C8.1 Stochastic Differential Equations

Zhongmin Qian
Mathematical Institute, University of Oxford

November 30, 2021

Contents

Chapter 1

Introduction

Stochastic differential equations become increasingly important not only in applied areas such as *quantitative finance*[1], *mathematical physics*[2], *turbulence*[3], and etc., but also in many research areas of pure mathematics like non-linear *partial differential equations*[4], *harmonic analysis*[5], *ergodic theory*[6], *differential geometry*[7] and etc. There are great names in modern science history which are associated with the development of the subject. Let us make a short list which is far from complete of course.

1) The major character in stochastic analysis is, without doubt, Brownian motion. The physical Brownian motion, which is the name given to the chaotic movements of small pollen grains of clarkia pulchella suspended in water, was observed and reported first to the Royal Society by the Scottish botanist Robert Brown in 1828 (he lived from 21 December 1773 to 10 June 1858).

2) Loius Bachelier, around 1900, took a far time-ahead step using Brownian motion to model stock market in his Ph. D thesis (submitted to the faculty of École Normale Supérieure Paris). His thesis was published: L. Bachelier: *Théorie de la spéculation*, in Ann. Sci. École Norm. Sup. 17, 21 - 86 (1900). L. Bachelier used the fundamental solution of the heat equation (i.e. the heat kernel, Green's function or called the source function) to describe the transition laws of Brownian motion. In a delighted book by Mark Davis and Alison Etheridge: *Louis Bachelier's Theory of Speculation – the origins*

[1]Steven E. Shreve: *Stochastic Calculus for Finance II – Continuous-Time Models,* Springer 2004. Also Robert C. Merton: *Continuous-Time Finance*, Blackwell 1990.

[2]Barry Simon: *Functional Integration and Quantum Physics*, Academic Press 1979.

[3]Stephen B. Pope: *Turbulent Flows*, Cambridge University Press 2000.

[4]Daniel W. Stroock and S.R.S. Varadhan: *Multidimensional Diffusion Processes*, Springer 1979 and 1997 (reprinted), and also Daniel W. Stroock: *Partial Differential Equations for Probabilists*, Cambridge University Press 2008.

[5]Ross G. Pinsky: *Positive Harmonic Functions and Diffusion.* Cambridge University Press 1995. For a discrete case refer to Tullio Ceccherini-Silberstein, Fabio Scarabotti and Filippo Tolli: *Harmonic Analysis on Finite Groups.* Cambridge University Press 2008.

[6]H. P. Mckean, Jr.: *Stochastic Integrals*, Academic Press 1969. This is the little book beautifully written which educated a generation of scholars in stochastic analysis. Also his recent very nice exercise book Henry McKean: *Probability – The Classical Limit Theorems.* Cambridge University Press 2014.

[7]Jean-Michel Bismut and Gilles Lebeau: *The Hypoelliptic Laplacian and Ray-Singer Metrics*, Princeton University Press 2008, and of course Nobuyuki Ikeda and Shinzo Watanabe: *Stochastic Differential Equations and Diffusion Processes*, North-Holland 1981.

of Modern Finance, Princeton University Press (2006), one may find a good account about this in English with detailed explanations, also the Xrox copy of the original print, and a Forward penned by the first Nobel economics laureate Paul A. Samuelson (May 15, 1915 to December 13, 2009).

3) Albert Einstein (14 March 1879-18 April 1955), well known for his special and general theories of relativity, in 1905, the year he created his fundamental paper on special relativity, published his study of Brownian motion [A. Einstein: *On the movement of small particles suspended in a stationary liquid demanded by the molecular-kinetic theory of heat*, Annalen der Physik 17 (1905), 549-560. English translation of the paper may be found in *Einstein's Miraculous Year – five papers that changed the face of physics*, Princeton University Press (1998)]. Einstein's Brownian motion paper and related papers are collected in a small book titled *Investigations on the theory of the Brownian movement*, edited by R. Fürth and translated by A. D. Cowper, Dover Publications, INC. (1956). In the same year A. Einstein published 5 fundamental papers. One is about *Special Relativity* [A. Einstein: Zur Elektrodynamik bewegter Körper, Annalen der Physik 17 (1905), 891 – 921. English translation including his fundamental paper on General Theory in 1916 and other key papers on gravitation may be found in *The Principe of Relativity – A collection of original papers on the special and general theory of relativity* by A. Einstein, H. A. Lorentz, H. Weyl and H. Minkowski (with notes by A. Sommerfeld), Dover Publications, INC (1952)]. The another work published in the same year by A. Einstein is on the law of the photoelectric effect, a fundamental quantum hypothesis [A. Einstein: *On a heuristic point of view concerning the production and transformation of light*, Annalen der Physik 17 (1905), 132 - 148], which is the work A. Einstein was awarded the Nobel Prize in Physics in 1921.

4) Norbert Wiener (November 26, 1894 - March 18, 1964) in 1923 [N. Wiener: *Differential space*, J. Math. Phys. 2 (1923), 131 - 174] constructed the distribution of Brownian motion as a probability measure over the space of continuous paths, which is the archetypal example of a measure over an *infinite dimensional space* in the Lebesgue sense. This work marked the start of the research area of stochastic analysis.

5) Paul Lévy (1886 - 1971) was probably the longest ever lived mathematician who worked by him own and discovered rich and interesting unusual properties of the mathematical model of Brownian motion proposed by L. Bachilier and A. Einstein. His results were published in a book: *Processus stochastiques et mouvement Brownien*, Gauthier-Villars et Fils (1948). He studied a class of stochastic processes (he called them additive processes, his monograph *Théorie de l'addition des variables aléatories* was published in 1937) which have important applications, now named after him called *Lévy's processes*.

6) Kiyoshi Itô (September 7, 1915 - 10 November 2008) was mainly credited as the founder of the theory of stochastic differential equations. His foundation papers including the paper about Itô's lemma or Itô's formula which was published around 1942, and his theory of SDE was, with much delay due to the second war, published in 1951 [K. Itô: *On stochastic differential equations*, Mem. Amer. Math. Soc. 4 (1951)]. The reader may gather more or less a full picture about Itô's calculus in Nobuyuki Ikeda and Shinzo Watanabe: *Stochastic Differential Equations and Diffusion Processes*, North-Holland/Kodansha (1981). Itô's calculus was also explained in a master piece *Stochastic Integrals* by H. P. McKean published in 1969 (Academic Press, recently reissued by AMS). K. Itô and H.

P. McKean also wrote a comprehensive monograph "*Diffusion processes and their sample paths*" (Springer-Verlag, 1965) which gives a path-wise construction of a class of diffusion models determined by second-order elliptic differential operators on intervals of the real line. This book has been the aspiring source for mathematicians since then.

7) The general theory of stochastic processes – the stochastic calculus for semi-martingales which generalizes Itô's theory to a large class of stochastic processes, was established mainly by the *French School* under the leadership of Paul-André Meyer (21 August 1934 - 30 January 2003). Meyer and his co-authors created monumental volumes on semi-martingales, represented with the 5 volumes written by him and his co-authors [C. Dellacherie, P.A. Meyer: *Probabilités et potentiel.* Hermann, Paris, vol. I (1975), vol. II (Chapitres 5 à 8 : Théorie des martingales, 1980), vol III (Chapitres 9 à 11, Théorie discrète du potentiel, 1983), vol. IV (Théorie du potentiel associée à une résolvante, Théorie des processus de Markov, 1987), vol. V (with one more co-author B. Maisonneuve : Processus de Markov (fin) : Compléments de calcul stochastique, 1992)].

8) Paul Malliavin (September 10, 1925 - June 3, 2010) developed a *calculus of variations for Wiener functionals.* His book *Stochastic Analysis* (Springer, 1997) soon becomes a classic. He even wrote a book with Anton Thalmaiera on the applications of his calculus to quantitative finance titled *Stochastic calculus of variations in mathematical finance* (Springer-Verlag, Berlin, 2006). The latter book should be useful for those interested in the applications of viscosity solutions to PDEs too.

9) There are many excellent books which appeared during 1970's to 1990's, such as (1) D. W. Stroock and S. R. S Varadhan: *Multidimensional diffusion processes* (Springer-Verlag 1979), (2) L. C. G. Rogers and D. Williams: *Diffusions, Markov processes, and martingales Vol. 2: Itô Calculus* (Wiley, New York, 1987) (reissued by Cambridge University Press), (3) D. Revuz and M. Yor: *Continuous martingales and Brownian motion* (Springer-Verlag, Berlin 1991), (4) I. Karatzas and S. E. Shreve: *Brownian motion and stochastic calculus* (Springer 1998). This list is still a small sample of books useful for learning stochastic analysis.

10) On the applied fronts, besides those traditional applications in pure mathematics and in engineering, Itô's calculus has been found to be the perfect tool for quantitative finance. In a paper published in 1971 by Fischer Sheffey Black (January 11, 1938 - August 30, 1995, an economist) and Myron Scholes (born July 1, 1941, and economist who was awarded the Nobel Memorial Prize in Economic Sciences in 1997) titled *The Pricing of options and corporate liabilities* (J. Polit. Economy) in which they derived the famous PDE now named after them called the Black-Scholes equation, which allow them to obtain an explicit formula for pricing the European options. Another economist Robert C. Merton (born July 31, 1944, awarded Nobel Prize in Economic in 1997) made the pioneering contributions to continuous-time finance by using Brownian motion in the "*Theory of rational option pricing*". Bell Journal of Economics and Management Science (The RAND Corporation) 4 (1),141-183 (1973). Here, after about 70 years of Bachilier's work, we witness the rebirth of the subject with modern and matured tools towards the understanding of financial markets. The reader can learn a lot from the collection of Robert C. Merton: *Continuous-time Finance* (Blackwell, Cambridge MA & Oxford UK, 1990).

Finally let me point out that a major problem in science, which is largely still open, is to describe and to construct laws (equivalently distributions) of random fields which are probability measures on certain infinite dimensional spaces of mappings between spaces, such as spaces of continuous mapping from $\mathbb{T}$ to M, where $\mathbb{T} \subset \mathbb{R}^d$ a subset and $M \subseteq \mathbb{R}^n$ (a sub-manifold of $\mathbb{R}^n$). In modern science, fields such as gauge fields in quantum field theories, velocity fields in turbulence, are fundamental objects, and fields appearing in many applications (such as turbulence, quantum field theory, statistical mechanics, condensed matter physics and so on) tend to be random. A theory in science should acquires at least two functions. First it should provide a good description of observed phenomena in terms of certain language, and second it should have the power of making certain predictions. Being a successful scientific theory, it is essential to build good mathematical models and theories of random fields. In this course we develop a general theory for constructing a special class of random fields, called stochastic processes.

Chapter 2

Martingales, Brownian motion and Itô calculus

In this chapter we recall the fundamental concepts and results about martingales (in continuous-time), Brownian motion, and the elements about stochastic integrals (Itô's integrals).

2.1 Some notions about stochastic processes

Stochastic processes are mathematical models used to describe random phenomena evolving in time. Let $\boldsymbol{T}$ denote the range of time-parameter t, which is $[0,\infty)$ in these lectures, though other choices are also allowed. $\boldsymbol{T}$ is thus an ordered set endowed with the usual topology.

Definition 2.1.1 *A* stochastic process *is a parameterized family $X=(X_t)_{t\in\boldsymbol{T}}$ of random variables on a probability space $(\Omega,\mathcal{F},\mathbb{P})$ taking values in a topological space S (called the state space). Unless otherwise specified, in this course, $S=\mathbb{R}$, or $\mathbb{R}^d$.*

A stochastic process $X=(X_t)_{t\in\boldsymbol{T}}$ may be considered as a function from $\boldsymbol{T}\times\Omega\to\mathbb{R}^d$, which is the reason why a stochastic process is also called a random function.

For each $\omega\in\Omega$, the mapping $t\to X_t(\omega)$ from $\boldsymbol{T}$ to S is called a *sample path* (or a trajectory, or a sample function). Naturally, a stochastic process $X=(X_t)_{t\in\boldsymbol{T}}$ is continuous (resp. right-continuous, right-continuous with left-limits) if sample paths $t\to X_t(\omega)$ are continuous (resp. right-continuous, right-continuous with left-limits) for almost all $\omega\in\Omega$.

If $X=(X_t)_{t\geq 0}$ is a stochastic process with values in $\mathbb{R}^d$, and $0\leq t_1<t_2<\cdots<t_n$, the joint distribution of random variables $(X_{t_1},\cdots,X_{t_n})$ given by

$$\mu_{t_1,t_2,\cdots,t_n}(\mathrm{d}x_1,\cdots,\mathrm{d}x_n)=\mathbb{P}\left(X_{t_1}\in\mathrm{d}x_1,\cdots,X_{t_n}\in\mathrm{d}x_n\right),$$

a probability measure on $\mathbb{R}^d\times\cdots\times\mathbb{R}^d$, is called a finite-dimensional distribution of $X=(X_t)_{t\geq 0}$. If $d=1$, the measure $\mu_{t_1,t_2,\cdots,t_n}$ may be determined from its distribution function

$$F_{t_1,t_2,\cdots,t_n}(x_1,\cdots,x_n)=\mathbb{P}\left(X_{t_1}\leq x_1,\cdots,X_{t_n}\leq x_n\right).$$

In order to overcome some technical difficulties arising from measurability, in particular in dealing with stochastic processes in continuous-time, a common condition, which is good enough to include a large class of interesting stochastic processes, is that almost all sample paths $X = (X_t)_{t\geq 0}$ are right-continuous.

Exercise 2.1.2 *Let* $(X_t)_{t\geq 0}$ *be a stochastic process in* $\mathbb{R}^d$ *on* $(\Omega, \mathcal{F}, \mathbb{P})$*, and* $B \subset \mathbb{R}^d$ *be Borel measurable. If* $F \subseteq [0, \infty)$ *is finite or countable, then* $\{\omega : X_t(\omega) \in B \textit{ for anyt} \in F\}$ *and* $\sup_{t\in F} |X_t|$ *are measurable.*

The main task of stochastic analysis is to study the probability (or distribution) properties of random functions determined by their families of finite-dimensional distributions.

Definition 2.1.3 *Two stochastic processes* $X = (X_t)_{t\geq 0}$ *and* $Y = (Y_t)_{t\geq 0}$ *are equivalent (distinguishable) if* $\mathbb{P}(X_t = Y_t) = 1$ *for every* $t \geq 0$*. In this case,* $(Y_t)_{t\geq 0}$ *is called a version of* $(X_t)_{t\geq 0}$.

By definition, the family of finite-dimensional distributions of a stochastic process $X = (X_t)_{t\geq 0}$ is unique up to equivalence of processes.

On the other hand, in many practical situations, we are given a collection of *compatible* finite dimensional distributions $\mathcal{D} = \{\mu_{t_1,\cdots,t_n}$, for $t_1 < \cdots < t_n$, $t_j \in \boldsymbol{T}$, we would like to construct a stochastic model $(X_t)_{t\in\boldsymbol{T}}$ on some probability space $(\Omega, \mathcal{F}, \mathbb{P})$ so that the family of finite dimensional distributions determined by $(X_t)_{t\in\boldsymbol{T}}$ coincides with the family $\mathcal{D}$ of distributions. In this case, $(X_t)_{t\in\boldsymbol{T}}$ is called a *realization* of $\mathcal{D}$.

2.2 Martingales in continuous time

The definition of martingales (super- and sub-martingales) and Doob's fundamental inequalities in discrete-time may be extended to martingales in continuous time, after necessary and obvious modifications. The only new result in the theory of martingales in continuous time is a regularity result about their sample paths.

Let $(\Omega, \mathcal{F}, \mathcal{F}_t, \mathbb{P})$ be a filtered probability space. A $(\mathcal{F}_t)$-adapted (real valued) process $(X_t)_{t\geq 0}$ is called a martingale (resp. super-martingale; resp. sub-martingale), if $\mathbb{E}(X_t|\mathcal{F}_s) = X_s$ (resp. $\mathbb{E}(X_t|\mathcal{F}_s) \leq X_s$; resp. $\mathbb{E}(X_t|\mathcal{F}_s) \geq X_s$) almost surely for any $t \geq s \geq 0$. Similarly, the concept of stopping times can be stated in this setting as well, namely, a function $T : \Omega \to [0, \infty]$ is an $(\mathcal{F}_t)$-stopping time if for every $t \geq 0$, the event $\{T \leq t\}$ belongs to $\mathcal{F}_t$. A new kind of stopping times called predictable times which has no interest in discrete-time case will play a role if the underlying stochastic processes have jumps. A stopping time $T : \Omega \to [0, \infty]$ is *predictable* if there is an increasing sequence (T_n) of $(\mathcal{F}_t)$-stopping times such that for each n, $T_n < T$ and $\lim_{n\to\infty} T_n = T$.

If T is a stopping time, then

$$\mathcal{F}_T = \{A \in \mathcal{F} : \text{ for any } t \geq 0,\ [T \leq t] \cap A \in \mathcal{F}_t\}$$

is the σ-algebra representing the information available up to the random time T, and

$$\mathcal{F}_{T-} = \{A \in \mathcal{F} : \text{ for any } t \geq 0,\ [T < t] \cap A \in \mathcal{F}_t\}$$

represents information known strictly before time T.

Let $\mathcal{F}_{t+} = \cap_{s>t}\mathcal{F}_s$ for every $t \geq 0$, and define $\mathcal{F}_{t-} = \sigma\{\mathcal{F}_s : s < t\}$ for $t > 0$. Then $(\mathcal{F}_{t+})$ is again a filtration and $\mathcal{F}_{t+} \supseteq \mathcal{F}_t$. If $T : \Omega \to [0,\infty]$ is an $(\mathcal{F}_{t+})$-stopping time then

$$\mathcal{F}_{T+} = \{A \in \mathcal{G} : \text{ for any } t \geq 0, \ [T \leq t] \cap A \in \mathcal{F}_{t+}\}$$

is a σ-algebra.

A filtration $(\mathcal{G}_t)$ is right-continuous if $\mathcal{G}_{t+} = \mathcal{G}_t$ for each $t \geq 0$. Clearly $(\mathcal{F}_{t+})$ is right-continuous.

Theorem 2.2.1 *If $(X_t)_{t\geq 0}$ is a martingale (resp. super-martingale, resp. sub-martingale) on $(\Omega, \mathcal{F}, \mathcal{F}_t, \mathbb{P})$ with right-continuous sample paths almost surely, then $(X_t)_{t\geq 0}$ is a martingale (resp. super-martingale, resp. sub-martingale) on $(\Omega, \mathcal{F}, \mathcal{F}_{t+}, \mathbb{P})$.*

One would ask when a martingale (super-martingale) has right-continuous sample paths almost surely? The question can be answered via Doob's convergence theorem for super-martingales.

Let $X = (X_t)_{t\geq 0}$ be a real valued stochastic process, and let $a < b$. If

$$F = \{0 \leq t_1 < t_2 < \cdots < t_N\}$$

is a finite subset of $[0,\infty)$, then $U_a^b(X,F)$ denotes the number of up-crossings by $\{X_{t_1}, \cdots, X_{t_N}\}$, and if $D \subset [0,\infty)$, then $U_a^b(X,D)$ denotes the superemum of $U_a^b(X,F)$ over F being finite subset of D. Obviously $D \to U_a^b(X,D)$ is increasing with respect to the inclusion $\subset$. In particular, if $X = (X_t)_{t\geq 0}$ is $(\mathcal{F}_t)$-adapted and if D is a countable subset of $[0,\infty)$ then for every $t \geq 0$, $U_a^b(X, D \cap [0,t])$ is measurable with respect to $\mathcal{F}_t$. We may apply Doob's up-crossing number inequality to $(X_t)_{t\in F}$ where F is a finite subset, and therefore establish the following

Theorem 2.2.2 *(Doob's up-crossing number inequality). If $X = (X_t)_{t\geq 0}$ is a super-martingale, then*

$$\mathbb{E}\left[U_a^b(X,D)\right] \leq \frac{1}{b-a}\mathbb{E}\left(X_t - a\right)^-$$

for any $a < b$, $t > 0$ and any countable subset D of $[0,t]$, where $x^- = (-x) \vee 0$.

Proof. Let $t > 0$ be any but fixed. List all elements in D as $\{t_1, t_2, \cdots\}$, and for each $n = 1, 2, \cdots$, $F_n = \{t_1, \cdots, t_n, t\}$. Then clearly $F_n \uparrow D \cap [0,t]$, and therefore

$$U_a^b(X, F_n) \uparrow U_a^b(X, D \cup \{t\}).$$

Each $U_a^b(X,F_n)$ is $\mathcal{F}_t$-measurable, and therefore $U_a^b(X, D\cup\{t\})$ is $\mathcal{F}_t$-measurable too. By the same reasoning (dropping t in the definition F_n) $U_a^b(X,D)$ is $\mathcal{F}_t$-measurable. By MCT

$$\mathbb{E}\left[U_a^b(X,D)\right] \leq U_a^b(X, D\cup\{t\}) = \lim_{n\to\infty} U_a^b(X,F_n).$$

While by Doob's up-crossing lemma for S-martingales in discrete-time

$$U_a^b(X,F_n) \leq \frac{1}{b-a}\mathbb{E}\left(X_t - a\right)^-$$

and the claim follows. ■

It follows thus the following version of the super-martingale convergence theorem.

Corollary 2.2.3 *Let $X = (X_t)_{t\geq 0}$ be a super-martingale, and let D be a countable dense subset of $[0,\infty)$. Then for almost all $\omega \in \Omega$, the right limit of $(X_t)_{t\geq 0}$ along the countable dense set D, $\lim_{s\in D, s>t, s\downarrow t} X_s$ exist at all $t \geq 0$, and the left limit along D, $\lim_{s\in D, s<t, s\uparrow t} X_s$ exist at $t > 0$.*

Here is the proof [also the proof of 1) and 2) in Föllmer's lemma below] which is quite typical of this kind of arguments using discrete results. For simplicity, let $D_t = D \cap [0, 2t]$ for $t > 0$. Then, by using the definition of limits, both limits

$$\lim_{s\in D, s>t, s\downarrow t} X_s \quad \text{and} \quad \lim_{s\in D, s<t, s\downarrow t} X_s$$

exist on

$$\left\{U_a^b(X, D_t) < \infty : a < b, a, b \in \mathbb{Q}\right\} = \bigcap_{a,b\in\mathbb{Q}, a<b} \left\{U_a^b(X, D_t) < \infty\right\}$$

which is $\mathcal{F}_{2t}$-measurable (in particular it is measurable). By Doob's up-crossing number lemma

$$\mathbb{E}\left[U_a^b(X, D)\right] \leq \frac{1}{b-a}\mathbb{E}\left(X_t - a\right)^- < \infty$$

so that $\left\{U_a^b(X, D_t) = \infty\right\}$ has probability zero, and therefore

$$N_t \equiv \bigcup_{a,b\in\mathbb{Q}, a<b} \left\{U_a^b(X, D_t) = \infty\right\}$$

has probability zero for every t. Now observe that N_t is actually increasing in t, so that $N = \cup_{t>0} N_t = \lim_{k\to\infty} N_k$ is measurable, and has probability zero. That is $\bigcup_{t\geq 0} N_t$ has probability zero and therefore

$$\left\{U_a^b(X, D_t) < \infty : \text{ for all } a < b, a, b \in \mathbb{Q} \text{ and } t > 0\right\} = \Omega \setminus N$$

has probability one. By construction of N, both limits

$$\lim_{s\in D, s>t, s\downarrow t} X_s \quad \text{and} \quad \lim_{s\in D, s<t, s\downarrow t} X_s$$

exist on $\Omega \setminus N$ for all $t > 0$.

The following is the main regularity result, which is called Föllmer's lemma.

Theorem 2.2.4 *Let $(X_t)_{t\geq 0}$ be a super-martingale (resp. martingale) on $(\Omega, \mathcal{F}, \mathcal{F}_t, \mathbb{P})$, and $D \subseteq [0,\infty)$ be countable. Let N be constructed in the previous proof. Then*

1. *For every $\omega \in \Omega \setminus N$, $Z_t(\omega) = \lim_{s\in D, s>t, s\downarrow t} X_s(\omega)$ exists, and Z_t is $\mathcal{F}_{t+}$-measurable for all $t \geq 0$.*

2. *For every $\omega \in \Omega \setminus N$ and for all $t > 0$ the left limit $Z_{t-}(\omega) = \lim_{s<t, s\uparrow} Z_s(\omega)$, and $(Z_t)_{t\geq 0}$ is a $(\mathcal{F}_{t+})$-adapted process with right-continuous sample paths and left limits.*

3. *For any $t \geq 0$, $\mathbb{E}\left(Z_t | \mathcal{F}_t\right) \leq X_t$ (resp. $\mathbb{E}\left(Z_t | \mathcal{F}_t\right) = X_t$) almost surely.*

4. *$(Z_t)_{t\geq 0}$ is a super-martingale (resp. martingale) on $(\Omega, \mathcal{F}, \mathcal{F}_{t+}, \mathbb{P})$.*

In general however $(Z_t)_{t\geq 0}$ constructed as above may not be not a version of $(X_t)_{t\geq 0}$, and therefore they may have different distributions. The following is the most useful regularity result about martingales.

Corollary 2.2.5 *Under the same assumptions and notations as in Theorem 2.2.4. Assume that $(\mathcal{F}_t)_{t\geq 0}$ is right-continuous. Then $(Z_t)_{t\geq 0}$ is a version of $(X_t)_{t\geq 0}$, that is, for each $t\geq 0$, $Z_t = X_t$ almost surely, if and only if $t \to \mathbb{E}X_t$ is right-continuous.*

Corollary 2.2.6 *Under the same assumptions and notations as in Theorem 2.2.4. If $(\mathcal{F}_t)_{t\geq 0}$ is right continuous, and if $(X_t)_{t\geq 0}$ is a martingale on $(\Omega, \mathcal{F}, \mathcal{F}_t, \mathbb{P})$, then the process $(Z_t)_{t\geq 0}$ defined in 2.2.4 is a version of $(X_t)_{t\geq 0}$.*

This is because for a martingale $(X_t)_{t\geq 0}$, $t \to \mathbb{E}X_t = \mathbb{E}X_0$ is a constant.

From now we will work on a filtered probability space $(\Omega, \mathcal{F}, \mathcal{F}_t, \mathbb{P})$ which satisfies the following *the usual conditions.*

1. $(\Omega, \mathcal{F}, \mathbb{P})$ is a complete probability space.

2. The filtration $(\mathcal{F}_t)_{t\geq 0}$ is right-continuous: $\mathcal{F}_t = \mathcal{F}_{t+} \equiv \cap_{s>t}\mathcal{F}_s$ for every $t \geq 0$.

3. Each $\mathcal{F}_t$ contains all null sets in $\mathcal{F}$.

Remark 2.2.7 *If $X = (X_t)_{t\geq 0}$ is a* ***right-continuous*** *stochastic process on a complete probability space $(\Omega, \mathcal{F}, \mathbb{P})$, then its natural filtration $(\mathcal{F}_t)_{t\geq 0}$ satisfies the usual conditions.*

Theorem 2.2.8 *If $X = (X_t)_{t\geq 0}$ is a right-continuous stochastic process adapted to $(\mathcal{F}_t)_{t\geq 0}$ (recall that our filtration $(\mathcal{F}_t)_{t\geq 0}$ satisfies the usual conditions), and if $T : \Omega \to [0, \infty]$ is a stopping time, then the random variable $X_T 1_{\{T<\infty\}}$ is measurable with respect to σ-algebra $\mathcal{F}_T$, where*

$$\begin{aligned} X_T 1_{\{T<\infty\}}(\omega) &= X_{T(\omega)}(\omega) 1_{\{\omega: T(\omega)<\infty\}}(\omega) \\ &= \begin{cases} X_{T(\omega)}(\omega)\ ; & \text{if } T(\omega) < \infty\ , \\ 0; & \text{if } T(\omega) = \infty\ . \end{cases} \end{aligned}$$

The following theorem provides us with a class of interesting stopping times.

Theorem 2.2.9 *Let $X = (X_t)_{t\geq 0}$ be an $\mathbb{R}^d$-valued, adapted stochastic process that is* right-continuous and has left-limits. *Then $T = \inf\{t \geq t_0 : X_t \in D\}$ is a stopping time, where $D \subset \mathbb{R}^d$ is Borel measurable and $t_0 \geq 0$, with the convention that $\inf \emptyset = \infty$. T is called the* hitting time *of D by the process X.*

Example 2.2.10 *If $X = (X_t)_{t\geq 0}$ is an adapted, continuous process on $(\Omega, \mathcal{F}, \mathcal{F}_t, \mathbb{P})$ and if $D \in \mathbb{R}^d$ is a bounded closed subset of $\mathbb{R}^d$, then $T = \inf\{t \geq 0 : X_t \in D\}$ is a stopping time. If $X_0 \in D^c$, then $X_T 1_{\{T<\infty\}} \in \partial D$. In particular, if $d = 1$ and b is a real number, then $T_b = \inf\{t \geq 0 : X_t = b\}$ is a stopping time. $\sup_{t\in[0,N]} X_t$ is a random variable (where $N > 0$ is any number),*

$$\left\{ \sup_{t\in[0,N]} X_t < b \right\} = \{T_b > N\}$$

and

$$\left\{ \sup_{t\in[0,N]} X_t \geq b \right\} = \{T_b \leq N\} \ .$$

The concept of stopping times provides us with a means of "localizing" quantities. Suppose $(X_t)_{t\geq 0}$ is a stochastic process, and T is a stopping time, then $X^T = (X_{t\wedge T})_{t\geq 0}$ is a stochastic process stopped at (random) time T, where

$$X_{t\wedge T}(\omega) = \begin{cases} X_t(\omega) & \text{if } \ t \leq T(\omega) \ ; \\ X_{T(\omega)}(\omega) & \text{if } \ t \geq T(\omega) \ . \end{cases}$$

Another interesting stopped process at random time T associated with X is the process $X1_{[0,T]}$ which is by definition

$$\begin{aligned} \left(X1_{[0,T]}\right)_t(\omega) &= X_t 1_{\{t\leq T\}}(\omega) \\ &= \begin{cases} X_t(\omega) & \text{if } \ t \leq T(\omega) \ ; \\ 0 & \text{if } \ t > T(\omega) \ . \end{cases} \end{aligned}$$

It is obvious that $X_t^T = X_t 1_{\{t\leq T\}} + X_T 1_{\{t>T\}}$. If $(X_t)_{t\geq 0}$ is adapted to the filtration $(\mathcal{F}_t)_{t\geq 0}$, so are the process $(X_{t\wedge T})_{t\geq 0}$ stopped at stopping time T and $X_t 1_{\{t\leq T\}}$.

Definition 2.2.11 *An adapted stochastic process $X = (X_t)_{t\geq 0}$ is called a* local martingale *if there is an increasing family $\{T_n\}$ of finite stopping times such that $T_n \uparrow \infty \quad$ as $n \to \infty$ and that $(X_{t\wedge T_n})_{t\geq 0}$ is a martingale for each n.*

Similarly, we may define local super- or sub-martingales etc.

2.3 Brownian motion

Let us begin with the definition of Brownian motion as a continuous stochastic process.

A stochastic process $B = (B_t)_{t\geq 0}$ on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ with values in $\mathbb{R}^d$ is called a Brownian motion (BM) in $\mathbb{R}^d$, if

1. $(B_t)_{t\geq 0}$ possesses independent increments: for any $0 \leq t_0 < t_1 < \cdots < t_n$ random variables
$$B_{t_0}, \ B_{t_1} - B_{t_0}, \ \cdots \ , \ B_{t_n} - B_{t_{n-1}}$$
are independent,

2. for any $t > s \geq 0$, random variable $B_t - B_s$ has a normal distribution $N(0, t-s)$, that is, $B_t - B_s$ has pdf (probability density function)
$$p(t-s, x) = \frac{1}{(2\pi(t-s))^{d/2}} \mathrm{e}^{-\frac{|x|^2}{2(t-s)}} \ ; \quad x \in \mathbb{R}^d \ .$$
In other words
$$\mathbb{P}\left(B_t - B_s \in dx\right) = p(t-s, x)\mathrm{d}x \ ,$$

3. almost all sample paths of $(B_t)_{t\geq 0}$ are continuous.

If, in addition, $\mathbb{P}\{B_0 = x\} = 1$ where $x \in \mathbb{R}^d$, then we say $(B_t)_{t\geq 0}$ is a Brownian motion starting at x. $\mathbb{P}\{B_0 = 0\} = 1$ where 0 is the origin of $\mathbb{R}^d$, then we say $(B_t)_{t\geq 0}$ is a standard Brownian motion.

We will see that the condition 3) is not nontrivial, which ensure that many interesting functionals of Brownian motion, such as the running maximum $B_t^\star = \sup_{s\leq t} B_s$, are indeed random variables. Let $p(t,x,y) = p(t,x-y)$, and define

$$P_t f(x) = \int_{\mathbb{R}^d} f(y) p(t,x,y) \mathrm{d}y \qquad \forall f \in C_b(\mathbb{R}^d) \ .$$

for every $t > 0$. Since

$$p(t+s,x,y) = \int_{\mathbb{R}^d} p(t,x,z) p(s,z,y) dz$$

therefore $(P_t)_{t\geq 0}$ is a semigroup on $C_b(\mathbb{R}^d)$. $(P_t)_{t\geq 0}$ is called the heat semigroup in $\mathbb{R}^d$: if $f \in C_b^2(\mathbb{R}^d)$, then $u(t,x) = (P_t f)(x)$ solves the heat equation

$$\left(\frac{1}{2}\Delta + \frac{\partial}{\partial t}\right) u(t,x) = 0 \ ; \quad u(0,\cdot) = f \ ,$$

where $\Delta = \sum_i \frac{\partial^2}{\partial x_i^2}$ is the Laplace operator.

