, \cdots F_d)$, each F_i is real-valued and $F_i \in \cap_{k,p\geqslant 1}\mathscr{W}^{k,p}$, 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

$$\left\langle D(\varphi(F)), DF^j\right\rangle_{\mathcal{H}} = \partial_i\varphi(F)\gamma_F^{ij}.$$

This yields that

$$\partial_i\varphi(F) = \left\langle D(\varphi(F)), DF^j\right\rangle_{\mathcal{H}}\left(\gamma_F^{-1}\right)^{ji}, \tag{5.7}$$

provided that γ_F is invertible.

Exercise 5.2.1. *Assume $F \in \cap_{k,p\geqslant 1}\mathscr{W}^{k,p}$ and $\gamma_F^{-1} \in \cap_{p\geqslant 1}L^p$. Then*

$$\gamma_F^{-1} \in \cap_{k,p\geqslant 1}\mathscr{W}^{k,p} \quad \text{and } D\gamma_F^{-1} = -\gamma_F^{-1}(D\gamma_F)\gamma_F^{-1}.$$

Theorem 5.2.1. *Suppose that $F \in \cap_{k,p\geqslant 1}\mathscr{W}^{k,p}$ and $\gamma_F^{-1} \in \cap_{p\geqslant 1}L^p$. Then*

$$|\mathbf{E}\nabla\varphi(F)| \leqslant C\|\varphi\|_\infty.$$

Proof. Thanks to (5.7) and Theorem 5.1.6,

$$\begin{aligned}
&\mathbf{E}\nabla^T\varphi(F) = \mathbf{E}\left[\gamma_F^{-1}D_k(\varphi(F))D_kF\right] \\
=&\mathbf{E}\left[D_k(\gamma_F^{-1}\varphi(F)D_kF)\right] - \mathbf{E}\left[D_k(\gamma_F^{-1})\varphi(F)D_kF\right] - \mathbf{E}\left[\gamma_F^{-1}\varphi(F)D_kD_kF\right] \\
=&\mathbf{E}\left\{\varphi(F)\gamma_F^{-1}\left[D_kF\int_0^1 v_k(s)\mathrm{d}\omega_s - D_kD_kF\right]\right\} - \mathbf{E}\left[\varphi(F)D_k(\gamma_F^{-1})D_kF\right] \\
=&-\mathbf{E}\left[\varphi(F)\gamma_F^{-1}LF\right] + \mathbf{E}\left[\varphi(F)\gamma_F^{-1}(D_k\gamma_F)\gamma_F^{-1}(D_kF)\right].
\end{aligned}$$

This yields that

$$\begin{aligned}
\left|\mathbf{E}\nabla^T\varphi(F)\right| \leqslant&\|\varphi\|_\infty\left\{\mathbf{E}[\gamma_F^{-1}LF] + \mathbf{E}\left[\gamma_F^{-1}(D_k\gamma_F)\gamma_F^{-1}(D_kF)\right]\right\} \\
\leqslant&C(p, \|F\|_{\mathscr{W}^{2,p}}, \|\gamma_F^{-1}\|_p)\|\varphi\|_\infty, \quad p \gg 1.
\end{aligned}$$

□

We need to calculate D_kX_t, where X_t solves (5.1). By definition,

$$X_t(\omega) = x + \int_0^t V(X_s) \circ \mathrm{d}\omega_s + \int_0^t V_0(X_s)\mathrm{d}s$$

and

$$\begin{aligned}
X_t(\omega + \varepsilon H_k) =&x + \int_0^t V(X_s(\omega + \varepsilon H_k)) \circ \mathrm{d}(\omega_s + \varepsilon H_k(s)) + \int_0^t V_0(X_s(\omega + \varepsilon H_k))\mathrm{d}s \\
=&x + \int_0^t V(X_s(\omega + \varepsilon H_k)) \circ \mathrm{d}\omega_s + \varepsilon\int_0^t V(X_t(\omega + \varepsilon H_k))\, h_k(s)\mathrm{d}s \\
&+ \int_0^t V_0(X_s(\omega + \varepsilon H_k))\mathrm{d}s
\end{aligned}$$

Taking the difference, dividing by ε, and letting $\varepsilon \to 0$, we obtain that

$$D_kX_t = \int_0^t \partial_jV_l(X_s)D_kX_s^j \circ \mathrm{d}\omega_s^l + \int_0^t \partial_jV_0(X_s)D_kX_s^j\mathrm{d}s + \int_0^t V_l(X_s)h_k^l(s)\mathrm{d}s. \tag{5.8}$$

Recall that $J(t) = \nabla X_t$, then

$$\mathrm{d}J_j^i(t) = \partial_lV_k^i(X_t)\, J_j^l(t) \circ \mathrm{d}\omega_t^k + \partial_lV_0^i(X_t)J_j^l(t)\mathrm{d}t, \quad J(0) = I. \tag{5.9}$$

Let $Z(t) : \Omega \to \mathbb{R}^{d\times d}$ be the solution to

$$\mathrm{d}Z_j^i(t) = -Z_l^i(t)\partial_jV_k^l(X_t) \circ \mathrm{d}\omega_t^k - Z_l^i(t)V_0^l(X_t)\,\mathrm{d}t, \quad Z(0) = I, \tag{5.10}$$

By Itô's formula, one can verify that

$$\mathrm{d}(Z(t)J(t)) = Z(t) \circ \mathrm{d}J(t) + \mathrm{d}Z(t) \circ J(t) = 0,$$

which yields that $Z(t) = J^{-1}(t)$.

