

UNIVERSIDADE FEDERAL DO RIO GRANDE
DO SUL

Notes on Probabilistic Analysis

Mark Thompson

Marcelo Bohrer

Porto Alegre, July 2013

Contents

0.1	General Notation	2
0.2	Introduction	3
1	Basic Definitions and Theorems	4
1.0.1	Conditional Expectation and Regular Conditional Probabilities . .	9
2	Stochastic Processes	12
2.0.2	Stopping Times	14
2.1	Martingales	16
2.1.1	Discrete Parameter Martingales	16
2.1.2	Continuous Parameter Martingales	20
2.1.3	The Doob-Meyer Decomposition	22
2.2	Brownian Motion	24
2.2.1	One Dimensional Brownian Motion	24
2.2.2	Brownian motion in higher dimension	34
3	Stochastic Integrals	40
3.1	The Itô Formula	49
4	Semi-groups and Generators	54
5	Stochastic Differential Equations and Diffusions	61
6	The Dirichlet Problem	68

0.1 General Notation

- \mathcal{B}, \mathcal{F} both denote σ -algebras;
- $a \wedge b$ denotes minimum of a and b ;
- $a \vee b$ denotes the maximum of a and b ;
- I_A and χ_A denote the indicator function of a set;
- \Re denotes the real part of a complex number;
- $\langle \cdot, \cdot \rangle$ denotes inner product;
- $\| \cdot \|$ denotes the (Hilbert) norm;
- $| \cdot |$ denotes absolute value;
- $V(\cdot, \cdot)$ denotes variance;
- $C(\cdot, \cdot)$ denotes covariance;
- $[\cdot]$ denotes both the integer part of a number and the quadratic variation. The meaning will be clear from the use.

Additional notation is mostly standard and anything out of the ordinary will be explained in the text.

0.2 Introduction

The first interest in diffusion process can be traced back to Robert Brown, a botanist who in 1827 observed that certain colloidal particles (by this is commonly understood particles of one component with diameters between 10^{-7} and 10^{-9} m, suspended in a continuous phase of another component) make irregular and apparently spontaneous movements. In fact his first observations were with pollen grains suspended in water but after making observations of other materials he thought that he had discovered active molecules in organic and inorganic bodies. However, it was only towards the end of the last century that the modern view, attributing the phenomenon to molecular agitation was put forward by Delsaux in 1877 and later by Gouy in 1888. The experimental observations subsequent to those of Brown established that finer particles move more rapidly; that the motion is stimulated by heat, and as was shown by Gouy, that the motion becomes more active with a decrease in the viscosity of the suspending fluid. The first theoretical and quantitative approach was given by Albert Einstein in 1905, the same year in which he published his theory of special relativity (see [4]).

Since Einstein the study of diffusion processes, or more generally, of random process has developed enormously. Applications have much surpassed the original realms of mathematics and Physics, finding place in a variety of fields such as finance, economics and life sciences.

These notes are an enlargement the first part of [14]. An introductory chapter on basic definitions and theorems of probability theory has been added in order to make the texts accessible to those with only a small amount of knowledge of probability theory. Additional topics have been added to facilitate the understanding of more advanced topics and to increase the range of possible applications. Most topics discussed here can be found on [9] and [7].

Chapter 1

Basic Definitions and Theorems

In this section we state a number of results needed to develop the theory. These results stem both from measure and probability theories and further relevant results will have indicated references. We also present some results that are not directly used in the theory that follows, but are themselves of interest and amount a great deal to a general comprehension of probabilistic analysis. More results can be obtained in [1] and [12].

The space of elementary outcomes is denoted by Ω . It is simply a non-empty set.

Definition 1.1. *A measure space is a pair (Ω, \mathcal{F}) , where Ω is the space of elementary outcomes and \mathcal{F} is a σ -algebra of subsets of Ω .*

Definition 1.2. *A measure P on a measure space (Ω, \mathcal{F}) is called a probability measure or a probability distribution if $P(\Omega) = 1$*

Definition 1.3. *A probability space is a triplet (Ω, \mathcal{F}, P) , where (Ω, \mathcal{F}) is a measure space and P is a probability measure. If $C \in \mathcal{F}$, then the number $P(C)$ is called the probability of C .*

The next definition characterizes random variables in the setting of measure theory.

Definition 1.4. *A measurable function defined on a probability space is called a Random Variable. That is, a function $X : \Omega \rightarrow \mathbb{R}$ is called a random variable if the event*

$$X^{-1}((-\infty, a]) \equiv \{\omega : X(\omega) \leq a\} \in \mathcal{F}$$

for each a in \mathbb{R} .

Proposition 1.1. *Let $(\Omega_i, \mathcal{F}_i), i = 1, 2$ be measure spaces and let $T : \Omega_1 \rightarrow \Omega_2$ be a $(\mathcal{F}_1, \mathcal{F}_2)$ -measurable mapping from Ω_1 to Ω_2 . Then, for any measure μ on $(\Omega_1, \mathcal{F}_1)$ the set function μT^{-1} , defined by*

$$\mu T^{-1}(A) \equiv \mu (T^{-1}(A)), \quad A \in \mathcal{F}_2 \quad (1.1)$$

is a measure on \mathcal{F}_2 .

This result is known from real analysis and is of fundamental importance in the next two definitions, which play a major role in the ideas we will develop later.

Definition 1.5. *The measure μT^{-1} is called the measure induced by T (or the induced measure of T) on \mathcal{F}_2 .*

Definition 1.6. *For a random variable X defined on a probability space (Ω, \mathcal{F}, P) , the probability distribution of X (or the law of X), denoted by P_X , is the induced measure of X under P on \mathbb{R} , as defined in (1.1). That is*

$$P_X(A) = P (X^{-1}(A)) \text{ for all } A \in \mathcal{B}(\mathbb{R})$$

Note that the same framework treats both discrete and continuous random variables.

Definition 1.7. *The cumulative distribution function (CDF) of a random variable X is defined as*

$$F_X(x) \equiv P_X ((-\infty, x])$$

These definitions can be easily generalized to higher dimensions.

Definition 1.8. *Let (Ω, \mathcal{F}, P) be a probability space, $k \in \mathbb{N}$ and $X : \Omega \rightarrow \mathbb{R}^k$ be $(\mathcal{F}, \mathcal{B}(\mathbb{R}^k))$ -measurable. Then X is called a (k -dimensional) random vector on (Ω, \mathcal{F}, P) .*

Definition 1.9. *Let X be a k -dimensional random vector on (Ω, \mathcal{F}, P) for some $k \in \mathbb{N}$.*

Let

$$F_X(x) \equiv P (\{\omega : X_1(\omega) \leq x_1, X_2(\omega) \leq x_2, \dots, X_k(\omega) \leq x_k\})$$

for $x = (x_1, x_2, \dots, x_k) \in \mathbb{R}^k$. Then $F_X(\cdot)$ is called the joint cumulative distribution function (joint CDF) of the random vector X .

Definition 1.10. *Let X be a k -dimensional random vector on (Ω, \mathcal{F}, P) for some $k \in \mathbb{N}$.*

Let

$$P_X(A) = P (X^{-1}(A)) \text{ for all } A \in \mathcal{B}(\mathbb{R}^k)$$

The probability measure P_X is called the joint probability distribution of X .

Definition 1.11. Let $X = (X_1, X_2, \dots, X_k)$ be a random vector on (Ω, \mathcal{F}, P) . Then, for each $i = 1, \dots, k$ the CDF F_{X_i} and the probability distribution P_{X_i} of the random variable X_i are called the marginal CDF and the marginal distribution of X_i respectively.

We now proceed to give a measure theoretic definition of the expected value of a random variable.

Definition 1.12. Let X be a random variable on (Ω, \mathcal{F}, P) . The expected value of X , denoted by $E(X)$ or EX , is defined as

$$E(X) = \int_{\Omega} X dP$$

provided the integral is well defined.

That is, the expectation of an arbitrary random variable is the Lebesgue integral of the \mathcal{F} -measurable function $X = X(\omega)$ with respect to the measure P . Note that this is a very natural definition, since in the discrete case the expected value is computed by a summation.

Just as in the case of an arbitrary measurable function, the integral may be defined through an increasing sequence of nonnegative random variables. It needs then to be shown that the limit (i.e, the expected value of the random variable) is independent of the choice of the approximating sequence. This approach can be seen in the general case of Lebesgue integration theory in [13]

We now list the main properties of the expectation $E(X)$, the proof of which are straightforward and can be found in [12]

- Let c be a constant and let $E(X)$ exist. Then $E(cX)$ exists and

$$E(cX) = cE(X)$$

- Let $X \leq Y$, then

$$E(X) \leq E(Y)$$

- If $E(X)$ exists then

$$|E(X)| \leq E|X|$$

- If $E(X)$ exists then $E(XI_A)$ exists for each $A \in \mathcal{F}$, if $E(X)$ is finite, $E(XI_A)$ is finite.

- If X and Y are nonnegative random variables such that $E|X| < \infty$ and $E|Y| < \infty$, then

$$E(X + Y) = E(X) + E(Y).$$

- If $X = 0$ a.s. then $E(X) = 0$.
- If $X = Y$ a.s. and $E|X| < \infty$, then $E|Y| < \infty$ and $E(X) = E(Y)$.
- Let $X \geq 0$ and $E(X) = 0$. Then $X = 0$.
- Let X and Y be such that $E|X| < \infty$, $E|Y| < \infty$ and $E(XI_A) \leq E(YI_A)$ for all $A \in \mathcal{F}$. Then $X \leq Y$ a.s.
- Let X be an extended random variable and $E|X| < \infty$. Then $|X| < \infty$.