The connection between Brownian motion and the Laplace operator Δ (hence the harmonic analysis) is demonstrated through the following identity:

$$\mathbb{E}\left(f(B_t + x)\right) = \left(P_t f\right)(x) = \frac{1}{(2\pi t)^{d/2}} \int_{\mathbb{R}^d} f(y) \mathrm{e}^{-\frac{|y-x|^2}{2t}} \mathrm{d}y$$

where B_t is a standard Brownian motion.

Example 2.3.1 *If $B = (B_t)_{t\geq 0}$ is a BM in $\mathbb{R}$, then*

$$\mathbb{E}|B_t - B_s|^p = c_p |t-s|^{p/2} \quad \text{for all } s,t \geq 0 \tag{2.1}$$

for $p \geq 0$, where c_p is a constant depending only on p. Indeed

$$\mathbb{E}|B_t - B_s|^p = \frac{1}{\sqrt{2\pi|t-s|}} \int_{\mathbb{R}} |x|^p \exp\left(-\frac{|x|^2}{2|t-s|}\right) dx \ .$$

Making change of variable

$$\frac{x}{\sqrt{|t-s|}} = y \ ; \quad dx = \sqrt{|t-s|} dy$$

we thus have

$$\begin{aligned} \mathbb{E}|B_t - B_s|^p &= \frac{(\sqrt{|t-s|})^p}{\sqrt{2\pi}} \int_{\mathbb{R}} |x|^p \exp\left(-\frac{|x|^2}{2}\right) dx \\ &= c_p |t-s|^{p/2} \end{aligned}$$

where

$$c_p = \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}^d} |x|^p \exp\left(-\frac{|x|^2}{2}\right) dx \ .$$

(2.1) remains true for BM in $\mathbb{R}^d$ with a constant c_p depending on p and d.

Remark 2.3.2 *If $d = 1$, then $B_t - B_s \sim N(0, t-s)$. It is an easy exercise to show that for every $n \in \mathbb{Z}^+$*

$$\mathbb{E}(B_t - B_s)^{2n} = \frac{(2n)!}{2^n n!}|t-s|^n \ .$$

Let $B = (B_t)_{t\geq 0}$ be a standard BM in $\mathbb{R}$. Then B is a centered Gaussian process with co-variance function $C(s,t) = s \wedge t$. Indeed, any finite-dimensional distribution of B is Gaussian [Exercise], so that B is a centered Gaussian process, and its co-variance function (if $s < t$)

$$\begin{aligned}
\mathbb{E}(B_t B_s) &= \mathbb{E}((B_t - B_s)B_s + B_s^2) \\
&= \mathbb{E}((B_t - B_s)B_s) + \mathbb{E}B_s^2 \\
&= \mathbb{E}(B_t - B_s)\mathbb{E}B_s + \mathbb{E}B_s^2 \\
&= s \ .
\end{aligned}$$

Theorem 2.3.3 *(N. Wiener) There is a standard Brownian motion in $\mathbb{R}^d$.*

Let $B = (B_t)_{t\geq 0}$ be a standard BM in $\mathbb{R}^d$.

(1) BM $B = (B_t)_{t\geq 0}$ is stationary, in the sense that for any fixed time S, $\tilde{B}_t = B_{t+S} - B_S$ is again a standard Brownian motion. This statement is true indeed for any finite stopping time S.

(2) *Scaling invariance, self-similarity.* For any real number $\lambda \neq 0$, $M_t \equiv \lambda B_{t/\lambda^2}$is a standard BM in $\mathbb{R}^d$.

(3) *Isotropic property.* If U is an $d \times d$ orthonormal matrix, then $UB = (UB_t)_{t\geq 0}$ is a standard BM in $\mathbb{R}^d$. That is, BM is invariant under the action of orthogonal group of $\mathbb{R}^d$. This says a little bit more than the usual definition of isotropic property in that the invariance is valid for all $U \in \mathrm{O}(d)$, not only for its sub-group $\mathrm{SO}(d)$.

(4) Let $B = (B_t)_{t\geq 0}$ be a standard BM in $\mathbb{R}$, and define $M_0 = 0$ and $M_t = tB_{1/t}$ for $t > 0$, is a standard BM in $\mathbb{R}$.

Brownian motion as a Markov process. Recall that

$$p(t,x) = \frac{1}{(2\pi t)^{d/2}} \mathrm{e}^{-\frac{|x|^2}{2t}}$$

in $\mathbb{R}^d$, and $(P_t)_{t\geq 0}$ the heat semigroup $P_t f(x) = \int_{\mathbb{R}^d} f(y)p(t, x-y)\mathrm{d}y$ for every $t > 0$.

Let $(\mathcal{F}_t^0)_{t\geq 0}$ denote the filtration generated by a standard Brownian motion $(B_t)_{t\geq 0}$, and $\mathcal{F}_\infty^0 = \cup_{t\geq 0}\mathcal{F}_t^0$.

(5) For any $t > s \geq 0$, the increment $B_t - B_s$ is independent of $\mathcal{F}_s^0$.

(6) If $t > s$, then the joint distribution of B_s and B_t is given by

$$\mathbb{P}\left(B_s \in \mathrm{d}x, \ B_t \in \mathrm{d}y\right) = p(s,x)p(t-s, y-x)\mathrm{d}x\mathrm{d}y \ .$$

Indeed, since B_s and $B_t - B_s$ are independent, so that $(B_s, B_t - B_s)$ has a pdf $p(s, x_1)p(t - s, x_2)$, thus, for any bounded Borel measurable function f

$$\begin{aligned} \mathbb{E}f(B_s, B_t) &= \mathbb{E}f(B_s, B_t - B_s + B_s) \\ &= \iint f(x_1, x_2 + x_1)p(s, x_1)p(t - s, x_2)\mathrm{d}x_1\mathrm{d}x_2 \ . \end{aligned}$$

Making change of variables $x_1 = x$ and $x_2 + x_1 = y$ in the last double integral, the induced Jacobi is 1 so that $\mathrm{d}x_1\mathrm{d}x_2 = \mathrm{d}x\mathrm{d}y$ (as measures), and therefore

$$\mathbb{E}f(B_s, B_t) = \iint f(x, y)p(s, x)p(t - s, y - x)\mathrm{d}x\mathrm{d}y$$

which implies that the pdf of (B_s, B_t) is $p(s, x)p(t - s, y - x)$.

(7) Let $t > s$, and f a bounded Borel measurable function. Then

$$\mathbb{E}\left(f(B_t)|\mathcal{F}_s^0\right) = P_{t-s}f(B_s) \quad \text{a.s.} \tag{2.2}$$

In particular $\mathbb{E}\left(f(B_t)|\mathcal{F}_s^0\right) = \mathbb{E}\left(f(B_t)|B_s\right)$ (which is called the simple Markov property).

(8) For any $0 < t_1 < t_2 < \cdots < t_n$, the ($\mathbb{R}^{n\times d}$-valued) random variable $(B_{t_1}, \cdots, B_{t_n})$ possesses the following probability density function

$$p(t_1, x_1)p(t_2 - t_1, x_2 - x_1) \cdots p(t_n - t_{n-1}, x_n - x_{n-1}).$$

That is, the joint distribution of $(B_{t_1}, \cdots, B_{t_n})$ is given by

$$\begin{aligned} &\mathbb{P}\left\{B_{t_1} \in \mathrm{d}x_1, \ \cdots, \ B_{t_n} \in \mathrm{d}x_n\right\} \\ = \ &p(t_1, x_1)p(t_2 - t_1, x_2 - x_1) \cdots p(t_n - t_{n-1}, x_n - x_{n-1})\mathrm{d}x_1 \cdots \mathrm{d}x_n \ . \end{aligned} \tag{2.3}$$

(9) Let $B_t = (B_t^1, \cdots, B_t^d)$ be a d-dimensional standard Brownian motion. Then for each j, B_t^j is a standard BM in $\mathbb{R}$, and $(B_t^j)_{t\geq 0}$ $(j = 1, \cdots, d)$ are mutually independent. Therefore a d-dimensional BM is d independent copies of BM in $\mathbb{R}$.

(10) Brownian motion starts afresh at a stopping time, and the Markov property for Brownian motion remains valid at a stopping time. Therefore Brownian motion possesses the *strong Markov property*, a very important property which had been used by Paul Lévy in the form of the *reflection principle*, long before the concept of strong Markov property had been properly defined. We will exhibit this principle by computing the distribution of *the running maximum* of a Brownian motion. Let $B = (B_t)_{t\geq 0}$ be a standard one dimensional Brownian motion on $(\Omega, \mathcal{F}, \mathcal{F}_t, \mathbb{P})$ in $\mathbb{R}$. Let $b > 0$, $b > a$ and $T_b = \inf\{t > 0 : B_t = b\}$. Then T_b is a stopping time, and the Brownian motion starts afresh at T_b as a Brownian motion (starting at b) after hitting the level b, and therefore

$$\begin{aligned} \mathbb{P}\left(\sup_{s\in[0,t]} B_s \geq b, B_t \leq a\right) &= \mathbb{P}\left(\sup_{s\in[0,t]} B_s \geq b, B_t \geq 2b - a\right) \\ &= \mathbb{P}\left(B_t \geq 2b - a\right) \end{aligned}$$

where the first equality follows from the "fact" that the Brownian motion starting at T_b (in position b): $B_{T_b} = b$, runs afresh like a Brownian motion starting at b, so that it moves with equal probability about the line $y = b$. The second equality follows from $2b - a = b + (b-a) > b$.

The above equation may be written as

$$\begin{aligned} \mathbb{P}\left(T_b \leq t, B_t \leq a\right) &= \mathbb{P}\left(T_b \leq t, B_t \geq 2b - a\right) \\ &= \mathbb{P}\left(B_t \geq 2b - a\right), \end{aligned}$$

which can be justified by the *Strong Markov Property* of Brownian motion, a topic that will not pursue here. Therefore

$$\mathbb{P}\left(\sup_{s\in[0,t]} B_s \geq b, B_t \leq a\right) = \frac{1}{\sqrt{2\pi t}} \int_{2b-a}^{+\infty} e^{-\frac{x^2}{2t}} \mathrm{d}x \ ,$$

which gives us the joint distribution of a Brownian motion and its maximum at a fixed time t. By differentiating in a and in b to obtain the pdf of the joint distribution of random variables $(M_t = \sup_{s\in[0,t]} B_s, B_t)$:

$$\mathbb{P}\left(M_t \in \mathrm{d}b, B_t \in \mathrm{d}a\right) = \frac{2(2b-a)}{\sqrt{2\pi t^3}} \exp\left\{-\frac{(2b-a)^2}{2t}\right\} \mathrm{d}a\mathrm{d}b$$

over the region $\{(b,a) : a \leq b, b \geq 0\}$ in $\mathbb{R}^2$.

Brownian motion as a martingale.

Let $B = (B_t^i)_{t\geq 0}$ $(i = 1, \cdots, d)$ be a standard BM in $\mathbb{R}^d$, with its natural filtration $(\mathcal{F}_t^0)_{t\geq 0}$.

(11) Each B_t is p-th integrable for any $p > 0$, and for $t > s$

$$\mathbb{E}(|B_t - B_s|^p) = c_{p,d}|t-s|^{p/2} \ . \tag{2.4}$$

$(B_t)_{t\geq 0}$ is a continuous, square-integrable martingale, and for each pair i, j, $M_t = B_t^i B_t^j - \delta_{ij} t$ is a continuous martingale.

(12) Let $B = (B_t)_{t\geq 0}$ be a continuous stochastic process in $\mathbb{R}$ such that $B_0 = 0$. Then $(B_t)_{t\geq 0}$ is a standard BM in $\mathbb{R}$, if and only if for any $\xi \in \mathbb{R}$ and $t > s$

$$\mathbb{E}\left\{\exp\left(i\langle \xi, B_t - B_s\rangle\right) | \mathcal{F}_s^0\right\} = \exp\left(-\frac{(t-s)|\xi|^2}{2}\right) \ . \tag{2.5}$$

(13) Let (B_t) be a standard BM in $\mathbb{R}$. If $\xi \in \mathbb{R}$, then $M_t \equiv \exp\left(i\langle \xi, B_t\rangle + \frac{|\xi|^2}{2} t\right)$ is a martingale.

(14) Let $B = (B_t)_{t\geq 0}$ be a standard BM in $\mathbb{R}$. Then $M_t \equiv B_t^2 - t$ are martingales and $\lim_{m(D)\to 0} \sum_l |B_{t_l} - B_{t_{l-1}}|^2 = t$ in $L^2(\Omega, \mathbb{P})$ for any t, where D runs over all finite partitions of interval $[0,t]$, and $m(D) = \max_l |t_l - t_{l-1}|$. Therefore $\lim_{m(D)\to 0} \sum_l |B_{t_l} - B_{t_{l-1}}|^2 = t$ in probability.

(15) Let $(B_t)_{t\geq 0}$ be a standard BM in $\mathbb{R}$. Then for any $t > 0$ we have

$$\sum_{j=1}^{2^n} \left|B_{\frac{j}{2^n}t} - B_{\frac{j-1}{2^n}t}\right|^2 \to t \quad \text{a.s.} \tag{2.6}$$

as $n \to \infty$.

It can be shown (not easy) that $\sup_D \sum_l |B_{t_l} - B_{t_{l-1}}|^p < \infty$ almost surely for any $p > 2$, where sup is taken over all finite partitions of $[0,1]$, and $\sup_D \sum_l |B_{t_l} - B_{t_{l-1}}|^2 = \infty$ almost surely. That is to say, Brownian motion has finite p-variation for any $p > 2$. Indeed almost all Brownian motion sample paths are α-Hölder continuous for any $\alpha < 1/2$ but not for $\alpha = 1/2$. It follows that almost all Brownian motion paths are nowhere differentiable. We will not go into a deep study about the sample paths of BM, which are not needed in order to develop Itô's calculus for Brownian motion.

Definition 2.3.4 *Let $p > 0$ be a constant. A path $f(t)$ in $\mathbb{R}^d$ [a function on $[0,T]$ valued in $\mathbb{R}^d$] is said to have finite p-variation on $[0,T]$, if*

$$\sup_D \sum_l |f(t_i) - f(t_{i-1})|^p < \infty$$

where D runs over all finite partitions of $[0,T]$. $f(t)$ in $\mathbb{R}^d$ has finite (total) variation if it has finite 1-variation.

A function with finite variation must be a difference of two increasing functions. It particular, it has at most countably many discontinuous points.

A stochastic process $V = (V_t)_{t\geq 0}$ is called a *variational process*, if for almost all $\omega \in \Omega$, the sample path $t \to V_t(\omega)$ possesses finite variation on any finite interval. A Brownian motion is not a variational process.

2.4 Itô's calculus

In this section we recall the basic definitions and results about Itô's integrals with respect to continuous martingales.

2.4.1 Quadratic processes

Let $M = (M_t)_{t\geq 0}$ be a continuous, square integrable martingale. Then

$$\langle M \rangle_t = \lim_{m(D)\to 0} \sum_l \left| M_{t_l} - M_{t_{l-1}} \right|^2$$

exists both in probability and in the L^2-norm for every $t \geq 0$, where the limit takes over all finite partitions D of the interval $[0,t]$. $\langle M \rangle$ is called the (quadratic) variational process of $(M_t)_{t\geq 0}$, or simply *the bracket process* of $(M_t)_{t\geq 0}$. The quadratic variational process $t \to \langle M \rangle_t$ is an adapted, continuous, increasing stochastic process [and therefore has finite variation] with initial zero. The following theorem demonstrates the importance of $\langle M \rangle_t$.

Theorem 2.4.1 (The quadratic variational process) *Let $M = (M_t)_{t\geq 0}$ be a continuous, square integrable martingale. Then $\langle M \rangle_t$ is the unique continuous, adapted and increasing process with initial zero, such that $M_t^2 - \langle M \rangle_t$ is a martingale.*

The theorem is a special case of the Doob-Meyer decomposition for sub-martingales: any sub-martingale can be decomposed into a sum of a martingale and a predictable, increasing process with initial value zero. The decomposition was conjectured by L. Doob, and proved by P. A. Meyer in the 60's, which opened the new era of stochastic calculus.

Theorem 2.4.2 *Let $(M_t)_{t\geq 0}$ and $(N_t)_{t\geq 0}$ be two continuous, square integrable martingales, and*

$$\langle M,N\rangle_t = \frac{1}{4}\left(\langle M+N\rangle_t - \langle M-N\rangle_t\right)$$

called the bracket process of M and N. Then $\langle M,N\rangle_t$ is the unique adapted, continuous, variational process with initial zero, such that $M_tN_t - \langle M,N\rangle_t$ is a martingale. Moreover

$$\lim_{m(D)\to 0}\sum_{l=1}^{n}(M_{t_l}-M_{t_{l-1}})(N_{t_l}-N_{t_{l-1}}) = \langle M,N\rangle_t \ , \quad \textit{in prob.} \tag{2.7}$$

where $D=\{0=t_0<\cdots<t_n=t\}$ and $m(D)=\max_l(t_l-t_{l-1})$.

Let $T>0$ be a fixed but arbitrary number, and $\mathcal{M}_0^2$ be the vector space of all continuous, square-integrable martingales up to time T on a probability space $(\Omega,\mathcal{F},\mathcal{F}_t,\mathbb{P})$ with *initial value zero*, equipped with the distance

$$d(M,N)=\sqrt{\mathbb{E}|M_T-N_T|^2} \quad \text{for } M,N\in\mathcal{M}_0^2 \ .$$

By definition, a sequence of square-integrable martingales $(M(k)_t)_{t\geq 0}$ $(k=1,\cdots,)$ converges to M in $\mathcal{M}_0^2$, if and only if $M(k)_T\to M_T$ in $L^2(\Omega,\mathcal{F},\mathbb{P})$ as $k\to\infty$. The following maximal inequality, which is the "martingale version" of the Markov inequality, allows us to show that $(\mathcal{M}_0^2,d)$ indeed is complete.

Theorem 2.4.3 *(Kolmogorov's inequality) Let $M\in\mathcal{M}_0^2$. Then for any $\lambda>0$*

$$\mathbb{P}\left(\sup_{0\leq t\leq T}|M_t|\geq\lambda\right)\leq\frac{1}{\lambda^2}\mathbb{E}\left(M_T^2\right) \ .$$

Theorem 2.4.4 $(\mathcal{M}_0^2,d)$ *is a complete metric space.*

Proof. Let $M(k)\in\mathcal{M}_0^2$ ($k=1,2,\cdots$) be a Cauchy sequence in $\mathcal{M}_0^2$. Then

$$\mathbb{E}|M(k)_T-M(l)_T|^2\to 0 \ , \quad \text{as } k,l\to\infty \ .$$

According to Kolmogorov's inequality

$$\mathbb{P}\left(\sup_{0\leq t\leq T}|M(k)_t-M(l)_t|\geq\lambda\right)\leq\frac{1}{\lambda^2}\mathbb{E}|M(k)_T-M(l)_T|^2 \ ,$$

so that, $M(k)$ uniformly converges to a limit M on $[0,T]$ in probability. Therefore there exists a stochastic process $M\equiv(M_t)$ such that $\sup_{0\leq t\leq T}|M(k)_t-M_t|\to 0$ in probability. Obviously $(M_t)_{t\geq 0}$ is a continuous and square -integrable martingale (up to time T) as the uniform limit of a sequence of continuous martingales. ∎

2.4.2 Itô's integrals

Let us now recall the definition of Itô's integral $\int_0^t F_s dM_s$ which is again a (local) martingale. This local martingale is also denoted by $F \cdot M$, where F is called an integrand.

The definition is divided into three steps.

Step 1. Assume that M is a continuous square integrable martingale, i.e. M is a continuous martingale on $(\Omega, \mathcal{F}, \mathcal{F}_t, \mathbb{P})$ such that $\mathbb{E}(M_t^2) < \infty$ for every $t > 0$.

If $F = (F_t)$ is a simple stochastic process if

$$F_t = \xi_0 1_{\{0\}}(t) + \sum_{i=1}^{\infty} \xi_i 1_{(t_i, t_{i+1}]}(t)$$

for some partition $0 \le t_1 < t_2 < \cdots$, where $t_i \to \infty$, F is adapted so that $\xi_i \in \mathcal{F}_{t_i}$, and ξ_i is bounded for $i = 1, 2, \cdots$. For such F the Itô's integral

$$(F \cdot M)_t = \int_0^t F_s dM_s = \sum_{i=1}^{\infty} \xi_i \left(M_{t_{i+1} \wedge t} - M_{t_i \wedge t} \right)$$

for every $t \ge 0$. By definition

$$(F \cdot M)_t = \int_0^t F_s dM_s = \lim_{m(D) \to 0} \sum_i F_{s_i} \left(M_{s_{i+1}} - M_{s_i} \right)$$

for every $t > 0$, where the limit takes over all finite partitions $D : 0 = s_0 < s_1 < \ldots < s_k = t$.

$F \cdot M$ is a continuous square integrable martingale and

$$\langle F \cdot M \rangle_t = \int_0^t |F_s|^2 d \langle M \rangle_s \tag{2.8}$$

which yields the Itô's isometry

$$\mathbb{E} \left[\langle F \cdot M \rangle_t \right] = \mathbb{E} \left[\int_0^t |F_s|^2 d \langle M \rangle_s \right] \tag{2.9}$$

for every $t \ge 0$, which allows us to generalize the definition to more general integrands.

An adapted process $F = (F_t)$ is said to be in $\mathcal{L}^2(M)$ if there is a sequence of simple adapted processes $F(n)$ (for $n = 1, 2, \ldots$)

$$\mathbb{E} \left(\int_0^T |F(n)_t - F_t|^2 d \langle M \rangle_t \right) \to 0$$

as $n \to \infty$ for every $T > 0$. The Itô's isometry implies that

$$(F \cdot M)_t = \lim_{n \to \infty} (F_n \cdot M)_t$$

in $\mathcal{M}_0^2$. $F \cdot M$ is also continuous square martingale and (2.8) holds. Moreover for any $F \in \mathcal{L}^2(M)$ then

$$(F \cdot M)_t = \lim_{m(D) \to 0} \sum_l F_{t_{l-1}} \left(M_{t_l} - M_{t_{l-1}} \right) \quad \text{in probability}$$

where the limit takes over all finite partitions of $[0, t]$.

By the use of the polarization identity, if M, $N \in \mathcal{M}_0^2$ and $F \in \mathcal{L}^2(M)$, $G \in \mathcal{L}^2(N)$, then

$$\langle F.M, G.N \rangle_t = \int_0^t F_s G_s \mathrm{d} \langle M, N \rangle_s$$

and $F.(G.M) = (FG).M$, as far as these stochastic integrals make sense. That is,

$$\int_0^t F_s \mathrm{d} \left(\int_0^s G_u \mathrm{d}M_u \right)_s = \int_0^t F_s G_s \mathrm{d}M_s \ .$$

Step 2. The Itô integration may be extended to local martingales. Let me briefly describe the idea. Suppose $M = (M_t)_{t\geq 0}$ is a continuous, local martingale with initial zero, then we may choose a sequence (T_n) of stopping times such that $T_n \uparrow \infty$ a.s. and for each n, $M^{T_n} = (M_{t\wedge T_n})_{t\geq 0}$ is a continuous, square integrable martingale with initial zero. In this case we may define $\langle M \rangle_t = \langle M^{T_n} \rangle_t$ for $t \leq T_n$, which is an adapted, continuous, increasing process with initial zero such that $M_t^2 - \langle M \rangle_t$ is a local martingale.

Let $F = (F_t)_{t\geq 0}$ be a left-continuous, adapted process such that for each $T > 0$

$$\int_0^T F_s^2 \mathrm{d}\langle M \rangle_s < \infty \qquad \text{a.s.} \tag{2.10}$$

and define

$$S_n = \inf \left\{ t \geq 0 : \int_0^t F_s^2 \mathrm{d}\langle M \rangle_s \geq n \right\} \wedge n$$

which is a sequence of stopping times. Condition (2.10) ensures that $S_n \uparrow \infty$. Let $\tilde{T}_n = T_n \wedge S_n$. Then $\tilde{T}_n \uparrow \infty$ almost surely, and for each n, $M^{\tilde{T}_n} \in \mathcal{M}_0^2$. Let $F(n)_t = F_t 1_{\{t\leq \tilde{T}_n\}}$. Then

$$\int_0^\infty F(n)_s^2 \mathrm{d}\langle M \rangle_s = \int_0^{\tilde{T}_n} F_s^2 \mathrm{d}\langle M \rangle_s \leq n$$

so that $F(n) \in \mathcal{L}_2(M^{\tilde{T}_n})$. We may define

$$(F.M)_t = \int_0^t F(n)_s \mathrm{d} \left(M^{\tilde{T}_n} \right)_s \qquad \text{if} \;\; t \leq \tilde{T}_n \uparrow \infty$$

for $n = 1, 2, 3, \cdots$, called the Itô integral of F with respect to local martingale M. It can be shown that $F.M$ does not depend on the choice of stopping times T_n. By definition, both $F.M$ and

$$(F.M)_t^2 - \int_0^t F_s^2 \mathrm{d}\langle M \rangle_s$$

are continuous, local martingales with initial zero.

Step 3. Finally let us extend the theory of stochastic integrals to the most useful class of (continuous) semi-martingales. An adapted, continuous stochastic process $X = (X_t)_{t\geq 0}$ is a semi-martingale if X possesses a decomposition $X_t = M_t + V_t$, where $(M_t)_{t\geq 0}$ is a continuous local martingale, and $(V_t)_{t\geq 0}$ is stochastic processes with finite variation on any finite interval.

If $f(t)$ is a function on $[0, T]$ *having finite variation*:

$$\sup_D \sum_l |f(t_l) - f(t_{l-1})| < +\infty$$

where D runs over all finite partitions of $[0, t]$ (for any fixed t), then $\int_0^t g(s)\mathrm{d}f(s)$ is understood as the Lebesgue-Stieltjes integral. If in addition $s \to f(s)$ is continuous, then

$$\int_0^t g(s)\mathrm{d}f(s) = \lim_{m(D)\to 0} \sum_l g(t_{l-1})(f(t_l) - f(t_{l-1})) \ .$$

Therefore, if $V = (V_t)_{t\geq 0}$ is a continuous stochastic process with finite variation, then $\int_0^t F_s \mathrm{d}V_s$ is a stochastic process defined path-wisely as the Lebesgue-Stieltjes integral

$$\begin{aligned} \int_0^t F_s \mathrm{d}V_s(\omega) &\equiv \int_0^t F_s(\omega)\mathrm{d}V_s(\omega) \\ &= \lim_{m(D)\to 0} \sum_l F_{t_{l-1}}(\omega)(V_{t_l}(\omega) - V_{t_{l-1}}(\omega)) \ . \end{aligned}$$

The definition of stochastic integrals may be extended to any continuous *semi-martingale* in an obvious way, namely

$$\int_0^t F_s \mathrm{d}X_s = \int_0^t F_s \mathrm{d}M_s + \int_0^t F_s \mathrm{d}V_s$$

where, the first term on the right-hand side is the Itô's integral with respect to local martingale M defined in probability sense, which is again a local martingale, the second term is the usual Lebesgue-Stieltjes integral which is defined path-wisely. Moreover

$$\int_0^t F_s \mathrm{d}X_s = \lim_{m(D)\to 0} \sum_l F_{t_{l-1}} \left(X_{t_l} - X_{t_{l-1}} \right) \qquad \text{in probab.}$$

2.4.3 Itô's formula

Ito's formula was established by K. Itô in 1944. Since Itô's stated it as a lemma in his seminal paper [ ], Itô's formula is also refereed in literature as Itô's Lemma. Itô's Lemma is indeed the Fundamental Theorem in stochastic calculus.

We have used in many occasions the following elementary formula

$$X_{t_j}^2 - X_{t_{j-1}}^2 = \left(X_{t_j} - X_{t_{j-1}} \right)^2 + 2X_{t_{j-1}} \left(X_{t_j} - X_{t_{j-1}} \right) \ .$$

If in addition $(X_t)_{t\geq 0}$ is a continuous square integrable martingale, then, by adding up the above identity over $j = 1, \cdots, n$, where $0 = t_0 < t_1 < \cdots < t_n = t$ is an arbitrary finite partition, one obtains

$$X_t^2 - X_0^2 = 2\sum_{j=1}^n X_{t_{j-1}} \left(X_{t_j} - X_{t_{j-1}} \right) + \sum_{j=1}^n \left(X_{t_j} - X_{t_{j-1}} \right)^2 \ .$$

Letting $m(D) \to 0$, we obtain

$$X_t^2 - X_0^2 = 2\int_0^t X_s dX_s + \langle X \rangle_t \ .$$

which is the Itô formula for the martingale $(X_t)_{t\geq 0}$ applying to $f(x) = x^2$. By using polarization and localization, we establish the following integration by parts formula.

Lemma 2.4.5 *If X, Y are two continuous semi-martingales: $X = M + A$ and $Y = N + B$, where M and N are two continuous local martingales, A and B are two adapted variational process, then*

$$X_t Y_t - X_0 Y_0 = \int_0^t Y_s dX_s + \int_0^t X_s dY_s + \langle M, N \rangle_t \ .$$

Corollary 2.4.6 *(Integration by parts) Let $X = M + A$ and $Y = N + B$ be a continuous semi-martingale: M and N are continuous local martingales, and A, B are continuous, adapted processes with finite variations. Then*

$$X_t Y_t - X_0 Y_0 = \int_0^t X_s dY_s + \int_0^t Y_s dX_s + \langle M, N \rangle_t \ .$$

The following is the fundamental theorem in stochastic calculus.

Theorem 2.4.7 *(Itô's formula) Let $X = (X_t^1, \cdots, X_t^d)$ be a continuous semi-martingale in $\mathbb{R}^d$ with decomposition $X_t^i = M_t^i + A_t^i$: $M_t^1, \cdots, M_t^d$ are continuous local martingales, and $A_t^1, \cdots, A_t^d$ are continuous, locally integrable, adapted processes with finite variations. Let $f \in C^2(\mathbb{R}^d, \mathbb{R})$. Then*

$$\begin{aligned} f(X_t) - f(X_0) &= \sum_{i=1}^d \int_0^t \frac{\partial f}{\partial x_i}(X_s) dX_s^i \\ &+ \frac{1}{2}\sum_{i,j=1}^d \int_0^t \frac{\partial^2 f}{\partial x_i \partial x_j}(X_s) d\langle M^i, M^j \rangle_s \ . \end{aligned} \tag{2.11}$$

The first term on the right-hand side of (2.11) can be decomposed into

$$\sum_{j=1}^d \int_0^t \frac{\partial f}{\partial x_i}(X_s) dM_s^j + \sum_{j=1}^d \int_0^t \frac{\partial f}{\partial x_i}(X_s) dA_s^j$$

so that $f(X_t) - f(X_0)$ is again a semi-martingale with its martingale part given by

$$M_t^f = \sum_{j=1}^d \int_0^t \frac{\partial f}{\partial x_i}(X_s) dM_s^j \ .$$

It follows that

$$\langle M^f, M^g \rangle_t = \int_0^t \sum_{i,j=1}^d \frac{\partial f}{\partial x_i}(X_s) \frac{\partial g}{\partial x_j}(X_s) \mathrm{d}\langle M^i, M^j \rangle_s \ .$$

If $B = (B_t^1, \cdots, B_t^d)_{t \geq 0}$ is Brownian motion in $\mathbb{R}^d$, then, for $f \in C^2(\mathbb{R}^d, \mathbb{R})$

$$f(B_t) - f(B_0) = \int_0^t \nabla f(B_s).\mathrm{d}B_s + \int_0^t \frac{1}{2}\Delta f(B_s)\mathrm{d}s \ .$$

Let

$$M_t^{[f]} = f(B_t) - f(B_0) - \int_0^t \frac{1}{2}\Delta f(B_s)\mathrm{d}s \ .$$

Then $M^{[f]}$ is a local martingale and

$$\langle M^{[f]}, M^{[g]}\rangle_t = \int_0^t \langle \nabla f, \nabla g\rangle(B_s)\mathrm{d}s \ .$$

Chapter 3

Selected applications of Itô's formula

We present several applications of Itô's lemma.

3.1 Lévy's characterization of Brownian motion

Our first application is Lévy's martingale characterization of Brownian motion. Let $(\Omega, \mathcal{F}, \mathcal{F}_t, \mathbb{P})$ be a filtered probability space satisfying the usual condition.