Proposition 5.2.2.

$$D_k X_t = J(t)\int_0^t J^{-1}(s)V(X_s)h_k(s)\mathrm{d}s \tag{5.11}$$

Consequently,

$$\gamma_{X_t} = \langle DX_t, DX_t^T\rangle = J(t)\int_0^t J^{-1}(s)V(X_s)V^T(X_s)[J^{-1}(s)]^T\mathrm{d}s\, J^T(t). \tag{5.12}$$

Proof. By Itô's formula and (5.9),

$$\begin{aligned}
&\mathrm{d}\left[J(t)\int_0^t J^{-1}(s)V(X_s)h_k(s)\mathrm{d}s\right]\\
=&\mathrm{d}J(t)\int_0^t J^{-1}(s)V(X_s)h_k(s)\mathrm{d}s + V(X_t)h_k(t)\mathrm{d}t\\
\overset{(5.9)}{=}&\nabla V(X_t)\left[J(t)\int_0^t J^{-1}(s)V(X_s)h_k(s)\mathrm{d}s\right]\circ \mathrm{d}\omega_t\\
&+\nabla V_0(X_t)\left[J(t)\int_0^t J^{-1}(s)V(X_s)h_k(s)\mathrm{d}s\right]\mathrm{d}t + V(X_t)h_k(t)\mathrm{d}t
\end{aligned}$$

Therefore, $t\mapsto J(t)\int_0^t J^{-1}(s)V(X_s)h_k(s)\mathrm{d}s$ satisfies the same equation as D_kX_t, which yields (5.11). □

Set

$$C(t) := \int_0^t J^{-1}(s)V(X_s)V^T(X_s)[J^{-1}(s)]^T\mathrm{d}s.$$

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:

$$[U,V](x) := U^i(x)\partial_i(V^j(x))\partial_j - V^i\partial_i(U^j(x))\partial_j = [U^i\partial_iV^j - V^i\partial_iU^j](x)\partial_j$$

Define

$$S_0 = \{V_i : i>0\}, S_{k+1} = S_k\cup\{[U,V_j] : U\in S_k, j\geqslant 0\},$$

and

$$\mathscr{V}^k = \operatorname{span}\{V : V\in S_k\}.$$

We say that the vector fields $V_0, V_1\cdots, V_n$ satisfy the parabolic Hörmander condition

$$\bigcup_{k\geqslant 0}\mathscr{V}^k(x) = \mathbb{R}^d,\quad \forall x\in\mathbb{R}^d \tag{\textbf{H}}$$

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* $\{\hat{V}_0(x,t), \cdots, \hat{V}_N(x,t)\}$*, where we set*

$$\hat{V}_0(x,t) = \begin{pmatrix} V_0(x) \\ 1 \end{pmatrix}, \quad \hat{V}_j(x,t) = \begin{pmatrix} V_j(x) \\ 0 \end{pmatrix}, \quad j = 1, 2 \cdots N.$$

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 $x \to E_x$ 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 $E(x,t) = \bigcup_{k \geqslant 0} \mathscr{V}^k(x,t)$, 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 $(y,s) \subset \mathcal{M}^-$. 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*

$$\sup_{|\xi|=1} \mathbf{P}(\xi^T M \xi \leqslant \varepsilon) \leqslant C_p \varepsilon^p.$$

Then $(\det M)^{-1} \in L^{\infty-}$.

Proof. $\forall t > 1$, Choose $\{\xi_1, \cdots, \xi_m\} \subset S^N$, such that $\sup_{|\xi|=1} \min_{k \leqslant m} |\xi - \xi_k| \leqslant t^{-2}$ and $m \leqslant Ct^{2N}$.