We now state the three main convergence theorems in Lebesgue integration applied to the case of arbitrary random variables. The proofs are completely analogous to the general case and are omitted.

Theorem 1.1 (Monotone Convergence). *Let Y, X, X_1, \dots be random variables.*

- If $X_n \geq Y$ for all $n \geq 1$, $E(Y) > -\infty$, and $X_n \uparrow X$, then

$$E(X_n) \uparrow E(X).$$

- If $X_n \leq Y$ for all $n \geq 1$, $E(Y) < \infty$, and $X_n \downarrow X$, then

$$E(X_n) \downarrow E(X).$$

Corollary 1.1. *Let $\{Y_n\}_{n \geq 1}$ be a sequence of nonnegative random variables. Then*

$$E\left(\sum_{n=1}^{\infty} Y_n\right) = \sum_{n=1}^{\infty} EY_n.$$

Theorem 1.2 (Fatou's Lemma). *Let Y, X_1, X_2, \dots be random variables.*

1. If $X_n \geq Y$ for all $n \geq 1$ and $E(Y) > -\infty$, then

$$E(\liminf X_n) \leq \liminf E(X_n).$$

2. If $X_n \leq Y$ for all $n \geq 1$ and $E(Y) < \infty$, then

$$\limsup E(X_n) \leq E(\limsup X_n).$$

3. If $|X_n| \leq Y$ for all $n \geq 1$ and $E(Y) < \infty$, then

$$E(\liminf X_n) \leq \liminf E(X_n) \leq \limsup E(X_n) \leq E(\limsup X_n)$$

Theorem 1.3 (Lebesgue Dominated Convergence). *Let Y, X, X_1, X_2, \dots be random variables such that $|X_n| \leq Y$, $E(Y) < \infty$ and $X_n \rightarrow X$ a.s.. Then $E|X| < \infty$,*

$$E(X_n) \rightarrow E(X) \quad \text{and} \quad E|X_n - X| \rightarrow 0.$$

Definition 1.13. *Let (Ω, \mathcal{F}) and (E, \mathcal{E}) be measure spaces. We say that a function $X = X(\omega)$, defined on Ω and taking values in E , is \mathcal{F}/\mathcal{E} -measurable, or is a random element (with values in E), if*

$$\{\omega : X(\omega) \in B\} \in \mathcal{F}$$

for any $B \in \mathcal{E}$.

We now state and prove a result from Lebesgue theory of integration which is of great importance in the development of conditional probabilities.

Theorem 1.4 (Change of Variables in the Lebesgue Integral). *Let (Ω, \mathcal{F}) and (E, \mathcal{E}) be measure spaces and $X = X(\omega)$ an \mathcal{F}/\mathcal{E} -measurable function with values in E . Let P be a probability measure on (Ω, \mathcal{F}) and P_X the probability measure on (E, \mathcal{E}) induced by $X = X(\omega)$:*

$$P_X(A) = P\{\omega : X(\omega) \in A\}, \quad A \in \mathcal{E}. \quad (1.2)$$

Then

$$\int_A g(x) P_X dx = \int_{X^{-1}(A)} g(X(\omega)) P d\omega \quad A \in \mathcal{E} \quad (1.3)$$

for every \mathcal{E} -measurable function $g = g(x), x \in \mathcal{E}$.

Proof. Let $A \in \mathcal{E}$ and $g(x) = I_B(x)$, where $B \in \mathcal{E}$. Then (1.3) becomes

$$P_X(AB) = P(X^{-1}(A) \cap X^{-1}(B)) \quad (1.4)$$

which follows from (1.2) and the observation that $(X^{-1}(A) \cap X^{-1}(B)) = X^{-1}(A \cap B)$

It follows from (1.4) that (1.3) is valid for nonnegative simple function $g = g(x)$, and therefore, by the Monotone Convergence Theorem, also for all nonnegative \mathcal{E} -measurable functions.

In the general case we need only represent g as $g^+ - g^-$. Then, since (1.3) is valid for g^+ and g^- , if $\int_A g^+ P_X dx < \infty$, we also have

$$\int_{X^{-1}(A)} g^+(X(\omega)) P(d\omega) < \infty$$

and therefore the existence of $\int_A g(x) P_X dx$ implies the existence of $\int_{X^{-1}(A)} g(X(\omega)) P(d\omega)$ □

The next definition is also analogous to the classic analysis result.

Definition 1.14. *The sequence P_n converges weakly to the probability measure P if, for each $f \in C_b(X)$ ¹,*

$$\lim_{n \rightarrow \infty} \int_X f(x) dP_n(x) = \int_X f(x) dP(x)$$

The weak convergence is sometimes denoted as $P_n \Rightarrow P$.

1.0.1 Conditional Expectation and Regular Conditional Probabilities

Definition 1.15. *Two random variables X and Y are identified if $P[\omega; X(\omega) \neq Y(\omega)] = 0$.*

Let X be an integrable random variable and $\mathcal{C} \subset \mathcal{F}$ be a sub σ -field of \mathcal{F} . Then $\mu(B) = E(X : B) := \int_B X(\omega) P(dx)$, $B \in \mathcal{C}$, defines a σ -additive set function on \mathcal{C} with finite total variation and is clearly absolutely continuous with respect to $\nu = P|_{\mathcal{C}}$. The Radon-Nikodym derivative $\frac{d\mu}{d\nu}$ is denoted by $E(X|\mathcal{C})(\omega)$; thus $E(X|\mathcal{C})$ is the unique (up to identification) \mathcal{C} -measurable integrable random variable Y such that $E(Y : B) = E(X : B)$ for every $B \in \mathcal{C}$

Definition 1.16. *$E(X|\mathcal{C})(\omega)$ is called the conditional expectation of X given \mathcal{C} . It has the following properties:*

1. $E(ax + by|\mathcal{C}) = aE(X|\mathcal{C}) + bE(Y|\mathcal{C})$ *a.s.*
2. If $X \leq 0$ *a.s.* then $E(X|\mathcal{C}) \leq 0$ *a.s.*
3. $E(1|\mathcal{C}) \leq 0$ *a.s.*

¹ $C_b(X)$ denotes the space of bounded continuous functions on X

4. If X is \mathcal{C} -measurable, then $E(X|\mathcal{C}) = X$ a.s., more generally, if XY is integrable and X is \mathcal{C} -measurable

$$E(XY|\mathcal{C}) = XE(Y|\mathcal{C}) \quad a.s.$$

5. If \mathcal{H} is a sub σ -field of \mathcal{C} , then

$$E(E(X|\mathcal{C})|\mathcal{H}) = E(X|\mathcal{H})$$

6. If $X_n \rightarrow X$ in $\mathcal{L}_1(\Omega)$, then $E(X_n|\mathcal{C}) \rightarrow E(X|\mathcal{C})$ in $\mathcal{L}_1(\Omega)$.

7. (Jensen's inequality) If $\Psi : \mathbb{R}^1 \rightarrow \mathbb{R}^1$ is convex and $\Psi(X)$ is integrable, then

$$\Psi(E(X|\mathcal{C})) \leq E(\Psi(X)|\mathcal{C}) \quad a.s.$$

In particular, $|E(X|\mathcal{C})| \leq E(|X|\mathcal{C})$ and, if X is square integrable, $|E(X|\mathcal{C})|^2 \leq E(|X|^2|\mathcal{C})$.

8. X is independent of \mathcal{C} if and only if every Borel measurable function f such that $f(X)$ is integrable, $E(f(X)|\mathcal{C}) = E(f(X))$ a.s.

Let ξ be a mapping from Ω into a measurable space (S, \mathcal{B}) such that it is \mathcal{F}/\mathcal{B} -measurable. Then $\mu(B) = E(X : \{\omega; \xi(\omega) \in B\})$ is a σ -additive set function on \mathcal{B} which is absolutely continuous with respect to the image measure $\nu = P^\xi$. The Radon-Nidokym density $d\mu/d\nu(x)$ is denoted by $E(X|\xi = x)$ and is called the conditional expectation of X given $\xi = x$.

Definition 1.17. Let (Ω, \mathcal{F}, P) be a probability space and \mathcal{C} be a sub σ -field of \mathcal{F} . A system $\{p(\omega; A)\}_{\omega \in \Omega, A \in \mathcal{F}}$ is called a regular conditional probability given \mathcal{C} if it satisfies the following conditions:

- for fixed ω , $A \mapsto p(\omega, A)$ is a probability on (Ω, \mathcal{F}) ;
- for fixed $A \in \mathcal{F}$, $\omega \mapsto p(\omega, A)$ is \mathcal{C} -measurable;
- for every $A \in \mathcal{F}$ and $B \in \mathcal{C}$,

$$P(A \cap B) = \int_B p(\omega, A) P(d\omega)$$

clearly this property is equivalent to

- for every non-negative random variable X and $B \in \mathcal{C}$

$$E(X : B) = \int_{\Omega} \left\{ I_B(\omega) \int_{\Omega} X(\omega') p(\omega, d\omega') \right\} P(d\omega)$$

that is, $\int_{\Omega} p(\omega, d\omega')$ coincides with $E(X|\mathcal{C})(\omega)$ a.s.