Theorem 3.1.1 *Let $M_t = (M_t^1, \cdots, M_t^d)$ be an adapted, continuous stochastic process on $(\Omega, \mathcal{F}, \mathcal{F}_t, \mathbb{P})$ taking values in $\mathbb{R}^d$ with initial zero. Then $(M_t)_{t\geq 0}$ is a Brownian motion if and only if*

1. *Each M_t^i is a continuous square-integrable martingale.*

2. *$M_t^i M_t^j - \delta_{ij} t$ is a martingale, that is, $\langle M^i, M^j \rangle_t = \delta_{ij} t$ for every pair (i, j).*

Proof. We only need to prove the sufficient part. Recall that, under the assumption, $(M_t)_{t\geq 0}$ is a Brownian motion if and only if

$$\mathbb{E}\left[e^{\sqrt{-1}\langle \xi, M_t - M_s \rangle} \middle| \mathcal{F}_s \right] = e^{-\frac{|\xi|^2}{2}(t-s)} \tag{3.1}$$

for any $t > s$ and $\xi = (\xi_i) \in \mathbb{R}^d$. We thus consider the adapted process

$$Z_t = \exp\left(\sqrt{-1} \sum_{i=1}^d \xi_i M_t^i + \frac{|\xi|^2}{2} t \right)$$

and we show that it is a martingale. To this end, we apply Itô's formula to $f(x) = e^x$ (in this case $f' = f'' = f$) and semi-martingale

$$X_t = \sqrt{-1} \sum_{i=1}^d \xi_i M_t^i + \frac{|\xi|^2}{2} t \ ,$$

and obtain

$$\begin{aligned}
Z_t &= Z_0 + \int_0^t Z_s \mathrm{d}\left(\sqrt{-1}\sum_{i=1}^d \xi_i M_s^i + \frac{|\xi|^2}{2}s\right) \\
&\quad + \frac{1}{2}\int_0^t Z_s \mathrm{d}\langle\sqrt{-1}\sum_{i=1}^d \xi_i M^i\rangle_s \\
&= 1 + \sqrt{-1}\sum_{i=1}^d \xi_i \int_0^t Z_s \mathrm{d}M_s^i + \frac{|\xi|^2}{2}\int_0^t Z_s \mathrm{d}s \\
&\quad - \frac{1}{2}\int_0^t \sum_{i,j=1}^d \xi_i\xi_j Z_s \mathrm{d}\langle M^i, M^j\rangle_s \\
&= 1 + \sqrt{-1}\sum_{i=1}^d \xi_i \int_0^t Z_s \mathrm{d}M_s^i
\end{aligned}$$

the last equality follows from

$$\frac{1}{2}\int_0^t \sum_{i,j=1}^d \xi_i\xi_j Z_s \mathrm{d}\langle M^i, M^j\rangle_s = \frac{1}{2}|\xi|^2 \int_0^t Z_s \mathrm{d}s \ .$$

due to the assumption that $\langle M^i, M^j\rangle_s = \delta_{ij}s$. Since $|Z_s| = e^{|\xi|^2 s/2}$, so that for any $T > 0$

$$\mathbb{E}\int_0^T |Z_s|^2 \mathrm{d}s = \int_0^T e^{|\xi|^2 s}\mathrm{d}s < +\infty$$

and therefore $(Z_t) \in \mathcal{L}^2(M^i)$ for $i = 1, \cdots, d$ as $\langle M^i\rangle_t = t$. It follows that $\int_0^t Z_s \mathrm{d}M_s^i \in \mathcal{M}_2^c$. That is, Z_s is a continuous, square-integrable martingale with initial value 1. (3.1) follows from the martingale property

$$\mathbb{E}\left\{ e^{i\langle\xi, M_t\rangle + \frac{|\xi|^2}{2}t} \middle| \mathcal{F}_s \right\} = e^{i\langle\xi, M_s\rangle + \frac{|\xi|^2}{2}s}$$

for $t > s$. ■

3.2 Time-changes of Brownian motion

A continuous local martingale is actually a time change of Brownian motion.

Theorem 3.2.1 (Dambis, Dubins, Schwarz) *Let $M = (M_t)_{t\geq 0}$ be a continuous, local martingale on $(\Omega, \mathcal{F}, \mathcal{F}_t, \mathbb{P})$ with initial value zero satisfying $\langle M\rangle_\infty = \infty$, and let*

$$T_t = \inf\left\{s : \langle M\rangle_s > t\right\}.$$

Then T_t is a stopping time for each $t \geq 0$, $B_t = M_{T_t}$ is an $(\mathcal{F}_{T_t})$-Brownian motion, and $M_t = B_{\langle M\rangle_t}$.

Proof. The family $T = (T_t)_{t\geq 0}$ is called a time-change, each T_t is a stopping time, and obviously $t \to T_t$ is increasing. The function $t \to T_t$ is right-continuous, and is the right-continuous inverse of $t \to \langle M\rangle_t$, so that $\langle M\rangle_{T_t} = t$ for all $t \geq 0$. Each T_t is finite $\mathbb{P}$-a.e. because $\langle M\rangle_\infty = \infty$ a.e. By continuity of $\langle M\rangle_t$ we have

$$\langle M\rangle_{T_t} = t$$

for every $t \geq 0$. Applying Doob's optional sampling theorem for the square integrable martingale $(M_{s\wedge T_t})_{s\geq 0}$ (which is a square integrable martingale) and stopping times $T_t \geq T_s$ $(t \geq s)$, we obtain that

$$\mathbb{E}\left(M_{T_t}|\mathcal{F}_{T_s}\right) = M_{T_s}$$

i.e. B_t is a $(\mathcal{F}_{T_t})$-martingale. By the same argument but to the martingale $(M^2_{s\wedge T_t} - \langle M\rangle_{s\wedge T_t})_{s\geq 0}$ we have

$$\mathbb{E}\left(M^2_{T_t} - \langle M\rangle_{T_t}|\mathcal{F}_{T_s}\right) = M^2_{T_s} - \langle M\rangle_{T_s}.$$

Hence $(B_t^2 - t)$ is an $(\mathcal{F}_{T_t})$-martingale. We may verify that $t \to B_t$ is continuous (which is actually not trivial, see the comment below). Therefore $B = (B_t)_{t\geq 0}$ is an $(\mathcal{F}_{T_t})$ Brownian motion according to Lévy's characterization of Brownian motion ∎

Remark 3.2.2 *The continuity of $t \to B_t$ can be argued as the following (outline). Since $t \to T_t$ is right continuous, so B is righy continuous at least. Note that $t \to \langle M\rangle_t$ is continuous and increasing, and*

$$\langle M\rangle_t = \lim_{m(D)\to 0} \sum_i \left(M_{t_i} - M_{t_{i-1}}\right)^2$$

where the limit runs over any finite partitions of $[0,t]$. Let us assume that the previous equality holds for every $\omega \in \Omega$. Let $\omega \in \Omega$ be any but fixed. Now observe that if t is an increasing point of $\langle M\rangle$, that is, there is a $\varepsilon > 0$ such that $s \to \langle M\rangle_s(\omega)$ is strictly increasing on $(t-\varepsilon, t+\varepsilon)$, then by inverse function theorem, $s \to T_s(\omega)$ is the inverse of $\langle M\rangle$ and therefore is continuous at t. Hence $B_s = M_{T_s}$ is continuous at t. On the other hand, if $s_1 < s_2$ such that $\langle M\rangle_{s_1} = \langle M\rangle_{s_2}$ at ω, then $M_s = M_{s_1}$ (at ω) for all $s \in [s_1, s_2]$, and therefore B_s is continuous on $[s_1, s_2]$. Therefore B is continuous.

Exercise 3.2.3 *Let f be a continuous increasing function on $[0,\infty)$ with $f(0) = 0$ and $f(\infty) = \infty$. Define*

$$f^{-1}(x) = \inf\{y : f(y) > x\}.$$

Prove that $f(f^{-1}(x)) = x$ for all $x \geq 0$.

3.3 Burkhölder-Davis-Gundy inequality

Recall Doob's L^p-inequality: if X is a martingale or a non-negative sub-martingale, and $X_t^\star = \sup_{s\leq t}|X_s|$, then for every $p > 1$ and $T > 0$ we have

$$\|X_T^\star\|_p \leq q\,\|X_T\|_p$$

where $\frac{1}{p}+\frac{1}{q}=1$ (so that $q=\frac{p}{p-1}$). In particular if X is a martingale and $p=2$ then $\|X_T\|_2^2=\mathbb{E}\langle X\rangle_T$. Hence

$$\mathbb{E}\langle X\rangle_T \leq \mathbb{E}|X_T^\star|^2 \leq 4\mathbb{E}\langle X\rangle_T .$$

In this section we prove the B-D-G inequality by using Itô's formula.

Theorem 3.3.1 *Let X be a continuous local martingale with $X_0=0$, and $X_t^\star=\sup_{s\leq t}|X_s|$. Let $T>0$. Then*

1) If $p>1$

$$(4p)^{-p}\mathbb{E}\langle X\rangle_T^p \leq \mathbb{E}|X_T^\star|^{2p} \leq \left(2ep^2\right)^p \mathbb{E}\langle X\rangle_T^p . \tag{3.2}$$

2) If $0<p<1$

$$p^{2p}\mathbb{E}\langle X\rangle_T^p \leq \mathbb{E}|X_T^\star|^{2p} \leq \left(\frac{16}{p}\right)^p \mathbb{E}\langle X\rangle_T^p . \tag{3.3}$$

3) In the case $p=1$

$$\mathbb{E}\langle X\rangle_T \leq \mathbb{E}|X_T^\star|^2 \leq 4\mathbb{E}\langle X\rangle_T. \tag{3.4}$$

Proof. The conclusion 3) follows from Kolmogorov's inequality, so we prove the case that $p\neq 1$. By stopping time techniques, we may assume that X is a bounded continuous martingale. First consider the case that $p>1$. In this case, according to Doob's inequality

$$\mathbb{E}|X_T^\star|^{2p} \leq \left(\frac{2p}{2p-1}\right)^{2p}\mathbb{E}|X_T|^{2p}.$$

On the other hand, applying Itô's formula to $f(x)=(x^2+\varepsilon)^p$, where $\varepsilon>0$ is an arbitrary constant, we have

$$\begin{aligned} f'(x) &= 2px(x^2+\varepsilon)^{p-1}, \\ f''(x) &= 2p\left((2p-1)x^2+\varepsilon\right)(x^2+\varepsilon)^{p-2} \end{aligned}$$

and

$$\begin{aligned} (|X_T|^2+\varepsilon)^p &= \varepsilon^p+\int_0^T f'(X_t)dX_t \\ &\quad +p\int_0^T \left((2p-1)X_t^2+\varepsilon\right)(X_t^2+\varepsilon)^{p-2}d\langle X\rangle_t. \end{aligned}$$

Integrating both sides then letting $\varepsilon\downarrow 0$ one obtains

$$\mathbb{E}|X_T|^{2p}=p(2p-1)\mathbb{E}\int_0^T X_t^{2(p-1)}d\langle X\rangle_t$$

and therefore

$$\begin{aligned} \mathbb{E}|X_T|^{2p} &\leq p(2p-1)\mathbb{E}\left(|X_T^\star|^{2(p-1)}\langle X\rangle_T\right) \\ &\leq p(2p-1)\left(\mathbb{E}|X_T^\star|^{2p}\right)^{1-\frac{1}{p}}\left(\mathbb{E}\langle X\rangle_T^p\right)^{1/p}. \end{aligned}$$

Combining with the Doob's inequality we obtain

$$\begin{aligned}\mathbb{E}|X_T^\star|^{2p} &\le \left(1+\frac{1}{2p-1}\right)^{2p-1} 2p^2\left(\mathbb{E}|X_T^\star|^{2p}\right)^{1-\frac{1}{p}}\left(\mathbb{E}\langle X\rangle_T^p\right)^{1/p} \\ &\le 2ep^2\left(\mathbb{E}|X_T^\star|^{2p}\right)^{1-\frac{1}{p}}\left(\mathbb{E}\langle X\rangle_T^p\right)^{1/p}\end{aligned}$$

so that

$$\mathbb{E}|X_T^\star|^{2p} \le \left(2ep^2\right)^p \mathbb{E}\langle X\rangle_T^p .$$

To prove the other direction of the inequality, we begin with

$$\begin{aligned}\mathbb{E}\langle X\rangle_T^p &= p\mathbb{E}\int_0^T \langle X\rangle_t^{p-1} d\langle X\rangle_t \\ &= p\mathbb{E}\left(\int_0^T \langle X\rangle_t^{\frac{p-1}{2}} dX_t\right)^2 ,\end{aligned}$$

and by integration by parts gives

$$\begin{aligned}\int_0^T \langle X\rangle_t^{\frac{p-1}{2}} dX_t &= X_T\langle X\rangle_T^{\frac{p-1}{2}} - \int_0^T X_t d\langle X\rangle_t^{\frac{p-1}{2}} \\ &\le 2X_T^\star\langle X\rangle_T^{\frac{p-1}{2}}\end{aligned}$$

we thus obtain

$$\begin{aligned}\mathbb{E}\langle X\rangle_T^p &\le 4p\mathbb{E}\left[\left(X_T^\star\right)^2\langle X\rangle_T^{p-1}\right] \\ &\le 4p\left(\mathbb{E}\left(X_T^\star\right)^{2p}\right)^{\frac{1}{p}}\left(\mathbb{E}\langle X\rangle_T^p\right)^{1-\frac{1}{p}}\end{aligned}$$

which implies that

$$\mathbb{E}\left(X_T^\star\right)^{2p} \ge (4p)^{-p}\mathbb{E}\langle X\rangle_T^p$$

and therefore (3.2).

Next we assume that $0 < p < 1$. Let $M_t = \int_0^t \langle X\rangle_s^{\frac{p-1}{2}} dX_s$, which is a continuous local martingale with bracket process

$$\langle M\rangle_t = \int_0^t \langle X\rangle_s^{p-1} d\langle X\rangle_s = \frac{1}{p}\langle X\rangle_t^p .$$

In terms of M we may recover X by formula

$$\begin{aligned}X_t &= \int_0^t \langle X\rangle_s^{-\frac{p-1}{2}} dM_s = \int_0^t \langle X\rangle_s^{\frac{1-p}{2}} dM_s \\ &= M_t\langle X\rangle_t^{\frac{1-p}{2}} - \int_0^t M_s d\langle X\rangle_s^{\frac{1-p}{2}} \\ &\le 2M_t^*\langle X\rangle_t^{\frac{1-p}{2}}\end{aligned}$$

where the third equality follows from the integration by parts, so that

$$X_T^* \leq 2M_T^* \langle X \rangle_T^{\frac{1-p}{2}}$$

and therefore

$$\begin{aligned}\mathbb{E}|X_T^*|^{2p} &\leq 2^{2p}\mathbb{E}\left[(M_T^*)^{2p}\langle X \rangle_T^{p(1-p)}\right] \\ &\leq 2^{2p}\left(\mathbb{E}\left(M_T^*\right)^2\right)^p \left(\mathbb{E}\langle X \rangle_T^p\right)^{1-p}\end{aligned}$$

where the second inequality follows from the Hölder inequality as $0 < p < 1$. Since

$$\begin{aligned}\mathbb{E}|M_T^*|^2 &\leq 4\mathbb{E}|M_T|^2 = 4\mathbb{E}\langle M \rangle_T \\ &= \frac{4}{p}\mathbb{E}\langle X \rangle_T^p\end{aligned}$$

thus we have

$$\begin{aligned}\mathbb{E}|X_T^*|^{2p} &\leq 4^p\left(\frac{4}{p}\mathbb{E}\langle X \rangle_T^p\right)^p \left(\mathbb{E}\langle X \rangle_T^p\right)^{1-p} \\ &= \left(\frac{16}{p}\right)^p \mathbb{E}\langle X \rangle_T^p.\end{aligned}$$

To prove the other direction, we consider the continuous martingale

$$Z_t = \int_0^t \frac{1}{(\varepsilon + X_s^*)^{1-p}} dX_s$$

where $\varepsilon > 0$ is a constant, whose bracket process

$$\langle Z \rangle_t = \int_0^t \frac{1}{(\varepsilon + X_s^*)^{2(1-p)}} d\langle X \rangle_s.$$

Since $p < 1$ so that

$$\langle Z \rangle_t \geq \frac{1}{(\varepsilon + X_t^*)^{2(1-p)}} \langle X \rangle_t.$$

On the other hand, by integration by parts we have

$$\begin{aligned}Z_t &= \frac{X_t}{(\varepsilon + X_t^*)^{1-p}} - \int_0^t X_s d(\varepsilon + X_s^*)^{p-1} \\ &= \frac{X_t}{(\varepsilon + X_t^*)^{1-p}} + (1-p)\int_0^t \frac{X_s}{(\varepsilon + X_s^*)^{2-p}} dX_s^*\end{aligned}$$

and therefore

$$
\begin{aligned}
|Z_t| &\leq \frac{X_t^*}{(\varepsilon + X_t^*)^{1-p}} + (1-p)\int_0^t \frac{X_s^*}{(\varepsilon + X_s^*)^{2-p}} dX_s^* \\
&= \frac{X_t^*}{(\varepsilon + X_t^*)^{1-p}} + (1-p)\int_0^t \frac{d(X_s^* + \varepsilon)}{(\varepsilon + X_s^*)^{1-p}} \\
&\quad -(1-p)\varepsilon \int_0^t \frac{d(X_s^* + \varepsilon)}{(\varepsilon + X_s^*)^{2-p}} \\
&= \frac{X_t^*}{(\varepsilon + X_t^*)^{1-p}} + \frac{1-p}{p}((\varepsilon + X_t^*)^p - \varepsilon^p) \\
&\quad -\varepsilon\left(\frac{1}{(\varepsilon + X_t^*)^{1-p}} - \varepsilon^{p-1}\right) \\
&\leq \frac{1}{p}(\varepsilon + X_t^*)^p - 2\varepsilon(\varepsilon + X_t^*)^{p-1} + \left(2 - \frac{1}{p}\right)\varepsilon^p \\
&\leq \frac{1}{p}(\varepsilon + X_t^*)^p + \varepsilon^p
\end{aligned}
$$

for every $\varepsilon > 0$. Hence

$$
\begin{aligned}
\mathbb{E}\frac{\langle X\rangle_T}{(\varepsilon + X_T^*)^{2(1-p)}} &\leq \mathbb{E}\langle Z\rangle_T \\
&= \mathbb{E}Z_T^2 \\
&\leq \mathbb{E}\left[\frac{1}{p}(\varepsilon + X_T^*)^p + \varepsilon^p\right]^2.
\end{aligned}
$$

On the other hand

$$
\begin{aligned}
\mathbb{E}\langle X\rangle_T^p &= \mathbb{E}\left[\left(\frac{\langle X\rangle_T}{(\varepsilon + X_T^*)^{2(1-p)}}\right)^p \left((\varepsilon + X_T^*)^{2p}\right)^{1-p}\right] \\
&= \left[\mathbb{E}\left(\frac{\langle X\rangle_T}{(\varepsilon + X_T^*)^{2(1-p)}}\right)\right]^p \left\{\mathbb{E}(\varepsilon + X_T^*)^{2p}\right\}^{1-p}
\end{aligned}
$$

together with the previous inequality we obtain

$$
\mathbb{E}\langle X\rangle_T^p \leq \left\{\mathbb{E}\left[\frac{1}{p}(\varepsilon + X_T^*)^p + \varepsilon^p\right]^2\right\}^p \left\{\mathbb{E}(\varepsilon + X_T^*)^{2p}\right\}^{1-p}
$$

for every $\varepsilon > 0$. Letting $\varepsilon \downarrow$ one obtains

$$
\mathbb{E}\langle X\rangle_T^p \leq \frac{1}{p^{2p}}\mathbb{E}|X_T^*|^{2p}.
$$

■

3.4 Stochastic exponentials

In this section we consider a simple stochastic differential equation

$$dZ_t = Z_t dX_t \ , \quad Z_0 = 1 \tag{3.5}$$

where $X_t = M_t + A_t$ is a continuous semi-martingale. The solution of (3.5) is called the *stochastic exponential* of X. The equation (3.5) should be understood as an integral equation

$$Z_t = 1 + \int_0^t Z_s dX_s \tag{3.6}$$

where the integral is taken as Itô's integral. To find the solution to (3.6) we may try $Z_t = \exp(X_t + V_t)$, where $(V_t)_{t\geq 0}$ to be determined as a "correction" term (which has finite variation) due to the Itô's integration. Applying Itô's formula we obtain

$$Z_t = 1 + \int_0^t Z_s \mathrm{d}(X_s + V_s) + \frac{1}{2}\int_0^t Z_s \mathrm{d}\langle M\rangle_s$$

and therefore, in order to match the equation (3.6) we must choose $V_t = -\frac{1}{2}\langle M\rangle_t$.

Lemma 3.4.1 *Let $X_t = M_t + A_t$ (where M is a continuous local martingale, A is an adapted continuous process with finite total variation) with $X_0 = 0$. Then*

$$\mathcal{E}(X)_t = \exp\left(X_t - \frac{1}{2}\langle M\rangle_t\right)$$

is the solution to (3.6).

$\mathcal{E}(X)$ is called the *stochastic exponential* of $X = (X_t)_{t\geq 0}$.

Proposition 3.4.2 *Let $(M_t)_{t\geq 0}$ be a continuous local martingale with $M_0 = 0$. Then the stochastic exponential $\mathcal{E}(M)$ is a continuous, non-negative local martingale.*

Remark 3.4.3 *According to definition of Itô's integrals, if $T > 0$ such that*

$$\mathbb{E}\int_0^T e^{2M_t - \langle M\rangle_t} d\langle M\rangle_t < +\infty \tag{3.7}$$

then the stochastic exponential

$$\mathcal{E}(M)_t = \exp\left(M_t - \frac{1}{2}\langle M\rangle_t\right)$$

is a non-negative, continuous martingale. For $n = 1, 2, \cdots$, define

$$T_n = \inf\left\{t \geq 0 : M_t > n \text{ or } \langle M\rangle_t > n\right\}.$$

Then every T_n is a stopping time and $T_n \uparrow \infty$. Moreover

$$\mathbb{E}\int_0^T e^{2M_{t\wedge T_n} - \langle M\rangle_{t\wedge T_n}} d\langle M\rangle_t \leq e^{2n} n < \infty$$

so that $\mathcal{E}(M)_{t\wedge T_n}$ is a martingale, and therefore $\mathcal{E}(M)$ is a non-negative local martingale.

The remarkable fact is that, although $\mathcal{E}(M)$ may fail to be a martingale, but it is nevertheless a super-martingale.

Lemma 3.4.4 *Let $X = (X_t)_{t\geq 0}$ be a **non-negative**, continuous local martingale. Then $X = (X_t)_{t\geq 0}$ is a super-martingale: $\mathbb{E}(X_t|\mathcal{F}_s) \leq X_s$ for any $t < s$. In particular $t \to \mathbb{E}X_t$ is decreasing, and therefore $\mathbb{E}X_t \leq \mathbb{E}X_0$ for any $t > 0$.*

Proof. Recall Fatou's lemma: if $\{f_n\}$ is a sequence of non-negative, integrable functions on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$, such that $\underline{\lim}_{n\to\infty}\mathbb{E}(f_n) < \infty$, then $\underline{\lim}_{n\to\infty} f_n$ is integrable and

$$\mathbb{E}\left(\underline{\lim}_{n\to\infty} f_n|\mathcal{G}\right) \leq \underline{\lim}_{n\to\infty}\mathbb{E}\left(f_n|\mathcal{G}\right)$$

for any sub σ-algebra $\mathcal{G}$ (see page 88, D. Williams: Probability with Martingales).

By definition, there is a sequence of finite stopping times $T_n \uparrow +\infty$ P-a.e. such that $X^{T_n} = (X_{t\wedge T_n})_{t\geq 0}$ is a martingale for each n. Hence

$$\mathbb{E}\left(X_{t\wedge T_n}|\mathcal{F}_s\right) = X_{s\wedge T_n}, \quad \forall t \geq s, n = 1, 2, \cdots .$$

In particular $\mathbb{E}\left(X_{t\wedge T_n}\right) = \mathbb{E}X_0$, so that, by Fatou's lemma, $X_t = \lim_{n\to\infty} X_{t\wedge T_n}$ is integrable. Applying Fatou's lemma to $X_{t\wedge T_n}$ and $\mathcal{G} = \mathcal{F}_s$ for $t > s$ we have

$$\begin{aligned}\mathbb{E}(X_t|\mathcal{F}_s) &= \mathbb{E}\left[\lim_{n\to\infty} X_{t\wedge T_n}|\mathcal{F}_s\right] \leq \underline{\lim}_{n\to\infty}\mathbb{E}(X_{t\wedge T_n}|\mathcal{F}_s) \\ &= \underline{\lim}_{n\to\infty} X_{s\wedge T_n} = X_s\end{aligned}$$

According to definition, $X = (X_t)_{t\geq 0}$ is a super-martingale. ∎

Corollary 3.4.5 *Let $M = (M_t)_{t\geq 0}$ be a continuous, local martingale with $M_0 = 0$. Then $\mathcal{E}(M)$ is a super-martingale. In particular,*

$$\mathbb{E}\exp\left(M_t - \frac{1}{2}\langle M\rangle_t\right) \leq 1 \quad \textit{for all} \quad t \geq 0 .$$

Clearly, a continuous super-martingale $X = (X_t)_{t\geq 0}$ is a martingale if and only if its expectation $t \to \mathbb{E}(X_t)$ is constant. Therefore

Corollary 3.4.6 *Let $M = (M_t)_{t\geq 0}$ be a continuous, local martingale with $M_0 = 0$. Then $\mathcal{E}(M)$ is a martingale up to time T, if and only if*

$$\mathbb{E}\exp\left(M_T - \frac{1}{2}\langle M\rangle_T\right) = 1 . \tag{3.8}$$

Stochastic exponentials of local martingales play an important rôle in probability transformations. It is vital in many applications to know whether the stochastic exponential of a given martingale $M = (M_t)_{t\geq 0}$ is indeed a martingale. A simple sufficient condition to ensure (3.8) is the so-called Novikov's condition stated in Theorem 3.4.7 below (A. A. Novikov: On moment inequalities and identities for stochastic integrals, *Proc. second Japan-USSR Symp. Prob. Theor., Lecture Notes in Math.*, **330**, 333-339, Springer-Verlag, Berlin 1973).

Theorem 3.4.7 *(A. A. Novikov 1973) Let $M=(M_t)_{t\geq 0}$ be a continuous local martingale with $M_0=0$. If*

$$\mathbb{E}\exp\left(\frac{1}{2}\langle M\rangle_T\right)<\infty \ , \tag{3.9}$$

then $\mathcal{E}(M)$ is a martingale up to time T.

To prove this, we need the following fact about uniform integrability.

Lemma 3.4.8 *Suppose $\mathcal{A}$ be a family of integrable random variables on $(\Omega,\mathcal{F},\mathbb{P})$ and suppose there is an integrable random variable η such that $\mathbb{E}\left[1_D|\xi|\right]\leq \mathbb{E}\left[1_D|\eta|\right]$ for any $D\in\mathcal{F}$ and $\xi\in\mathcal{A}$. Then $\mathcal{A}$ is uniformly integrable.*

Proof. *of Theorem 3.4.7.* (The following proof is due to J. A. Yan: *Critères d'intégrabilité uniforme des martingales exponentielles, Acta. Math. Sinica* **23**, 311-318 (1980).) The idea is the following, first show that, under the Novikov condition (3.9), for any $0<\alpha<1$

$$\mathcal{E}(\alpha M)_t\equiv\exp\left(\alpha M_t-\frac{1}{2}\alpha^2\langle M\rangle_t\right)$$

is a uniformly integrable martingale up to time T. By Corollary (3.4.5), $\mathbb{E}\left(\mathcal{E}(\alpha M)_t\right)\leq 1$ for every α. While

$$\begin{aligned}\mathcal{E}(\alpha M)_t &\equiv \exp\left\{\alpha\left(M_t-\frac{1}{2}\langle M\rangle_t\right)-\frac{1}{2}\alpha\left(\alpha-1\right)\langle M\rangle_t\right\}\\ &= \left(\mathcal{E}(M)_t\right)^{\alpha}\exp\left\{\frac{1}{2}\alpha\left(1-\alpha\right)\langle M\rangle_t\right\},\end{aligned}$$

so that for every $t\leq T$ and for every $A\in\mathcal{F}_T$

$$\mathbb{E}\left(1_A\mathcal{E}(\alpha M)_t\right)=\mathbb{E}\left\{1_A\left(\mathcal{E}(M)_t\right)^{\alpha}\exp\left[\frac{1}{2}\alpha\left(1-\alpha\right)\langle M\rangle_t\right]\right\} \ . \tag{3.10}$$

If $\alpha\in(0,1)$, by using Hölder's inequality with $p=\frac{1}{\alpha}>1$ and $q=\frac{1}{1-\alpha}$ in (3.10) one obtains

$$\begin{aligned}\mathbb{E}\left\{1_A\mathcal{E}(\alpha M)_t\right\} &= \mathbb{E}\left\{\left(\mathcal{E}(M)_t\right)^{\alpha}\exp\left[\frac{1}{2}\alpha\left(1-\alpha\right)\langle M\rangle_t\right]\right\}\\ &\leq \left\{\mathbb{E}\left(\mathcal{E}(M)_t\right)\right\}^{\alpha}\left\{\mathbb{E}\left[1_A\exp\left(\frac{1}{2}\alpha\langle M\rangle_t\right)\right]\right\}^{1-\alpha}\\ &\leq \left\{\mathbb{E}\left(\mathcal{E}(M)_t\right)\right\}^{\alpha}\left\{\mathbb{E}\left[1_A\exp\left(\frac{1}{2}\alpha\langle M\rangle_t\right)\right]\right\}^{1-\alpha}\\ &\leq \left\{\mathbb{E}\left[1_A\exp\left(\frac{1}{2}\alpha\langle M\rangle_T\right)\right]\right\}^{1-\alpha}\\ &\leq \mathbb{E}\left\{1_A\exp\left(\frac{1}{2}\langle M\rangle_T\right)\right\} \ .\end{aligned} \tag{3.11}$$

Therefore by Lemma (3.4.8) we deduce that $(\mathcal{E}(\alpha M)_t)_{t\le T}$ is, up to time T, uniformly integrable local martingale, and therefore $\mathcal{E}(\alpha M)$ must be a martingale on $[0,T]$. In particular

$$\mathbb{E}\left(\mathcal{E}(\alpha M)_T\right) = \mathbb{E}\left(\mathcal{E}(\alpha M)_0\right) = 1, \quad \forall \alpha \in (0,1).$$

Set $A = \Omega$ in (3.11), the first inequality of (3.11) becomes

$$\begin{aligned} 1 &= \mathbb{E}\left(\mathcal{E}(\alpha M)_t\right) \\ &\le \left(\mathbb{E}\left(\mathcal{E}(M)_t\right)\right)^{\alpha} \left\{\mathbb{E}\left(\exp\left(\frac{1}{2}\langle M\rangle_T\right)\right)\right\}^{1-\alpha} \end{aligned}$$

for every $\alpha \in (0,1)$. Letting $\alpha \uparrow 1$ we thus obtain that $\mathbb{E}\left(\mathcal{E}(M)_t\right) \ge 1$ for every $t \le T$, so that $\mathbb{E}\left(\mathcal{E}(M)_t\right) = 1$ for any $t \le T$. Hence $\mathcal{E}(M)_t$ is a martingale up to T. ■

Consider a standard Brownian motion $B = (B_t)$, and $F = (F_t)_{t\ge 0} \in \mathcal{L}_2$. If $\mathbb{E}\exp\left[\frac{1}{2}\int_0^T F_t^2 \mathrm{d}t\right] < \infty$ then

$$X_t = \exp\left\{\int_0^t F_s \mathrm{d}B_s - \frac{1}{2}\int_0^t F_s^2 \mathrm{d}s\right\} \tag{3.12}$$

is a positive martingale on $[0,T]$. For example, for any bounded process $F = (F_t)_{t\ge 0} \in \mathcal{L}_2$: $|F_t(\omega)| \le C$ (for all $t \le T$ and $\omega \in \Omega$), where C is a constant, then

$$\mathbb{E}\left\{\exp\left(\frac{1}{2}\int_0^T F_t^2 \mathrm{d}t\right)\right\} \le \exp\left(\frac{1}{2}C^2 T\right) < \infty$$

so that, in this case, $X = (X_t)$ defined by (3.12) is a martingale up to time T.

Novikov's condition is very nice, it is however not easy to verify in many interesting cases. For example, consider the stochastic exponential of the martingale $\int_0^t B_s \mathrm{d}B_s$, the Novikov condition requires to estimate the integral $\mathbb{E}\left\{\exp\left[\frac{1}{2}\int_0^T B_t^2 \mathrm{d}t\right]\right\}$, which is already not an easy task.

3.5 Exponential inequality

We are going to present three significant applications of stochastic exponentials: a sharp improvement of Doob's maximal inequality for martingales, Girsanov's theorem, and the martingale representation theorem (in the next section). Additional applications will be discussed in the next chapter.

Recall that, according to Doob's maximal inequality, if $(X_t)_{t\ge 0}$ is a continuous super-martingale on $[0,T]$, then for any $\lambda > 0$

$$\mathbb{P}\left\{\sup_{t\in[0,T]} |X_t| \ge \lambda\right\} \le \frac{1}{\lambda}\left(\mathbb{E}(X_0) + 2\mathbb{E}(X_T^-)\right)$$

where $x^- = -x$ if $x < 0$ and $= 0$ if $x \ge 0$. In particular, if $(X_t)_{t\ge 0}$ is a *non-negative*, continuous super-martingale on $[0,T]$, then

$$\mathbb{P}\left\{\sup_{t\in[0,T]} X_t \ge \lambda\right\} \le \frac{1}{\lambda}\mathbb{E}(X_0) . \tag{3.13}$$

This inequality has a significant improvement stated as follows.