$\forall \xi \in S^N$, we can find a vector ξ_k such that,

$$\xi^T M \xi = \xi_k^T M \xi_k + \xi_k^T M(\xi - \xi_k) + (\xi - \xi_k)^T M \xi \geqslant \xi_k^T M \xi_k - 2\|M\| t^{-2}$$

So we get

$$\{\inf_{|\xi|=1} \xi^T M \xi < t^{-1}\} \setminus \bigcup_{k=1}^{m} \{\xi_k^T M \xi_k < 3t^{-1}\} \subset \{\|M\| > t\}$$

Now,

$$\begin{aligned}
\mathbf{P}(\|M^{-1}\| > t) =& \mathbf{P}(\inf_{|\xi|=1} \xi^T M \xi < t^{-1}) \\
\leqslant& \mathbf{P}\Big(\cup_{k=1}^m \{\xi_k^T M \xi_k < 3t^{-1}\} \Big) + \mathbf{P}(\|M\| > t) \\
\leqslant& C C_{2N+p} t^{2N} t^{-2N-p} + \mathbf{E}(\|M\|^p) t^{-p} \\
\leqslant& C_p t^{-p}
\end{aligned}$$

$\forall q \geqslant 1$,

$$\begin{aligned}
\mathbf{E}(\det M^{-1})^q \leqslant& \mathbf{E}\|M^{-1}\|^q \leqslant C \int_0^\infty t^{q-1} \mathbf{P}(\|M^{-1}\| > t) \mathrm{d}t \\
\leqslant& \int_0^\infty 1 \wedge t^{q-1} C_q t^{-q-1} \mathrm{d}t < \infty.
\end{aligned}$$

□

Fixed $\xi \in S^N$, define $Z_U(t) = \xi^T J_t^{-1} U(X_t)$. Using Ito's formula,

$$\begin{aligned}
\mathrm{d}Z_U(t) =& -\xi^T J_t^{-1} V_0'(X_t) U(X_t) \mathrm{d}t - \xi^T J_t^{-1} V_j'(X_t) U(X_t) \circ \mathrm{d}W_t^j \\
& + \xi^T J_t^{-1}(X_t) U'(X_t) V_0(X_t) \mathrm{d}t + \xi^T J_t^{-1}(X_t) U'(X_t) V_j(X_t) \circ \mathrm{d}W_t^j \\
=& \xi^T J_t^{-1} [V_0, U](X_t) \mathrm{d}t + \xi^T J_t^{-1} [V_j, U](X_t) \circ \mathrm{d}W_t^j \\
=& Z_{[V_0,U]}(t) + Z_{[V_j,U]}(X_t) \circ \mathrm{d}W_t^j \\
=& \Big[Z_{[V_0,U]}(t) + \frac{1}{2} Z_{[V_j,[V_j,U]]} \Big] \mathrm{d}t + Z_{[V_j,U]}(X_t) \cdot \mathrm{d}W_t^j.
\end{aligned} \tag{5.13}$$

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*

$$\|f\|_{C^1} \leqslant C \|f\|_{C^0}^{\frac{\alpha}{1+\alpha}} \cdot \|f\|_{C^{1+\alpha}}^{\frac{1}{1+\alpha}}.$$

Lemma 5.3.6. *Suppose $Y_t = \int_0^t \sigma_s \mathrm{d}B_s$, $\mathbf{E}\|\sigma\|_\infty^p \leqslant K_p < \infty$, then $\forall \alpha \in (0, \frac{1}{2})$, $p \geqslant 1$, $\mathbf{E}\|Y\|_{C^\alpha}^p \leqslant C(K, \alpha, p)$, i.e. $Y \in L^{\infty-}$.*

Proof. Choose β, p such that
$$\alpha + \frac{1}{p} < \beta < \frac{1}{2},$$
by Sobolev embedding theorem,
$$\mathbf{E}\|Y\|_{C^\alpha}^p \leqslant C_{\alpha,\beta,p}\mathbf{E}\|Y\|_{W_p^\beta}^p = C\mathbf{E}\left(\int_0^1\int_0^1 \frac{|Y_s - Y_t|^p}{|s-t|^{1+\beta p}}\mathrm{d}s\mathrm{d}t\right) = C\int_0^1\int_0^1 \frac{\mathbf{E}|Y_s - Y_t|^p}{|s-t|^{1+\beta p}}\mathrm{d}s\mathrm{d}t.$$
By BDG inequality,
$$\mathbf{E}|Y_s - Y_t|^p \leqslant C\mathbf{E}\Big(\int_s^t \|\sigma_r\|_\infty^2 \mathrm{d}r\Big)^{p/2} \leqslant C_p|s-t|^{p/2}.$$
Combining the above inequalities,
$$\mathbf{E}\|Y\|_{C^\alpha}^p \leqslant C\int_0^1\int_0^1 |s-t|^{-1-\beta p+p/2}\mathrm{d}s\mathrm{d}t < \infty.$$
□

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.*

$A_\varepsilon \to_\varepsilon B_\varepsilon$ means $\forall p >> 1$, there exists a constant C_p, such that,
$$\mathbf{P}(A_\varepsilon \setminus B_\varepsilon) \leqslant C_p\varepsilon^p.$$

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*
$$Z_t = Z_0 + \int_0^t a(s)\mathrm{d}s + \int_0^t b_j(s)\mathrm{d}W_s^j.$$
Then there exists a constant $r \in (0,1)$ such that
$$\{\|Z\|_\infty < \varepsilon\} \to_\varepsilon \{\|a\|_\infty < \varepsilon^r\} \cap \{\|b\|_\infty < \varepsilon^r\}.$$