We say that a the regular conditional probability is unique if whenever $\{p(\omega, A)\}$ and $\{p'(\omega, A)\}$ possess the above properties , then there exists a set $N \in \mathcal{C}$ of P-measure 0 such that, if $\omega \notin N$ then $p(\omega, A) = p'(\omega, A)$ for all $A \in \mathcal{F}$.

Chapter 2

Stochastic Processes

A stochastic process is a mathematical model for the occurrence, at each moment after the initial time, of a random phenomenon. The randomness is introduced by the sample space (ω, \mathcal{F}) . Thus, a stochastic process is a collection of random variables $X = \{X_t; 0 \leq t < \infty\}$ on (Ω, \mathcal{F}) , which takes values in a second measurable space (S, \mathcal{S}) , called the state space. For a fixed sample point $\omega \in \Omega$, the function $t \mapsto X_t(\omega); \quad t \geq 0$ is the sample path (realization, trajectory) of processes X associated with ω .

Definition 2.1. Y is a modification of X if, for every $t \geq 0$, we have $P[X_t = Y_t] = 1$.

Definition 2.2. X and Y are said to have the same finite dimensional distribution if, for any integer $n \geq 1$, real numbers $0 \leq t_1 < t_2 < \dots < t_n < \infty$ and $A \in \mathcal{B}(\mathbb{R}^{nd})$, we have:

$$P[(X_{t_1}, \dots, X_{t_n}) \in A] = P[(Y_{t_1}, \dots, Y_{t_n}) \in A]$$

Definition 2.3. X and Y are called indistinguishable if almost all their sample paths agree:

$$P[X_t = Y_t; \forall 0 \leq t < \infty] = 1.$$

Definition 2.4. The stochastic process X is called measurable if, for every $A \in \mathcal{B}(\mathbb{R}^d)$, the set $\{(t, \omega); X_t(\omega) \in A\}$ belongs to the product σ -field $\mathcal{B}([0, \infty)) \otimes \mathcal{F}$, in other words, if the mapping

$$(t, \omega) \mapsto X_t(\omega) : ([0, \infty) \times \Omega, \mathcal{B}([0, \infty)) \otimes \mathcal{F}) \rightarrow (\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d))$$

is measurable.

It is an immediate consequence of Fubini's theorem that the trajectories of such process are Borel-measurable functions of $t \in [0, \infty)$.

In order to talk about past, present and future we need to equip our sample space (Ω, \mathcal{F}) with a filtration.

Definition 2.5. *A filtration is a nondecreasing family $\{\mathcal{F}_t; t \geq 0\}$ of sub σ -fields of $\mathcal{F} : \mathcal{F}_s \subseteq \mathcal{F}_t \subseteq \mathcal{F}$ for $0 \neq s \neq t < \infty$. We set $\mathcal{F}_\infty = \sigma(\cup_{t \geq 0} \mathcal{F}_t)$ (σ -field generated).*

Given a stochastic process, the simplest choice of a filtration is that generated by the process itself, *i.e.*,

$$\mathcal{F}_t^X = \sigma(X_s; 0 \leq s \leq t)$$

the smallest σ -field with respect to which X_s is measurable for every $s \in [0, t]$. We interpret $A \in \mathcal{F}_t^X$ to mean that by time t , an observer of X knows whether or not A has occurred.

Let $\{\mathcal{F}_t; t \geq 0\}$ be a filtration. We define $\mathcal{F}_{t-} = \sigma(\cup_{s < t} \mathcal{F}_s)$ to be the σ -field of events strictly prior to $t > 0$ and $\mathcal{F}_{t+} = \cap_{\epsilon > 0} \mathcal{F}_{t+\epsilon}$ to be the σ -field of events immediately after $t \geq 0$. We define $\mathcal{F}_{0-} = \mathcal{F}_0$ and say that the filtration $\{\mathcal{F}_t\}$ is right (left) continuous if $\mathcal{F}_t = \mathcal{F}_{t+}$ (resp. $\mathcal{F}_t = \mathcal{F}_{t-}$) holds for every $t \geq 0$.

Definition 2.6. *The stochastic process X is adapted to the filtration $\{\mathcal{F}_t\}$ if, for each $t \geq 0$, X_t is an \mathcal{F}_{t-} measurable random variable.*

Note that every process X is adapted to $\{\mathcal{F}_t^X\}$.

Definition 2.7. *The stochastic process X is called progressively measurable with respect to the filtration \mathcal{F}_t if, for each $t \geq 0$ and $A \in \mathcal{B}(\mathbb{R}^d)$, the set $\{(s, \omega); 0 \leq s \leq t, \omega \in \Omega, X_s(\omega) \in A\}$ belongs to the product σ -field $\mathcal{B}([0, t]) \otimes \mathcal{F}_t$; in other words, if the mappings*

$$(s, \omega) \mapsto X_s(\omega) : ([0, t] \times \Omega, \mathcal{B}([0, t]) \otimes \mathcal{F}_t) \rightarrow (\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d))$$

is measurable for each $t \geq 0$.

Proposition 2.1. *If the stochastic process X is measurable and adapted to the filtration \mathcal{F}_t , then it has a progressively measurable modification.*

Proposition 2.2. *If the stochastic process X is right or left continuous and adapted to the filtration \mathcal{F}_t , then it is also progressively measurable with respect to \mathcal{F}_t .*

Definition 2.8. A stochastic process X on a probability space (Ω, \mathcal{F}, P) and $D \subset \mathbb{R}$ is said to be stochastically continuous at $t_0 \in D$ if $X(t, \cdot)$ converges to $X(t_0, \cdot)$ in probability as $t \rightarrow t_0$ in the sense that

$$\lim_{t \rightarrow t_0} P \{ \omega \in \Omega \mid |X(t, \omega) - X(t_0, \omega)| \geq \epsilon \} = 0 \quad \text{for every } \epsilon > 0$$

We say that X is stochastically continuous when it is stochastically continuous at every $t_0 \in D$.

Definition 2.9. Given a stochastic process X on a probability space and $D \subset \mathbb{R}$. Let \mathcal{F}' be an arbitrary sub-collection of the σ -field of Borel sets in \mathbb{R} and let \mathcal{F}_0 be the collection of open intervals in \mathbb{R} . The process X is said to be separable with respect to \mathcal{F}' if there exists a countable dense subset $S \subset D$ such that for every $A \in \mathcal{F}'$ and $I \in \mathcal{F}_0$ we have

$$\{ \omega \in \Omega; X(t, \omega) \in A \text{ for all } t \in D \cap I \} = \{ \omega \in \Omega; X(s, \omega) \in A \text{ for all } s \in S \cap I \}$$

when X is separable with respect to the collection of closed intervals \mathcal{F}_C in \mathbb{R} we say simply that X is separable.

2.0.2 Stopping Times

In this section we keep in mind the interpretation of the parameter t as time, and of the σ -field \mathcal{F}_t as the accumulated information up to t . We are interested in the instant $T(\omega)$, at which a phenomenon manifests itself for the first time.

Definition 2.10. A random time T is an \mathcal{F} -measurable random variable with values in $[0, \infty]$. It is simply a possible but random time for an event to occur.

Definition 2.11. Let us consider a measurable space (Ω, \mathcal{F}) equipped with a filtration \mathcal{F}_t . A random time T is a stopping time of the filtration, if the event $T \leq t$ belongs to the σ -field \mathcal{F}_t , for every $t \geq 0$. A random time T is an optimal time of the filtration if $T < t \in \mathcal{F}_t$ for every $t \geq 0$. Intuitively, a stopping time is a rule to stop a gamble, for example.

Proposition 2.3. Every random time equal to a nonnegative constant is a stopping time. Every stopping time is optional, and the two concepts coincide if the filtration is right-continuous.

Proof. The first statement is trivial; the second is based on the observation $T < t = \bigcup_{n=1}^{\infty} \{T \leq t - (1/n)\} \in \mathcal{F}_t$, because if T is a stopping time, then $\{T \leq t - (1/n)\} \in \mathcal{F}_{t-(1/n)} \subseteq \mathcal{F}_t$ for $n \geq 1$.

For the third claim, suppose that T is an optional time of the right-continuous filtration \mathcal{F}_t . Since $\{T \leq t\} = \bigcap_{\epsilon > 0} \{T < t + \epsilon\}$ we have $T \leq t \in \mathcal{F}_{t+\epsilon}$ for every $t \geq 0$ and every $\epsilon > 0$; whence $T \leq t \in \mathcal{F}_{t+} = \mathcal{F}_t$. \square

Corollary 2.1. *T is an optional time of the filtration \mathcal{F}_t if and only if it is a stopping time of the right-continuous filtration \mathcal{F}_{t+}*

Lemma 2.1. *If T is optional and θ is a positive constant, then $T + \theta$ is a stopping time.*

Proof. If $0 \leq t < \theta$, then $T + \theta \leq t = \emptyset \in \mathcal{F}_t$. If $t \geq 0$ then $T + \theta \leq t = T \leq t - \theta \in \mathcal{F}_{(t-\theta)+} \subseteq \mathcal{F}_t$. \square

Lemma 2.2. *If T, S are stopping times then so are $T \wedge S, T \vee S, T + S$.*

Proof. The first two assertions are trivial. For the third, start with the decomposition, valid for $t > 0$:

$$\{T + S > t\} = \{T = 0, S > t\} \cup \{0 < T < t, T + S > t\} \cup \{T > t, S = 0\} \cup \{T \geq t, S > 0\}$$

The first, third and fourth events in this decomposition are in \mathcal{F}_t , either trivially or by virtue of the proposition. As for the second event, we rewrite it as

$$\bigcup_{\substack{r \in \mathbb{Q}^+ \\ 0 < r < t}} t > T > r, S > t - r$$

\square

Lemma 2.3. *Let $T_{n=1}^{\infty}$ be a sequence of optional times, then the random times*

$$\sup_{n \geq 1} T_n, \inf_{n \geq 1} T_n, \limsup T_n, \liminf T_n$$

are all optional. Furthermore, if the T_n 's are stopping times, then so is $\sup_{n \geq 1} T_n$.