Theorem 3.5.1 *Let $M = (M_t)_{t\geq 0}$ be a continuous square integrable martingale with $M_0 = 0$. Suppose there is a (deterministic) continuous, increasing function $a = a(t)$ such that $a(0) = 0$, $\langle M \rangle_t \leq a(t)$ for all $t \in [0, T]$. Then*

$$\mathbb{P}\left\{ \sup_{t\in[0,T]} M_t \geq \lambda a(T) \right\} \leq e^{-\frac{\lambda^2}{2}a(T)} \ . \tag{3.14}$$

Proof. For every $\alpha > 0$ and $t \leq T$

$$\begin{aligned} \alpha M_t - \frac{\alpha^2}{2}\langle M \rangle_t &\geq \alpha M_t - \frac{\alpha^2}{2}\langle M \rangle_T \\ &\geq \alpha M_t - \frac{\alpha^2}{2}a(T) \end{aligned}$$

so that

$$\mathcal{E}(\alpha M)_t \geq e^{\alpha M_t - \frac{\alpha^2}{2}a(T)} \qquad \text{for } \alpha > 0 \ .$$

Hence, by applying Doob's maximal inequality to the non-negative super-martingale $\mathcal{E}(\alpha M)$ we obtain

$$\begin{aligned} \mathbb{P}\left\{ \sup_{t\in[0,T]} M_t \geq \lambda a(T) \right\} &\leq \mathbb{P}\left\{ \sup_{t\in[0,T]} \mathcal{E}(\alpha M)_t \geq e^{\alpha\lambda a(T) - \frac{\alpha^2}{2}a(T)} \right\} \\ &\leq e^{-\alpha\lambda a(T) + \frac{\alpha^2}{2}a(T)} \mathbb{E}\left\{ \mathcal{E}(\alpha M)_0 \right\} \\ &= e^{-\alpha\lambda a(T) + \frac{\alpha^2}{2}a(T)} \end{aligned}$$

for any $\alpha > 0$. The exponential inequality follows by setting $\alpha = \lambda$. ■

In particular, by applying the exponential inequality to a standard Brownian motion $B = (B_t)_{t\geq 0}$,

$$\mathbb{P}\left\{ \sup_{t\in[0,T]} B_t \geq \lambda T \right\} \leq e^{-\frac{\lambda^2}{2}T} \ . \tag{3.15}$$

3.6 Girsanov's theorem

This is an important tool in stochastic analysis. To state it let us introduce several notions and notations.

Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space. If $\mathcal{G} \subset \mathcal{F}$ is a sub σ-algebra, and ξ is an non-negative, integrable and $\mathcal{G}$-measurable with $\mathbb{E}[\xi] = 1$, then define a probability measure $\mathbb{Q}$ on $(\Omega, \mathcal{F})$ by

$$\mathbb{Q}(A) = \mathbb{E}[\xi 1_A] = \int_A \xi d\mathbb{P}$$

for every A. Of course $\mathbb{Q}$ restricted on $(\Omega, \mathcal{G})$ is also a probability measure. In this case we use the notation that

$$\left.\frac{d\mathbb{Q}}{d\mathbb{P}}\right|_{\mathcal{G}} = \xi$$

to represent the Radon-Nikodym derivative of $\mathbb{Q}$ with respect to $\mathbb{P}$ as measures on $(\Omega, \mathcal{G})$.

Let us work with a filtered probability space $(\Omega, \mathcal{F}, \mathcal{F}_t, \mathbb{P})$. Let $T > 0$, and $\mathbb{Q}$ be a probability measure on $(\Omega, \mathcal{F}_T)$ such that

$$\left.\frac{d\mathbb{Q}}{d\mathbb{P}}\right|_{\mathcal{F}_T} = \xi$$

for some non-negative random variable $\xi \in L^1(\Omega, \mathcal{F}_T, \mathbb{P})$. By definition, for any bounded $\mathcal{F}_T$-measurable random variable X

$$\int_\Omega X(\omega)\mathbb{Q}(d\omega) = \int_\Omega X(\omega)\xi(\omega)\mathbb{P}(d\omega)$$

or simply written as $\mathbb{E}^{\mathbb{Q}}(X) = \mathbb{E}^{\mathbb{P}}(\xi X)$. If, however, X is $\mathcal{F}_t$-measurable, $t \leq T$, then

$$\mathbb{E}^{\mathbb{Q}}(X) = \mathbb{E}^{\mathbb{P}}(\mathbb{E}^{\mathbb{P}}(\xi X|\mathcal{F}_t)) = \mathbb{E}^{\mathbb{P}}(\mathbb{E}^{\mathbb{P}}(\xi|\mathcal{F}_t)X) .$$

That is, for every $t \leq T$

$$\left.\frac{d\mathbb{Q}}{d\mathbb{P}}\right|_{\mathcal{F}_t} = \mathbb{E}^{\mathbb{P}}(\xi|\mathcal{F}_t)$$

which is a non-negative martingale up to T under the probability $\mathbb{P}$.

Conversely, if $T > 0$ and $Z = (Z_t)_{t\geq 0}$ is a continuous, positive martingale up to time T, with $Z_0 = 1$, on a filtered probability space $(\Omega, \mathcal{F}, \mathcal{F}_t, \mathbb{P})$. We define a measure $\mathbb{Q}$ on $(\Omega, \mathcal{F}_T)$ by

$$\mathbb{Q}(A) = \mathbb{P}(Z_T A) \quad \text{if } A \in \mathcal{F}_T \quad . \tag{3.16}$$

That is, $\frac{d\mathbb{Q}}{d\mathbb{P}}\big|_{\mathcal{F}_T} = Z_T$. $\mathbb{Q}$ is a probability measure on $(\Omega, \mathcal{F}_T)$ as $\mathbb{E}(Z_T) = 1$. Since $(Z_t)_{t\leq T}$ is a martingale up to time T, so that $\frac{d\mathbb{Q}}{d\mathbb{P}}\big|_{\mathcal{F}_t} = Z_t$ for all $t \leq T$.

If $(Z_t)_{t\geq 0}$ is a positive martingale with $Z_0 = 1$, then there is a probability measure $\mathbb{Q}$ on $(\Omega, \mathcal{F}_\infty)$, where $\mathcal{F}_\infty \equiv \sigma\{\mathcal{F}_t : t \geq 0\}$, such that $\frac{d\mathbb{Q}}{d\mathbb{P}}\big|_{\mathcal{F}_t} = Z_t$ for all $t \geq 0$.

We are now in a position to prove Girsanov's theorem.

Theorem 3.6.1 *(Girsanov's theorem) Let $(M_t)_{t\geq 0}$ be a continuous local martingale on $(\Omega, \mathcal{F}, \mathcal{F}_t, \mathbb{P})$ up to time T. Then*

$$X_t = M_t - \int_0^t \frac{1}{Z_s} d\langle M, Z\rangle_s$$

is a continuous local martingale on $(\Omega, \mathcal{F}, \mathcal{F}_t, \mathbb{Q})$ up to time T.

Proof. Using localization technique, we may assume that $M, Z, 1/Z$ are all bounded. In this case M, Z are bounded martingales. We want to prove that X is a martingale under the probability $\mathbb{Q}$:

$$\mathbb{E}^{\mathbb{Q}}\{X_t|F_s\} = X_s \quad \text{for all} \quad s < t \leq T ,$$

that is,

$$\mathbb{E}^{\mathbb{Q}}\{1_A(X_t - X_s)\} = 0 \quad \text{for all} \quad s < t \leq T ,\ A \in \mathcal{F}_s .$$

By definition

$$\mathbb{E}^{\mathbb{Q}}\left\{1_A\left(X_t - X_s\right)\right\} = \mathbb{E}^{\mathbb{P}}\left\{(Z_t X_t - Z_s X_s) 1_A\right\}$$

thus we only need to show that $(Z_t X_t)$ is a martingale up to time T under probability measure $\mathbb{P}$. By use of integration by parts, we have

$$\begin{aligned} Z_t X_t &= Z_0 X_0 + \int_0^t Z_s \mathrm{d}X_s + \int_0^t X_s \mathrm{d}Z_s + \langle Z, X\rangle_t \\ &= Z_0 X_0 + \int_0^t Z_s \left(\mathrm{d}M_s - \frac{1}{Z_s}\mathrm{d}\left\langle M, Z\right\rangle_s\right) + \int_0^t X_s \mathrm{d}Z_s + \langle Z, X\rangle_t \\ &= Z_0 X_0 + \int_0^t Z_s \mathrm{d}M_s + \int_0^t X_s \mathrm{d}Z_s \end{aligned}$$

which is a local martingale. ■

Since $Z_t > 0$ is a positive martingale up to time T, we may apply the Itô formula to $\log Z_t$, to obtain

$$\log Z_t - \log Z_0 = \int_0^t \frac{1}{Z_s}\mathrm{d}Z_s - \int_0^t \frac{1}{Z_s^2}\mathrm{d}\langle Z\rangle_s \ ,$$

that is, $Z_t = \mathcal{E}(N)_t$ with $N_t = \int_0^t \frac{1}{Z_s}\mathrm{d}Z_s$ is a continuous local martingale. Hence $Z_t = \mathcal{E}(N)_t$ solves the Itô integral equation

$$Z_t = 1 + \int_0^t Z_s \mathrm{d}N_s \ ,$$

and therefore

$$\langle M, Z\rangle_t = \langle \int_0^t \mathrm{d}M_s, \int_0^t Z_s \mathrm{d}N_s\rangle = \int_0^t Z_s \mathrm{d}\langle N, M\rangle_s \ .$$

It follows thus that

$$\int_0^t \frac{1}{Z_s}\mathrm{d}\left\langle M, Z\right\rangle_s = \langle N, M\rangle_t \ .$$

Corollary 3.6.2 *Let N_t be a continuous local martingale on $(\Omega, \mathcal{F}, \mathcal{F}_t, \mathbb{P})$, $N_0 = 0$, such that its stochastic exponential $\mathcal{E}(N)_t$ is a continuous martingale up to T. Define a probability measure $\mathbb{Q}$ on the measurable space $(\Omega, \mathcal{F}_T)$ by*

$$\left.\frac{d\mathbb{Q}}{d\mathbb{P}}\right|_{\mathcal{F}_t} = \mathcal{E}(N)_t \quad \text{for all} \quad t \le T.$$

If $M = (M_t)_{t\ge 0}$ is a continuous local martingale under the probability P, then $X_t = M_t - \langle N, M\rangle_t$ is a continuous, local martingale under $\mathbb{Q}$ up to time T. (You should carefully define the concept of a local martingale up to time T).

3.7 The martingale representation theorem

The martingale representation theorem is a deep result about Brownian motion. There is a natural version for multi-dimensional Brownian motion, for simplicity of notations, we however concentrate on one-dimensional Brownian motion.

Let $B = (B_t)_{t\geq 0}$ be a standard Brownian motion in $\mathbb{R}$ on a complete probability space $(\Omega, \mathcal{F}, \mathbb{P})$, and $(\mathcal{F}_t^0)_{t\geq 0}$ (together with $\mathcal{F}_\infty^0 = \cup \mathcal{F}_t^0$) be the filtration generated by the Brownian motion $(B_t)_{t\geq 0}$. Let $\mathcal{F}_t$ be the completion of $\mathcal{F}_t^0$, and $\mathcal{F}_\infty = \sigma(\cup \mathcal{F}_t)$. As a matter of fact, $(\mathcal{F}_t)_{t\geq 0}$ is continuous.

Theorem 3.7.1 *Let $M = (M_t)_{t\geq 0}$ be a square-integrable martingale on $(\Omega, \mathcal{F}, \mathcal{F}_t, \mathbb{P})$. Then there is a stochastic process $F = (F_t)_{t\geq 0}$ in $\mathcal{L}_2$, such that*

$$M_t = \mathbb{E}M_0 + \int_0^t F_s dB_s \qquad a.s.$$

for any $t \geq 0$. In particular, any martingale with respect to the Brownian filtration $(\mathcal{F}_t)_{t\geq 0}$ has a continuous version.

The proof of this theorem relies on the following several lemmas. There are several facts from analysis we are going to use, which we state them here.

1. For $p \in [1, \infty]$, $L^p(\mathbb{R}^d, \mu)$ denotes the L^p-space on the measure space $(\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d), \mu)$ where μ is a measure which is absolutely continuous with respect to the Lebesgue measure. Then $C_0^\infty(\mathbb{R}^d)$ is dense in $L^p(\mathbb{R}^d, \mu)$, where $C_0^\infty(\mathbb{R}^d)$ denotes the space of all smooth functions on $\mathbb{R}^d$ with compact supports.

2. If $\phi \in C_0^\infty(\mathbb{R}^d)$ then its Fourier transform $\hat{\phi}$ is defined to be

$$\hat{\phi}(x) = \frac{1}{(2\pi)^{n/2}} \int_{\mathbb{R}^n} \phi(x) e^{-i\langle z,x\rangle} \mathrm{d}x$$

which is smooth too (while its support may be not compact) and f is recovered by the inverse Fourier transform:

$$\phi(x) = \frac{1}{(2\pi)^{n/2}} \int_{\mathbb{R}^n} \hat{\phi}(z) e^{i\langle z,x\rangle} \mathrm{d}z.$$

3. This is a theorem from Functional Analysis. Suppose H is a Hilbert space (i.e. a complete metric space whose metric is induced by the inner product $\langle \cdot, \cdot \rangle$). Let $X \subset H$ be a subset of H and span$X$ be the linear subspace generated by X, that is the linear span of X, i.e. the linear space of all finite linear combinations of elements in X. Then span$X$ is dense in H if and only if the only element $h \in H$ satisfying the condition $\langle x, h \rangle = 0$ for all $x \in X$ is the zero element.

Let $T > 0$ be any fixed time.

Lemma 3.7.2 *The following collection of random variables on $(\Omega, \mathcal{F}_T, \mathbb{P})$*

$$\left\{\phi(B_{t_1}, \cdots, B_{t_k}) : \forall k \in \mathbb{Z}_+, \;\; t_j \in [0,T] \text{ and } \phi \in C_0^\infty(\mathbb{R}^k)\right\}$$

is dense in $L^2(\Omega, \mathcal{F}_T, \mathbb{P})$.

Proof. If $X \in L^2(\Omega, \mathcal{F}_T, \mathbb{P})$, then, by definition, there is an $\mathcal{F}_T^0$-measurable function which equals X almost surely. Therefore, without losing generality, we may assume that $X \in L^2(\Omega, \mathcal{F}_T^0, \mathbb{P})$. According to definition, $\mathcal{F}_T^0 = \sigma\{B_t : t \leq T\}$. Let $D = \mathbb{Q} \cap [0,T]$ the set of all rational numbers in the interval $[0,T]$. Since D is dense in $[0,T]$, so that $\mathcal{F}_T^0 = \sigma\{B_t : t \in D\}$. Moreover D countable, so that we may write $D = \{t_1, \cdots, t_n, \cdots\}$. Let $D_n = \{t_1, \cdots, t_n\}$ for each n, and $\mathcal{G}_n = \sigma\{B_{t_1}, \cdots, B_{t_n}\}$. Then $\{\mathcal{G}_n\}$ is increasing, and $\mathcal{G}_n \uparrow \mathcal{F}_T^0$. Let $X_n = \mathbb{E}(X|\mathcal{G}_n)$. Then $(X_n)_{n\geq 1}$ is square integrable martingale, thus, according to the martingale convergence theorem $X_n \to X$ almost surely. Moreover $X_n \to X$ in L^2. While, for each n, X_n is measurable with respect to $\mathcal{G}_n$, so that $X_n = f_n(B_{t_1}, \cdots, B_{t_n})$ for some Borel measurable function $f_n : \mathbb{R}^n \to \mathbb{R}$. Since $X_n \in L^2$, so that $f_n \in L^2(\mathbb{R}^n, \mu)$ where μ is a Gaussian measure such that $\mathbb{E}X_n^2 = \int_{\mathbb{R}^n} f_n(x)^2 \mu(dx)$. Since $C_0^\infty(\mathbb{R}^n)$ is dense in $L^2(\mathbb{R}^n, \mu)$, for each n, there is a sequence $\{\phi_{nk}\}$ in $C_0^\infty(\mathbb{R}^n)$ such that $\phi_{nk} \to f_n$ in $L^2(\mathbb{R}^n, \mu)$. It follows that $\phi_{nn}(B_{t_1}, \cdots, B_{t_n}) \to X$ in L^2. ■

If $I \subset \mathbb{R}$ is an interval, then we use $L^2(I)$ to denote the Hilbert space of all functions h on I which are square integrable.

Lemma 3.7.3 *Let $T > 0$. For any $h \in L^2([0,T])$, we associate with an exponential martingale up to time T:*

$$M(h)_t = \exp\left\{\int_0^t h(s)dB_s - \frac{1}{2}\int_0^t h(s)^2 ds\right\} \ ; \quad t \in [0,T]. \tag{3.17}$$

Then $\mathbb{L}$ =span$\{M(h)_T : h \in L^2([0,T])\}$ is dense in $L^2(\Omega, \mathcal{F}_T, \mathbb{P})$.

Proof. For any $0 = t_0 < t_1 < \cdots < t_n = T$ and $c_i \in \mathbb{R}$, consider a step function $h(t) = c_i$ for $t \in (t_i, t_{i+1}]$. Then

$$M(h)_T = \exp\left\{\sum_i c_i(B_{t_{i+1}} - B_{t_i}) - \frac{1}{2}\sum_i c_i^2(t_{i+1} - t_i)\right\} .$$

Suppose that $H \in L^2$ such that $\int_\Omega H\Phi \mathrm{d}\mathbb{P} = 0$ for any $\Phi \in \mathbb{L}$, then

$$\int_\Omega H \exp\left\{\sum_i c_i(B_{t_{i+1}} - B_{t_i}) - \frac{1}{2}\sum_i c_i^2(t_{i+1} - t_i)\right\} \mathrm{d}\mathbb{P} = 0 .$$

The deterministic, positive term $e^{-\frac{1}{2}\sum_i c_i^2(t_{i+1}-t_i)}$ can be removed from the integrand, and it follows therefore that

$$\int_\Omega H \exp\left\{\sum_i c_i(B_{t_{i+1}} - B_{t_i})\right\} \mathrm{d}\mathbb{P} = 0 .$$

Since c_i are arbitrary numbers, hence

$$\int_\Omega H \exp\left\{\sum_i c_i B_{t_i}\right\} \mathrm{d}\mathbb{P} = 0$$

for any c_i and $t_i \in [0,T]$. Since the left-hand is analytic in c_i, so that the equality remains true for any complex numbers c_i. If $\phi \in C_0^\infty(\mathbb{R}^n)$, then

$$\phi(x) = \frac{1}{(2\pi)^{n/2}} \int_{\mathbb{R}^n} \hat{\phi}(z) e^{i\langle z,x\rangle} \mathrm{d}z$$

where
$$\hat{\phi}(z) = \frac{1}{(2\pi)^{n/2}} \int_{\mathbb{R}^n} \phi(x) e^{-i\langle z,x\rangle} \mathrm{d}x$$
is the Fourier transform of ϕ. Hence

$$\begin{aligned}
\int_\Omega H\phi(B_{t_1}, \cdots, B_{t_n}) \mathrm{d}\mathbb{P} &= \frac{1}{(2\pi)^{n/2}} \int_\Omega \left\{ H \int_{\mathbb{R}^n} \hat{\phi}(z) \exp\left(i \sum_j z_j B_{t_j} \right) \right\} \mathrm{d}z \mathrm{d}\mathbb{P} \\
&= \frac{1}{(2\pi)^{n/2}} \int_{\mathbb{R}^n} \left\{ \hat{\phi}(z) \int_\Omega H \exp\left(i \sum_i z_i B_{t_i} \right) \mathrm{d}\mathbb{P} \right\} \mathrm{d}z \\
&= 0 .
\end{aligned}$$

Therefore, for any $\phi \in C_0^\infty(\mathbb{R}^n)$,
$$\int_\Omega H\phi(B_{t_1}, \cdots, B_{t_n}) \mathrm{d}\mathbb{P} = 0 . \tag{3.18}$$

By Lemma 3.7.2, the collection of all functions like $\phi(B_{t_1}, \cdots, B_{t_n})$ is dense in $L^2(\Omega, \mathcal{F}_T, \mathbb{P})$, so that
$$\int_\Omega HG\mathrm{d}\mathbb{P} = 0 \quad \text{for any} \quad G \in L^2(\Omega, \mathcal{F}_T, \mathbb{P}) .$$
In particular, $\int_\Omega H^2 \mathrm{d}\mathbb{P} = 0$ so that $H = 0$. ■

Theorem 3.7.4 *(Itô's representation theorem) Let $\xi \in L^2(\Omega, \mathcal{F}_T, \mathbb{P})$. Then there is a $F = (F_t)_{t\geq 0} \in \mathcal{L}_2$, such that*
$$\xi = \mathbb{E}\xi + \int_0^T F_t dB_t .$$

Proof. By Lemma 3.7.3 we only need to show this lemma for $\xi = X(h)_T$ (where $h \in L^2([0,T])$) defined by (3.17). While, $X(h)_t$ is an exponential martingale so that it must satisfy the following integral equation

$$\begin{aligned}
X(h)_T &= 1 + \int_0^T X(h)_t \mathrm{d}\left(\int_0^t h(s) \mathrm{d}B_s \right) \\
&= \mathbb{E}(X(h)_T) + \int_0^T X(h)_t h(t) \mathrm{d}B_t .
\end{aligned}$$

Therefore $F_t = X(h)_t h(t)$ will do. ■

Chapter 4

Stochastic differential equations

The main goal of the chapter is to establish the basic existence and uniqueness theorem for a class of stochastic differential equations which are important in applications.

4.1 Introduction

Stochastic differential equations (SDE) are ordinary differential equations perturbed by noises. We will consider a simple class of noises modeled by Brownian motion. Thus we consider the following type of equation

$$dX_t^j = \sum_{i=1}^n f_i^j(t, X_t)dB_t^i + f_0^j(t, X_t)dt \ , \quad j = 1, \cdots, N \tag{4.1}$$

where $B_t = (B_t^1, \cdots, B_t^n)_{t\geq 0}$ is a standard Brownian motion in $\mathbb{R}^n$ on a filtered probability space $(\Omega, \mathcal{F}, \mathcal{F}_t, \mathbb{P})$, and

$$f_i^j : [0, \infty) \times \mathbb{R}^N \to \mathbb{R}^N$$

are Borel measurable functions. Of course, differential equation (4.1) should be interpreted as an integral equation in terms of Itô's integration. More precisely, an adapted, continuous, $\mathbb{R}^N$-valued stochastic process $X_t \equiv (X_t^1, \cdots, X_t^N)$ is a solution of (4.1), if

$$X_t^j = X_0^j + \sum_{k=1}^n \int_0^t f_k^j(s, X_s)dB_s^k + \int_0^t f_0^j(s, X_s)ds \tag{4.2}$$

for $j = 1, \cdots, N$. Since we are concerned only with the distribution determined by the solution $(X_t)_{t\geq 0}$ of (4.1), we therefore expect that any solution of SDE (4.1) should have the same distribution for *any* Brownian motion $B = (B_t)_{t\geq 0}$. It thus leads to different concepts of solutions and uniqueness: strong solutions and weak solutions, path-wise uniqueness and uniqueness in law.

Definition 4.1.1 *1) An adapted, continuous, $\mathbb{R}^N$-valued stochastic process $X = (X_t)_{t\geq 0}$ on $(\Omega, \mathcal{F}, \mathcal{F}_t, \mathbb{P})$ is a (weak) solution of (4.1), if there is a Brownian motion $W = (W_t)_{t\geq 0}$ in $\mathbb{R}^n$, adapted to the filtration $(\mathcal{F}_t)$, such that*

$$X_t^j - X_0^j = \sum_{l=1}^n \int_0^t f_l^j(s, X_s)dW_s^l + \int_0^t f_0^j(s, X_s)ds, \quad j = 1, \cdots, N.$$

In this case we also call the pair (X, W) a (weak) solution of (4.1).

2) Given a standard Brownian motion $B = (B_t)_{t\geq 0}$ in $\mathbb{R}^n$ on $(\Omega, \mathcal{F}, \mathbb{P})$ with its natural filtration $(\mathcal{F}_t)_{t\geq 0}$, an adapted, continuous stochastic process $X = (X_t)_{t\geq 0}$ on $(\Omega, \mathcal{F}, \mathcal{F}_t, \mathbb{P})$ is a strong solution *of (4.1), if*

$$X_t^j - X_0^j = \sum_{i=1}^{n} \int_0^t f_i^j(s, X_s) dB_s^i + \int_0^t f_0^j(s, X_s) ds \ .$$

We also have different concepts of uniqueness.

Definition 4.1.2 *Consider SDE (4.1).*

1. *We say that the **path-wise uniqueness** holds for (4.1), if whenever (X, B) and $(\widetilde{X}, B)$ are two solutions defined on the same filtered space and the same Brownian motion B, and $X_0 = \widetilde{X}_0$, then $X = \widetilde{X}$.*

2. *It is said that **uniqueness in law** holds for (4.1), if (X, B) and $(\widetilde{X}, \widetilde{B})$ are two solutions (with possibly different Brownian motions B and $\widetilde{B}$, even can be on different probability spaces), and X_0 and $\widetilde{X}_0$ possess the same distribution, then X and $\widetilde{X}$ have same distribution.*

Theorem 4.1.3 *(Yamada-Watanabe) Path-wise uniqueness implies uniqueness in law.*

The following is a simple example of SDE for which has no strong solution, but possesses weak solutions and uniqueness in law holds.

Example 4.1.4 *(H. Tanaka) Consider 1-dimensional stochastic differential equation:*

$$X_t = \int_0^t sgn(X_s) dB_s \ , \qquad 0 \leq t < \infty$$

where $sgn(x) = 1$ if $x \geq 0$, and equals -1 for negative value of x.

1. Uniqueness in law holds, since X is a standard Brownian motion (Lévy's Theorem).

2. There is a weak solution. Let W_t be a one-dimensional Brownian motion, and $B_t = \int_0^t \mathrm{sgn}(W_s) dW_s$. Then B is a one-dimensional Brownian motion, and

$$W_t = \int_0^t \mathrm{sgn}(W_s) dB_s \ ,$$

so that (W, B) is a solution.

3. Path-wise uniqueness does not hold.

4. There is no any strong solution.

4.2 Several examples

4.2.1 Linear-Gaussian diffusions

Linear stochastic differential equations can be solved explicitly. Consider

$$dX_t^j = \sum_{i=1}^n \sigma_i^j dB_t^i + \sum_{k=1}^N \beta_k^j X_t^k dt \tag{4.3}$$

($j = 1, \cdots, N$), where B is a Brownian motion in $\mathbb{R}^n$, $\sigma = (\sigma_i^j)$ a constant $N \times n$ matrix, and $\beta = (\beta_k^j)$ a constant $N \times N$ matrix. (4.3) may be written as

$$dX_t = \sigma dB_t + \beta X_t dt \ .$$

Let

$$e^{\beta t} = \sum_{k=0}^\infty \frac{t^k}{k!}\beta^k$$

be the exponential of the square matrix β. Using Itô's formula, we have

$$\begin{aligned} e^{-\beta t}X_t - X_0 &= \int_0^t e^{-\beta s} dX_s - \int_0^t e^{-\beta s}\beta X_s ds \\ &= \int_0^t e^{-\beta s}(dX_s - \beta X_s ds) \\ &= \int_0^t e^{-\beta s}\sigma dB_s \end{aligned}$$

so that

$$X_t = e^{\beta t}X_0 + \int_0^t e^{\beta(t-s)}\sigma \mathrm{d}B_s \ .$$

In particular, if $X_0 = x$, then X_t has a normal distribution with mean $e^{\beta t}x$. For example, if $n = N = 1$, then

$$X_t \sim N(e^{\beta t}x, \frac{\sigma^2}{2\beta}\left(e^{2\beta t} - 1\right)) \ .$$

It can be shown that (X_t) is a diffusion process, and its distribution can be described by its transition probability $P_t(x, dz)$. According to definition

$$(P_t f)(x) \equiv \int_{\mathbb{R}^N} f(z)P_t(x, dz) = \mathbb{E}\left(f(X_t)|X_0 = x\right) \ ,$$

thus

$$\begin{aligned} (P_t f)(x) &= \mathbb{E}\left(f(X_t)|X_0 = x\right) \\ &= \int_{\mathbb{R}} f(z)\frac{1}{\sqrt{2\pi\frac{\sigma^2}{2}\left(e^{2\beta t} - 1\right)}} \exp\left(-\frac{|z - e^{\beta t}x|^2}{\frac{\sigma^2}{2}\left(e^{2\beta t} - 1\right)}\right) dz \\ &= \int_{\mathbb{R}} f(e^{\beta t}x + \sqrt{\frac{\sigma^2}{2}\left(e^{2\beta t} - 1\right)}z)\frac{1}{\sqrt{2\pi}} \exp\left(-\frac{|z|^2}{2}\right) dz \\ &= \mathbb{E}f(e^{\beta t}x + \sqrt{\frac{\sigma^2}{2}\left(e^{2\beta t} - 1\right)}\xi) \end{aligned}$$

where ξ has the standard normal distribution $N(0,1)$. From the second line of the above formula, and compare to the definition of $P_t(x, dz)$, we may conclude that $P_t(x, dz) = p(t, x, z)dz$ with

$$p(t,x,z) = \frac{1}{\sqrt{2\pi\frac{\sigma^2}{2}\left(e^{2\beta t}-1\right)}} \exp\left(-\frac{|z-e^{\beta t}x|^2}{\frac{\sigma^2}{2}\left(e^{2\beta t}-1\right)}\right) .$$

$p(t, x, z)$ is called the transition density of the diffusion process $(X_t)_{t\geq 0}$. We thus, again following from the above formula, have a probability representation

$$(P_t f)(x) = \mathbb{E} f(e^{\beta t}x + \sqrt{\frac{\sigma^2}{2}\left(e^{2\beta t}-1\right)}\xi)$$

which is quite useful in some computations.

Remark 4.2.1 *It is easy to see from the above representation that*

$$\frac{d}{dx}(P_t f) = e^{\beta t} P_t \left(\frac{d}{dx} f\right) .$$

The distribution of (X_t) is determined by the transition density $p(t, x, z)$. Indeed, for any $0 < t_1 < \cdots < t_k$, the joint distribution of $(X_{t_1}, \cdots, X_{t_k})$ is Gaussian, and indeed its pdf is

$$p(t_1, x, z_1)p(t_2 - t_1, z_1, z_2)\cdots p(t_k - t_{k-1}, z_{k-1}, z_k) .$$

If $B = (B_t^1, \cdots, B_t^n)_{t\geq 0}$ is a Brownian motion in $\mathbb{R}^n$, then the solution X_t of the SDE:

$$dX_t = dB_t - (AX_t)\, dt$$

is called the Ornstein-Uhlenbeck process, where $A \geq 0$ is a $d \times d$ matrix called the drift matrix. Hence we have

$$X_t = e^{-At}X_0 + \int_0^t e^{-(t-s)A} dB_s .$$

Exercise 4.2.2 *If $X_0 = x \in \mathbb{R}^n$, compute $\mathbb{E}f(X_t)$, where X_t is the Ornstein-Uhlenbeck process with drift matrix A.*

4.2.2 Geometric Brownian motion

The Black-Scholes model is the stochastic differential equation

$$dS_t = S_t\left(\mu dt + \sigma dB_t\right) \tag{4.4}$$

whose solution to (4.4) is the stochastic exponential of

$$\int_0^t \mu ds + \int_0^t \sigma dB_s .$$

Hence

$$S_t = S_0 \exp\left(\int_0^t \sigma dB_s + \int_0^t \left(\mu - \frac{1}{2}\sigma^2\right) ds\right) .$$

In the case σ and μ are constants, then

$$S_t = S_0 \exp\left(\sigma B_t + \left(\mu - \frac{1}{2}\sigma^2\right) t\right)$$

which is called the geometric Brownian motion. If $S_0 = x > 0$, then S_t remains positive, and

$$\log S_t = \log x + \sigma B_t + \left(\mu - \frac{1}{2}\sigma^2\right) t$$

has a normal distribution with mean $\log x + \left(\mu - \frac{1}{2}\sigma^2\right) t$ and variance σ^2. Again, as a solution to the stochastic differential equation (4.4), $(S_t)_{t\geq 0}$ is a diffusion process, its distribution is determined by its transition function $P_t(x, dz)$ (unfortunately we have to use the same notations as in the last sub-section), and according to definition

$$\begin{aligned}
\int_{\mathbb{R}} f(z)P_t(x,dz) &= \mathbb{E}\left(f(S_t)|X_0 = x\right) \\
&= \mathbb{E}\left(f(xe^{\sigma B_t + \left(\mu - \frac{1}{2}\sigma^2\right)t})\right) \\
&= \int_{\mathbb{R}} f(xe^{\sigma z + \left(\mu - \frac{1}{2}\sigma^2\right)t})\frac{1}{\sqrt{2\pi t}}e^{-\frac{z^2}{2\pi t}}dz \\
&= \int_0^\infty f(y)\frac{1}{\sqrt{2\pi t}}\frac{1}{\sigma y}e^{-\frac{1}{2\pi t}\left(\frac{1}{\sigma}\log\frac{y}{x} - \left(\frac{\mu}{\sigma} - \frac{1}{2}\sigma\right)\right)^2}dy
\end{aligned}$$

where we assume that $\sigma > 0$ and have made the change of variable

$$xe^{\sigma z + \left(\mu - \frac{1}{2}\sigma^2\right)t} = y .$$

As usual, we define $(P_t f)(x) = \int_{\mathbb{R}} f(z)P_t(x,dz)$. By the third line of the previous formula

$$\begin{aligned}
(P_t f)(x) &= \int_{\mathbb{R}} f(xe^{\sigma z + \left(\mu - \frac{1}{2}\sigma^2\right)t})\frac{1}{\sqrt{2\pi t}}e^{-\frac{z^2}{2\pi t}}dz \\
&= \int_{\mathbb{R}} f(xe^{\sigma\sqrt{t}y + \left(\mu - \frac{1}{2}\sigma^2\right)t})\frac{1}{\sqrt{2\pi}}e^{-\frac{y^2}{2\pi}}dy \\
&= \mathbb{E}\left(f(xe^{\sigma\sqrt{t}\xi + \left(\mu - \frac{1}{2}\sigma^2\right)t})\right)
\end{aligned}$$

[we have made a change variable z into $\sqrt{t}y$], where $\xi \sim N(0,1)$. Compare with the definition of $P_t(x, dy)$ we have

$$P_t(x,dy) = \frac{1}{\sqrt{2\pi t}}\frac{1}{\sigma y}e^{-\frac{1}{2\pi t}\left(\frac{1}{\sigma}\log\frac{y}{x} - \left(\frac{\mu}{\sigma} - \frac{1}{2}\sigma\right)\right)^2}dy \quad \text{on } (0,\infty)$$

That is, (S_t) has the transition density

$$p(t,x,y) = \frac{1}{\sqrt{2\pi t}}\frac{1}{\sigma y}e^{-\frac{1}{2\pi t}\left(\frac{1}{\sigma}\log\frac{y}{x} - \left(\frac{\mu}{\sigma} - \frac{1}{2}\sigma\right)\right)^2} \quad \text{on } (0,\infty) .$$

and, therefore, for geometric Brownian motion

$$(P_t f)(x) = \int_0^\infty \frac{1}{\sqrt{2\pi t}} \frac{1}{\sigma y} e^{-\frac{1}{2\pi t}\left(\frac{1}{\sigma}\log\frac{y}{x} - \left(\frac{\mu}{\sigma} - \frac{1}{2}\sigma\right)\right)^2} f(y)dy$$

for any $x > 0$.