Proof.
$$Z_t^2 = Z_0^2 + \int_0^t (2a_sZ_s + |b(s)|^2)\mathrm{d}s + \int_0^t 2b_j(s)Z_s\mathrm{d}W_s^j. \tag{5.14}$$
$\|a\|_\alpha \in L^{\infty-} \to \{\|a\|_\infty \leqslant \varepsilon^{-1/4}\} \to_\varepsilon \varnothing$. Hence, $\{\|Z\|_\infty \leqslant \varepsilon\} \to_\varepsilon \{\|\int_0^\cdot 2a_sZ_s\mathrm{d}s\|_\infty \leqslant \varepsilon^{3/4}\}$. Similarly $\{\|Z\|_\infty \leqslant \varepsilon\} \to_\varepsilon \{\|\int_0^1 |b_j(s)Z_s|^2\mathrm{d}s\|_\infty \leqslant \varepsilon^{3/2}\}$. Using exponential martingale inequality,
$$\left\{\Big\|\int_0^1 |b_j(s)Z_s|^2\mathrm{d}s\Big\|_\infty \leqslant \varepsilon^{3/2}\right\} \to_\varepsilon \left\{\Big\|\int_0^\cdot 2b_j(s)\mathrm{d}W_s^j\Big\|_\infty \leqslant \varepsilon^{2/3}\right\}.$$

From (5.14), we get

$$\{\|Z\|_\infty \leqslant \varepsilon\} \to_\varepsilon \left\{\int_0^1 |b(s)|^2 \mathrm{d}s_\infty \leqslant \varepsilon^{2/3}\right\} \to_\varepsilon \left\{\int_0^1 |b(s)| \mathrm{d}s \leqslant \varepsilon^{1/3}\right\}.$$

Combining the above relation, (5.3.5) and $\{\|b\|_{1/3} \leqslant \varepsilon^{-1/4}\} \to_\varepsilon \varnothing$, we get

$$\{\|Z\|_\infty \leqslant \varepsilon\} \to_\varepsilon \{\|b\|_\infty \leqslant \varepsilon^{1/16}\}.$$

Using the same argument, we can prove

$$\{\|Z\|_\infty \leqslant \varepsilon\} \to_\varepsilon \{\|a\|_\infty \leqslant \varepsilon^{1/80}\}.$$

□

Proof of Theorem. Notice

$$\xi^T C_t \xi = \sum_{j=1}^d \int_0^t |Z_{V_j}(s)|^2 \mathrm{d}s. \tag{5.15}$$

Using Lemma 5.3.5, we get

$$\{\xi^T C_t \xi \leqslant \varepsilon\} \to_\varepsilon \{\|Z_{H_k}\|_\infty \leqslant \varepsilon^q\}$$

By Lemma 5.3.8,

$$\{\xi^T C_t \xi \leqslant \varepsilon\} \to_\varepsilon \bigcap_{V \in \mathscr{H}_k} \{\|Z_V\|_\infty \leqslant \varepsilon^{q_k}\}$$

for suitable $q_k > 0$. Now observe that $Z_V(0) = \langle x, U(x_0)\rangle$. By Hörmander's condition, $\mathscr{V}^{k'}(x_0) = \mathbb{R}^N$ for k large enough. However, if $\mathscr{V}^{k'}(x_0) = \mathbb{R}^N$, we can pick $V \in \mathscr{V}^{k'}(x_0)$ such that $|Z_V(0)| = |\langle x, V(x_0)\rangle| \geqslant \varepsilon_0$, so that the right-hand-side in the above equation is the empty set. We have thus proved $\{\xi^T C_t \xi \leqslant \varepsilon\} \to_\varepsilon \varnothing$. Now, using Lemma 5.3.4, we complete the proof. □

Appendix A

Useful facts

Lemma A.0.1 (Area formula). *Consider a locally Lipschitz function $f : \mathbb{R}^d \to \mathbb{R}^d$ and a Borel set $A \subseteq \mathbb{R}^d$. Then the function $y \mapsto N_A(y) := \operatorname{card}\{f^{-1}(y) \cap A\}\}$ is measurable and*

$$\int_A |\det(\nabla f(x))| \mathrm{d}x = \int_{\mathbb{R}^n} N_A(y) \mathrm{d}y \geqslant \mathscr{L}^d(f(A)).$$

Consequently, for any $g \geqslant 0$,

$$\int_{f(A)} g(y) \mathrm{d}y \leqslant \int_A g(f(x)) |\det \nabla f(x)| \mathrm{d}x. \tag{A.1}$$

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 $1 \leqslant p < \infty$. The first part of the Sobolev embedding theorem states that