Proof. Obvious from the corollary to the proposition and from the identities

$$\left\{ \sup_{n > 1} T_n \leq t \right\} = \bigcap_{n > 1} \{T_n \leq t\} \quad \text{and} \quad \left\{ \inf_{n \geq 1} T_n < t \right\} = \bigcup_{n=1}^{\infty} \{T_n < t\}$$

\square

Let T be a stopping time of the filtration \mathcal{F}_t . The σ -field \mathcal{F}_t of events determined prior to the stopping time T consists of those events $A \in \mathcal{F}$ for which $A \cap T \leq t \in \mathcal{F}_t$ for every $t \geq 0$.

Lemma 2.4. *For any two stopping times T and S , and for any $A \in \mathcal{F}_S$, we have $A \cap s \leq T \in \mathcal{F}_T$. In particular, if $S \leq T$ on Ω , we have $\mathcal{F}_S \subseteq \mathcal{F}_T$.*

Proof. It is not hard to verify that, for every stopping time T and positive constant t , $T \wedge t$ is an \mathcal{F}_t -measurable random variable. With this in mind, the claim follows from the decomposition

$$A \cap \{S \leq T\} \cap \{T \leq t\} = [A \cap \{S \leq t\}] \cap \{T \leq t\} \cap \{S \vee t \leq T \vee t\}$$

which shows readily that the left-hand side is an event in \mathcal{F}_t . □

Lemma 2.5. *Let T and S be stopping times. Then $\mathcal{F}_{T \wedge S} = \mathcal{F}_T \cap \mathcal{F}_S$, and for each of the events*

$$\{T < S\}, \{S < T\}, \{T \leq S\}, \{S \leq T\}, \{T = S\}$$

belongs to $\mathcal{F}_t \cap \mathcal{F}_s$

2.1 Martingales

We proceed now to the study of an important class of stochastic processes, namely Martingales. We will present two version of martingales, the first in discrete time, and in the second we will consider continuous time. The study of discrete parameter martingales is presented here in order to strengthen the readers intuition on the subject.

2.1.1 Discrete Parameter Martingales

Definition 2.12. *Let (Ω, \mathcal{F}, P) be a probability space and $N = 1, \dots, n_0$ be a nonempty subset of $\mathbb{N} = 1, 2, \dots, n_0 \leq \infty$. Given a filtration $\mathcal{F}_n : n \in N$ and random variables $X_n : n \in N$, the collection $(X_n, \mathcal{F}_n) : n \in N$ is called a martingale if:*

1. $\{X_n : n \in N\}$ is adapted to $\{\mathcal{F}_n : n \in N\}$;
2. $E|X_n| < \infty$ for all $n \in N$; and

3. for all $1 \leq n < n_0$

$$E(X_{n+1}|\mathcal{F}_n) = X_n$$

In the context of a gamble discrete parameter Martingales can be interpreted in the following manner. Let X_n represent the fortune of a gambler at the end of the n th play and let \mathcal{F}_n be the information available to the gambler up to and including the n th play. Then, \mathcal{F}_n contains the knowledge of all events like $X_j \leq r$ for $r \in \mathbb{R}$, $j \leq n$, making X_n measurable with respect to \mathcal{F}_n . And condition 3 above says that given all the information up until the end of the n th play, the expected fortune of the gambler at the end of the $(n + 1)$ th play remains unchanged. Thus a Martingale represents a fair game.

Definition 2.13. Let $\{\mathcal{F}_n : n \in N\}$ be a filtration and $\{X_n : n \in N\}$ be a collection of random variables in $L^1(\Omega, \mathcal{F}, P)$ adapted to $\{\mathcal{F}_n : n \in N\}$. Then $\{(X_n, \mathcal{F}_n) : n \in N\}$ is called a sub-martingale if

$$E(X_{n+1}|\mathcal{F}_n) \geq X_n \text{ for all } 1 \leq n < n_0 \quad (2.1)$$

and a super-martingale if

$$E(X_{n+1}|\mathcal{F}_n) \leq X_n \text{ for all } 1 \leq n < n_0 \quad (2.2)$$

For a martingale we have

$$E(X_{n+1}|\mathcal{F}_n) = E(X_n|\mathcal{F}_n) = X_n \quad (2.3)$$

where, in this case, \mathcal{F}_n represents the filtration generated by the process.

Suppose that $\{(X_n, \mathcal{F}_n) : n \in N\}$ is a sub-martingale. Then $A \in \mathcal{F}_n$ implies that $A \in \mathcal{F}_{n+1} \subset \dots \subset \mathcal{F}_{n+k}$ for every $k \geq 1$, $n + k \in N$ and hence, by (2.1):

$$\int_A X_n dP \leq \int_A E(X_{n+1}|\mathcal{F}_n) dP = \int_A X_{n+1} dP \dots \leq \int_A X_{n+k} dP \quad (2.4)$$

Therefore, $E(X_{n+k}|\mathcal{F}_n) \geq X_n$ and, by taking $A = \Omega$ in (1.1), $E(X_{n+k}) \geq E(X_n)$. Thus, the expected value of a sub-martingale is non-decreasing. For a martingale, by (2.3), equality holds at every step of (2.4) and hence,

$$E(X_{n+k}|\mathcal{F}_n) = X_n, \quad EX_{n+k} = EX_n$$

for all $k \geq 1$, $n, n + k \in N$. Thus, in a fair game, the expected fortune of the gambler remains constant over time.

Proposition 2.4 (Convex Functions of Martingales and Sub-martingales). *Let $\phi : \mathbb{R} \rightarrow \mathbb{R}$ be a convex function and let $N = 1, 2, \dots, n_0 \subset \mathbb{N}$ be a nonempty subset.*

- *If $(X_n, \mathcal{F}_n) : n \in N$ is a martingale with $E|\phi(X_n)| < \infty$ for all $n \in N$, then $(\phi(X_n), \mathcal{F}_n) : n \in N$ is sub-martingale.*
- *If $(X_n, \mathcal{F}_n) : n \in N$ is a sub-martingale, $E|\phi(X_n)| < \infty$ for all $n \in N$, and in addition, ϕ is nondecreasing, then $\{(\phi(X_n)), \mathcal{F}_n : n \in N\}$ is a sub-martingale.*

Proof. Follows directly from Jensen's inequality. □

Proposition 2.5 (Doob's Decomposition of a Sub-martingale). *Let $\{(X_n, \mathcal{F}_n) : n \in N\}$ be a sub-martingale for some $N = 1, \dots, n_0 \subset \mathbb{N}$. Then, there exist two sets of random variables $\{Y_n : n \in N\}$ and $\{Z_n : n \in N\}$ satisfying $X_n = Y_n + Z_n$, $n \in N$ such that*

1. *$\{(X_n, \mathcal{F}_n) : n \in N\}$ is a martingale;*
2. *For all $n \in N$, $Z_{n+1} \geq Z_n \geq 0$ with probability 1 and Z_n is \mathcal{F}_{n-1} -measurable, where $\mathcal{F}_0 = \emptyset, \Omega$;*
3. *If $X_n : n \in N$ are L^1 -bounded, i.e., $M \equiv \text{Max}E|X_n| : n \in N < \infty$, then so are $\{Y_n : n \in N\}$ and $\{Z_n : n \in N\}$.*

Proof. Define the difference variables Δ 's by

$$\Delta_1 = X_1 \quad \text{and} \quad \Delta_n = X_n - X_{n-1}, \quad n \geq 2, \quad n \in N.$$

Note that $X_n = \sum_{j=1}^n \Delta_j$ (telescopic sum), $n \in N$, and $E(\Delta_n | \mathcal{F}_{n-1}) \geq 0$ for all $n \geq 2$, $n \in N$ (since X_n is a sub-martingale). Now, set

$$Y_1 = \Delta_1, \quad Y_n = X_n - \sum_{j=2}^n E(\Delta_j | \mathcal{F}_{j-1}), \quad n \geq 2, \quad n \in N.$$

and

$$Z_1 = \Delta_1, \quad Z_n = \sum_{j=2}^n E(\Delta_j | \mathcal{F}_{j-1}), \quad n \geq 2, \quad n \in N.$$

Requirements 1 and 2 follow since $X_n : n \in N$ is a sub-martingale. To establish the L^1 -boundedness, notice that 2, for any $n \geq 1$, $n \in N$,

$$\begin{aligned} E|Z_n| &= EZ_n = E \left[\sum_{j=2}^n E(\Delta_j | \mathcal{F}_{j-1}) \right] = \sum_{j=2}^n E(\Delta_j) \\ &= EX_n - EX_1 \leq 2M. \end{aligned}$$

Also, $X_n = Y_n + Z_n$ for all $n \in N$ implies that

$$|Y_n| \leq |X_n| + |Z_n|, \quad n \in N.$$

hence 3 follows. □

Theorem 2.1 (Optional Sampling). *If $(X_n, \mathcal{F}_n)_{n \in \mathbb{N}}$ is a sub-martingale and σ and τ are two stopping times such that $\sigma \leq \tau \leq k$ for some $k \in \mathbb{N}$, then*

$$X_\sigma \leq E(X_\tau | \mathcal{F}_\sigma)$$

If $(X_n, \mathcal{F}_n)_{n \in \mathbb{N}}$ is a martingale or a super-martingale, then the same statement holds with the \leq sign replaced by $=$ or \geq respectively.