4.2.3 Cameron-Martin's formula

Consider a simple stochastic differential equation

$$dX_t = dB_t + b(t, X_t)dt \tag{4.5}$$

where $b(t,x)$ is a bounded, Borel measurable function on $[0,\infty) \times \mathbb{R}$. We may solve (4.5) by means of *change of probabilities.*

Let $(W_t)_{t\geq 0}$ be a standard Brownian motion on $(\Omega, \mathcal{F}, \mathcal{F}_t, \mathbb{P})$, and define probability measure $\mathbb{Q}$ on $(\Omega, \mathcal{F}_\infty)$ by

$$\left.\frac{d\mathbb{Q}}{d\mathbb{P}}\right|_{\mathcal{F}_t} = \mathcal{E}(N)_t \qquad \text{for all} \quad t \geq 0$$

where $N_t = \int_0^t b(s, W_s)\mathrm{d}W_s$ is a martingale (under the probability $\mathbb{P}$), with $\langle N \rangle_t = \int_0^t b(s, W_s)^2 \mathrm{d}s$, which is bounded on any finite interval. Hence

$$\mathcal{E}(N)_t = \exp\left(\int_0^t b(s, W_s)\mathrm{d}W_s - \frac{1}{2}\int_0^t b(s, W_s)^2 \mathrm{d}s\right)$$

is a martingale. According to Girsanov's theorem

$$B_t \equiv W_t - W_0 - \langle W, N \rangle_t$$

is a martingale under the new probability $\mathbb{Q}$, and $\langle B \rangle_t = \langle W \rangle_t = t$. By Lévy's martingale characterization of Brownian motion, $(B_t)_{t\geq 0}$ is a Brownian motion. Moreover

$$\begin{aligned} \langle W, N \rangle_t &= \langle \int_0^t \mathrm{d}W_s, \int_0^t b(s, W_s)\mathrm{d}W_s \rangle \\ &= \int_0^t b(s, W_s)\mathrm{d}s \end{aligned}$$

and therefore

$$W_t - W_0 - \int_0^t b(s, W_s)\mathrm{d}s = B_t$$

is a standard Brownian motion on $(\Omega, \mathcal{F}, \mathbb{Q})$. Thus

$$W_t = W_0 + B_t + \int_0^t b(s, W_s)\mathrm{d}s \tag{4.6}$$

so that $(W_t)_{t\geq 0}$ on $(\Omega, \mathcal{F}_\infty, \mathbb{Q})$ is a solution of (4.5). The solution we have just constructed is a weak solution of SDE (4.5).

Theorem 4.2.3 *(Cameron-Martin's formula) Let $b(t,x) = (b^1(t,x), \cdots, b^n(t,x))$ be bounded, Borel measurable functions on $[0,\infty) \times \mathbb{R}^n$. Let $W_t = (W^1, \cdots, W_t^n)$ be a standard Brownian motion on a filtered probability space $(\Omega, \mathcal{F}, \mathcal{F}_t, \mathbb{P})$, and let $\mathcal{F}_\infty = \sigma\{\mathcal{F}_t, t \geq 0\}$. Define probability measure Q on $(\Omega, \mathcal{F}_\infty)$ by*

$$\left.\frac{d\mathbb{Q}}{d\mathbb{P}}\right|_{\mathcal{F}_t} = e^{\sum_{k=1}^n \int_0^t b^k(s,W_s)dB_s^k - \frac{1}{2}\sum_{k=1}^n \int_0^t \left|b^k(s,W_s)\right|^2 ds} \quad \text{for} \quad t \geq 0 \ .$$

Then $(W_t)_{t\geq 0}$ under the probability measure $\mathbb{Q}$ is a solution to

$$dX_t^j = dB_t^j + b^j(t, X_t)\,dt \tag{4.7}$$

for some Brownian motion $(B_t^1, \cdots, B_t^n)_{t\geq 0}$ under probability $\mathbb{Q}$.

On the other hand, if (X_t) is a solution of SDE (4.7) on some probability space $(\Omega, \mathcal{F}, \mathcal{F}_t, \mathbb{P})$ and define $\tilde{\mathbb{P}}$

$$\left.\frac{d\tilde{\mathbb{P}}}{d\mathbb{P}}\right|_{\mathcal{F}_t} = \exp\left\{-\sum_{k=1}^n \int_0^t b^k(s,X_s)\mathrm{d}B_s^k - \frac{1}{2}\sum_{k=1}^n \int_0^t \left|b^k(s,X_s)\right|^2 \mathrm{d}s\right\} \quad \text{for} \quad t \geq 0$$

we may show that $(X_t)_{t\geq 0}$ under probability measure $\tilde{\mathbb{P}}$ is a Brownian motion. Therefore solutions to SDE (4.7) is unique in law: all solutions have the same distribution.

4.3 Existence and uniqueness

In this section we present a fundamental result about the existence and uniqueness of strong solutions.

4.3.1 Strong solutions: existence and uniqueness

By definition, any strong solution is a weak solution. We next prove a basic existence and uniqueness theorem for a stochastic differential equation under a global Lipschitz condition. Our proof will rely on two inequalities: The Gronwall inequality and Doob's L^p-inequality.

Lemma 4.3.1 *(The Gronwall inequality) If a non-negative function g satisfies the integral equation*

$$g(t) \leq h(t) + \alpha \int_0^t g(s)\,ds \ , \qquad 0 \leq t \leq T$$

where α is a constant and $h : [0,T] \to \mathbb{R}$ is an integrable function, then

$$g(t) \leq h(t) + \alpha \int_0^t e^{\alpha(t-s)} h(s)\,ds \quad , \qquad 0 \leq t \leq T \ .$$

Proof. Let $F(t) = \int_0^t g(s)\mathrm{d}s$. Then $F(0) = 0$ and

$$F'(t) \leq h(t) + \alpha F(t)$$

so that

$$\left(e^{-\alpha t}F(t)\right)' \leq e^{-\alpha t}h(t) \ .$$

Integrating the differential inequality we obtain

$$\int_0^t \left(e^{-\alpha s}F(s)\right)' \mathrm{d}s \leq \int_0^t e^{-\alpha s}h(s)\mathrm{d}s$$

and therefore

$$F(t) \leq \int_0^t e^{\alpha(t-s)}h(s)\mathrm{d}s$$

which yields Gronwall's inequality. ■

Consider the following stochastic differential equation

$$dX_t^j = \sum_{l=1}^n f_l^j(t, X_t)\mathrm{d}B_t^l + f_0^j(t, X_t)\mathrm{d}t \ ; \quad j = 1, \cdots, N \tag{4.8}$$

where $f_k^j(t,x)$ are Borel measurable functions on $\mathbb{R}_+ \times \mathbb{R}^N$, which are bounded on any compact subset in $\mathbb{R}^N$. We are going to show the existence and uniqueness by Picard's iteration. The main ingredient in the proof is a special case of Doob's L^p- inequality: if $(M_t)_{t\geq 0}$ is a square-integrable, continuous martingale with $M_0 = 0$, then for any $t > 0$

$$\mathbb{E}\left\{\sup_{s\leq t}|M_s|^2\right\} \leq 4\sup_{s\leq t} E\left(|M_s|^2\right) = 4E\langle M\rangle_t \ . \tag{4.9}$$

Lemma 4.3.2 *Let $(B_t)_{t\geq 0}$ be a standard BM in $\mathbb{R}$ on $(\Omega, \mathcal{F}_t, \mathcal{F}, \mathbb{P})$, and $(Z_t)_{t\geq 0}$ and $(\tilde{Z}_t)_{t\geq 0}$ be two continuous, adapted processes. Let $f(t,x)$ be a Lipschitz function*

$$|f(t,x) - f(t,y)| \leq C|x-y| \ ; \quad \forall t \geq 0, \ \ x, y \in \mathbb{R}$$

for some constant C.

1. *Let*

$$M_t = \int_0^t f(s, Z_s)dB_s - \int_0^t f(s, \tilde{Z}_s)dB_s \qquad \forall t \geq 0 \ .$$

Then

$$\mathbb{E}\sup_{s\leq t}|M_s|^2 \leq 4C^2\int_0^t \mathbb{E}\left|Z_s - \tilde{Z}_s\right|^2 ds$$

for all $t \geq 0$.

2. *If*

$$N_t = \int_0^t f(s, Z_s)ds - \int_0^t f(s, \tilde{Z}_s)ds \qquad \forall t \geq 0$$

then

$$\mathbb{E}\sup_{s\leq t}|N_s|^2 \leq C^2 t\int_0^t \mathbb{E}\left|Z_s - \tilde{Z}_s\right|^2 ds \qquad \forall t \geq 0 \ .$$

Proof. To prove the first statement, we notice that

$$\sup_{s\le t}|M_s|^2=\sup_{s\le t}\left|\int_0^s\left(f(u,Z_u)-f(u,\tilde{Z}_u)\right)\mathrm{d}B_u\right|^2$$

so that, by Doob's L^2-inequality

$$\begin{aligned}\mathbb{E}\sup_{s\le t}|M_s|^2 &= \mathbb{E}\sup_{s\le t}\left|\int_0^s\left(f(u,Z_u)-f(u,\tilde{Z}_u)\right)\mathrm{d}B_u\right|^2\\ &\le 4\mathbb{E}\left|\int_0^t\left(f(s,Z_s)-f(s,\tilde{Z}_s)\right)\mathrm{d}B_s\right|^2\\ &= 4\mathbb{E}\int_0^t\left|f(s,Z_s)-f(s,\tilde{Z}_s)\right|^2\mathrm{d}s\\ &\le 4C^2\int_0^t\mathbb{E}\left|Z_s-\tilde{Z}_s\right|^2\mathrm{d}s\ .\end{aligned}$$

Next we prove the second claim. Indeed

$$\begin{aligned}\sup_{s\le t}|N_s|^2 &= \sup_{s\le t}\left|\int_0^s\left(f(u,Z_u)-f(u,\tilde{Z}_u)\right)\mathrm{d}u\right|^2\\ &\le \left(\int_0^t\left|f(s,Z_s)-f(s,\tilde{Z}_s)\right|\mathrm{d}s\right)^2\\ &\le t\int_0^t\left|f(s,Z_s)-f(s,\tilde{Z}_s)\right|^2\mathrm{d}s\\ &\le C^2t\int_0^t\left|Z_s-\tilde{Z}_s\right|^2\mathrm{d}s\end{aligned}$$

where the second inequality follows from the Schwartz inequality. ∎

Theorem 4.3.3 *Consider SDE (4.8). Suppose that f_i^j satisfy the Lipschitz condition:*

$$|f(t,x)-f(t,y)|\le C|x-y| \tag{4.10}$$

and the linear-growth condition that

$$|f(t,x)|\le C(1+|x|) \tag{4.11}$$

for $t\in\mathbb{R}^+$ and $x,y\in\mathbb{R}^N$. Then for any $\eta\in L^2(\Omega,\mathcal{F}_0,\mathbb{P})$ and a standard Brownian motion $B_t=(B_t^i)$ in $\mathbb{R}^n$, there is a unique strong solution (X_t) of (4.8) with $X_0=\eta$.

Proof. *Proof of the existence of strong solutions.* The unique strong solution can be constructed by Picard's iteration. As the first step in the iteration, $X^{(0)}=\eta$, and define $X^{(n)}$ inductively by the following equation

$$X_t^{(n+1)}=\eta+\int_0^t f(s,X_s^{(n)})\cdot dB_s+\int_0^t f_0(s,X_s^{(n)})ds \tag{4.12}$$

for $n = 0, 1, 2, \cdots$. For simplifying our notations we introduce the following notations:

$$D_t^{(n)} = \left| X_t^{(n)} - X_t^{(n-1)} \right|$$

and

$$\rho^{(n)}(t) = \mathbb{E} \sup_{s \le t} \left| X_s^{(n)} - X_s^{(n-1)} \right|^2$$

for $n = 1, 2, \cdots$. Since f and f_0 are globally Lipschitz, so that

$$\begin{aligned} D_t^{(n+1)} &\le \left| \int_0^t \left(f(s, X_s^{(n)}) - f(s, X_s^{(n-1)}) \right) \cdot dB_s \right| \\ &+ \left| \int_0^t \left(f_0(s, X_s^{(n)}) - f_0(s, X_s^{(n-1)}) \right) ds \right| \\ &\le \left| \int_0^t \left(f(s, X_s^{(n)}) - f(s, X_s^{(n-1)}) \right) \cdot dB_s \right| \\ &+ C \left| \int_0^t D_s^{(n)} ds \right| . \end{aligned}$$

It follows that

$$\begin{aligned} \left(D_t^{(n+1)} \right)^2 &\le 2 \sup_{s \le t} \left| \int_0^s \left(f(r, X_r^{(n)}) - f(r, X_r^{(n-1)}) \right) \cdot dB_r \right|^2 \\ &+ 2C \left| \int_0^t D_s^{(n)} ds \right|^2 . \end{aligned}$$

After taking the expectations both sides of the previous inequality to obtain

$$\rho^{(n+1)}(s) \le 2\mathbb{E} \left[\sup_{s \le t} \left| \int_0^s \left(f(r, X_r^{(n)}) - f(r, X_r^{(n-1)}) \right) \cdot dB_r \right|^2 \right] + 2C\mathbb{E} \left| \int_0^t D_s^{(n)} ds \right|^2 . \quad (4.13)$$

The first expectation can be estimated by Doob's inequality to the martingale

$$M_t = \int_0^t \left(f(r, X_r^{(n)}) - f(r, X_r^{(n-1)}) \right) \cdot dB_r$$

where

$$\langle M \rangle_t = \int_0^t \left| f(r, X_r^{(n)}) - f(r, X_r^{(n-1)}) \right|^2 dr \le C \int_0^t \left(D_r^{(n)} \right)^2 dr$$

here the inequality follows from ((4.10)), so that

$$\mathbb{E} \sup_{s \le t} \left| \int_0^s \left(f(r, X_r^{(n)}) - f(r, X_r^{(n-1)}) \right) \cdot dB_r \right| \le 4C \int_0^t \rho^{(n)}(s) ds. \quad (4.14)$$

While the second expectation on the right-hand side of (4.13) can be handled as the following. By Cauchy-Schwartz inequality

$$\mathbb{E} \left| \int_0^t D_s^{(n)} ds \right|^2 \le t \int_0^t \mathbb{E} \left| D_s^{(n)} \right|^2 ds \le t \int_0^t \rho^{(n)}(s) ds. \quad (4.15)$$

Inserting (4.14) and (4.15) into (4.13) we thus obtain the following key inequality

$$\rho^{(n+1)}(t) \leq (8C + 2Ct) \int_0^t \rho^{(n)}(s) ds \tag{4.16}$$

for all $t \geq 0$ and $n = 0, 1, 2, \cdots$.

For any but fixed $T > 0$ we have

$$\rho^{(n+1)}(t) \leq (8C + 2CT) \int_0^t \rho^{(n)}(s) ds$$

and

$$\rho^{(n)}(s) \leq (8C + 2CT) \int_0^s \rho^{(n-1)}(r) dr,$$

hence

$$\rho^{(n+1)}(t) \leq (8C + 2CT)^2 \int_0^t \left(\int_0^s \rho^{(n-1)}(r) dr \right) ds.$$

By repeating the same procedure we obtain that

$$\rho^{(n+1)}(t) \leq (8C + 2CT)^n \int_0^t \int_0^{t_1} \cdots \int_0^{t_{n-1}} \rho^{(1)}(t_{n-1}) dt_{n-1} dt_{n-2} \cdots dt_1. \tag{4.17}$$

While

$$\begin{aligned}
\rho^{(1)}(t) &= \mathbb{E} \sup_{s \leq t} \left| \int_0^s f(r, \eta) dB_r + \int_0^s f_0(r, \eta) dr \right|^2 \\
&\leq 2\mathbb{E} \sup_{s \leq t} \left| \int_0^s f(r, \eta) dB_s \right|^2 + 2\mathbb{E} \sup_{s \leq t} \left| \int_0^s f_0(r, \eta) ds \right|^2 \\
&\leq 2\mathbb{E} \sup_{s \leq T} \left| \int_0^s f(r, \eta) B_r \right|^2 + 2\mathbb{E} \sup_{s \leq t} \left| \int_0^s f_0(r, \eta) ds \right|^2 \\
&\equiv C_2(T),
\end{aligned}$$

thus by inserting this estimate to (4.17) we conclude that

$$\begin{aligned}
\rho^{(n+1)}(t) &\leq C_2(T) \left(8C + 2CT\right)^n \int_0^t \int_0^{t_1} \cdots \int_0^{t_{n-1}} dt_{n-1} dt_{n-2} \cdots dt_1 \\
&= C_2(T) \frac{\left(8C + 2CT\right)^n T^n}{n!}
\end{aligned}$$

for every $n = 0, 1, 2, \cdots$ and for all $t \leq T$. Therefore

$$\mathbb{E} \sup_{t \leq T} \left| X_t^{(n+1)} - X_t^{(n)} \right|^2 \leq C_2(T) \frac{\left(8C + 2CT\right)^n T^n}{n!} \tag{4.18}$$

and therefore, by using Markov inequality,

$$\mathbb{P} \left\{ \sup_{t \leq T} \left| X_t^{(n+1)} - X_t^{(n)} \right|^2 > \frac{1}{2^n} \right\} \leq C_2(T) \frac{\left(8C + 2CT\right)^n (2T)^n}{n!}$$

for every $n = 0, 1, 2, \cdots$. Since

$$\sum_n C_2(T) \frac{(8C + 2CT)^n (2T)^n}{n!} < \infty$$

so that, according to Borel-Cantelli's lemma, with probability one, $X^{(n)}$ converges uniformly to a limit X on any $[0, T]$. It is easy to verify X is a strong solution.

Proof of uniqueness.

Let X and $\hat{X}$ be two solutions with same Brownian motion B. Then

$$X_t = \eta + \int_0^t f(s, X_s) \cdot \mathrm{d}B_s + \int_0^t f_0(s, X_s) ds$$

and

$$\hat{X}_t = \eta + \int_0^t f(s, \hat{X}_s) \cdot \mathrm{d}B_s + \int_0^t f_0(s, \hat{X}_s) ds$$

Then, as in the proof of the existence,

$$\mathbb{E}|X_t - \hat{X}_t|^2 \le C(T) \int_0^t \mathbb{E}|X_t - \hat{X}_t|^2 \mathrm{d}s$$

for all $t \le T$, where $C(T)$ is a constant. The Gronwall inequality implies thus that $\mathbb{E}\left(|Y_t - Z_t|^2\right) = 0$. ■

Remark 4.3.4 *The iteration $X^{(n)}$ constructed in the proof of Theorem 4.3.3 is a function of the Brownian motion B, and $X_t^{(n)}$ only depends on η and B_s, $0 \le s \le t$.*

4.3.2 Continuity in initial conditions

Theorem 4.3.5 *Under the same assumptions as in Theorem 4.3.3. Given a BM $B = (B_t)_{t \ge 0}$ in $\mathbb{R}^n$ on $(\Omega, \mathcal{F}, \mathcal{F}_t, \mathbb{P})$, let $(X^x(t))_{t \ge 0}$ be the unique strong solution of (4.8). Then $x \to X^x$ is uniformly continuous almost surely on any finite interval $[0, T]$:*

$$\lim_{\delta \downarrow 0} \sup_{|x-y|<\delta} \mathbb{E}\left\{ \sup_{0 \le t \le T} |X^x(t) - X^y(t)|^2 \right\} = 0 \ . \tag{4.19}$$

Proof. Let us only consider 1-dimensional case. Thus

$$X^x(t) = x + \int_0^t f_1(s, X^x(s)) \mathrm{d}B_s + \int_0^t f_0(s, X^x(s)) \mathrm{d}s$$

and

$$X^y(t) = y + \int_0^t f_1(s, X^y(s)) \mathrm{d}B_s + \int_0^t f_0(s, X^y(s)) \mathrm{d}s \ .$$

Therefore, by Doob's maximal inequality,

$$
\begin{aligned}
& \mathbb{E}\left\{\sup_{0\le t\le T}|X^x(t)-X^y(t)|^2\right\} \le 3|x-y|^2 \\
& +3\mathbb{E}\left\{\sup_{0\le t\le T}\left|\int_0^t (f_1(s,X^x(s))-f_1(s,X^y(s)))\mathrm{d}B_s\right|^2\right\} \\
& +3\mathbb{E}\left\{\sup_{0\le t\le T}\left|\int_0^t (f_0(s,X^x(s))-f_0(s,X^y(s)))\mathrm{d}s\right|^2\right\} \\
\le\ & 3|x-y|^2+12\mathbb{E}\left\{\left|\int_0^T (f_1(s,X^x(s))-f_1(s,X^y(s)))\,\mathrm{d}B_s\right|^2\right\} \\
& +3T\mathbb{E}\left\{\int_0^T |f_0(X^x(s))-f_0(X^y(s))|^2\,\mathrm{d}s\right\} \\
\le\ & 3|x-y|^2+12\mathbb{E}\left\{\int_0^T |f_1(s,X^x(s))-f_1(s,X^y(s))|^2\,\mathrm{d}s\right\} \\
& +3TC^2\mathbb{E}\left\{\int_0^T |X^x(s)-X^y(s)|^2\,ds\right\} \\
\le\ & 3|x-y|^2+3C^2(4+T)\int_0^T \mathbb{E}\left(|X^x(t)-X^y(t)|^2\right)\mathrm{d}t.
\end{aligned}
$$

Setting

$$\Delta(t)=\mathbb{E}\left\{\sup_{0\le s\le t}|X^x(s)-X^y(s)|^2\right\}\ ,$$

then we have

$$\Delta(T)\le 3|x-y|^2+3C^2(4+T)\int_0^T \Delta(t)dt$$

and therefore by Gronwall's inequality

$$\Delta(T)\le 6|x-y|^2\exp(12C^2+3TC^2)$$

which yields (4.19). ■

4.4 Martingale problem and weak solutions

In the lecture (lecture 13) I did the computation for one dimensional case, but the computations I did in the lecture do not depend on the dimension, so I supply the details for a general case here, though some technical conditions still can be weaken but this can be done better case by case.

Consider the following SDE

$$dX_t^i=\sum_{k=1}^d \sigma_k^i(X_t)\mathrm{d}B_t^k+b^i(X_t)\mathrm{d}t \tag{4.20}$$

where $i = 1, \cdots, d$. We assume that all coefficients σ_j^i and b^i are Borel measurable and are bounded (here the boundedness is brought in for simplicity, which can be replaced by other conditions). Let $X = (X_t)_{\geq 0}$ be a solution on a filtered probability space $(\Omega, \mathcal{F}, \mathcal{F}_t, \mathbb{P})$ and B is a d-dimensional Brownian motion. That is, X satisfies the following stochastic integral equation

$$X_t^i = X_0^i + \int_0^t \sum_{k=1}^d \sigma_k^i(X_s) dB_s^k + \int_0^t b^i(X_t) ds$$

for $t \geq 0$ and $i = 1, \cdots, d$. Then, by Itô's formula, for every $f \in C^2(\mathbb{R}^d)$,

$$f(X_t) - f(X_0) = \sum_{i=1}^d \int_0^t \frac{\partial f}{\partial x^i}(X_s) \mathrm{d}X_s^i + \frac{1}{2} \sum_{i,j=1}^d \int_0^t \frac{\partial^2 f}{\partial x^i \partial x^j}(X_s) \mathrm{d}\langle X^i, X^j \rangle_s.$$

Since

$$\begin{aligned}
\langle X^i, X^j \rangle_t &= \int_0^t \sum_{k,l=1}^d \sigma_k^i(X_s) \sigma_l^j(X_s) d \left\langle B^k, B^l \right\rangle_s \\
&= \int_0^t \sum_{k=1}^d \sigma_k^i(X_s) \sigma_k^j(X_s) ds.
\end{aligned}$$

Let $a^{ij}(x) = \sum_{k=1}^d \sigma_k^i(x) \sigma_k^j(x)$. Then $a(x) = (a^{ij}(x))$, as a square $d \times d$-matrix, is symmetric and non-negative (here non-negative means all its eigenvalues are non-negative). Then the previous equality shows that

$$\langle X^i, X^j \rangle_t = \int_0^t a^{ij}(X_s) ds$$

for all $t \geq 0$. Substituting this relation into the previous equation for $f(X_t)$ we obtain

$$f(X_t) - f(X_0) = \sum_{i=1}^d \int_0^t \frac{\partial f}{\partial x^i}(X_s) \mathrm{d}X_s^i + \frac{1}{2} \sum_{i,j=1}^d \int_0^t a^{ij}(X_s) \frac{\partial^2 f}{\partial x^i \partial x^j}(X_s) \mathrm{d}s.$$

Now using the fact that X is a solution to our SDE we thus deduce that

$$\begin{aligned}
f(X_t) - f(X_0) &= \sum_{i=1}^d \int_0^t \frac{\partial f}{\partial x^i}(X_s) \left(\sum_{k=1}^d \sigma_k^i(X_s) \mathrm{d}B_s^k + b^i(X_s) \mathrm{d}s \right) \\
&\quad + \frac{1}{2} \sum_{i,j=1}^d \int_0^t a^{ij}(X_s) \frac{\partial^2 f}{\partial x^i \partial x^j}(X_s) \mathrm{d}s \\
&= \int_0^t \sum_{k,i=1}^d \sigma_k^i(X_s) \frac{\partial f}{\partial x^i}(X_s) \mathrm{d}B_s^k + \int_0^t \sum_{i=1}^d b^i(X_s) \frac{\partial f}{\partial x^i}(X_s) \mathrm{d}s \\
&\quad + \frac{1}{2} \sum_{i,j=1}^d \int_0^t a^{ij}(X_s) \frac{\partial^2 f}{\partial x^i \partial x^j}(X_s) \mathrm{d}s.
\end{aligned}$$

Let us introduce the following linear operator

$$L=\frac{1}{2}\sum_{i,j=1}^{d}a^{ij}\frac{\partial^2}{\partial x^i\partial x^j}+\sum_{i=1}^{d}b^i\frac{\partial}{\partial x^i} \tag{4.21}$$

which is an elliptic differential operator of second-order. The linear operator L operating on a C^2 function f gives a new function Lf:

$$Lf(x)=\frac{1}{2}\sum_{i,j=1}^{d}a^{ij}(x)\frac{\partial^2 f(x)}{\partial x^i\partial x^j}+\sum_{i=1}^{d}b^i(x)\frac{\partial f(x)}{\partial x^i}$$

for every $x\in\mathbb{R}^d$. By our assumption Lf is Borel measurable and locally bounded, i.e. Lf is bounded on any compact subset of $\mathbb{R}^d$. Under these notations, the previous equation for $f(X_t)$ can be rearranged as the following

$$f(X_t)-f(X_0)-\int_0^t Lf(X_s)ds=\int_0^t\sum_{k,i=1}^{d}\sigma_k^i(X_s)\frac{\partial f}{\partial x^i}(X_s)\mathrm{d}B_s^k$$

which is a continuous local martingale. Now the crucial observation, made first by Stroock and Varadhan, is that the left-hand side is independent of an underlying Brownian motion B. Since its importance, we introduce the following notation:

$$M_t^{[f]}=f(X_t)-f(X_0)-\int_0^t Lf(X_s)ds \tag{4.22}$$

for $t\geq 0$, and $f\in C^2(\mathbb{R}^d)$.

As a by-product of the previous computation, we also have

$$M_t^{[f]}=\int_0^t\sum_{k,i=1}^{d}\sigma_k^i(X_s)\frac{\partial f}{\partial x^i}(X_s)\mathrm{d}B_s^k$$

and therefore

$$\begin{aligned}\langle M^{[f]},M^{[g]}\rangle_t&=\int_0^t\sum_{k,i=1}^{d}\sigma_k^i(X_s)\frac{\partial f}{\partial x^i}(X_s)\sum_{l,j=1}^{d}\sigma_l^j(X_s)\frac{\partial f}{\partial x^j}(X_s)\mathrm{d}\left\langle B^k,B^l\right\rangle s\\&=\int_0^t\sum_{j,i=1}^{d}\sum_{l,j=1}^{d}\sigma_k^j(X_s)\sigma_k^i(X_s)\frac{\partial f}{\partial x^i}(X_s)\frac{\partial f}{\partial x^j}(X_s)\mathrm{d}s\\&=\int_0^t\sum_{j,i=1}^{d}a^{ij}(X_s)\frac{\partial f}{\partial x^i}(X_s)\frac{\partial f}{\partial x^j}(X_s)\mathrm{d}s\end{aligned}$$

Lemma 4.4.1 *Suppose σ_j^i and b^i are bounded and Borel measurable, a^{ij} and L are defined by (4.21). If $(X_t)_{t\geq 0}$ is a (weak) solution to SDE (4.20) on $(\Omega,\mathcal{F},\mathcal{F}_t,\mathbb{P})$, then for any $f\in C^2(\mathbb{R})$*

$$M_t^{[f]}=f(X_t)-f(X_0)-\int_0^t(Lf)(X_s)ds$$

is a continuous local martingale on $(\Omega, \mathcal{F}, \mathcal{F}_t, \mathbb{P})$. *Moreover*

$$\langle M^{[f]}, M^{[g]}\rangle_t = \int_0^t \sum_{j,i=1}^d a^{ij}(X_s)\frac{\partial f}{\partial x^i}(X_s)\frac{\partial f}{\partial x^j}(X_s)ds.$$

The fact the statement that $M^{[f]}$ is a local martingale does not depend on the underlying Brownian motion allows us formulate the concept of weak solutions in terms of martingale problem.

Definition 4.4.2 *Let* $a(x) = (a^{ij}(x))_{i,j\leq d}$ *be symmetric square* $d \times d$ *matrix valued, Borel measurable and bounded function on* $\mathbb{R}^d$ *and* $b^i(x)$ *are bounded Borel measurable. Let* L *be defined by (4.21) which is a linear operator operating on* C^2 *functions. Let* $(X_t)_{t\geq 0}$ *be a continuous stochastic process on a filtered space* $(\Omega, \mathcal{F}, \mathcal{F}_t, \mathbb{P})$. *Then we say that* $(X_t)_{t\geq 0}$ *together with the probability* $\mathbb{P}$ *is a solution to the* L*-martingale problem, if for every* $f \in C^2(\mathbb{R})$

$$M_t^{[f]} = f(X_t) - f(X_0) - \int_0^t Lf(X_s)ds$$

is a local martingale under the probability $\mathbb{P}$.

Therefore a solution $(X_t)_{t\geq 0}$ of SDE (4.20) on $(\Omega, \mathcal{F}, \mathbb{P})$ is a solution to L-martingale problem. Conversely, we have the following theorem.