Theorem A.1.1 (Sobolev). *If $k > l$, and $1 \leqslant p < q < \infty$ 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

$$W^{k,p} \hookrightarrow C^{r,\alpha}.$$

A.2 Singular integral

Singular integrals are central to harmonic analysis and are intimately connected with the study of partial differential equations. Singular integral is an integral operator

$$T(f)(x) = \int K(x,y)f(y)\,\mathrm{d}y,$$

whose kernel function $K : \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}$ 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 $d \geqslant 2$. Specifically, the Riesz transforms of a complex-valued function f are defined by

$$R_i f(x) = c_d \lim_{\epsilon\to 0} \int_{\mathbf{R}^d \setminus B_\epsilon(x)} \frac{(x_i - y_i)f(y)}{|x-y|^{d+1}}\,\mathrm{d}y, \quad i = 1, \cdots d.$$

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 A.2.1. *For each $i \in \{1, \cdots, d\}$, R_i is bounded on L^p with $p \in (1,\infty)$ and satisfy weak-type (1,1) estimates:*

$$\left|\left\{x \in \mathbb{R}^d : |R_i f(x)| > \lambda\right\}\right| \leqslant C_d \|f\|_1/\lambda. \tag{A.2}$$

A.3 Interpolation Theorems

The following simple Interpolation theorem is useful.

Lemma A.3.1. *1. Let $p \in [1,\infty)$. There exits constant C such that*

$$\|\nabla u\|_p \leqslant C\|\nabla^2 u\|_p^{\frac{1}{2}} \|u\|_p^{\frac{1}{2}}. \tag{A.3}$$

Let $\alpha \in (0,\alpha)$. There exits constant C such that

2.

$$[\nabla u]_\alpha \leqslant C[\nabla^2 u]_\alpha^{\frac{1}{2}} [u]_\alpha^{\frac{1}{2}}. \tag{A.4}$$

Appendix B

Some basic results of PDEs

B.1 Monge-Ampère Equation

To motivate the definition of weak solutions to (2.6), given an open set $D \subset \mathbb{R}^n$, consider $u : D \to \mathbb{R}$ a convex function of class C^2 satisfying (2.6) for some $f : D \to \mathbb{R}^+$. Then given any Borel set $E \subset D$, it follows by the area formula that

$$\int_E f \, \mathrm{d}x = \int_E \det D^2 u \, \mathrm{d}x = |\nabla u(E)|.$$

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

$$\partial u(x) := \left\{ p \in \mathbb{R}^d : u(y) \geqslant u(x) + \langle p, y - x \rangle \quad \forall y \in D \right\}.$$

This motivates the following definition:

Definition B.1.1. *Given an open set $D \subset \mathbb{R}^n$ and a convex function $u : D \to \mathbb{R}$, 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 B.1.2. *Let $u, v : D \to \mathbb{R}$ be convex functions. Then*

$$\mu_{u+v} \geqslant \mu_u + \mu_v \quad \text{and} \quad \mu_{\lambda u} = \lambda^n \mu_u \quad \forall \lambda > 0.$$

The following result is the celebrated Alexandrov maximum principle.

Theorem B.1.3. *Let D be an open bounded convex set, and let $u : D \to \mathbb{R}$ be a convex function such that $u|_{\partial D} = 0$. Then there exists a dimensional constant $C = C(d)$ such that*

$$|u(x)| \leqslant C(d)\,\mathrm{diam}(D)^{\frac{d-1}{d}}\,\mathrm{dist}(x,\partial D)^{\frac{1}{d}}|\partial u(D)|^{\frac{1}{d}}, \quad \forall x \in D. \tag{B.1}$$

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 ∂D. Since $u \leqslant \widehat{C}_x$ in D (by the convexity of u), Lemma 2.7 implies that

$$\left|\partial \widehat{C}_x(x)\right| \leqslant \left|\partial \widehat{C}_x(D)\right| \leqslant |\partial u(D)|;$$

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_ρ with $\rho = |u(x)|/\mathrm{diam}(D)$
- $\partial \widehat{C}_x(x)$ contains a vector of norm $|u(x)|/\mathrm{dist}(x, \partial D)$

Thus,

$$\partial \widehat{C}_x(x) \supset B_\varrho(0) \cup \{q\}, \quad |q| = |u(x)|/\mathrm{dist}(x, \partial D).$$

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

$$c(d)\rho^{d-1}|q| = |\mathcal{C}| \leqslant |\partial u(D)|.$$

□

Theorem B.1.4. *Let D be an open bounded convex set, and let ν be a Borel measure on D with $\nu(D) < \infty$. Then there exists a unique convex function $u : D \to \mathbb{R}$ solving the Dirichlet problem*

$$\begin{cases} \mu_u = v & \text{in } D \\ u = 0 & \text{on } \partial D \end{cases}$$

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

$$\mathcal{S}[\nu] := \{v : \Omega \to \mathbb{R} \text{ convex : } \left.v\right|_{\partial\Omega} = 0, \mu_v \geqslant \nu \text{ in } \Omega\}$$