Proof. The case of $(X_n, \mathcal{F}_n)_{n \in \mathbb{N}}$ being a super-martingale is equivalent to considering the sub-martingale $(-X_n, \mathcal{F}_n)_{n \in \mathbb{N}}$. Then, without loss of generality, we may assume that $(X_n, \mathcal{F}_n)_{n \in \mathbb{N}}$ is a sub-martingale.

Let $A \in \mathcal{F}_\sigma$. For $1 \leq m \leq n$ we define

$$\begin{aligned} A_m &= A \cap \sigma = m, & A_{m,n} &= A_m \cap \tau = n \\ B_{m,n} &= A \cap \tau > n, & C_{m,n} &= A_m \cap \tau \geq n \end{aligned}$$

Note that $B_{m,n} \in \mathcal{F}_n$, since $\tau > n = \Omega \setminus \tau \leq n \in \mathcal{F}_n$. Therefore, by definition of a sub-martingale,

$$\int_{B_{m,n}} X_n dP \leq \int_{B_{m,n}} X_{n+1} dP.$$

Since $C_{m,n} = A_{m,n} \cup B_{m,n}$,

$$\int_{C_{m,n}} X_n dP \leq \int_{A_{m,n}} X_n dP + \int_{B_{m,n}} X_{n+1} dP$$

and thus, since $B_{m,n} = C_{m,n+1}$,

$$\int_{C_{m,n}} X_n dP - \int_{C_{m,n+1}} X_{n+1} dP \leq \int_{A_{m,n}} X_n dP.$$

By taking the sum from $n = m$ to k , and noting that we have a telescopic sum on the left-hand side, we obtain

$$\int_{A_m} X_m dP \leq \int_{A_m} X_\tau dP,$$

where we used that $A_m = C_{m,m}$. By taking the sum from $m = 1$ to k , we obtain

$$\int_A X_\sigma dP \leq \int_A X_\tau dP.$$

Since $A \in \mathcal{F}_\sigma$ was arbitrary, this completes the proof of the theorem □

We may interpret the result above in the following sense: in a fair game (modeled by a martingale), a gambler cannot increase or decrease the expectation of his fortune by entering the game at a point of time $\sigma(\omega)$, and then quitting the game at $\tau(\omega)$, provided that he decides to enter and leave the game based only on the information available by the time of decision.

2.1.2 Continuous Parameter Martingales

Definition 2.14. *The process $\{X_t, \mathcal{F}_t; 0 \leq t < \infty\}$ is said to be a sub-martingale (respct. a super-martingale) if, for every $0 \leq s \leq t < \infty$, we have a.s. $(P) E(X_t | \mathcal{F}_s) \geq X_s$ (respct. $E(X_t | \mathcal{F}_s) \leq X_s$). We shall say that $\{X_t, \mathcal{F}_t; 0 \leq t < \infty\}$ is a Martingale if it is both a sub-martingale and a super-martingale.*

Proposition 2.6. *Let $\{X_t, \mathcal{F}_t; 0 \leq t < \infty\}$ be a martingale (respct. a sub-martingale) and $\varphi : \mathbb{R} \rightarrow \mathbb{R}$ a convex (respct. convex nondecreasing) function, such that $E|\varphi(X_t)| < \infty$ holds for every $t \geq 0$. Then $\{\varphi(X_t), \mathcal{F}_t; 0 \leq t < \infty\}$ is a sub-martingale.*

Once again the proof follows from Jensen's inequality.

Let $\{X_t, \mathcal{F}_t; 0 \leq t < \infty\}$ be a real valued stochastic process. Consider two numbers $\alpha < \beta$ and a finite subset F of $[0, \infty)$. We define the number of upcrossings $U_F(\alpha, \beta; X(\omega))$ of the interval $[\alpha, \beta]$ (the number of times the sequence passes from under α to above β) by the restricted sample path $X_t; t \in F$ as follows. Set

$$\tau_1(\omega) = \min \{t \in F; X_t(\omega) \leq \alpha\}$$

and define recursively for $j = 1, 2, \dots$

$$\sigma_j(\omega) = \min \{t \in F, t \geq \tau_j(\omega), X_t(\omega) \geq \beta\},$$

$$\tau_{j+1}(\omega) = \min \{t \in F, t \geq \sigma_j(\omega), X_t(\omega) \leq \alpha\}$$

The convention here is that the minimum of the empty set is ∞ , and we denote by $U_F(\alpha, \beta; X(\omega))$ the largest integer j for which $\sigma_j(\omega)$. If $I \subset [0, \infty)$ is not necessarily finite, we define

$$U_I(\alpha, \beta; X(\omega)) = \sup \{U_F(\alpha, \beta; X(\omega)); F \subseteq I, F \text{ is finite}\}$$

The number of downcrossings $D_I(\alpha, \beta; X(\omega))$ is defined similarly.

Theorem 2.2. Let $X_t, \mathcal{F}_t; 0 \leq t < \infty$ be a right-continuous sub-martingale, $[\sigma, \tau]$ a subinterval, of $[0, \infty)$, and $\alpha < \beta, \lambda$ given real numbers. We have the following results:

- First sub-martingale inequality :

$$\lambda P \left[\sup_{\sigma \leq t \leq \tau} X_t \geq \lambda \right] \leq E(X_\tau^+)$$

- Second sub-martingale inequality:

$$\lambda P \left[\inf_{\sigma \leq t \leq \tau} X_t \leq -\lambda \right] \leq E(X_\tau^+) - E(X_\sigma)$$

- Upcrossings and downcrossings inequalities:

$$EU_{[\sigma, \tau]}(\alpha, \beta; X(\omega)) \leq \frac{E(X_\tau^+) + |\alpha|}{\beta - \alpha}$$

$$ED_{[\sigma, \tau]}(\alpha, \beta; X(\omega)) \leq \frac{E(X_\tau - \alpha)^+}{\beta - \alpha}$$

- Doob's maximal inequality:

$$E \left(\sup_{\sigma \leq t \leq \tau} X_t \right)^p \leq \left(\frac{p}{p-1} \right)^p E(X_\tau^p), \quad p > 1$$

provided $X_t \geq 0$ a.s. (P) for every $t \geq 0$, and $E(X_\tau^p) < \infty$.

- Regularity of paths:

Almost every sample path $\{X_t(\omega); 0 \leq t < \infty\}$ is bounded on compact intervals; is free of discontinuities of the second kind, i.e., admits left-hand limits everywhere on $(0, \infty)$, and thus has countably many jumps.

Theorem 2.3 (Sub-martingale convergence). Let $\{X_t, \mathcal{F}_t; 0 \leq t < \infty\}$ be a right-continuous sub-martingale and assume $C = \sup_{t \geq 0} E(X_t^+) < \infty$. Then $X_\infty(\omega) = \lim_{t \rightarrow \infty} X_t(\omega)$ exists for a.e. $\omega \in \Omega$, and $E|X_\infty| < \infty$.

For the proof of this and of the preceding theorem the reader is referred to [9]

Definition 2.15. A right-continuous, nonnegative super-martingale $\{Z_t, \mathcal{F}_t; 0 \leq t < \infty\}$ with $\lim_{t \rightarrow \infty} E(Z_t) = 0$ is called a potential.

Proposition 2.7. Suppose that the filtration $\{\mathcal{F}_t\}$ satisfies the usual conditions. Then every right-continuous, uniformly integrable super-martingale $\{X_t, \mathcal{F}_t; 0 \leq t < \infty\}$ admits the Riesz decomposition $X_t = M_t + Z_t$, a.s. (P), as the sum of a right-continuous uniformly integrable martingale M_t and a potential Z_t .

Theorem 2.4 (Optional sampling). *Let $\{X_t, \mathcal{F}_t; 0 \leq t < \infty\}$ be a right-continuous submartingale with a last element x_∞ , and let $S < T$ be two optional times of the filtration $\{\mathcal{F}_t\}$ we have*

$$E(X_T | \mathcal{F}_{S+}) \geq X_S \quad \text{a.s.}(P)$$

If S is a stopping time, then \mathcal{F}_S can replace \mathcal{F}_{S+} above. In particular, $EX_T \geq EX_0$, and for a martingale with a last element we have $EX_T = EX_0$

2.1.3 The Doob-Meyer Decomposition

In this section we follow closely [9], the proofs are here omitted and can be found on that text.