Theorem 4.4.3 *Let* a^{ij}, b^i *and* L *be given in Definition 4.4.2. Suppose there is* $(\sigma_j^i(x))$ *which is a symmetric matrix-valued, Borel measurable and*

$$\lambda^{-1} \leq \sum_{i,j}^d \xi^i\xi^j\sigma_j^i(x) \leq \lambda$$

for some constant $\lambda > 0$, *for all* $x \in \mathbb{R}^d$, *(this is equivalent to say all eigenvalues of the square matrix* $\left(\sigma_j^i(x)\right)$ *are between* λ^{-1} *and* λ *for all* x*), such that* $a^{ij} = \sum_{k=1}^d \sigma_k^i\sigma_k^j$ *for* $i, j = 1, \cdots, d$. *If* $(X_t)_{t\geq 0}$ *on* $(\Omega, \mathcal{F}, \mathcal{F}_t, \mathbb{P})$ *is a continuous process solving the* L*-martingale problem: for every* $f \in C^2(\mathbb{R}^d)$

$$M_t^{[f]} = f(X_t) - f(X_0) - \int_0^t Lf(X_s)ds$$

is a continuous local martingale on $(\Omega, \mathcal{F}, \mathcal{F}_t, \mathbb{P})$, *then* $(X_t)_{t\geq 0}$ *on* $(\Omega, \mathcal{F}, \mathcal{F}_t, \mathbb{P})$ *is a weak solution to SDE*

$$dX_t^i = \sum_{k=1}^d \sigma_k^i(X_t)dB_t^k + b^i(X_t)dt \tag{4.23}$$

for $i = 1, \cdots, d$. *Moreover*

$$\langle M^{[f]}, M^{[g]}\rangle_t = \int_0^t \{L(fg) - f(Lg) - g(Lf)\}(X_s)ds\ .$$

for every $f, g \in C^2(\mathbb{R}^d)$. *Of course*

$$L(fg) - f(Lg) - g(Lf) = \sum_{j,i=1}^d a^{ij}\frac{\partial f}{\partial x^i}\frac{\partial g}{\partial x^j}.$$

Proof. The key idea is to construct a Brownian motion B such that

$$X_t^i = X_0^i + \int_0^t \sum_{k=1}^d \sigma_k^i(X_s) dB_s^k + \int_0^t b^i(X_t) ds.$$

If we apply the local martingale property to functions $f^i(x) = x^i$ (the i-th coordinate of x), according to the definition of L-martingale problem,

$$\begin{aligned} M_t^i &= f^i(X_t) - f^i(X_0) - \int_0^t Lf^i(X_s) ds \\ &= X_t^i - X_0^i - \int_0^t b^i(X_s) ds \end{aligned}$$

where $M^i = M^{[f^i]}$ for simplicity are continuous martingale. Therefore the candidate for Brownian motion is defined by

$$B_t^i = \sum_{k=1}^d \int_0^t \left(\sigma^{-1}(X_s)\right)_k^i dM_s^k$$

which are continuous martingales with initial zero. Here σ^{-1} denotes the inverse matrix of σ with its entries denoted by $(\sigma^{-1})_j^i$. We need to show that $B = (B^1, \cdots, B^d)$ is a standard Brownian motion in $\mathbb{R}^d$. We will apply Lévy's characterization for Brownian motion to this end. Therefore we need to compute

$$\left\langle B^i, B^j \right\rangle_t = \int_0^t \sum_{k,l=1}^d \left(\sigma^{-1}(X_s)\right)_k^i \left(\sigma^{-1}(X_s)\right)_l^j d\left\langle M^k, M^l \right\rangle_s.$$

Let us compute $\langle M^{[f]}, M^{[g]} \rangle_t$. By the polarization identity, we only need to compute $\langle M^{[f]} \rangle_t$. By assumption

$$M_t^{[f^2]} = f^2(X_t) - f^2(X_0) - \int_0^t Lf^2(X_s) ds$$

is a local martingale. On the other hand, by integration by part

$$\begin{aligned} f^2(X_t) - f^2(X_0) &= 2\int_0^t f(X_s) df(X_s) + \left\langle M^{[f]} \right\rangle_t \\ &= 2\int_0^t f(X_s) \left[dM_s^{[f]} + Lf(X_s) \mathrm{d}s\right] + \left\langle M^{[f]} \right\rangle_t \\ &= 2\int_0^t f(X_s) dM_s^{[f]} + 2\int_0^t f(X_s) Lf(X_s) \mathrm{d}s + \left\langle M^{[f]} \right\rangle_t. \end{aligned}$$

Substitute it into the previous equality we obtain that

$$M_t^{[f^2]} = 2\int_0^t f(X_s) dM_s^{[f]} + \left\langle M^{[f]} \right\rangle_t - \int_0^t \left(Lf^2 - 2fLf\right)(X_s) ds.$$

Therefore we must have

$$\left\langle M^{[f]}\right\rangle_t - \int_0^t \left(Lf^2 - 2fLf\right)(X_s)ds = 0$$

that is

$$\left\langle M^{[f]}\right\rangle_t = \int_0^t \left(Lf^2 - 2fLf\right)(X_s)ds.$$

Therefore

$$\left\langle M^{[f]}, M^{[g]}\right\rangle_t = \int_0^t (L(fg) - fLg - gLf)(X_s)\mathrm{d}s.$$

It is an easy exercise to verify that

$$L(fg) - fLg - gLf = \sum_{j,i=1}^d a^{ij}\frac{\partial f}{\partial x^i}\frac{\partial g}{\partial x^j}.$$

In particular, for coordinate functions $f^k(x) = x^k$ and $f^l(x) = x^l$ we have

$$\left\langle M^k, M^l\right\rangle_t = \int_0^t a^{kl}(X_s)\mathrm{d}s = \int_0^t \sum_{m=1}^d \sigma_m^k(X_s)\sigma_m^l(X_s)\mathrm{d}s.$$

Thanks to this formula we can compute

$$\begin{aligned}
\left\langle B^i, B^j\right\rangle_t &= \int_0^t \sum_{k,l=1}^d \left(\sigma^{-1}(X_s)\right)_k^i \left(\sigma^{-1}(X_s)\right)_l^j d\left\langle M^k, M^l\right\rangle_s \\
&= \int_0^t \sum_{m,k,l=1}^d \left(\sigma^{-1}(X_s)\right)_k^i \left(\sigma^{-1}(X_s)\right)_l^j \sigma_m^k(X_s)\sigma_m^l(X_s)\mathrm{d}s \\
&= \int_0^t \sum_{m,l=1}^d \delta_{im}\left(\sigma^{-1}(X_s)\right)_l^j \sigma_m^l(X_s)\mathrm{d}s \\
&= \int_0^t \sum_{m=1}^d \delta_{im}\delta_{jm}\mathrm{d}s = \delta_{ij}t.
\end{aligned}$$

According to Lévy's theorem, B is a standard Brownian motion. By definition

$$\begin{aligned}
\int_0^t \sum_{k=1}^d \sigma_k^i(X_s)\mathrm{d}B_s^k &= \int_0^t \sum_{k,l=1}^d \sigma_k^i(X_s)\left(\sigma^{-1}(X_s)\right)_l^k dM_s^l \\
&= \int_0^t \sum_{l=1}^d \delta_{il}dM_s^l = M_t^i \\
&= X_t^i - X_0^i - \int_0^t b^i(X_s)ds.
\end{aligned}$$

Rearrange this equation to deduce that

$$X_t^i = X_0^i + \int_0^t \sum_{k=1}^d \sigma_k^i(X_s)\mathrm{d}B_s^k + \int_0^t b^i(X_s)ds$$

for $i = 1, \cdots, d$. That is to say that (X, B) is a weak solution to (4.23). ∎

Chapter 5

Local times

Itô's formula provides us a powerful semimartingale decomposition for $f(X_t) - f(X_0)$, where f is a C^2-function and X is a semimartingale. For example, if $X = (X_t)$ is a continuous semimartingale, then

$$f(X_t) - f(X_0) = \int_0^t f'(X_s)dX_s + \frac{1}{2}\int_0^t f''(X_s)d\langle X\rangle_s.$$

It has been recognized that such decomposition plays a central rôle in stochastic analysis, it is thus natural to search for a decomposition for functions f with less regularity.

The concept of "local times" of Brownian motion was introduced by Lévy in his study of fine properties of Brownian sample paths, which was further developed into an important tool in constructing diffusion processes on the real line in the classics Itô-McKean "Diffusion processes and their sample paths". In 1958, Tanaka established a decomposition for $|B_t - a|$ where B is a Brownian motion. One of the important contributions made by the French school is that if f is convex and X is a continuous semimartingale, then $f(X)$ is again a semimartingale, its decomposition sometimes is called the generalized Itô's formula.

5.1 Tanaka's formula and local time

H. Tanaka proved that $|X_t - a|$ is a semimartingale if X is a Brownian motion, established its semimartingale decomposition, and identified its variation process as the local time introduced by L. Lévy. These results have been extended to a general continuous semimartingales by M. Yor and etc.

To study these results, we need a few facts about convex functions which are discussed with details in the appendix.

Let $X = (X_t)_{t\geq 0}$ be a continuous semimartingale and $a \in \mathbb{R}$. Consider the convex function $f(x) = |x - a|$ on $\mathbb{R}$, which is smooth except at a. Its left derivative $f'_-(x) = -1$ for $x \leq a$ and $f'_-(x) = 1$ for $x > a$. Choose a function $\alpha \in C_0^\infty(\mathbb{R})$ with a compact support in $(0, T)$ (for some $T > 0$), such that $\int_{\mathbb{R}} \alpha(s)ds = 1$. For each $\varepsilon > 0$, define $f_\varepsilon = f * \alpha_\varepsilon$

where $\alpha_\varepsilon(z) = \frac{1}{\varepsilon}\alpha\left(\frac{x}{\varepsilon}\right)$, so that

$$\begin{aligned} f_\varepsilon(x) &= \int_{-\infty}^{\infty} f(x-z)\alpha_\varepsilon(z)dz \\ &= \int_{-\infty}^{\infty} f(z)\alpha_\varepsilon(x-z)dz. \end{aligned}$$

f_ε is smooth for every $\varepsilon > 0$ and

$$\frac{d^k}{dx^k} f_\varepsilon(x) = \int_{-\infty}^{\infty} f(z)\frac{d^k}{dx^k}\alpha_\varepsilon(x-z)dz$$

for $k = 0, 1, 2, \cdots$. Since f is continuous, so that $f_\varepsilon \to f$ uniformly on any compact subset, and

$$\begin{aligned} \frac{d}{dx} f_\varepsilon(x) &= \int_{-\infty}^{\infty} f(z)\frac{d}{dx}\alpha_\varepsilon(x-z)dz \\ &= -\int_{-\infty}^{\infty} f(z)\frac{d}{dz}\alpha_\varepsilon(x-z)dz \\ &= -\lim_{\delta\to 0}\frac{1}{\delta}\int_{-\infty}^{\infty} f(z)\left(\alpha_\varepsilon(x-z-\delta) - \alpha_\varepsilon(x-z)\right)dz \\ &= -\lim_{\delta\to 0}\int_0^{\varepsilon T} \frac{f(x-z-\delta) - f(x-z)}{\delta}\alpha_\varepsilon(z)dz. \end{aligned}$$

If $x - a \le 0$, then for $z \in (0, \infty)$ and $\delta > 0$

$$\frac{f(x-z-\delta) - f(x-z)}{\delta} = 1$$

so that

$$\begin{aligned} \frac{d}{dx} f_\varepsilon(x) &= -\lim_{\delta\to 0+}\int_0^{\infty} \frac{f(x-z-\delta) - f(x-z)}{\delta}\alpha_\varepsilon(z)dz \\ &= -\int_0^{\infty}\alpha_\varepsilon(z)dz = -1. \end{aligned}$$

While for $x - a > 0$, and for every $\varepsilon > 0$ such that $\varepsilon T < \frac{x-a}{2}$, then for $|\delta| < 0$ we have for any $z \in (0, \varepsilon T)$

$$\frac{f(x-z-\delta) - f(x-z)}{\delta} = \frac{x-z-\delta-(x-z)}{\delta} = -1$$

we therefore have

$$\begin{aligned} \frac{d}{dx} f_\varepsilon(x) &= -\lim_{\delta\to 0}\int_0^{\infty} \frac{f(x-z-\delta) - f(x-z)}{\delta}\alpha_\varepsilon(z)dz \\ &= -\lim_{\delta\to 0}\int_0^{\varepsilon T}(-1)\alpha_\varepsilon(z)dz = -1. \end{aligned}$$

Hence $f'_\varepsilon(x) \to f'_-(x)$ at every point $x \in \mathbb{R}$ including $x = a$. Applying Itô's formula to f_ε for each $\varepsilon > 0$ to obtain

$$f_\varepsilon(X_t) = f_\varepsilon(X_0) + \int_0^t f'_\varepsilon(X_s)dX_s + \frac{1}{2}\int_0^t f''_\varepsilon(X_s)d\langle X\rangle_s.$$

Since $f_\varepsilon(X_t) \to f(X_t)$, $f_\varepsilon(X_0) \to f(X_0)$ and

$$\int_0^t f'_\varepsilon(X_s)dX_s \to \int_0^t f'_-(X_s)dX_s$$

as $\varepsilon \downarrow 0$, therefore $\frac{1}{2}\int_0^t f''_\varepsilon(X_s)d\langle X\rangle_s$ has a limit as $\varepsilon \downarrow 0$, which is denoted by $2L_t^a$.

Tanaka's formula. Suppose X is a continuous semimartingale and $a \in \mathbb{R}$. Then there is an adapted, continuous increasing process $(L_t^a)_{t\geq 0}$ such that

$$|X_t - a| = |X_0 - a| + \int_0^t \operatorname{sgn}(X_s - a)dX_s + 2L_t^a. \tag{5.1}$$

where

$$\operatorname{sgn} = 1_{(0,\infty)} - 1_{(-\infty,0]}$$

is the sign function (which is left-hand continuous). This formula indeed can be considered as the definition of the local time L_t^a of X at a.

The important thing here is that, $t \to L_t^a$ is continuous, increasing, starting at 0, and adapted to the filtration generated by $(X_t)_{t\geq 0}$. Since

$$L_t^a = \frac{1}{2}|X_t - a| - \frac{1}{2}|X_0 - a| - \frac{1}{2}\int_0^t \operatorname{sgn}(X_s - a)dX_s$$

so that L_t^a is jointly measurable in (t, a) with respect to the Borel σ-algebra $\mathcal{B}([0,\infty)\times\mathbb{R})$.

Next we establish the Itô's formula for a general convex function f.

Suppose $f : \mathbb{R} \to \mathbb{R}$ is a convex function, then f must be continuous on $\mathbb{R}$, its right-derivative

$$f'_+(a) = \lim_{h\to 0+} \frac{f(a+h) - f(a)}{h}$$

and its left derivative

$$f'_-(a) = \lim_{h\to 0-} \frac{f(a+h) - f(a)}{h}$$

exist and are increasing on $\mathbb{R}$. f'_+ is right-continuous and f'_- is left-continuous. Moreover f'_+ is the right-continuous modification of f'_-, and similarly f'_- is the left-continuous modification of f'_+, in the sense that

$$f'_+(a) = \lim_{\varepsilon\to 0+} f'_-(a+\varepsilon)$$

and

$$f'_-(a) = \lim_{\varepsilon\to 0+} f'_+(a-\varepsilon)$$

for every $a \in \mathbb{R}$.

The Lebesgue-Stieltjies measure associated with the right derivative f'_+ is denoted by $\mu_{f'_+}$ which is the unique measure on $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ such that

$$\mu_{f'_+}((a,b]) = f'_+(b) - f'_+(a)$$

for any $a \leq b$. As a consequence we also have

$$\mu_{f'_+}([a,b)) = f'_-(b) - f'_-(a)$$

for any $a \leq b$. On can verify that if in addition $f \in C^2(\mathbb{R})$, then $\mu_{f'_+}(dx) = f''(x)dx$ which is absolutely continuous with respect to the Lebesgue measure. [Be careful, if the second derivative of f exists but not continuous, then $\mu_{f'_+}(dx)$ is in general different from the measure $f''(x)dx$!]

To establish the Itô's formula for $f(X_t)$ (where f is a convex function), let us compute the following integral

$$J(x) = \int_{(-n,n)} |x-z| \mu_{f'_+}(dz)$$

where $x \in (-n,n)$, where $n > 0$ is a positive number. Firstly we notice that

$$\begin{aligned} J &= \int_{(-n,x)} |x-z| \mu_{f'_+}(dz) + \int_{(x,n)} |x-z| \mu_{f'_+}(dz) \\ &= \int_{(-n,x)} (x-z) \mu_{f'_+}(dz) - \int_{(x,n)} (x-z) \mu_{f'_+}(dz) \\ &\equiv J_1 - J_2. \end{aligned}$$

Integrating by parts we obtain that

$$\begin{aligned} J_1 &= \int_{(-n,x)} (x-z) \mu_{f'_+}(dz) \\ &= -(x+n) f'_+(-n) + \int_{(-n,x)} f'_+(z) dz \\ &= -(x+n) f'_+(-n) + f(x) - f(-n) \end{aligned}$$

and similarly

$$\begin{aligned} J_2 &= \int_{(x,n)} (x-z) \mu_{f'_+}(dz) \\ &= (x-n) f'_-(n) + f(n) - f(x) \end{aligned}$$

so that

$$\begin{aligned} J &= -(x+n) f'_+(-n) + f(x) - f(-n) - (x-n) f'_-(n) - f(n) + f(x) \\ &= 2f(x) - \left(f'_+(-n) + f'_-(n) \right) x - n f'_+(-n) + n f'_-(n) - f(-n) - f(n) \end{aligned}$$

which yields that

$$f(x) = \alpha_n x + \beta_n + \frac{1}{2} \int_{(-n,n)} |x-z| \mu_{f'_+}(dz)$$

for any $x \in (-n, n)$, where

$$\alpha_n = \frac{f'_+(-n) + f'_-(n)}{2}$$

and

$$\beta_n = \frac{1}{2}\left(f(-n) + f(n) + nf'_+(-n) - nf'_-(n)\right).$$

By continuity in x, the formula holds for $x = \pm n$ as well. Now we have established all necessary facts about a convex function required to derive the Itô's formula.

Theorem 5.1.1 *Let $X = (X_t)_{t \geq 0}$ be a continuous semimartingale and $f : \mathbb{R} \to \mathbb{R}$ a convex function. Then*

$$f(X_t) = f(X_0) + \int_0^t f'_-(X_s)dX_s + \int_{\mathbb{R}} L_t^a \mu_{f'_+}(da) \tag{5.2}$$

where $\mu_{f'_+}$ is the Lebesgue-Stieltjies measure associated with the right derivative f'_+. The same formula is equally valid if $f = \sum_{j=1}^n c_j f_j$ where f_j are convex functions, c_j are constants, and $\mu_{f'_+}(da) = \sum_{j=1}^n c_j \mu_{f'_{j+}}(da)$ which is a signed measure.

If $f \in C^2(\mathbb{R})$ then

$$\int_{\mathbb{R}} f''(a) L_t^a da = \frac{1}{2}\int_0^t f''(X_s)d\langle X\rangle_s \tag{5.3}$$

for every $t \geq 0$.

Proof. For any $n > 0$ we have

$$f(x) = \alpha_n x + \beta_n + \frac{1}{2}\int_{(-n,n)} |x - z|\mu_{f'_+}(dz)$$

for every $x \in [-n, n]$, where α_n and β_n are given as above.

Let $T_n = \inf\{t \geq 0 : |X_t| \geq n\}$. Then $\{T_n : n \geq 1\}$ is an increasing sequence of stopping times and $T_n \uparrow \infty$. Applying Tanaka's formula to $X_t^n = X_{t\wedge T_n}$ one obtains

$$|X_t^n - a| = |X_0^n - a| + \int_0^t \mathrm{sgn}(X_s^n - a)dX_s^n + 2L_t^{n,a}$$

for each n, and as $n \to \infty$, $L_t^{n,a} = L_{t\wedge T_n}^a \uparrow L_t^a$. According to the previous representation for f one obtains

$$\begin{aligned}
f(X_t^n) &= \alpha_n X_t^n + \beta_n + \frac{1}{2}\int_{(-n,n)} |X_t^n - z|\mu_{f'_+}(dz) \\
&= \alpha_n X_t^n + \beta_n + \frac{1}{2}\int_{(-n,n)} |X_0^n - z|\mu_{f'_+}(dz) \\
&\quad + \int_{(-n,n)}\int_0^t \mathrm{sgn}(X_s^n - z)dX_s^n \mu_{f'_+}(dz) + \int_{(-n,n)} L_t^{n,z}\mu_{f'_+}(dz) \\
&= \alpha_n X_t^n + \beta_n + \frac{1}{2}\int_{(-n,n)} |X_0^n - z|\mu_{f'_+}(dz) \\
&\quad + \int_0^t \int_{(-n,n)} \mathrm{sgn}(X_s^n - z)\mu_{f'_+}(dz)dX_s^n + \int_{(-n,n)} L_t^{n,z}\mu_{f'_+}(dz).
\end{aligned}$$

According to definition of sgn and $\mu_{f'_+}$ we have

$$\begin{aligned}\int_{(-n,n)} \operatorname{sgn}(X_s^n - z)\mu_{f'_+}(dz) &= \int_{(-n,X_s^n)} \mu_{f'_+}(dz) - \int_{[X_s^n,n)} \mu_{f'_+}(dz) \\ &= f'_-(X_s^n) - f'_+(-n) - \big(f'_-(n) - f'_-(X_s^n)\big) \\ &= 2f'_-(X_s^n) - \big(f'_+(-n) + f'_-(n)\big)\end{aligned}$$

so after rearranging the terms to obtain

$$f'_-(X_s^n) = \frac{1}{2}\int_{(-n,n)} \operatorname{sgn}(X_s^n - z)\mu_{f'_+}(dz) + \alpha_n.$$

By the representation

$$f(X_0^n) = \alpha_n X_0^n + \beta_n + \frac{1}{2}\int_{(-n,n)} |X_0^n - z|\mu_{f'_+}(dz).$$

Putting these relations into the equation for $f(X_t^n)$ we thus have

$$\begin{aligned}f(X_t^n) &= f(X_0^n) + \alpha_n\left(X_t^n - X_0^n\right) + \int_0^t \left[f'_-(X_s^n) - \alpha_n\right] dX_s^n \\ &\quad + \int_{(-n,n)} L_t^{n,z}\mu_{f'_+}(dz) \\ &= f(X_0^n) + \int_0^t f'_-(X_s^n)dX_s^n + \int_{(-n,n)} L_t^{n,z}\mu_{f'_+}(dz).\end{aligned}$$

Letting $n \to \infty$ we obtain the formula. ■

For each a, $(L_t^a)_{t\geq 0}$ is called the local time (process) of X at site $a \in \mathbb{R}$. According to definition

$$L_t^a = \frac{1}{2}\left(|X_t - a| - |X_0 - a|\right) - \frac{1}{2}\int_0^t \operatorname{sgn}(X_s - a)dX_s \tag{5.4}$$

for $t \geq 0$. Our next task is to study the regularity of $(t,a) \to L_t^a$ as a random field.

Let us begin with the following

Lemma 5.1.2 *Let $b \in \mathbb{R}$, and $g(x) = \frac{1}{2}(x-b)^2$ if $x \leq b$ and $g(x) = 0$ if $x > b$.*

1) Then $g'(x) = x - b$ for $x \leq b$, $g'(x) = 0$ for $x > b$, and

$$\mu_{g'}(dx) = 1_{(-\infty,b]}(x)dx.$$

2) If X is a continuous semimartingale, then

$$g(X_t) - g(X_0) = \int_0^t g'(X_s)dX_s + \frac{1}{2}\int_0^t 1_{(-\infty,b]}(X_s)d\langle X\rangle_s.$$

Proof. Item 1) follows an easy computation, and the fact that g' is continuous. To show item 2), we choose a function $\alpha \in C_0^\infty(\mathbb{R})$ with compact support in $(0,\infty)$ such that $\int_{\mathbb{R}} \alpha(x)dx = 1$ and for $\varepsilon > 0$ set $\alpha_\varepsilon(x) = \varepsilon^{-1}\alpha\left(\varepsilon^{-1}x\right)$ and $g_\varepsilon = \alpha_\varepsilon * g$. Then, since g' is continuous, both g_ε, g'_ε converge to g, g' uniformly on any compact set, and g''_ε converges to $1_{(-\infty,b]}$ point-wise. Therefore, by passing to the limit as $\varepsilon \downarrow 0$ in the Itô's formula

$$g_\varepsilon(X_t) - g_\varepsilon(M_0) = \int_0^t g'_\varepsilon(M_s)dM_s + \frac{1}{2}\int_0^t g''_\varepsilon(M_s)d\langle M\rangle_s$$

we obtain item 2). ■

Lemma 5.1.3 *Let $X = X_0 + M + A$ be the Doob-Meyer decomposition of a continuous semimartingale, and $S_n = \inf\{t \geq 0 : |X_t| \geq n\}$. Then, for any $p \geq 1$ there is a constant C depending only on p such that*

$$\begin{aligned} & \mathbb{E}\left|\sup_{0\leq s\leq t}\int_0^{s\wedge S_n} 1_{(a,b]}(X_r)dM_r\right|^{2p} \\ \leq\ & C\left\{(L+n)^p + \mathbb{E}\left(\int_0^t |dA_s|\right)^p + \mathbb{E}\langle M\rangle_t^{p/2}\right\}|b-a|^p \end{aligned} \tag{5.5}$$

for any $n \geq 0$, $t \geq 0$ and any $a \leq b$ such that $|a|, |b| \leq L$.

Proof. Without losing a generality, we may assume that $S_n = \infty$. That is, $|X_t| \leq n$ for all t. Let $N_t = \int_0^t 1_{(a,b]}(X_r)dM_r$ for $t \geq 0$. According to Burkhölder's inequality

$$\mathbb{E}\sup_{0\leq s\leq t}\left|\int_0^s 1_{(a,b]}(X_r)dM_r\right|^{2p} \leq C\mathbb{E}\left|\int_0^t 1_{(a,b]}(X_r)d\langle M\rangle_r\right|^p$$

where C, and through the proof, denotes a constant depending only on p, which can be different from place to place. Let $g_c(x) = \frac{1}{2}(x-c)^2$ for $x \leq c$ and $g_c(x) = 0$ for $x > c$. By Lemma 5.1.2

$$\begin{aligned} & \int_0^t 1_{(a,b]}(X_s)d\langle M\rangle_s \\ =\ & 2\left(g_b(X_t) - g_a(X_t) - g_b(X_0) + g_a(X_0)\right) \\ & -2\int_0^t \left(g'_b - g'_a\right)(X_s)dM_s - 2\int_0^t \left(g'_b - g'_a\right)(X_s)dA_s. \end{aligned}$$

It is easy to see that

$$|g_b(x) - g_a(x)| \leq (L+n)|b-a|$$

and

$$|g'_b(x) - g'_a(x)| \leq |b-a|$$

so that

$$\begin{aligned} & \left|\int_0^t \mathcal{X}_{\{a<X_s\leq b\}}(X_s)d\langle M\rangle_s\right| \\ \leq\ & \left[4(L+n) + 2\int_0^t |dA_s|\right]|b-a| \\ & +2\left|\int_0^t \left(g'_b - g'_a\right)(X_s)dM_s\right|. \end{aligned}$$

Therefore

$$\begin{aligned}
& \mathbb{E}\left|\int_0^t \mathcal{X}_{\{a<X_s\leq b\}}(X_s)d\langle M\rangle_s\right|^p \\
\leq\ & 2^{p-1}\mathbb{E}\left[4(L+n)+2\int_0^t |dA_s|\right]^p |b-a|^p \\
& +2^p\mathbb{E}\left|\int_0^t \left(g_b'-g_a'\right)(X_s)dM_s\right|^p .
\end{aligned}$$

For the last term on the right hand side, we may use Burkhölder's inequality again, to obtain

$$\begin{aligned}
\mathbb{E}\left|\int_0^t \left(g_b'-g_a'\right)(X_s)dM_s\right|^p &\leq C\mathbb{E}\left|\int_0^t \left(g_b'-g_a'\right)^2(X_s)d\langle M\rangle_s\right|^{p/2} \\
&\leq C\mathbb{E}\langle M\rangle_t^{p/2}|b-a|^p
\end{aligned}$$

which completes the proof. ■

Theorem 5.1.4 *Let $X_t = X_0 + M_t + A_t$ be a continuous semimartingale, where M is a local continuous martingale, A is a continuous process with finite variation. Let*

$$L_t^a = \frac{1}{2}\left(|X_t-a|-|X_0-a|\right)-\frac{1}{2}\int_0^t sgn(X_s-a)dX_s$$

be the local time of X at a.

1) There is a set $N\in\mathcal{F}$ with probability zero, such that

$$(t,a)\to L_t^a(\omega)$$

is jointly continuous in $t\in\mathbb{R}_+$ and right-continuous in $a\in\mathbb{R}$ for any $\omega\in\Omega\setminus N$.

2) For each $\omega\in\Omega\setminus N$, $L_t^{a-}(\omega)$ exists and

$$L_t^a(\omega)-L_t^{a-}(\omega)=\int_0^t 1_{\{X_s(\omega)=a\}}dA_s(\omega)$$

and $(t,a)\to L_t^a(\omega)$ is jointly continuous almost surely if and only if $\int_0^t 1_{\{X_s=a\}}dA_s=0$.

Proof. For each n define

$$T_n=\inf\{t\geq 0: |M_t|\geq n,\ \langle M\rangle_t\geq n \text{ or } V(A)_t\geq n\}$$

where $V(A)_t$ is the total variation of A on $[0,t]$. Then $\{T_n : n\geq 1\}$ is an increasing sequence of stopping times, such that $T_n\uparrow\infty$. Let $X_t^n = X_{t\wedge T_n}$, and similar notations apply to other processes. Then

$$L_{t\wedge T_n}^a=\frac{1}{2}\left(|X_t^n-a|-|X_0^n-a|\right)-\frac{1}{2}\int_0^t \operatorname{sgn}(X_s^n-a)dX_s^n.$$

Hence we first consider a semimartingale such that all processes M, $\langle M\rangle$ and $V(A)$ are bounded. For this case, $t \to \langle M\rangle_t + V(A)_t + t$ is continuous and strictly increasing, its (right-continuous) inverse

$$\tau_t = \inf\{s : \langle M\rangle_s + V(A)_s + s > t\}.$$

Then each τ_t is a stopping, $t \to \tau_t$ is continuous and $\tau_t \uparrow \infty$. By definition

$$\langle M\rangle_{\tau_t} + V(A)_{\tau_t} + \tau_t = t, \ \ \tau_{\langle M\rangle_t + V(A)_{\tau_t} + t} = t \quad \forall t \geq 0. \tag{5.6}$$

In particular

$$\langle M\rangle_{\tau_t} - \langle M\rangle_{\tau_s} + V(A)_{\tau_t} - V(A)_{\tau_s} + (\tau_t - \tau_s) = t - s$$

for $t \geq s \geq 0$, which yields that

$$\langle M\rangle_{\tau_t} - \langle M\rangle_{\tau_s} \leq t - s \quad \forall t \geq s \geq 0 \tag{5.7}$$

and

$$V(A)_{\tau_t} - V(A)_{\tau_s} \leq t - s \quad \forall t \geq s \geq 0. \tag{5.8}$$

According to (5.4) one has

$$\begin{aligned} L^a_{\tau_t} &= \frac{1}{2}(|X_{\tau_t} - a| - |X_0 - a|) - \frac{1}{2}\int_0^{\tau_t} \mathrm{sgn}(X_s - a)dX_s \\ &= \frac{1}{2}(|Z_t - a| - |Z_0 - a|) - \frac{1}{2}\int_0^{t} \mathrm{sgn}(Z_s - a)dZ_s. \end{aligned}$$

where $Z_t = X_{\tau_t}$. Let $Z_t = Z_0 + \tilde{M}_t + \tilde{A}_t$.Then $\tilde{M}_t = M_{\tau_t}$ and $\tilde{A}_t = A_{\tau_t}$.

Since $t \to \tau_t$ is continuous and strictly increasing, so that

$$L^a_t = L^a_{\tau_{\tau_t^{-1}}}$$

where

$$\tau_t^{-1} = \langle M\rangle_t + V(A)_t + t$$

is continuous, so that we only need to prove the conclusions in the theorem for $L^a_{\tau_t}$.

Therefore we may further assume that processes M, $\langle M\rangle$ and $V(A)$ are bounded by some constant C_0, and

$$\langle M\rangle_t - \langle M\rangle_s \leq t - s \quad \forall t \geq s \geq 0,$$

and

$$V(A)_t - V(A)_s \leq t - s \quad \forall t \geq s \geq 0.$$

Let $N(a)_t = \frac{1}{2}\int_0^t \mathrm{sgn}(X_s - a)dM_s$ for simplicity. Then

$$\begin{aligned} L^a_t &= \frac{1}{2}(|X_t - a| - |X_0 - a|) - N(a)_t \\ &\quad - \frac{1}{2}\int_0^t \mathrm{sgn}(X_s - a)dA_s. \end{aligned}$$

We next prove that

$$(t,a) \to J_t^a \equiv \frac{1}{2}\left(|X_t - a| - |X_0 - a|\right) - N(a)_t \tag{5.9}$$

is jointly continuous. More precisely, there is a set $\mathcal{N} \subset \Omega$ with probability zero, such that $(t,a) \to J_t^a(\omega)$ is continuous on $\mathbb{R}_+ \times \mathbb{R}$ for any $\Omega \setminus \mathcal{N}$.