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 β_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 λ has to be chosen. We notice that $\bar{v}|_{\partial\Omega} = 0$. In addition, provided λ is sufficiently large, Lemma below implies that

$$\mu_{\bar{v}} \geqslant \sum_{i=1}^{N} \mu_{\lambda \hat{C}_{x_i}} = \sum_{i=1}^{N} \lambda^d \mu_{\hat{C}_{x_i}} = \sum_{i=1}^{N} \lambda^d \beta_i \delta_{x_i} \geqslant \sum_{i=1}^{N} \alpha_i \delta_{x_i} = \nu.$$

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 Ω_1 and Ω_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 Ω_0 and $w \geqslant v_1$ everywhere else, we have

$$\partial v_1(E) \subseteq \partial w(E_0).$$

Therefore,

$$\mu_w(E) \geqslant \mu_{v_1}(E_0 \cup E_1) + \mu_{v_2}(E_2) \geqslant \nu(E).$$

Step 3: $u := \sup_{v \in S[\nu]} v$ belongs to $S[\nu]$. 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 μ_u is supported at the points $\{x_1, \cdots x_N\}$. Otherwise, there exists a set $E \subseteq D$ such that

$$E \cap \{x_1, \ldots, x_N\} = \emptyset \quad \text{and} \quad |\partial u(E)| = \mu_u(E) > 0$$

Therefore,

$$\left|\partial u(E) \backslash [\cup_{i=1}^{N} \partial u(x_i) \cup \partial u(\partial D)]\right| = |\partial u(E)| > 0$$

Let $x_0 \in E$ and $p \in \partial u(x_0) \backslash [\cup_{i=1}^{N} \partial u(x_i) \cup \partial u(\partial D)]$. Then there exists $\delta > 0$ such that

$$u \geqslant \ell_{x_0,p} + 2\delta \quad \text{on } \{x_1, \ldots, x_N\} \cup \partial\Omega, \tag{B.2}$$

where $\ell_{x_0,p}(x) = u(x_0) + p \cdot (x - x_0)$. Set $\bar{u} := \max\{\ell_{x_0,p} + \delta, u\} \gneqq u$. Notice that $\bar{u}$ is convex, $\bar{u} \geqslant u$, and it follows by (B.2) that $\bar{u} = u$ in a neighborhood of $\{x_1, \ldots, x_N\} \cup \partial\Omega$. 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 $\beta_i \geqslant \alpha_i$. 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

$$\{U \leqslant U\,(x_j) + \delta\} \cap (\{x_1, \ldots, x_N\} \cup \partial\Omega) = \emptyset,$$

and define the function

$$\widetilde{U}(z) := \begin{cases} U(z) & \text{if } U > U(x_j) + \delta \\ (1-\delta)U(z) + \delta\left[U(x_j) + \delta\right] & \text{if } U \leqslant U(x_j) + \delta. \end{cases}$$

Taking $\delta > 0$ even smaller if necessary, we observe that

$$\left|\partial\widetilde{U}(x_j)\right| = (1-\delta)^n \left|\partial U(x_j)\right| = (1-\delta)^n \beta_j > \alpha_j.$$

Notice that $\widetilde{U} \geqslant U$ and $\widetilde{U} = U$ in a neighborhood of $\{x_1, \ldots, x_N\} \cup \partial\Omega$. Hence, considering $\tilde{u} := \widetilde{U} + \ell_{x_j,p}$ (see Figure 2.7), we see that $\tilde{u} \geqslant u$,

$$\tilde{u}|_{\partial\Omega} = u|_{\partial\Omega} = 0, \quad \partial\tilde{u}(x_i) = \partial u(x_i) \quad \forall i \neq j, \quad \text{and} \quad |\partial\tilde{u}(x_j)| > \alpha_j.$$

Thus, $\tilde{u} \in \mathcal{S}[v]$, but this contradicts the maximality of u and concludes the proof. □

B.2 Schauder estimate

Let $\mathscr{S}$ be the Schwartz space of all rapidly decreasing functions, and $\mathscr{S}'$ the dual space of $\mathscr{S}$ called Schwartz generalized function (or tempered distribution) space. Given $f \in \mathscr{S}$, let $\mathscr{F}f = \hat{f}$ be the Fourier transform defined by

$$\hat{f}(\xi) := \int_{\mathbb{R}^d} \mathrm{e}^{-\mathrm{i}2\pi\xi\cdot x} f(x)\mathrm{d}x.$$

Let $\chi : \mathbb{R}^d \to [0,1]$ be a smooth radial function with

$$\chi(\xi) = 1,\ |\xi| \leqslant 1,\ \chi(\xi) = 0,\ |\xi| \geqslant 3/2.$$

Define

$$\varphi(\xi) := \chi(\xi) - \chi(2\xi).$$

It is easy to see that $\varphi \geqslant 0$ and $\operatorname{supp} \varphi \subset B_{3/2} \setminus B_{1/2}$ and formally

$$\sum_{j=-k}^{k} \varphi(2^{-j}\xi) = \chi(2^{-k}\xi) - \chi(2^{k+1}\xi) \overset{k\to\infty}{\to} 1. \tag{B.3}$$

In particular, if $|j - j'| \geqslant 2$, then

$$\operatorname{supp}\varphi(2^{-j}\cdot) \cap \operatorname{supp}\varphi(2^{-j'}\cdot) = \varnothing.$$

From now on we shall fix such χ and φ and define

$$\Delta_j f := \mathscr{F}^{-1}(\varphi(2^{-j}\cdot)\mathscr{F}f), \quad j \in \mathbb{Z}.$$

We first recall the following useful lemmas.