Definition 2.16. *Consider a probability space (Ω, \mathcal{F}, P) and a random sequence $\{A_n\}_{n=0}^\infty$ adapted to the discrete filtration $\{\mathcal{F}_n^\infty\}$. The sequence is called increasing, if for P . a.e. $\omega \in \Omega$ we have $0 = A_0(\omega) \leq A_1(\omega) \leq \dots$, and $E(A_n) < \infty$ holds for every $n \geq 1$.*

Definition 2.17. *An increasing sequence is called integrable if $E(A_\infty) < \infty$, where $A_\infty = \lim_{n \rightarrow \infty} A_n$.*

Definition 2.18. *An arbitrary random sequence $\{\xi_n\}_{n=0}^\infty$ is called predictable for the filtration $\{\mathcal{F}_n\}_{n=0}^\infty$, if for every $n \geq 1$ the random variable ξ_n is \mathcal{F}_{n-1} -measurable.*

Note that if $A = \{A_n, \mathcal{F}_n; n = 0, 1, \dots\}$ is predictable with $E|A_n| < \infty$ for every n , and if $\{M_n, \mathcal{F}_n; n = 0, 1, \dots\}$ is a bounded martingale, then the martingale transform of A by M defined by

$$Y_0 = 0 \quad \text{and} \quad Y_n = \sum_{k=1}^n A_k (M_k - M_{k-1}), \quad n \geq 1$$

is itself a martingale.

Definition 2.19. *An increasing sequence $\{A_n, \mathcal{F}_n; n = 0, 1, \dots\}$ is called natural if for every bounded martingale $\{M_n, \mathcal{F}_n; n = 0, 1, \dots\}$ we have*

$$E(M_n, A_n) = E \sum_{k=1}^n M_{k-1} (A_k - A_{k-1})$$

An increasing sequence A is natural if and only if the martingale transform $Y = \{Y_n\}_{n=0}^\infty$ of A by every bounded martingale M satisfies $E(Y_n) = 0$, $n \geq 0$.

Proposition 2.8. *An increasing random sequence A is predictable if and only if, it is natural.*

Definition 2.20. *An adapted process A is called increasing if for P.a.e. $\omega \in \Omega$ we have*

- $A_0(\omega) = 0$
- $t \mapsto A_t(\omega)$ is a nondecreasing, right-continuous function, and $E(A_t) < \infty$ holds for every $t \in [0, \infty)$. An increasing process is called integrable if $E(A_\infty) < \infty$, where $A_\infty = \lim_{t \rightarrow \infty} A_t$.

Definition 2.21. *An increasing process A is called natural if for every bounded, right-continuous martingale $\{M_t, \mathcal{F}_t; 0 \leq t < \infty\}$ we have*

$$E \int_{(0,t]} M_s dA_s = E \int_{(0,t]} M_{s-} dA_s, \quad \text{for every } 0 < t < \infty$$

Lemma 2.6. *In the definition above, the condition is equivalent to*

$$E(M_t A_t) = E \int_{(0,t]} M_{s-} dA_s$$

Definition 2.22. *Let us consider the class $\mathcal{C}(\mathcal{C}_a)$ of all stopping times T of the filtration $\{\mathcal{F}_t\}$ which satisfy $P(T < \infty) = 1$ (respectively, $P(T \leq a) = 1$ for a given finite number $a > 0$). The right-continuous process $\{X_t, \mathcal{F}_t; 0 \leq t < \infty\}$ is said to be of class D , if the family $\{X_T\}_{T \in \mathcal{C}}$ is uniformly integrable; and of class DL , if the family $\{X_T\}_{T \in \mathcal{C}_a}$ is uniformly integrable for every $0 < a < \infty$.*

Theorem 2.5 (Doob-Meyer Decomposition). *Let $\{\mathcal{F}_t\}$ satisfy the usual conditions. If the right-continuous sub-martingale $X = \{X_t, \mathcal{F}_t; 0 \leq t < \infty\}$ is of class DL , then it admits the decomposition*

$$X_t = M_t + A_t, \quad 0 \leq t < \infty$$

as the assumption of a right-continuous martingale $M = \{M_t, \mathcal{F}_t; 0 \leq t < \infty\}$ and an increasing process $A = \{A_t, \mathcal{F}_t; 0 \leq t < \infty\}$. The latter can be taken to be natural; under this additional condition, the decomposition is unique (up to indistinguishability). Further, if X is of class D , then M is a uniformly integrable martingale and A is integrable.

Definition 2.23. *A sub-martingale $\{X_t, \mathcal{F}_t; 0 \leq t < \infty\}$ is called regular if for every $a > 0$ and every nondecreasing sequence of stopping times $\{T_n\}_{n=1}^\infty \subseteq \mathcal{C}_a$ with $T = \lim_{n \rightarrow \infty} T_n$, we have $\lim_{n \rightarrow \infty} E(X_{T_n}) = E(X_T)$.*

2.2 Brownian Motion

We will now begin the study of one of the most important classes of stochastic processes, named Brownian motion after Robert Brown, a botanist who in 1827 described the movement of particles suspended in a fluid. The first quantitative formulation is due to Einstein in 1905.

Our approach here is to present two versions of Brownian motion. The first is restricted to a one dimensional processes, in the second we generalize to process to higher dimensions. This approach should help the reader grasp the intuition in a one dimensional setting before attempting to understand a multidimensional process.

Before proceeding to either version we state a definition, which the reader is probably already familiar with.

Definition 2.24. *A real valued random variable $X(\omega)$ defined on a probability space is said to be Gaussian (normal) if its distribution $F(X)$ is of the form*

$$F(X) = (2\pi\sigma^2)^{-1/2} \int_{-\infty}^X \exp -\frac{(Y - m)^2}{2\sigma^2} dY$$

where σ^2 and m are constants. Equivalently, X has a characteristic function $\Phi(Z)$ of the form

$$\Phi(Z) = E(\exp iZX) = \exp(imZ - \frac{1}{2}Z^2\sigma^2), \quad Z \in \mathbb{R}$$

2.2.1 One Dimensional Brownian Motion

Definition 2.25. *A stochastic process X on a probability space (Ω, \mathcal{B}, P) and an interval $D \subset \mathbb{R}$ is called an additive process or a process with independent increments if for any $\{t_1, t_2, \dots, t_n\} \subset D$, $t_1 < t_2 < \dots < t_n$, the system of random variables $\{X(t_{i+1}, \cdot) - X(t_i, \cdot), i = 1, 2, \dots, n - 1\}$ is independent.*

If X is an additive process and $\Phi_{t't''}$ is the one dimensional probability distribution determined by the random variable $X(t'', \cdot) - X(t', \cdot)$, where $t', t'' \in D$, $t' < t''$, then for any $t_1, t_2, t_3 \in D$, $t_1 < t_2 < t_3$, the convolution of Φ_{t_1, t_2} and Φ_{t_2, t_3} satisfies the condition

$$\bar{\Phi}_{t_1, t_2} * \bar{\Phi}_{t_2, t_3} = \bar{\Phi}_{t_1, t_3}$$

Definition 2.26. *By a Brownian motion process we mean an additive process X on a probability space (Ω, \mathcal{B}, P) and an interval $D \subset \mathbb{R}$ in which the probability distribution*

of the random variable $X(t'', \cdot) - X(t', \cdot)$, $t', t'' \in D$, $t' < t''$, is the normal distribution $N(0, c(t'' - t'))$ where c is a fixed positive number for the process.

$N(m, v)(B)$, $B \in \mathcal{B}$, $m \in \mathbb{R}$, $v > 0$ is a one dimensional probability distribution which is absolutely continuous with respect to the Lebesgue measure m_L on $(\mathbb{R}, \mathcal{B})$ with density function (that is, Radom-Nikodym derivative with respect to m_L) given by

$$N'(m, v)(X) = \frac{dN(m, v)}{dm_L} = \frac{1}{\sqrt{2\pi v}} \exp \left\{ -\frac{1}{2} \frac{(X - m)^2}{v} \right\} \quad \text{for } X \in \mathbb{R} \quad (2.5)$$

Lemma 2.7. *Let Φ be the Normal distribution $N(m, v)$, $m \in \mathbb{R}$, $v > 0$, then the following hold*

- *The characteristic function φ of Φ is given by*

$$\varphi(Y) = \exp \left\{ imY - \frac{v}{2} Y^2 \right\} \quad \text{for } Y \in \mathbb{R}$$

- *All moments of Φ , i.e. $M^p(\Phi)$ where*

$$M^p(\Phi) = \int_{\mathbb{R}} X^p(\Phi) dX \quad \text{for } p = 1, 2, \dots$$

exist and are finite. For the first moment $M^1(\Phi)$ and the variance $V(\Phi)$ we have

$$M^1(\Phi) = m \quad \text{and} \quad V(\Phi) = v$$

- *All the central moments of Φ , i.e., $M_0^p(\Phi)$ where*

$$M_0^p(\Phi) = \int_{\mathbb{R}} \{X - M^1(\Phi)\}^p \Phi dX \quad \text{for } p = 1, 2, \dots$$

exist as finite numbers and are given by

$$M_0^p(\Phi) = \begin{cases} 0 & \text{for odd } p \\ \left(\frac{2^p v^p}{\pi}\right)^{1/2} \Gamma(p/2 + 1/2) & \text{for even } p \end{cases}$$

where $\Gamma(t)$ is the gamma function $\Gamma(t) = \int_0^\infty X^{t-1} e^{-X}$ and $\Gamma(t) = (t-1)\Gamma(t-1)$

with $\Gamma(1/2) = \sqrt{\pi}$.