Clearly we only need to show $(t,a) \to N(a)_t$ is continuous except a probability zero set. Let $a < b \in \mathbb{R}$, $0 \le s \le t$. Then

$$\begin{aligned} N(a)_t - N(b)_s &= \frac{1}{2}\int_0^t \operatorname{sgn}(X_r - a)dM_r - \frac{1}{2}\int_0^s \operatorname{sgn}(X_r - b)dM_r \\ &= \frac{1}{2}\int_s^t \operatorname{sgn}(X_r - a)dM_r + \int_0^s 1_{(a,b]}(X_r)dM_r. \end{aligned}$$

Therefore for any $p \ge 1$ we have

$$\begin{aligned} \mathbb{E}\,|N(a)_t - N(b)_s|^{2p} \ &\le \ 2^{2(p-1)}\mathbb{E}\left|\int_0^s 1_{(a,b]}(X_r)dM_r\right|^{2p} \\ &\quad +2^{2(p-1)}\mathbb{E}\left|\int_s^t \operatorname{sgn}(X_r - a)dM_r\right|^{2p}. \end{aligned}$$

The first integral on the right-hand side can be dominated via (5.5), so let us control the second expectation. Indeed, by using the Burkhölder's inequality

$$\begin{aligned} \mathbb{E}\left|\int_s^t \operatorname{sgn}(X_r - a)dM_r\right|^{2p} &\le C\mathbb{E}\left|\int_s^t d\langle M\rangle_r\right|^p \\ &= C\mathbb{E}(\langle M\rangle_t - \langle M\rangle_s)^p. \end{aligned}$$

Therefore

$$\begin{aligned} & \mathbb{E}\,|N(a)_t - N(b)_s|^{2p} \\ \le\ & C\mathbb{E}(\langle M\rangle_t - \langle M\rangle_s)^p \\ & +C\left\{(L+C_0)^p + \mathbb{E}\left(\int_s^t |dA_s|\right)^p + \mathbb{E}\langle M\rangle_{t-s}^{p/2}\right\}|b-a|^p \end{aligned}$$

for any a, b such that $|a|, |b| \le L$, any $t \ge 0$. Since, under our reduction, it holds that

$$\int_0^t |dA_s| \le t \quad \text{and} \quad \langle M\rangle_t - \langle M\rangle_s \le t - s$$

so that there is a constant C depending only on p, L and C_0

$$\mathbb{E}\,|N(a)_t - N(b)_s|^{2p} \le C\left[t^{\frac{p}{2}}(b-a)^p + (t-s)^p\right] \tag{5.10}$$

for any $s, t \ge 0$, $a, b \in \mathbb{R}$ such that $|a|, |b| \le L$. It follows that, according to the Kolmogorov-Čentsov theorem, $(t,a) \to N(a)_t$ is α_p-Hölder continuous on any compact subset of $[0,\infty)\times\mathbb{R}$, for any $p > 2$, where $\alpha_p = \frac{p-2}{2p}$.

Finally we treat the Riemann-Stieltjies integral

$$I_t^a \equiv \int_0^t \operatorname{sgn}(X_s - a) dA_s.$$

It is clear that

$$\begin{aligned} I_t^a - I_s^b &= \int_0^t (\operatorname{sgn}(X-a) - \operatorname{sgn}(X-b))\, dA \\ &\quad + \int_s^t \operatorname{sgn}(X-b) dA \end{aligned}$$

so that, since $a \to \operatorname{sgn}(X_s - a)$ is right-continuous,

$$\lim_{\substack{s\to t \\ b\downarrow a}} \left(I_t^a - I_s^b\right) = 0$$

and

$$\lim_{\substack{s\to t \\ b\uparrow a}} \left(I_t^a - I_s^b\right) = -2\int_0^t \mathcal{X}_{\{X_s - a = 0\}} dA_s.$$

Therefore $(t,a) \to L_t^a$ is jointly right continuous in a and continuous in t. Moreover it is jointly continuous if and only if $\int_0^t \mathcal{X}_{\{X_s=a\}} dA_s = 0$. In particular

$$L_t^a - L_t^{a-} = \int_0^t \mathcal{X}_{\{X_s=a\}} dA_s$$

which completes the proof. ■

From now on, for a continuous semimartingale $X = (X_t)_{t\geq 0}$, we always use a version of its local time L_t^a such that $(t,a) \to L_t^a$ is jointly continuous in t, right continuous in a and having left hand limits L_t^{a-}, without further qualification.

By checking the preceding proof, we find that it is possible to improve on the regularity for the local time of Brownian motion.

Theorem 5.1.5 *Let $\{L_t^a : t \geq 0,\ a \in \mathbb{R}\}$ be the local time of one-dimensional standard Brownian motion $(W_t)_{t\geq 0}$.*

1) For any $L, T > 0$ and $p > 1$ there is a constant C depending only on L, T, p such that

$$\mathbb{E}|L_t^a - L_s^b|^{2p} \leq C\left(|b-a|^p + |t-s|^p\right) \tag{5.11}$$

for $(t,a), (s,b) \in [0,T] \times [-L, L]$.

2) If $\alpha < \frac{1}{2}$, $(t,a) \to L_t^a$ is jointly α-Hölder continuous.

Proof. From the proof of the previous theorem, for Brownian motion $(W_t)_{t\geq 0}$, $\langle W\rangle_t = t$.

$$L_t^a = \frac{1}{2}\left(|W_t - a| - |a|\right) - N(a)_t$$

where

$$N(a)_t = \frac{1}{2}\int_0^t \operatorname{sgn}(W_s - a) dW_s$$

so that

$$N(b)_t - N(a)_t = \frac{1}{2}\int_0^t \mathcal{X}_{\{a<W_s\leq b\}}(W_s)dW_s.$$

Let $g_c(x) = \frac{1}{2}(x-c)^2$ if $x \leq c$ and $g_c(x) = 0$ if $x > c$. By Lemma 5.1.2 one has

$$\begin{aligned}
&\int_0^t \mathcal{X}_{\{a<W_s\leq b\}}(W_s)ds \\
= \; & (W_t - b)^2 1_{\{W_t\leq b\}} - (W_t - a)^2 1_{\{W_t \leq a\}} \\
& -2\int_0^t \left(g_b' - g_a'\right)(W_s)dW_s \\
= \; & (W_t - b)^2 1_{\{a<W_t\leq b\}} - 2\,(b-a)\,W_t 1_{\{W_t\leq a\}} \\
& + \left(b^2 - a^2\right) 1_{\{W_t\leq a\}} - 2\int_0^t \left(g_b' - g_a'\right)(W_s)dW_s.
\end{aligned}$$

While for any $p > 1$ one has the following elementary estimates

$$\mathbb{E}|W_t - b|^{2p} 1_{\{a<W_t\leq b\}} \leq |a-b|^{2p}$$

and

$$\begin{aligned}
\mathbb{E}|W_t|^{2p} 1_{\{W_t\leq a\}} &= \int_{-\infty}^a |z|^{2p}\frac{1}{\sqrt{2\pi t}}e^{-\frac{z^2}{2t}}dz \\
&\leq t^p \int_{-\infty}^{\infty} |z|^{2p}\frac{1}{\sqrt{2\pi}}e^{-\frac{z^2}{2}}dz \\
&= C_p t^p\ .
\end{aligned}$$

so that

$$\begin{aligned}
&\mathbb{E}\left|\int_0^t \mathcal{X}_{\{a<W_s\leq b\}}(W_s)ds\right|^p \\
\leq \; & C\left[(a-b)^p + (b+a)^p + t^p\right](b-a)^p \\
& +C\mathbb{E}\left|\int_0^t \left(g_b' - g_a'\right)(W_s)dW_s\right|^p \\
\leq \; & C\left[(a-b)^p + (b+a)^p + t^p\right](b-a)^p \\
& +C\mathbb{E}\left|\int_0^t (b-a)^2\, ds\right|^{p/2}
\end{aligned}$$

where we have used the fact that $|g_b'(x) - g_a'(x)| \leq |b-a|$ for any x. Therefore

$$\mathbb{E}\left|\int_0^t \mathcal{X}_{\{a<W_s\leq b\}}(W_s)ds\right|^p \leq C|b-a|^p$$

where

$$C = C_p\left[(a-b)^p + (b+a)^p + t^p + t^{p/2}\right]$$

where C_p depends only on p. It follows from B-D-G inequality that

$$\begin{aligned}\mathbb{E}|\sup_{s\leq t}(N(b)_s - N(a)_s)|^{2p} &\leq C\mathbb{E}\left|\int_0^t \mathcal{X}_{\{a<W_s\leq b\}}(W_s)ds\right|^p \\ &\leq C|b-a|^p\end{aligned}$$

where C as above. Therefore, for any $L, T > 0$ and $p > 1$, there is a constant $C = C(L, T, p)$ such that

$$\mathbb{E}|N(a)_t - N(b)_s|^{2p} \leq C\left(|b-a|^p + |t-s|^p\right)$$

for all $(s, b), (t, a) \in [0, T] \times [-L, L]$. Now, by the Tanaka formula

$$\begin{aligned}L_t^a - L_s^b &= \frac{1}{2}\left(|W_t - a| - |W_s - b|\right) \\ &\quad -\frac{1}{2}\left(|a| - |b|\right) - \left(N(a)_t - N(b)_s\right)\end{aligned}$$

so that

$$|L_t^a - L_s^b| \leq \frac{1}{2}|W_t - W_s| + |b-a| + |N(a)_t - N(b)_s|\,,$$

which, together with the estimate for N_t, yields the desired result. ∎

Remark 5.1.6 *The following formula (from the preceding proof) may be useful. For $t, s \geq 0$ and $a > b$ then*

$$\begin{aligned}L_t^a - L_s^b &= \frac{1}{2}\left(|W_t - a| - |W_s - b|\right) + \frac{1}{2}\left(|b| - |a|\right) \\ &\quad + (b-a)\left(\frac{b+a}{2} - W_t\right)1_{\{W_t\leq a\}} \\ &\quad +\frac{1}{2}(W_t - b)^2 1_{\{a<W_t\leq b\}} - (b-a)\int_0^t 1_{\{W_s\leq a\}}dW_s \\ &\quad -\int_0^t (W_s - a)1_{\{a<W_t\leq b\}}dW_s - \int_s^t sgn(W_r - b)dW_r.\end{aligned}$$

Let us prove important properties about the local time L_t^a of a continuous semimartingale $X_t = X_0 + M_t + A_t$.

Proposition 5.1.7 *The following Tanaka's formulas hold*

$$(X_t - a)^+ = (X_0 - a)^+ + \int_0^t 1_{(a,\infty)}(X_s)dX_s + L_t^a \tag{5.12}$$

and

$$(X_t - a)^- = (X_0 - a)^- - \int_0^t 1_{(-\infty,a]}(X_s)dX_s + L_t^a. \tag{5.13}$$

Proof. Apply (5.2) to $f(x) = (x-a)^+$ so that $f'_-(x) = 1_{(a,\infty)}(x)$ and $\mu_{f'_+}(dx) = \delta_a(dx)$, to obtain (5.12). Similarly, applying (5.2) to $f(x) = (x-a)^-$: $f'_-(x) = -1_{(-\infty,a]}$ and $\mu_{f'_+}(dx) = \delta_a(dx)$ so that (5.13) follows. ∎

Theorem 5.1.8 *The local time $t \to L_t^a$ increasing only on $\{s : X_s = a\}$. More precisely $\int_0^t |X_s - a| dL_s^a = 0$ for all $t > 0$.*

Proof. According to Itô's formula

$$|X_t - a|^2 = |X_0 - a|^2 + 2\int_0^t (X_s - a)dX_s + \langle M\rangle_t.$$

On the other hand, according to Tanaka's formula, $|X_t - a|$ is a continuous semimartingale with decomposition

$$|X_t - a| = |X_0 - a| + \int_0^t \operatorname{sgn}(X_s - a)dX_s + 2L_t^a.$$

Thus

$$\begin{aligned} \langle |X - a|\rangle_t &= \int_0^t (\operatorname{sgn}(X_s - a))^2 \, d\langle X\rangle_s \\ &= \int_0^t d\langle X\rangle_s = \langle M\rangle_t \end{aligned}$$

together with the integration by parts

$$\begin{aligned} |X_t - a|^2 &= |X_0 - a|^2 + 2\int_0^t |X_s - a| d|X_s - a| \\ &\quad + \langle |X - a|\rangle_t \\ &= |X_0 - a|^2 + 2\int_0^t (X_s - a)\, dX_s + \langle M\rangle_t \\ &\quad + 4\int_0^t |X_s - a| dL_s^a. \end{aligned}$$

Therefore we must have

$$\int_0^t |X_s - a| dL_s^a = 0.$$

■

Theorem 5.1.9 *(*Occupation time formula*) Let φ be a Borel measurable function on $\mathbb{R}$, bounded or non-negative.*

1) If X is a continuous semi-martingale, then

$$\int_0^t \varphi(X_s)d\langle X\rangle_s = 2\int_{\mathbb{R}} \varphi(a)L_t^a da \tag{5.14}$$

for all $t \geq 0$.

2) If X is a continuous local martingale, or if $X = M + A$ (where M is a continuous local martingale and A is adapted with finite variations) such that $\int_0^t 1_{\{X_s = a\}} dA_s = 0$ for all $t > 0$, then

$$\lim_{\varepsilon \downarrow 0} \frac{1}{4\varepsilon} \int_0^t 1_{(a-\varepsilon, a+\varepsilon)}(X_s) d\langle X\rangle_s = L_t^a$$

for $t \geq 0$*.*

3) If B *is a Brownian motion and* L_t^a *(for* $t \geq 0$*) is the local time process of* B *at* a*, then*

$$\lim_{\varepsilon \downarrow 0} \frac{1}{4\varepsilon} \int_0^t 1_{(a-\varepsilon, a+\varepsilon)}(X_s) dt = L_t^a$$

for all a *and* $t \geq 0$*.*

Proof. Suppose $\varphi = f''$ for some $f \in C^2(\mathbb{R})$, then by (5.3), (5.14) holds. In particular, (5.14) is valid for any continuous function φ, so it is true for any bounded or non-negative φ. ∎

Proposition 5.1.10 *Suppose* $f : \mathbb{R} \to \mathbb{R}$ *is continuous, and there are finite many points* $a_1 < \cdots < a_n$ *such that* $f \in C^2(a_i, a_{i+1})$ *(*$i = 0, \cdots, n$*, with* $a_0 = -\infty$*,* $a_{n+1} = \infty$*),* f *has left and right derivatives at* a_i*. Then*

$$\begin{aligned} f(X_t) &= f(X_0) + \int_0^t f'_-(X_s) dX_s + \frac{1}{2} \int_0^t f''(X_s) d\langle X \rangle_s \\ &\quad + \sum_{i=1}^n \left(f'(a_i+) - f'(a_i-) \right) L_t^{a_i}. \end{aligned} \tag{5.15}$$

Proof. We can apply Theorem 5.1.1 to f. In this case f'' exists except at a_i and therefore

$$\mu_{f'_+}(dx) = f''(x)dx + \sum_{i=1}^n \left(f'_+(a_i) - f'_-(a_i) \right) \delta_{a_i}(dx).$$

Hence

$$\begin{aligned} f(X_t) &= f(X_0) + \int_0^t f'_-(X_s) dX_s + \int_{\mathbb{R}} f''(a) L_t^a da \\ &\quad + \sum_{i=1}^n \left(f'_+(a_i) - f'_-(a_i) \right) L_t^{a_i}. \end{aligned}$$

and the formula follows from

$$\int_{\mathbb{R}} f''(a) L_t^a da = \frac{1}{2} \int_0^t f''(X_s) d\langle X \rangle_s.$$

∎

As an application we prove the comparison theorem (due to Yamada and etc.)

Theorem 5.1.11 *Let* σ*,* b^1 *and* b^2 *are real valued, Lipschitz continuous functions on* $\mathbb{R}$*. Let* X^1 *and* X^2 *be the strong solutions to the following SDEs respectively:*

$$X_t^i = X_0^i + \int_0^t \sigma(X_s^i) dB_s + \int_0^t b^i(X_s^i) ds$$

for $i = 1, 2$*, where* B *is a Brownian motion on* $(\Omega, \mathcal{F}, \mathbb{P})$*. Suppose*

(1) $X_0^1 \geq X_0^2$ *almost surely, and*

(2) $b^1(x) \geq b^2(x)$ *for all* $x \in \mathbb{R}$*.*

Then $X_t^1 \geq X_t^2$ *for all* t *almost surely.*

Proof. Under the assumption there are unique strong solutions X^1 and X^2 for a given Brownian motion B. Apply Tanaka 's formula to the difference $Y_t = X_t^2 - X_t^1$ we have

$$(X_t^2 - X_t^1)^+ = \int_0^t 1_{(0,\infty)}(X_s^2 - X_s^1) d\left(X_s^2 - X_s^1\right) + L_t^0$$

as $(X_0^2 - X_0^1)^+ = 0$, where L_t^a is the local time of $X^2 - X^1$ at a, so L_t^a is increasing with $L_0^a = 0$. Taking expectations both side to obtain that

$$\mathbb{E}(X_t^2 - X_t^1)^+ = \mathbb{E}(Z_t) + \mathbb{E}(L_t^0)$$

for all $t \geq 0$, where

$$Z_t = \int_0^t 1_{(0,\infty)}(X_s^2 - X_s^1) d\left(X_s^2 - X_s^1\right)$$

for simplicity. We first show that indeed $L_t^0 = 0$. In fact by the occupation formula we have

$$\int_0^t \frac{1}{Y_s} 1_{(0,\infty)}(Y_s) d\left\langle Y\right\rangle_s = 2\int_0^\infty \frac{1}{a} L_t^a da.$$

Since

$$\begin{aligned}\int_0^t \frac{1}{Y_s} 1_{(0,\infty)}(Y_s) d\left\langle Y\right\rangle_s &= \int_0^t \frac{1}{Y_s}\left(\sigma(X_s^2) - \sigma(X_s^1)\right)^2 1_{(0,\infty)}(Y_s) ds \\ &\leq C\int_0^t Y_s 1_{(0,\infty)}(Y_s) ds < \infty\end{aligned}$$

where C is the Lipschits constant of σ, so that

$$\int_0^\infty \frac{1}{a} L_t^a da < \infty.$$

Since $a \to L_t^a$ is right continuous, so that $L_t^0 = 0$ for all $t \geq 0$ almost surely. Therefore $\mathbb{E}(X_t^2 - X_t^1)^+ = \mathbb{E}(Z_t)$.

By definition

$$\begin{aligned}
Z_t &= \int_0^t 1_{(0,\infty)}(X_s^2 - X_s^1)\left(\sigma(X_s^2) - \sigma(X_s^1)\right) dB_s \\
&+ \int_0^t 1_{(0,\infty)}(X_s^2 - X_s^1)\left(b^2(X_s^2) - b^1(X_s^1)\right) ds \\
&\leq \int_0^t 1_{(0,\infty)}(X_s^2 - X_s^1)\left(\sigma(X_s^2) - \sigma(X_s^1)\right) dB_s \\
&+ \int_0^t 1_{(0,\infty)}(X_s^2 - X_s^1)\left(b^1(X_s^2) - b^1(X_s^1)\right) ds \\
&\leq \int_0^t 1_{(0,\infty)}(X_s^2 - X_s^1)\left(\sigma(X_s^2) - \sigma(X_s^1)\right) dB_s \\
&+ C\int_0^t 1_{(0,\infty)}(X_s^2 - X_s^1)|X_s^2 - X_s^1| ds \\
&\leq \int_0^t 1_{(0,\infty)}(X_s^2 - X_s^1)\left(\sigma(X_s^2) - \sigma(X_s^1)\right) dB_s \\
&+ C\int_0^t (X_s^2 - X_s^1)^+ ds
\end{aligned}$$

where C is the Lipschitz constant, which yields that

$$\mathbb{E}(X_t^2 - X_t^1)^+ = \mathbb{E}(Z_t) \leq C\int_0^t \mathbb{E}(X_s^2 - X_s^1)^+ ds$$

for all $t \geq 0$. Therefore, by the Gronwall inequality, we may conclude that

$$\mathbb{E}(X_t^2 - X_t^1)^+ = 0$$

so that $X_t^2 \leq X_t^1$ for all t almost surely. ∎

5.2 The Skorohod equation

It is possible to represent the local time L_t^a of a continuous semimartingale X as a functional of X. Skorohod has established an explicit formula for L_t^a in terms of the running maximum of $-X$, and his formula is completely deterministic without any stochastic content. It is a theorem which should be an exercise in Prelim Analysis!

Theorem 5.2.1 *(Skorohod's equation) Suppose $y \in C[0,\infty)$ is a continuous path in $\mathbb{R}$ with initial value $y(0) \geq 0$. Let*

$$k(t) = \max\left\{0, \max_{0\leq s\leq t}(-y(s))\right\} \qquad \text{for } t \geq 0. \tag{5.16}$$

Then k is the unique continuous and increasing function on $[0,\infty)$ with initial zero, such that

$$x(t) = y(t) + k(t) \geq 0 \qquad \text{for } t \geq 0$$

and $t \to k(t)$ increases only on $\{t : x(t) = 0\}$, that is, $\int_0^\infty 1_{\{x(t)\neq 0\}} dk(t) = 0$.

Proof. Let us follow the proof in Karatzas-Shreve, page 210. Suppose $y \in C[0,\infty)$ is a continuous path with $y(0) \geq 0$. Define a path k by (5.16). Since y is a continuous path so is k, $t \to k(t)$ is increasing by definition. Since $-y(0) \leq 0$, so that $k(0) = 0$. Moreover

$$x(t) = y(t) + k(t) = \max\left\{y(t), \max_{0\leq s\leq t}(y(t) - y(s))\right\} \geq 0.$$

We next show that k increases only on $I = \{t : x(t) = 0\}$. To this end we only need to show that k does not increase on $I_\varepsilon = \{t \geq 0 : x(t) > \varepsilon\}$ for every $\varepsilon > 0$. Since x is continuous, I_ε is open, so that I_ε is countable union of disjoint open intervals (a_i, b_i) where $b_i > a_i$ (where $i \in \Lambda$ an index set which may be empty). For every $s \in [a_i, b_i]$

$$-y(s) = k(s) - x(s) \leq k(b_i) - \varepsilon$$

so that by definition

$$\begin{aligned} k(b_i) &= \max\left\{0, \max_{0\leq s\leq b_i}(-y(s))\right\} \\ &= \max\left\{k(a_i), \max_{a_i<s\leq b_i}(-y(s))\right\} \\ &\leq \max\{k(a_i), k(b_i) - \varepsilon\}. \end{aligned}$$

Since k is increasing, $k(a_i) \leq k(b_i)$, so we must have

$$k(a_i) = k(b_i)$$

for all i. Therefore k is constant on I_ε for every $\varepsilon > 0$, hence k must be flat on $\mathbb{R} \setminus I = \{t \geq 0 : x(t) > 0\}$.

Suppose there are two continuous increasing functions k_1 and k_2 with initial value 0 at $t = 0$, such that

$$x_i(t) = y(t) + k_i(t) \geq 0 \quad \text{for all } t \geq 0$$

and k_i increases only on $\{t \geq 0 : x_i(t) = 0\}$, where $i = 1, 2$.

By definition $x_i(0) = y(0)$. Suppose there is $b > 0$ such that $x_1(b) > x_2(b)$. Let

$$a = \sup\{t \leq b : x_1(t) = x_2(t)\}.$$

Then $0 \leq a < b$ by the continuity of x_i. Hence $x_1(s) > x_2(s) \geq 0$ by definition of a, for all $s \in (a, b]$. Since k_1 increases only on $\{t \geq 0 : x_1(t) = 0\}$, so that

$$k_1(a) = k_1(b)$$

and therefore

$$\begin{aligned} 0 &< x_1(b) - x_2(b) = k_1(b) - k_2(b) \\ &= k_1(a) - k_2(b) \leq k_1(a) - k_2(a) \\ &= x_1(a) - x_2(a) = 0 \end{aligned}$$

which is a contradiction. Hence $x_1(t) \leq x_2(t)$ for all $t \geq 0$. By symmetry we also have $x_1(t) \geq x_2(t)$ for all $t \geq 0$. The uniqueness now follows immediately. ■

If X is a continuous semimartingale, then according to Tanaka's formula, for every $a \in \mathbb{R}$

$$|X_t - a| = |X_0 - a| + \int_0^t \operatorname{sgn}(X_s - a) dX_s + 2L_t^a$$

where $t \to L_t^a$ is continuous, initial zero, and increasing only on $\{t : X_t = a\}$, which is equivalent to that L_t^a increases only on $\{t \geq 0 : |X_t - a| = 0\}$, hence, according to Skorohod's equation, we have the following corollary.

Corollary 5.2.2 *If X is a continuous semimartingale, $a \in \mathbb{R}$. Then its local time at a is given by*

$$L_t^a = \frac{1}{2} \max\left\{0, \max_{0 \leq s \leq t}\left[-\int_0^s sgn(X_r - a) dX_r\right] - |X_0 - a|\right\}. \tag{5.17}$$

We end this section to propose the following question which is mentally challenging. Skorohod's equation indeed gives rise to a mapping which sends a continuous path $y(t)$ in $\mathbb{R}$ with initial $y(0) \geq 0$ (with running time $t \in [0, \infty)$) to a continuous and increasing path $k(t)$ with $k(0) = 0$ so that the path $x(t) = y(t) + k(t)$ is a continuous path in $\mathbb{R}^+$, and k increase only on $\{t \geq 0 : x(t) = 0\}$. Are there similar mappings for continuous paths in $\mathbb{R}^d$? There are of course trivial mappings which essentially one dimensional, are there true multi-dimensional versions of Skorohod's equation?

5.3 Local time for Brownian motion

According to Wiener, for every $z \in \mathbb{R}$, there is a unique probability measure $\mathbb{P}^z$ on the Borel σ-algebra over $\Omega = C[0, \infty)$ the path space of all continuous paths in $\mathbb{R}$, such that the coordinate process $\{x_t : t \geq 0\}$ is Brownian motion started at z: $\mathbb{P}^z\{x_0 = z\} = 1$. The theory established in the previous section may be applied to one-dimensional Brownian motion. Therefore there is a random field $\{L_t^a : t \geq 0, a \in \mathbb{R}\}$ jointly Hölder continuous in (t, a), such that

$$|x_t - a| = |z - a| + \int_0^t \operatorname{sgn}(x_s - a) dx_s + 2L_t^a \quad \mathbb{P}^z\text{-a.s.} \tag{5.18}$$

for any z.

According to Skorohod's equation

$$2L_t^a = \max\left\{\max_{0 \leq s \leq t} \beta_s - |z - a|, 0\right\} \quad \mathbb{P}^z\text{-a.s.} \tag{5.19}$$

where

$$\beta_t = -\int_0^t \operatorname{sgn}(x_s - a) dx_s \tag{5.20}$$

is Brownian motion started at 0 under $\mathbb{P}^z$, and in terms of $\{\beta_t : t \geq 0\}$ defined by (5.20), we may rewrite Tanaka's formula

$$|x_t - a| = \max\left\{\max_{0\leq s\leq t} \beta_s, |z-a|\right\} - \beta_t \quad \mathbb{P}^z\text{-a.s.} \tag{5.21}$$

Therefore

Theorem 5.3.1 *(P. Lévy, 1948) Let $z, a \in \mathbb{R}$. Let $W = (W_t)_{t\geq 0}$ be a Brownian motion on $(\Omega, \mathcal{F}, \mathbb{P})$ started at $z \in \mathbb{R}$, and $(L_t^a)_{t\geq 0}$ be the local time of W at $a \in \mathbb{R}$. Then $(2L_t^a, |W_t - a|)_{t\geq 0}$ on $(\Omega, \mathcal{F}, \mathbb{P})$ has the same distribution as that of*

$$\left((\max_{0\leq s\leq t} \beta_s - |z-a|)^+, |z-a| \vee \max_{0\leq s\leq t} \beta_s - \beta_t\right)_{t\geq 0} \tag{5.22}$$

where $\{\beta_t : t \geq 0\}$ is Brownian motion starting at 0.

The previous conclusion can be stated in terms of a Brownian motion $\{\beta_t : t \geq 0\}$ is Brownian motion starting at 0 in $\mathbb{R}$.

Theorem 5.3.2 *(P. Lévy) Let $\{\beta_t : t \geq 0\}$ be a Brownian motion starting at 0 in $\mathbb{R}$. For any $a \in \mathbb{R}$, and L_t^a be the local time of β at a, that is*

$$2L_t^a = |\beta_t - a| - |\beta_0 - a| - \int_0^t sgn\,(\beta_s - a)\, d\beta_s.$$

Then $(2L_t^a, |\beta_t - a|)_{t\geq 0}$ and

$$\left((\max_{0\leq s\leq t} \beta_s - |a|)^+, |a| \vee \max_{0\leq s\leq t} \beta_s - \beta_t\right)_{t\geq 0} \tag{5.23}$$

have the same distribution.

In particular, if $\{\beta_t : t \geq 0\}$ is Brownian motion starting at 0, and $L_t = L_t^0$ is the local time of $\{\beta_t : t \geq 0\}$ at 0, then the pair of processes $\{(2L_t, |\beta_t|) : t \geq 0\}$ has the same distribution as that of

$$\left(\max_{0\leq s\leq t} \beta_s, \max_{0\leq s\leq t} \beta_s - \beta_t\right)_{t\geq 0}$$

which was discovered by P. Lévy in 1948.

Next we would like to look at the distribution of the local time L_t^a.

The central problem is that whether one is able to describe the distribution of the random field $\{(L_t^a, \beta_t) : t \geq 0, a \in \mathbb{R}\}$?

Firstly, for fixed $t > 0$ and $z \in \mathbb{R}$, the joint distribution $(L_t^z, |\beta_t - z|)$ can be determined according to Lévy's theorem. Indeed, by the reflection principle,

$$\mathbb{P}\left\{\beta_t \in da, \max_{0\leq s\leq t} \beta_s \in db\right\} = \frac{2(2b-a)}{\sqrt{2\pi t^3}} e^{-\frac{(2b-a)^2}{2t}} \tag{5.24}$$

on $\{b \geq 0, a \leq b\}$, from which we can work out the distribution of L_t^z for fixed $t > 0$ and $z \in \mathbb{R}$.