Lemma B.2.1. *Let $\alpha \in (0,1)$. For any $u \in C^\alpha$, it holds that*

$$\frac{1}{C}\sup_{j\in\mathbb{Z}} 2^{j\alpha}\|\Delta_j u\|_\infty \leqslant [u]_\alpha \leqslant C\sup_{j\in\mathbb{Z}} 2^{j\alpha}\|\Delta_j u\|_\infty, \tag{B.4}$$

where

$$[u]_\alpha := \sup_{x\neq y}\frac{|u(x)-u(y)|}{|x-y|^\alpha},$$

and C only depends on d and α.

Proof. □

Lemma B.2.2. *There is a constant $C = C(d,\alpha)$, such that for any $u \in C^{2,\alpha}$,*

$$[\nabla^2 u]_\alpha \leqslant C[\Delta u]_\alpha. \tag{B.5}$$

Proof. Define

$$\varphi^{kl}(\xi) := \frac{\xi_k\xi_l}{|\xi|^2}\varphi(\xi), \quad h^{kl}(x) := \mathscr{F}^{-1}(\varphi^{kl})(x); \quad \varphi^{kl}_j(\xi) := \varphi^{kl}(2^{-j}\xi), \quad h^{kl}_j(x) := 2^{jd}h^{kl}(2^j x).$$

It is easy to see

$$\partial_{kl}u = \sum_{j\in\mathbb{Z}} u^{kl}_j := \sum_{j\in\mathbb{Z}}\varphi^{kl}_j(D)f = \sum_{j\in\mathbb{Z}} h^{kl}_j * f,$$

For any $k,l \in \{1,2,\cdots,d\}$, there is a constant C only depending on α, d such that

$$\|u^{kl}_j\|_\infty \leqslant C2^{-j\alpha}[f]_\alpha, \quad \forall j \in \mathbb{Z}. \tag{B.6}$$

For any $x \in \mathbb{R}^d$, noticing $h^{kl} \in \mathscr{S}(\mathbb{R}^d)$ and $\int h^{kl} = \varphi(0) = 0$, we get

$$\begin{aligned}|u^{kl}_j(x)| &= \left|\int_{\mathbb{R}^d} h^{kl}_j(y)f(x-y)\mathrm{d}y\right| = \left|\int_{\mathbb{R}^d} h^{kl}(z)(f(x-2^{-j}z)-f(x))\mathrm{d}z\right| \\ &\leqslant \int_{\mathbb{R}^d}|h^{kl}(z)|\cdot[f]_\alpha|2^{-j}z|^\alpha\mathrm{d}z \leqslant C[f]_\alpha 2^{-j\alpha}.\end{aligned}$$

By this,

$$\begin{aligned}|\partial_{kl}u(x)-\partial_{kl}u(y)| \leqslant& \sum_{j\leqslant K}|u^{kl}_j(x)-u^{kl}_j(y)| + \sum_{j>K}|u^{kl}_j(x)-u^{kl}_j(y)| \\ \leqslant& |x-y|\cdot\sum_{j\leqslant K}\|\nabla u^{kl}_j\|_\infty + 2\sum_{j>K}\|u^{kl}_j\|_\infty \\ \leqslant& C[f]_\alpha|x-y|\sum_{j\leqslant K}2^{(1-\alpha)j} + C[f]_\alpha\sum_{j>K}2^{-j\alpha} \\ \leqslant& C[f]_\alpha\left(|x-y|2^{(1-\alpha)K}+2^{-\alpha K}\right)\end{aligned}$$

Choosing $K \in \mathbb{Z}$ such that $2^{-K} \leqslant |x-y| < 2^{-K+1}$, we obtain

$$|\partial_{kl}u(x)-\partial_{kl}u(y)| \leqslant C_1[f]_\alpha\cdot|x-y|^\alpha.$$

So we complete our proof. □

Lemma B.2.3. *Suppose $f:\mathbb{R}_+\to\mathbb{R}_+$, and for any $0\leqslant s<t\leqslant 1$,*

$$f(s)\leqslant \theta f(t)+A(t-s)^{-\gamma}+B,$$

for some $\gamma>0$, then

$$f(s)\leqslant C(\gamma,\theta)\big(A(t-s)^{-\gamma}+B\big) \tag{B.7}$$

Proof. Let $t_0=t$, $t_i=s+(1-\tau)\tau^i(t-s)$, where $\tau\in(\theta^{1/\gamma},1)$. By iteration,

$$\begin{aligned} f(s)=&f(t_0)\leqslant \theta f(t_1)+A(t_i-t_{i+1})^{-\gamma}+B\\ \leqslant&\theta^2 f(t_2)+\theta A(t_2-t_1)^{-\gamma}+\theta B+A(t_1-t_0)^{-\gamma}+B\\ \leqslant&\cdots\leqslant\theta^k f(t_k)+A(1-\tau)^{-\gamma}(t-s)^{-\gamma}\sum_{i=0}^{k-1}(\theta\tau^{-\gamma})^i+B\sum_{i=0}^{k-1}\theta^i\\ \leqslant&C\left(\frac{A}{(t-s)^\gamma}+B\right). \end{aligned}$$