Lemma 2.8. *The convolution of two normal distributions $N(m_1, v_1)$ and $N(m_2, v_2)$ is a normal distribution and*

$$N(m_1, v_1) * N(m_2, v_2) = N(m_1 + m_2, v_1 + v_2).$$

Proof. Let φ_1, φ_2 and φ be the characteristic functions of the normal distributions $N(m_1, v_1)$, $N(m_2, v_2)$ and $N(m_1 + m_2, v_1 + v_2)$. Then

$$\varphi_1(Y)\varphi_2(Y) = \exp \left\{ i(m_1 + m_2)Y - \frac{v_1 v_2}{2} Y^2 \right\} = \varphi(Y) \text{ for } Y \in \mathbb{R}$$

□

Theorem 2.6. *Given an interval $D \subset \mathbb{R}$ and a system of one dimensional probability distributions $\{\Phi_{t', t''}, t', t'' \in D, t' < t''\}$ where $\Phi_{t', t''} = N(0, c(t'' - t))$ with fixed $c > 0$. There exist a Brownian motion process X on a probability space (Ω, \mathcal{B}, P) and the interval D in which $\Phi_{t', t''}$ is the probability distribution of the random variable $X(t'', \cdot) - X(t', \cdot)$ for any $t', t'' \in D, t' < t''$. If D contains its left (right) endpoint and $\Phi_a(\Phi_b)$ is an arbitrary one dimensional probability distribution then there exists a Brownian motion process X with the additional property that $\Phi_a(\Phi_b)$ is the initial (final) probability distribution of X .*

Proof. See [17]

□

To discuss the continuity of a Brownian motion process we need the following lemma and theorem, whose proofs are omitted.

Lemma 2.9. *Let f and F be the density and the distribution functions of the normal distribution $N(0, 1)$, i.e.*

$$f(X) = \frac{1}{\sqrt{2\pi}} \exp \left\{ -\frac{X^2}{2} \right\} \quad \text{for } X \in \mathbb{R}$$

$$F(X) = \int_{-\infty}^X f(u) du \quad \text{for } X \in \mathbb{R}$$

Then on $(0, \infty)$ and for every $X \in (0, \infty)$

$$1 - F(X) < f(X) \frac{1}{X}$$

Theorem 2.7. *Let X be an additive process on a probability space (Ω, \mathcal{B}, P) and an interval $D \subset \mathbb{R}$. If X is stochastically continuous at $t_0 \in D$ then for every sequence $\{S_n, n = 1, 2, \dots\} \subset D$ such that $\lim_{n \rightarrow \infty} S_n = t$ we have $P \{\omega \in \Omega; \lim_{n \rightarrow \infty} X(S_n, \omega) = X(t_0, \omega)\} = 1$. If in addition X is separable then $P \{\omega \in \Omega; \lim_{t \rightarrow t_0} X(t, \omega) = X(t_0, \omega)\} = 1$.*

Theorem 2.8. *Let X be a Brownian motion process on a probability space (Ω, \mathcal{B}, P) and an interval D in which the probability distribution of the random variable $X(t'', \cdot) -$*

$X(t', \cdot)$, $t', t'' \in D$, $t' < t''$, is given by $N(0, c(t'' - t'))$, $c > 0$. Then X is stochastically continuous at every $t_0 \in D$ and in fact, for every $\epsilon > 0$,

$$P \{ \omega \in \Omega; |X(t_0 + h, \omega) - X(t_0, \omega)| \geq \epsilon \} \leq \frac{1}{\epsilon} \sqrt{\frac{2c|h|}{\pi}} \exp \left\{ -\frac{\epsilon^2}{2c|h|} \right\}$$

If X is separable, then at every $t_0 \in D$

$$P \left\{ \omega \in \Omega; \lim_{t \rightarrow t_0} X(t, \omega) = X(t_0, \omega) \right\} = 1$$

Proof. $X(t_0 + h, \cdot) - X(t_0, \cdot)$ is a random variable with probability distribution $N(0, c|h|)$. Then

$$P \{ \omega \in \Omega; |X(t_0 + h, \omega) - X(t_0, \omega)| \geq \epsilon \} = \frac{2}{\sqrt{2\pi c|h|}} \int_{\epsilon}^{\infty} \exp \left\{ -\frac{u^2}{2c|h|} \right\} du$$

making the change of variables $\frac{u}{\sqrt{c|h|}} \rightarrow v$

$$= \sqrt{\frac{2}{\pi}} \int_{\left\{ \frac{\epsilon}{\sqrt{c|h|}} \right\}}^{\infty} \exp \left\{ -\frac{v^2}{2} \right\} dv$$

Since $\int_{-\infty}^{\infty} f(u) du = 1$ $\int_{-\infty}^X f(u) du + \int_X^{\infty} f(u) du = 1 \Rightarrow \int_X^{\infty} f(u) du = 1 - \int_{-\infty}^X f(u) du$

So

$$\begin{aligned} &= 2 \left\{ 1 - F \left(\frac{\epsilon}{\sqrt{c|h|}} \right) \right\} < 2f \left(\frac{\epsilon}{\sqrt{c|h|}} \right) \frac{\sqrt{c|h|}}{\epsilon} \\ &= \frac{1}{\epsilon} \sqrt{\frac{2c|h|}{\pi}} \exp \left\{ -\frac{\epsilon^2}{2c|h|} \right\} \end{aligned}$$

which establishes the first part. The second part follows from the preceding theorem. \square

Theorem 2.9. *Let X be a Brownian motion process on a probability space (Ω, \mathcal{B}, P) and an interval $D \subset \mathbb{R}$. Then there exists a continuous process Y on (Ω, \mathcal{B}, P) and D which is equivalent to X (and is hence a Brownian motion).*

To establish this theorem we need a result due to Loève:

Theorem 2.10 (Loève). *Let X be a stochastic process on a probability space (Ω, \mathcal{B}, P) and $D = [0, 1]$. Let $g(h)$ and $q(h)$ be arbitrary nonnegative even functions defined for $0 < |h| < h_0$ with some $h_0 > 0$ which are monotone increasing for $h \in (0, h_0)$ and satisfy*

$$\sum_{n=1}^{\infty} g \left(\frac{1}{2^n} \right) < \infty; \quad \sum_{n=1}^{\infty} 2^n q \left(\frac{1}{2^n} \right) < \infty$$

If X satisfies the condition that

$$P\{\omega \in \Omega; |X(t+h, \omega) - X(t, \omega)| \geq g(h)\} \leq q(h)$$

whenever $t, t+h \in [0, 1]$, $0 < |h| < h_0$.

Then there exists a stochastic process Y on (Ω, \mathcal{B}, P) and D which is equivalent to X and is continuous.

Proof. (Theorem 2.9) To prove the existence of Y we show that X satisfies the condition of Loève's theorem. If we take $\epsilon = |h|^a$ where $0 < a < 1/2$ in the first result of the preceding theorem, then

$$P\{\omega \in \Omega; |X(t+h, \omega) - X(t, \omega)| \geq |h|^a\} \leq \sqrt{\frac{2c}{\pi}} |h|^{1/2-a} \exp\left\{-\frac{|h|^{2a-1}}{2c}\right\}$$

Let

$$g(h) = |h|^a \quad \text{and} \quad q(h) = \sqrt{\frac{2c}{\pi}} |h|^{1/2-a} \exp\left\{-\frac{|h|^{2a-1}}{2c}\right\}$$

For $h > 0$, $g(h)$ is monotone increasing, and

$$q'(h) = \sqrt{\frac{2c}{\pi}} \exp\left\{-\frac{h^{2a-1}}{2c}\right\} \left\{\left(\frac{1}{2}\right) h^{-1/2-a} - \frac{1}{2c}(2a-1)h^{-3/2+a}\right\} > 0$$

since $\frac{1}{2} - a > 0$ and $2a - 1 < 0$, so that $q(h)$ is also monotone increasing. Now

$$\sum_{n=1}^{\infty} g\left(\frac{1}{2^n}\right) = \sum_{n=1}^{\infty} \left(\frac{1}{2^n}\right)^a < \infty$$

and

$$\sum_{n=1}^{\infty} 2^n q\left(\frac{1}{2^n}\right) = \sqrt{\frac{2c}{\pi}} (2^{1/2+a})^n \exp\left\{-\frac{1}{2c}(2^{1-2a})^n\right\}$$

By applying the logarithmic ratio test we see that this series converges. Thus the conditions of Loève's theorem are satisfied by X , and Y exists. \square

Theorem 2.11. *A separable Brownian motion is almost surely continuous.*

Reflection Principle

Definition 2.27. *A random variable X on a probability space (Ω, \mathcal{B}, P) is said to be symmetrically distributed if*

$$P\{\omega \in \Omega; X(\omega) \leq x\} = P\{\omega \in \Omega; X(\omega) \geq -x\} \quad \text{for every } x \in \mathbb{R}$$

the condition is equivalent to

$$P\{\omega \in \Omega; X(\omega) < x\} = P\{\omega \in \Omega; X(\omega) > -x\} \quad \text{for every } x \in \mathbb{R}$$

it is also equivalent to

$$F(x+0) + F(-x-0) = 1$$

and

$$\Phi(B) = \Phi(-B) \quad \text{for every } B \in \mathcal{B} \quad \text{where } -B = \{x \in \mathbb{R}; -x \in B\}$$