Corollary 5.3.3 *Same assumptions as in Theorem 5.3.2. Let $t > 0$ and $z \in \mathbb{R}$.*

1) The law of $(L_t^0, |\beta_t|)$ has a PDF given by

$$p(x, y) = \frac{4(2x + y)}{\sqrt{2\pi t^3}} e^{-\frac{(2x+y)^2}{2t}}; \quad x \geq 0, \ y \geq 0. \tag{5.25}$$

2) If $z \neq 0$, $(L_t^z, |\beta_t - z|)$ has a distribution in $\mathbb{R}^2$ with a PDF given by

$$\begin{aligned} & \mathbb{P}\left[L_t^z \in dx, |\beta_t - z| \in dy\right] \\ = \ & \frac{1}{\sqrt{2\pi t}} \left(e^{-\frac{(y-|z|)^2}{2t}} - e^{-\frac{(y+|z|)^2}{2t}} \right) 1_{\{y \geq 0\}} dy \delta_0(dx) \\ & + \frac{4}{\sqrt{2\pi t^3}} (2x + y + |z|) e^{-\frac{(2x+y+|z|)^2}{2t}} 1_{\{x \geq 0, y \geq 0\}} dy dx. \end{aligned} \tag{5.26}$$

Proof. Let $\{\beta_t : t \geq 0\}$ is Brownian motion starting at 0 in $\mathbb{R}$ on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$. Let $\beta_t^M = \max_{0 \leq s \leq t} \beta_s$ is the running maximum of the Brownian motion. Then, according to Lévy's theorem, $(2L_t^z, |\beta_t - z|)_{t \geq 0}$ has the same distribution as that of $\big((\beta_t^M - |z|)^+, |z| \vee \beta_t^M - \beta_t\big)_{t \geq 0}$. In particular, for a fixed $t > 0$

$$\mathbb{E} f(L_t^0, |\beta_t|) = \mathbb{E} f(\frac{1}{2}\beta_t^M, \beta_t^M - \beta_t),$$

together with (5.24), we have

$$\mathbb{E} f(L_t^0, |\beta_t|) = \frac{2}{\sqrt{2\pi t^3}} \iint_{b \geq 0, a \leq b} f(\frac{1}{2}b, b - a)(2b - a) e^{-\frac{(2b-a)^2}{2t}} da db.$$

Making change of variables: $x = \frac{1}{2}b$ and $y = b - a$ we obtain

$$\mathbb{E} f(L_t^0, |\beta_t|) = \frac{4}{\sqrt{2\pi t^3}} \iint_{x \geq 0, y \geq 0} f(x, y)(2x + y) e^{-\frac{(2x+y)^2}{2t}} dx dy.$$

In the case that $|z| > 0$, one has

$$\begin{aligned} & \mathbb{E} f(L_t^z, |\beta_t - z|) \\ = \ & \mathbb{E} f(\frac{1}{2}(\beta_t^M - |z|)^+, |z| \vee \beta_t^M - \beta_t) \\ = \ & \frac{2}{\sqrt{2\pi t^3}} \iint_{b \geq 0, a \leq b} f(\frac{1}{2}(b - |z|)^+, |z| \vee b - a)(2b - a) e^{-\frac{(2b-a)^2}{2t}} da db \\ = \ & \frac{2}{\sqrt{2\pi t^3}} \iint_{|z| \geq b \geq 0, a \leq b} f(0, |z| - a)(2b - a) e^{-\frac{(2b-a)^2}{2t}} da db \\ & + \frac{2}{\sqrt{2\pi t^3}} \iint_{b \geq |z|, a \leq b} f(\frac{1}{2}(b - |z|), b - a)(2b - a) e^{-\frac{(2b-a)^2}{2t}} da db, \end{aligned}$$

where the first integral

$$\begin{aligned}
&\iint_{|z|\geq b\geq 0, a\leq b} f(0,|z|-a)(2b-a)e^{-\frac{(2b-a)^2}{2t}}dadb \\
= & \int_{-\infty}^{0}\int_{0}^{|z|} f(0,|z|-a)(2b-a)e^{-\frac{(2b-a)^2}{2t}}dbda \\
&+\int_{0}^{|z|}\int_{a}^{|z|} f(0,|z|-a)(2b-a)e^{-\frac{(2b-a)^2}{2t}}dbda \\
= & \frac{t}{2}\int_{-\infty}^{|z|} f(0,|z|-a)\left(e^{-\frac{a^2}{2t}}-e^{-\frac{(2|z|-a)^2}{2t}}\right)da \\
= & \frac{t}{2}\int_{0}^{\infty} f(0,y)\left(e^{-\frac{(|z|-y)^2}{2t}}-e^{-\frac{(|z|+y)^2}{2t}}\right)dy,
\end{aligned}$$

and in the second integral we make change of variables: $x=\frac{1}{2}(b-|z|)$ and $y=b-a$, so that

$$\begin{aligned}
&\iint_{b\geq |z|, a\leq b} f(\frac{1}{2}(b-|z|), b-a)(2b-a)e^{-\frac{(2b-a)^2}{2t}}dadb \\
= & 2\iint_{x\geq 0, y\geq 0} f(x,y)(2x+y+|z|)e^{-\frac{(2x+y+|z|)^2}{2t}}dxdy
\end{aligned}$$

Putting together we have

$$\begin{aligned}
&\mathbb{E}f(L_t^z, |\beta_t - z|) \\
= & \frac{1}{\sqrt{2\pi t}}\int_{0}^{\infty} f(0,y)\left(e^{-\frac{(y-|z|)^2}{2t}}-e^{-\frac{(y+|z|)^2}{2t}}\right)dy \\
&+\frac{4}{\sqrt{2\pi t^3}}\iint_{x\geq 0, y\geq 0} f(x,y)(2x+y+|z|)e^{-\frac{(2x+y+|z|)^2}{2t}}dxdy.
\end{aligned}$$

This completes the proof. ∎

Chapter 6

Appendix: Convex functions

In this chapter we collect some results about monotonic functions and convex functions.

6.1 Monotonic functions, Lebesgue-Stieltjes measures

A few elementary properties about about monotonic functions may be found in W. Rudin's Principles, page 95. If $g : (a,b) \to \mathbb{R}$ is an increasing (or called non-decreasing) function, then the left and right limits $g(t-)$ and $g(t+)$ of g at any $t \in (a,b)$ exist. In fact, $g(t-) = \lim_{s\uparrow t} g(s) = \sup_{s<t} g(s)$, and similarly $g(t+) = \lim_{s\downarrow t} g(s) = \inf_{s>t} g(s)$. Clearly $g(t-) \leq g(t+)$. g is continuous at $t \in (a,b)$ if and only if $g(t+) = g(t-)$. The difference $g(t+) - g(t-)$ is called the jump of g at $t \in (a,b)$. Define $g_-(t) = g(t-)$ and $g_+(t) = g(t+)$. Then g_- and g_+ are increasing on (a,b) as well, and $g_- \leq g \leq g_+$. Moreover $g_- = g = g_+$ except at most countable many points in (a,b), thus $g_- = g = g_+$ almost everywhere on (a,b) with respect to the Lebesgue measure. g_- is left-continuous and g_+ is right-continuous on (a,b). We call that g_- is the left-continuous modification of g, and that g_+ is the right-continuous modification of g.

Let g be a *right-continuous* increasing function on (a,b) with values in $\mathbb{R}$. Then $g(a) = g(a+) = \lim_{t\downarrow a} g(t)$ and $g(b) = g(b-) = \lim_{t\uparrow b} g(t)$ exist, $g(a) \in [-\infty, \infty)$ and $g(b) \in (-\infty, \infty]$. The following convention is used: $-(-\infty) = \infty$, $-\infty < t < \infty$ for every $t \in \mathbb{R}$ and $t + \infty = \infty$. Therefore, g is naturally extended to an increasing function on $[a,b]$ taking values in $[-\infty, \infty]$, and g is finite on (a,b).

The Lebesgue-Stieltjes measure μ_g (also denoted by $g(dt)$) on (a,b) may be constructed as the following. First of all, define the length of $(s,t] \subseteq (a,b)$ by $\mu_g((s,t]) = g(t) - g(s)$. Then we construct the outer measure μ_g^* by

$$\mu_g^*(A) = \inf\left\{ \sum_{i=1}^{\infty} \mu_g((s_i,t_i]) : \cup_{i=1}^{\infty}(s_i,t_i] \supseteq A \right\}$$

where the inf runs over all possible countable coverings of A via intervals of form $(s,t] \subset A$. μ_g^* is an outer measure on (a,b):

1) $\mu_g^*(A) \geq 0$ for every $A \subseteq (a,b)$,

2) $\mu_g^*(A) \leq \mu_g^*(B)$ if $A \subset B \subset (a,b)$, and

3) μ_g^* is countably sub-additive, that is,

$$\mu_g^*\left(\cup_{i=1}^{\infty} A_i\right) \leq \sum_{i=1}^{\infty} \mu_g^*(A_i) \text{ for any } A_i \subset (a,b).$$

The Caratheodory extension theorem allows us to construct the Lebesgue-Stieltjes measure μ_g out off the outer measure μ_g^*. Namely, we say a subset $A \subset (a,b)$ is μ_g-measurable if the Caratheodory condition holds:

$$\mu_g^*(F) = \mu_g^*(F \cap A) + \mu_g^*(F \cap A^c)$$

for every subset $F \subset (a,b)$. The totality of all μ_g-measurable subsets of (a,b) is denoted by $\mathcal{M}_g$. Caratheodory's extension theorem says that

1) $\mathcal{M}_g$ is a σ-algebra on (a,b). Any subset in $\mathcal{M}_g$ is called μ_g-measurable, or simply a measurable set if g is specified.

2) The outer measure μ_g^* restricted on $\mathcal{M}_g$ is a measure, the restriction of μ_g^* is denoted by μ_g, dg or by $g(dt)$ if no confusion may arise.

3) $((a,b),\mathcal{M}_g,\mu_g)$ is complete, that is, every μ_g^*-null set belongs to $\mathcal{M}_g$.

4) Every Borel subset of (a,b) is μ_g-measurable, that is, $\mathcal{B}(a,b) \subset \mathcal{M}_g$, and finally

5) If $(s,t] \subset (a,b)$, then $\mu_g^*((s,t]) = g(t) - g(s)$.

In particular, if $a < t < b$, then $\{t\} = [t,t]$ is measurable and

$$\mu_g(\{t\}) = \lim_{n\to\infty} \mu_g\left((t - \frac{1}{n}, t]\right) = g(t) - \lim_{n\to\infty} g\left(t - \frac{1}{n}\right) = g(t) - g(t-)$$

which is the jump of g at t as g is right continuous. It follows that, if $[s,t] \subset (a,b)$, then

$$\mu_g([s,t]) = \mu_g(\{s\}) + \mu_g((s,t]) = g(s) - g(s-) + g(t) - g(s) = g(t) - g(s-),$$

and similarly

$$\mu_g([s,t)) = g(t-) - g(s-).$$

If g is increasing function, but not necessary right-continuous on (a,b), then we define μ_g to be the Lebesgue-Stieltjes measure generated by its right-continuous modification g_+. That is, $\mu_g = \mu_{g_+}$. Thus

$$\mu_g((s,t]) = g(t+) - g(s+) \qquad \forall (s,t] \subset (a,b) \tag{6.1}$$

and

$$\mu_g(\{t\}) = g(t+) - g(t-) \tag{6.2}$$

is the jump of g at t.

Exercise 6.1.1 *Let $g : (a,b) \to \mathbb{R}$ be increasing and right-continuous. For $\varepsilon > 0$ set $g_\varepsilon(t) = \frac{g(t+\varepsilon)-g(t)}{\varepsilon}$ if $t, t+\varepsilon \in (a,b)$ otherwise $g_\varepsilon(t) = 0$. Though g_ε may not be increasing, but it is non-negative Borel measurable function, thus we can construct the measure $\mu_\varepsilon(dt) = g_\varepsilon(t)dt$ on (a,b), then*

$$\int_{(a,b)} \varphi d\mu_\varepsilon \to \int_{(a,b)} \varphi d\mu_g$$

as $\varepsilon \downarrow 0$ for any $\varphi \in C_0^\infty(a,b)$.

One of the famous theorems of Lebesgue is that, if g is increasing on (a,b), then g has finite derivative almost everywhere on (a,b), denoted by g' (see for example Theorem 17.12, of E. Hewitt and K. Stromberg, page 264, or F. Riesz and B. Sz-Nagy: Leçons d'Analyse Fonctionnelle, page 5). The measure $m(dt) = g'(t)dt$ in general does not coincide with the Lebesgue-Stieltjes measure μ_g. This is obvious in the case that g is discontinuous at some point $t \in (a,b)$ as then $\mu_g(\{t\}) = g(t+) - g(t-) \neq 0$ and thus μ_g is not absolutely continuous with respect to the Lebesgue measure.

Next we discuss several versions of Lebesgue-Stieltjes measures, which are most useful forms in the study of stochastic processes.

If g is defined on $[a,b)$ or $[a,b]$, then the natural Lebesgue-Stieltjes measure associated g should be

$$m(dt) = (g(a+) - g(a))\delta_{\{a\}} + 1_{(a,b)}\mu_g$$

on $[a,b)$, and

$$m(dt) = (g(a+) - g(a))\delta_{\{a\}} + 1_{(a,b)}\mu_g + (g(b) - g(b-))\delta_{\{b\}}$$

for $[a,b]$. The drawback for this convention lies in that the restriction of μ_g on $[a_1,b_1] \subset (a,b)$ may not coincide with the measure m on $[a_1,b_1]$. There is no better way to formulate a unified way to define Lebesgue-Stieltjes measures on an interval including at least end point. The best we can do is to specify the measure at the end point(s) in each concrete case.

If $g : [0,b) \to [0,\infty)$ is increasing. Then g is increasing on $(0,b)$, it defines the Lebesgue-Stieltjes measure μ_g on $(0,b)$. We define a measure, denoted by m_g on $[0,b)$ by

$$m_g(A) = g(0+)\delta_{\{0\}}(A \cap \{0\}) + \mu_g(A \cap (0,b)),$$

which is a measure on the σ-algebra of all subset A of $[0,b)$ such that $A \cap (0,b) \in \mathcal{M}_g$. This σ-algebra is still denoted by $\mathcal{M}_g$, if no confusion may arise. Thus, by definition

$$\int_{[0,b)} f(t)m_g(dt) = f(0)g(0+) + \int_{(0,b)} f(t)g(dt).$$

In literature, m_g is also denoted by $g(dt)$ (as a measure on $[0,b)$), and the equality above becomes

$$\int_{[0,b)} f(t)g(dt) = f(0)g(0+) + \int_{(0,b)} f(t)g(dt).$$

Unfortunately, both integrals $\int_{[0,b)}$ and $\int_{(0,b)}$ could be written as $\int_0^b$, thus confusion may be produced. Therefore, we will, as long as is there is a risk for confusion, write Lebesgue integrals as $\int_E$, rather than using lower and upper limits. However, if $g(0+) = 0$, or $g(0) = 0$ and g is right continuous, then the contribution of m_g from $\{0\}$ vanishes, in this case both integrals $\int_{[0,b)} f dg$ and $\int_{(0,b)} f dg$ coincide, and $\int_0^b f(t)g(dt)$ can be used without ambiguous.

6.2 Right continuous inverse

Let $g : [0, b) \to [0, \infty)$ is right continuous and increasing. If $b < \infty$ then we set $g(t) = \lim_{s\uparrow b} g(s) = \sup g$ (which may be ∞) for $t \geq b$, thus g is extended to a right continuous increasing function on $[0, \infty)$ taking values in $[0, \infty]$. The right-continuous inverse of g, denoted by g^{-1} is defined by

$$g^{-1}(t) = \inf\{s \geq 0 : g(s) > t\} \qquad \text{for } t \geq 0$$

then $g^{-1} : [0, \infty) \to [0, \infty]$ is *right-continuous* and increasing. Moreover, $g^{-1}(t) < \infty$ if and only if $t < \lim_{s\uparrow b} g(s)$. In particular, if $\lim_{s\uparrow b} g(s) = \infty$ then g^{-1} is finite. For $t > 0$ the left limit

$$\begin{aligned} g_-^{-1}(t) &= \lim_{s<t, s\uparrow t} g^{-1}(s) = \inf\{s \geq 0 : g(s) \geq t\} \\ &= \sup\{s \geq 0 : g(s) < t\} \end{aligned}$$

and set $g_-^{-1}(0) = 0$.

Lemma 6.2.1 *g and g^{-1} as above. Then*
1) For any $t \geq 0$, $g_-(g_-^{-1}(t)) \leq g_-(g^{-1}(t)) \leq t$, and
2) For $t < \lim_{s\uparrow b} g(s)$, $g(g^{-1}(t)) \geq g(g_-^{-1}(t)) \geq t$.
3) If g is a continuous *then $g(g^{-1}(t)) = g(g_-^{-1}(t)) = t$ for any $t < \lim_{s\uparrow b} g(s)$.*
4) g is the right-continuous inverse of g^{-1} so that

$$g(t) = \inf\{s \geq 0 : g^{-1}(s) > t\} \qquad \text{for } t \geq 0$$

and

$$g(t) = \sup\{s \geq 0 : g^{-1}(s) \leq t\} \qquad \text{for } t \geq 0.$$

These facts can be easily derived from definitions.

6.3 Convex functions

References:

1. G. H. Hardy, J. E. Littlewood and G. Pólya's classic "Inequalities' (which has been published by Cambridge University Press is and has been in print since 1934).

2. L. Hörmander: "Notions of Convexity", published by Birkhäuser in 1994. Birkhäuser has recently re-issued Hörmander's book in paperback and listed it as one of Birkhäuser's classics.

We however follow the conventional definition of convex functions which is slightly different from but equivalent to that of Hörmander's one.

Definition 6.3.1 *Let $f : (a, b) \to \mathbb{R}$ (where a or/and b may be infinity). f is called a convex function on (a, b) if*

$$f(\lambda_1 x_1 + \lambda_2 x_2) \leq \lambda_1 f(x_1) + \lambda_2 f(x_2) \tag{6.3}$$

for any x_1, $x_2 \in (a, b)$ and $\lambda_j \geq 0$ such that $\lambda_1 + \lambda_2 = 1$.

Remark 6.3.2 *1) Some authors allow convex functions taking values ∞, but, for our propose, we only concern with real valued functions.*

2) Some authors also define convex functions on closed (or half closed) intervals,which do bring some difference. For example, the continuity at the end points will be not guarranted.

The following proposition summarize some elementary properties of convex functions.

Proposition 6.3.3 *Let $f : (a, b) \to \mathbb{R}$ be convex on (a, b).*

1) For any $x_1, x_2 \in (a, b)$, and any $c \in \mathbb{R}$

$$\sup_{x \in [x_1, x_2]} (f(x) - cx) = \max \{f(x_1) - cx_1, f(x_2) - cx_2\} . \tag{6.4}$$

2) For any $x_j \in (a, b)$ and $\lambda_j \geq 0$ such that $\sum_j \lambda_j = 1$, then

$$f(\sum_j \lambda_j x_j) \leq \sum_j \lambda_j f(x_j) . \tag{6.5}$$

3) It holds that

$$\frac{f(x) - f(x_1)}{x - x_1} \leq \frac{f(x_2) - f(x)}{x_2 - x} \tag{6.6}$$

for any $x_1 < x < x_2$ such that $x_j \in (a, b)$.

4) If $x \in (a, b)$, then $h \to \frac{f(x+h)-f(x)}{h}$ for any $h \neq 0$ such that $x+h \in (a, b)$ is increasing.

5) f is (locally) Lipschitz continuous. That is for any closed interval $[x_1, x_2] \subset (a, b)$ there is a constant C depending only on $f(x_1)$ and $f(x_2)$ such that

$$|f(x) - f(y)| \leq C|x - y| \quad \forall x, y \in [x_1, x_2]. \tag{6.7}$$

The first item appears as the definition property in Hörmander's book, 2) is the generalization of our definition for convexity, and is a special case of Jensen's inequality (see below). 3) is the reformulation of (6.3). Item 4) follows 3) and is the most useful form to us. 5) comes a little bit over awarded.

Exercise 6.3.4 *If $f : (a, b) \to \mathbb{R}$ is convex, then both limits $\lim_{x \downarrow a} f(x)$ and $\lim_{x \uparrow b} f(x)$ exist [as reals or ∞ !]*

Exercise 6.3.5 *$f : (a, b) \to \mathbb{R}$ is convex, if and only if*

$$\begin{vmatrix} 1 & 1 & 1 \\ x_1 & x_2 & x_3 \\ f(x_1) & f(x_2) & f(x_3) \end{vmatrix} \geq 0$$

for any $x_i \in (a, b)$ such that $x_1 \leq x_2 \leq x_3$.

Exercise 6.3.6 *If $f : (a,b) \to \mathbb{R}$ is convex, and $t \in (a,b)$, then*

$$f(s) \geq \beta(s-t) + f(t)$$

for all $s \in (a,b)$, and for any β such that

$$\sup_{s<t} \frac{f(t)-f(s)}{t-s} \leq \beta \leq \inf_{s>t} \frac{f(s)-f(t)}{s-t}.$$

That is, the graph of f is above the line defined by the equation $s = \beta(s-t) + f(t)$ (and thus we call it a tangent line of f at t).

Proposition 6.3.7 *Let $f : (a,b) \to \mathbb{R}$ be convex. Then*

1) At any $t \in (a,b)$, the left derivative $f'_-(t)$ and right derivative $f'_+(t)$ exist. Moreover

$$f'_-(t) = \sup_{s<t} \frac{f(t)-f(s)}{t-s} \qquad \text{and} \qquad f'_+(t) = \inf_{s>t} \frac{f(s)-f(t)}{s-t}.$$

2) If $s, t \in (a,b)$ and $s < t$ then

$$f'_-(s) \leq f'_+(s) \leq \frac{f(t)-f(s)}{t-s} \leq f'_-(t) \leq f'_+(t). \tag{6.8}$$

In particular, both f'_- and f'_+ are increasing on (a,b).

3) f'_- is left-continuous and f'_+ is right-continuous on (a,b), respectively. Moreover for any $t \in (a,b)$

$$f'_-(t) = f'_-(t-) = \lim_{\varepsilon \downarrow 0} f'_+(t-\varepsilon) \tag{6.9}$$

and

$$f'_+(t) = f'_+(t+) = \lim_{\varepsilon \downarrow 0} f'_-(t+\varepsilon). \tag{6.10}$$

4) Let $[s,t] \subset (a,b)$. Then

$$\left| \frac{f(y)-f(x)}{y-x} \right| \leq \max\{|f'_+(s)|, |f'_-(t)|\}$$

for any $x, y \in (s,t)$ and $x \neq y$.

5) Let $t \in (a,b)$. Then f is differentiable at t if and only if f'_- is right continuous at t, i.e. $f'_-(t) = \lim_{\varepsilon\downarrow 0} f'_-(t+\varepsilon)$, which is also equivalent to that f'_+ is left continuous at t, i.e. $f'_+(t) = \lim_{\varepsilon\downarrow 0} f'_+(xt-\varepsilon)$. That is, f'_+ (resp. f'_-) is the right-continuous (resp. left-continuous) modification of f'_- (resp. f'_+). In particular, f'_+ and f'_- generate the same Lebesgue-Stieltjes measure on Borel sets of (a,b).

6) f is differentiable on (a,b) except at most countable many points.

All these conclusions follow easily from the item 4 in Proposition 6.3.3.

Next we turn to integral representations of convex functions. We need the following integration by parts.

Lemma 6.3.8 *Let $g : (a,b) \to \mathbb{R}$ be increasing, and $[x_1, x_2] \subset (a,b)$ be a bounded, closed interval. Suppose $\varphi \in C^1[x_1, x_2]$. Then*

$$\begin{aligned}\int_{(x_1,x_2)} \varphi(x)\mu_g(dx) &= \varphi(x_2)g(x_2-) - \varphi(x_1)g(x_1+) \\ &\quad - \int_{x_1}^{x_2} g(x)\varphi'(x)dx.\end{aligned}$$

Proof. Under the conditions, $\int_{(x_1,x_2)} \varphi(x)\mu_g(dx)$ exists as Riemann-Stieltjes' integral, and the integration by parts formula is well known, see Theorem 7.6, in T. M. Apostol: Mathematical Analysis, page 144. ■

If $I = [x_1, x_2]$ is a compact interval, then define the Green function

$$G_I(x,y) = \begin{cases} \frac{(y-x_2)(x-x_1)}{x_2-x_1} & \text{if } y \geq x, \\ \frac{(x-x_2)(y-x_1)}{x_2-x_1} & \text{if } y \leq x. \end{cases}$$

Note that $G \leq 0$, symmetric and continuous on $I \times I$.

Theorem 6.3.9 *Let $f : (a,b) \to \mathbb{R}$ be convex, and $I = [x_1, x_2] \subset (a,b)$ be a bounded closed interval. Then*

$$\begin{aligned}f(x) &= \frac{x-x_1}{x_2-x_1}f(x_2) - \frac{x-x_2}{x_2-x_1}f(x_1) \\ &\quad + \int_{(x_1,x_2)} G_I(x,y)\mu_{f'_+}(dy)\end{aligned} \tag{6.11}$$

for $x \in [x_1, x_2]$.

Proof. Let us consider the case $x \in (x_1, x_2)$. Let us compute the integral

$$\begin{aligned}J(x) &= \int_{(x_1,x_2)} G_I(x,y)\mu_{f'_+}(dy) \\ &= \frac{x-x_2}{x_2-x_1}\int_{(x_1,x)} (y-x_1)\mu_{f'_+}(dy) \\ &\quad + \frac{x-x_1}{x_2-x_1}\int_{(x,x_2)} (y-x_2)\mu_{f'_+}(dy) \\ &\quad + \frac{(x-x_2)(x-x_1)}{x_2-x_1}\left(f'_+(x) - f'_-(x)\right).\end{aligned}$$

The two integrals are Riemann-Stieltjes integrals which can be worked out by means of integration by parts

$$\begin{aligned}\int_{(x_1,x)} (y-x_1)\mu_{f'_+}(dy) &= (x-x_1)f'_-(x) - \int_{(x_1,x)} f'_+(y)dy \\ &= (x-x_1)f'_-(x) - f(x) + f(x_1)\end{aligned}$$

and

$$\int_{(x,x_2)} (y-x_2)\mu_{f'_+}(dy) = -(x-x_2)f'_+(x) - f(x_2) + f(x).$$

Therefore

$$J(x) = f(x) + \frac{x-x_2}{x_2-x_1}f(x_1) - \frac{x-x_1}{x_2-x_1}f(x_2)$$

which proves the representation. ■

The following representation is most useful in our study of local times.

Theorem 6.3.10 *Let $f:(a,b)\to\mathbb{R}$ be convex, and $[x_1,x_2]\subset(a,b)$ be any bounded, closed interval. Then*

$$f(x) = \alpha x + \beta + \frac{1}{2}\int_{(x_1,x_2)} |x-y|\mu_{f'_+}(dy)$$

for any $x\in[x_1,x_2]$, where

$$\alpha = \frac{f'_+(x_1)+f'_-(x_2)}{2}$$

and

$$\beta = \frac{1}{2}\left(f(x_1)+f(x_2)-x_1f'_+(x_1)-x_2f'_-(x_2)\right).$$

Proof. Let us consider the following integral

$$J(x) = \int_{(x_1,x_2)} |x-y|\mu_{f'_+}(dy)$$

where $[x_1,x_2]\subset(a,b)$ and $x\in[x_1,x_2]$. Let us consider the case that $x\in(x_1,x_2)$. Using integration by parts we have

$$\begin{aligned}
J(x) &= \int_{(x_1,x)} (x-y)\mu_{f'_+}(dy) + \int_{(x,x_2)} (y-x)\mu_{f'_+}(dy) \\
&= (x-y)f'_+(y)\Big|_{(x_1,x)} + (y-x)f'_+(y)\Big|_{(x,x_2)} \\
&\quad + \int_{(x_1,x)} f'_+(y)dy - \int_{(x,x_2)} f'_+(y)dy \\
&= -(x-x_1)f'_+(x_1) + (x_2-x)f'_-(x_2) \\
&\quad + f(x) - f(x_1) - f(x_2) + f(x_2)
\end{aligned}$$

which yields the conclusion. ■

Similarly we have

Theorem 6.3.11 *Let $f:(a,b)\to\mathbb{R}$ be convex, and $[x_1,x_2]\subset(a,b)$ be any bounded, closed interval. Let $h(x)=1$ for $x\geq 0$ and $h(x)=-1$ for $x<0$. Then*

$$f'_+(x) = \frac{1}{2}\int_{(x_1,x_2)} h(x-y)\mu_{f'_+}(dy) + \frac{f'_+(x_1)+f'_-(x_2)}{2}$$

for $x\in[x_1,x_2]$, and, if $\tilde{h}(x)=1$ for $x>0$ and $\tilde{h}(x)=-1$ for $x\leq 0$ then

$$f'_-(x) = \frac{1}{2}\int_{(x_1,x_2)} \tilde{h}(x-y)\mu_{f'_+}(dy) + \frac{f'_+(x_1)+f'_-(x_2)}{2}$$

for $x\in[x_1,x_2]$.

Proof. Again, compute the integral

$$J(x)=\int_{(x_1,x_2)}h(x-y)\mu_{f'_+}(dy)$$

for $x\in(x_1,x_2)$. Since $h(0)=1$ so that

$$\int_{\{x\}}h(x-y)\mu_{f'_+}(dy)=f'_+(x)-f'_-(x)$$

and integrating by parts one has

$$\int_{(x_1,x)}h(x-y)\mu_{f'_+}(dy)=f'_-(x)-f'_+(x_1)$$

and

$$\int_{(x,x_2)}h(x-y)\mu_{f'_+}(dy)=-f'_-(x_2)+f'_+(x)$$

so that

$$2f'_+(x)=J(x)+f'_+(x_1)+f'_-(x_2).$$

■

Our last topic is about approximations to convex functions by C^2-functions. Recall the standard approach. Suppose α is a C^∞-function with a compact support in $\mathbb{R}$, such that $\int_{\mathbb{R}}\alpha(x)dx=1$, and for each $\varepsilon>0$ we consider $\alpha_\varepsilon(x)=\frac{1}{\varepsilon}\alpha\left(x/\varepsilon\right)$. If u is a function on $\mathbb{R}$ which is locally integrable, then we set

$$\begin{aligned}u_\varepsilon(x) &= \int_{\mathbb{R}}u(x-y)\alpha_\varepsilon(y)dy\\ &= \int_{\mathbb{R}}u(y)\alpha_\varepsilon(x-y)dy.\end{aligned}$$

The right-hand side is the convolution of u and α_ε, denoted by $u*\alpha_\varepsilon$. For each $\varepsilon>0$, u_ε is a smooth function, and if u is continuous then $u_\varepsilon\to u$ uniformly on any compact set. If $u\in L^p(\mathbb{R})$ where $1\le p<\infty$, then $u_\varepsilon\to u$ in $L^p(\mathbb{R})$ (the conclusion is not true for $p=\infty$).

Suppose $f:\mathbb{R}\to\mathbb{R}$ is convex, then f must be continuous, its first derivative in distribution sense is f'_+ (and f'_-), and its second derivative in distribution sense is the Lebesgue-Stieltjes measure $\mu_{f'_+}$ generated by the increasing function f'_+. Therefore

$$\begin{aligned}f'_\varepsilon(x) &= \int_{\mathbb{R}}f(y)\frac{d}{dx}\alpha_\varepsilon(x-y)dy\\ &= -\int_{\mathbb{R}}f(y)\frac{d}{dy}\alpha_\varepsilon(x-y)dy \qquad (6.12)\\ &= \int_{\mathbb{R}}f'_-(y)\alpha_\varepsilon(x-y)dy \qquad (6.13)\end{aligned}$$

the last equation follows from the fact that f'_- is the distribution derivative of f. In particular $f'_\varepsilon=f'_-*\alpha_\varepsilon$ for each $\varepsilon>0$.

Similarly

$$\begin{aligned} f''_\varepsilon(x) &= \int_{\mathbb{R}} f(y)\alpha''_\varepsilon(x-y)dy \\ &= \int_{\mathbb{R}} \alpha_\varepsilon(x-y)\mu_{f'_+}(dy) \end{aligned} \tag{6.14}$$

which implies in particular that each f_ε is convex.

Lemma 6.3.12 *Suppose $\alpha \in C_0^\infty(\mathbb{R})$ has a compact support in $(0,\infty)$ such that $\int_{\mathbb{R}} \alpha(x)ds = 1$.*

1) If $f_\varepsilon = f * \alpha_\varepsilon$ then

$$f'_\varepsilon(x) \to f'_-(x) \quad \forall x \in \mathbb{R}.$$

2) If $f_\varepsilon = f * \tilde{\alpha}_\varepsilon$ where $\tilde{\alpha}(x) = \alpha(-x)$ (so its support is in $(-\infty, 0)$), then

$$f'_\varepsilon(x) \to f'_+(x) \quad \forall x \in \mathbb{R}.$$

Proof. Suppose α has a support in $(0,R)$ where $R > 0$ is a number. Then

$$\begin{aligned} f'_\varepsilon(x) - f'_-(x) &= \int_{\mathbb{R}} \left(f'_-(y) - f'_-(x)\right) \alpha_\varepsilon(x-y)dy \\ &= \int_{\mathbb{R}} \left(f'_-(x-y) - f'_-(x)\right) \alpha_\varepsilon(y)dy \\ &= \int_0^R \left(f'_-(x-y) - f'_-(x)\right) \frac{1}{\varepsilon}\alpha(y/\varepsilon)dy \\ &= \int_0^R \left(f'_-(x-\varepsilon z) - f'_-(x)\right) \alpha(z)dz. \end{aligned}$$

Since f'_- is left continuous, so that $f'_-(x-\varepsilon z) \to f'_-(x)$ uniformly in $z \in (0,R)$ as $\varepsilon \downarrow 0$. Therefore $f'_\varepsilon(x) - f'_-(x) \to 0$. ■

Clearly we can not expect that f''_ε converges to the derivative of f'_+, but the family of measures $f''_\varepsilon(x)dx$ does converges (in distribution sense).

Lemma 6.3.13 *Let $\alpha \in C_0^\infty(\mathbb{R})$ has a compact support such that $\int_{\mathbb{R}} \alpha(x)ds = 1$, and $f_\varepsilon = f * \alpha_\varepsilon$. Then*

$$\int_{\mathbb{R}} \varphi(x) f''_\varepsilon(x)dx \to \int_{\mathbb{R}} \varphi(x)\mu_{f'_+}(dx)$$

as $\varepsilon \downarrow 0$ for any $\varphi \in C_0^\infty(\mathbb{R})$.

Proof. By definition

$$\begin{aligned} \int_{\mathbb{R}} \varphi(x) f''_\varepsilon(x)dx &= \int_{\mathbb{R}} \left(\varphi(x)\int_{\mathbb{R}} \alpha_\varepsilon(x-y)\mu_{f'_+}(dy)\right)dx \\ &= \int_{\mathbb{R}} \left(\int_{\mathbb{R}} \varphi(x)\alpha_\varepsilon(x-y)dx\right)\mu_{f'_+}(dy) \\ &= \int_{\mathbb{R}} \varphi * \tilde{\alpha}_\varepsilon(y)\mu_{f'_+}(dy) \end{aligned}$$

where $\tilde{\alpha}(x) = \alpha(-x)$. Since φ has a compact support, so is $\varphi * \tilde{\alpha}_\varepsilon$ and therefore $\varphi * \tilde{\alpha}_\varepsilon \to \varphi$ uniformly. It follows that

$$\int_{\mathbb{R}} \varphi(x) f''_\varepsilon(x) dx \to \int_{\mathbb{R}} \varphi(x) \mu_{f'_+}(dx).$$

■