□

Theorem B.2.4. *Suppose*

$$Lu=a^{ij}\partial_{ij}u+b^i\partial_i u+cu=f,\quad \text{in } B_R, \tag{B.8}$$

and

$$\delta|\xi|^2\leqslant a^{ij}\xi_i\xi_j, \tag{B.9}$$

$$\begin{aligned}(\|a\|_{L^\infty(B_R)}+R^\alpha[a]_{\alpha;B_R})+(R\|b\|_{L^\infty(B_R)}+R^{1+\alpha}[b]_{\alpha;B_R})&\\ +(R^2\|c\|_{L^\infty(B_R)}+R^{2+\alpha}[c]_{\alpha;B_R})&\leqslant\Lambda.\end{aligned} \tag{B.10}$$

Then,

$$[\nabla^2 u]_{\alpha;B_{R/2}}\leqslant C\Big\{[f]_{\alpha;B_R}+R^{-\alpha}\|f\|_{L^\infty(B_R)}+R^{-2-\alpha}\|u\|_{L^\infty(B_R)}\Big\}.$$

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

$$(r-\rho)^k\|\nabla^k\eta\|_\infty+(r-\rho)^{k+\alpha}[\nabla^k\eta]_\alpha\leqslant C(d,k).$$

Let $v:=u\eta$, then

$$a_o^{ij}\partial_{ij}v=(a_o^{ij}-a^{ij})\eta\cdot\partial_{ij}u+2a_o^{ij}\partial_i u\partial_j\eta+a_o^{ij}u\partial_{ij}\eta+(b^i\partial_i u)\cdot\eta+(cu)\cdot\eta+f\eta.$$

$$\begin{aligned}[\nabla^2 u]_{\alpha;B_\rho}\leqslant&[\nabla^2 v]_\alpha\leqslant C_1\Big\{r^\alpha[\nabla^2 u]_{\alpha;B_r}+(r-\rho)^{-\alpha}\|\nabla^2 u\|_{L^\infty(B_r)}\\ &+(r-\rho)^{-1}[\nabla u]_{\alpha;B_r}+(r-\rho)^{-1-\alpha}\|\nabla u\|_{L^\infty(B_r)}+(r-\rho)^{-2}[u]_{\alpha;B_r}\\ &+(r-\rho)^{-2-\alpha}\|u\|_{L^\infty(B_r)}+[f]_{\alpha;B_r}+(r-\rho)^{-\alpha}\|f\|_{L^\infty(B_r)}\Big\}\end{aligned}$$

There is a constant $r_0\in(0,1)$ such that $C_1r_0^\alpha\leqslant 1/4$. By interpolation, there is a constant C such that

$$C_1\Big\{(r-\rho)^{-\alpha}\|\nabla^2 u\|_{L^\infty(B_r)}+(r-\rho)^{-1}[\nabla u]_{\alpha;B_r}+(r-\rho)^{-1-\alpha}\|\nabla u\|_{L^\infty(B_r)}+(r-\rho)^{-2}[u]_{\alpha;B_r}$$

$$+ (r-\rho)^{-2-\alpha}\|u\|_{L^\infty(B_r)}\Big\} \leqslant \frac{1}{4}[\nabla^2 u]_{\alpha;B_r} + C(r-\rho)^{-2-\alpha}\|u\|_{L^\infty(B_r)}$$

Combing the above inequalities and B.2.3, we obtain that for any $0 \leqslant \rho < r \leqslant r_0$,

$$[\nabla^2 u]_{\alpha;B_\rho} \leqslant C\Big\{(r-\rho)^{-2-\alpha}\|u\|_{L^\infty(B_r)} + [f]_{\alpha;B_r} + (r-\rho)^{-\alpha}\|f\|_{L^\infty(B_r)}\Big\}.$$

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

$$\lambda|\xi|^2 \leqslant a^{ij}\xi_i\xi_j; \quad \|a\|_{C^\alpha(B_1)} + \|b\|_{C^\alpha(B_1)} + \|c\|_{C^\alpha(B_1)} \leqslant \Lambda,$$

then

$$[\nabla^2 u]_{\alpha;B_{1/2}} \leqslant C\Big\{[f]_{\alpha;B_1} + \|f\|_{L^\infty(B_1)} + \|u\|_{L^\infty(B_1)}\Big\}.$$

By rescaling, it is easy to obtain our result. $\square$

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.

[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.