Lemma 2.10. *A random variable X on a probability space (Ω, \mathcal{B}, P) is symmetrically distributed if and only if its characteristic function φ is real valued.*

Proof. Let Φ be the one dimensional probability distribution determined by X . If X is symmetrically distributed then, by the relation $\Phi(B) = \Phi(-B)$,

$$\varphi(y) = \int_{\mathbb{R}} e^{iyx} \Phi dx = \int_{\mathbb{R}} e^{-iyu} \Phi(-du) = \int_{\mathbb{R}} e^{-iyu} \Phi du = \varphi(\bar{y})$$

so that φ is real valued. Conversely, if φ is real valued then $\varphi(y) = \varphi(\bar{y})$ for every $y \in \mathbb{R}$,
i.e

$$\int_{\mathbb{R}} e^{iyx} \Phi dx = \int_{\mathbb{R}} e^{-iyx} \Phi dx$$

or

$$\int_{\mathbb{R}} e^{iyx} \Phi dx = \int_{\mathbb{R}} e^{iyu} \Phi(-du)$$

From the one-to-one correspondence between characteristic functions and probability distributions (Bochner's theorem) we conclude that $\Phi(B) = \Phi(-B)$ holds so that X is symmetrically distributed. \square

Lemma 2.11. *Let $\{X_j, j = 1, 2, \dots, n\}$ be an independent system of symmetrically distributed random variables on a probability space (Ω, \mathcal{B}, P) and let*

$$S_j = X_1 + \dots + X_j \quad \text{for } j = 1, 2, \dots, n \quad \text{and} \quad S_0 = 0.$$

Then for every $\lambda, \epsilon > 0$

$$\begin{aligned} 2P\{\omega \in \Omega, S_n(\omega) \geq \lambda\} &\geq P\{\omega \in \Omega; \max_{1 \leq j \leq n} S_j(\omega) \geq \lambda\} \\ &\geq 2P\{\omega \in \Omega; S_n(\omega) \geq \lambda + 2\epsilon\} - \sum_{j=1}^n P\{\omega \in \Omega, X_j(\omega) \geq \epsilon\} \end{aligned}$$

Theorem 2.12 (Reflexion Principle). *Let X be a separable Brownian motion process on a probability space (Ω, \mathcal{B}, P) and an interval $D \subset \mathbb{R}$ in which the probability distribution of*

the random variable $X(t'', \cdot) - X(t', \cdot)$, $t', t'' \in D$, $t' < t''$, is given by $N(0, c(t'' - t'))$, $c >$

0. Then for any $\alpha, \beta \in D$, $\alpha < \beta$ and $\lambda > 0$

$$P \left\{ \omega \in \Omega; \sup_{[\alpha, \beta]} \{X(t, \omega) - X(\alpha, \omega)\} \geq \lambda \right\} = 2P \{ \omega \in \Omega; X(\beta, \omega) - X(\alpha, \omega) \geq \alpha \}$$

$$\leq \sqrt{\frac{2c(\beta - \alpha)}{\pi}} \frac{1}{\alpha} \exp \left\{ -\frac{\lambda^2}{2c(\beta - \alpha)} \right\}$$

Proof. Since X is separable there exists a countable dense subset $S \subset D$ such that, for every open interval $I \subset \mathbb{R}$,

$$\sup_{D \cap I} X(t, \omega) = \sup_{S \cap I} X(s, \omega)$$

Let $\alpha, \beta \in D$, $\alpha < \beta$. Adjoin α and β to S if they are not already in S . With $I = (\alpha, \beta)$, the above equality implies

$$\sup_{[\alpha, \beta]} X(t, \omega) = \sup_{S \cap [\alpha, \beta]} X(s, \omega) \quad (2.6)$$

Let $S \cap (\alpha, \beta) = \{S_n, n = 1, 2, \dots\}$. For any n , let the rearrangement of $\{t_{n,k}, k = 0, 1, 2, \dots, n+1\}$, or simply $\{t_k, k = 0, 1, 2, \dots, n+1\}$ for brevity. For arbitrary $\lambda, \epsilon > 0$ we have

$$\left\{ \omega \in \Omega; \sup_{S \cap [\alpha, \beta]} \{X(S, \omega) - X(\alpha, \omega)\} \geq \lambda \right\} \subset$$

$$\bigcup_{n=1}^{\infty} \left\{ \omega \in \Omega, \max_{1 \leq k \leq n+1} \{X(t_k, \omega) - X(t_0, \omega)\} \geq \lambda - \epsilon \right\} \quad (2.7)$$

By (2.6) and the fact that the sequence of sets in (2.7) is monotone increasing we have

$$P \left\{ \omega \in \Omega; \sup_{S \cap [\alpha, \beta]} \{X(S, \omega) - X(\alpha, \omega)\} \geq \lambda \right\}$$

$$\leq \lim_{n \rightarrow \infty} P \left\{ \omega \in \Omega; \max_{1 \leq k \leq n+1} \{X(t_k, \omega) - X(t_0, \omega)\} \geq \lambda - \epsilon \right\} \quad (2.8)$$

Now $\{X(t_k, \cdot) - X(t_{k-1}, \cdot), k = 1, 2, \dots, n+1\}$ is an independent system of random variables, each of which is normally and hence symmetrically distributed. By the first inequality in the lemma above,

$$P \left\{ \omega \in \Omega; \max_{1 \leq k \leq n+1} \{X(t_k, \omega) - X(t_0, \omega)\} \geq \lambda - \epsilon \right\} \leq 2P \{ \omega \in \Omega; X(\beta, \omega) - X(\alpha, \omega) \geq \lambda - \epsilon \} \quad (2.9)$$

From (2.8) and (2.9) we have

$$P \left\{ \omega \in \Omega; \sup_{[\alpha, \beta]} \{X(t, \omega) - X(\alpha, \omega)\} \geq \lambda \right\} \leq 2P \{ \omega \in \Omega; \{X(\beta, \omega) - X(\alpha, \omega)\} \geq \lambda - \epsilon \} \quad (2.10)$$

since $X(\beta, \omega) - X(\alpha, \omega)$ is normally distributed its distribution function is continuous. Then letting $\epsilon \downarrow 0$ in (2.10) we obtain

$$P \left\{ \omega \in \Omega; \sup_{[\alpha, \beta]} \{X(t, \omega) - X(\alpha, \omega)\} \geq \lambda \right\} \leq 2P \{ \omega \in \Omega; \{X(\beta, \omega) - X(\alpha, \omega)\} \geq \lambda \} \quad (2.11)$$

To obtain the reverse inequality to (2.11), let $t_k = \alpha + k(\beta - \alpha)/n$, $k = 0, 1, 2, \dots, n$. For any $\lambda, \epsilon > 0$, according to the second inequality in the lemma above,

$$\begin{aligned} & P \left\{ \omega \in \Omega; \sup_{[\alpha, \beta]} \{X(t, \omega) - X(\alpha, \omega)\} \geq \lambda \right\} \\ & \geq P \left\{ \omega \in \Omega; \sup_{1 \leq k \leq n} \{X(t_k, \omega) - X(\alpha, \omega)\} \geq \lambda \right\} \\ & \geq 2P \{ \omega \in \Omega; X(\beta, \omega) - X(\alpha, \omega) \geq \lambda + 2\epsilon \} - \\ & \quad - 2 \sum_{k=1}^n P \{ \omega \in \Omega; X(\beta, \omega) - X(\alpha, \omega) \geq \lambda \} \end{aligned}$$

writing f and F for the density function and the distribution function of $N(0, 1)$, we have by lemma(2.9)

$$\begin{aligned} & P \{ \omega \in \Omega; X(t_k, \omega) - X(t_{k-1}, \omega) \geq \epsilon \} \\ & = \frac{1}{\sqrt{2\pi n^{-1}c(\beta - \alpha)}} \int_{\epsilon}^{\infty} \exp \left\{ -\frac{u^2}{2n^{-1}c(\beta - \alpha)} \right\} du \\ & = 1 - F \left(\sqrt{\frac{n}{c(\beta - \alpha)}} \epsilon \right) \leq f \left(\sqrt{\frac{n}{c(\beta - \alpha)}} \epsilon \right) \sqrt{\frac{c(\beta - \alpha)}{n}} \frac{1}{\epsilon} \\ & = \sqrt{\frac{c(\beta - \alpha)}{2\pi n}} \frac{1}{\epsilon} \exp \left\{ \frac{-n\epsilon^2}{2c(\beta - \alpha)} \right\} \end{aligned} \quad (2.12)$$

Thus

$$\begin{aligned} & \lim_{n \rightarrow \infty} \sum_{k=1}^n P \{ \omega \in \Omega; X(t_k, \omega) - X(t_{k-1}, \omega) \geq \epsilon \} \\ & \leq \lim_{n \rightarrow \infty} \sqrt{\frac{c(\beta - \alpha)n}{2\pi}} \frac{1}{\epsilon} \exp \left\{ \frac{-n\epsilon^2}{2c(\beta - \alpha)} \right\} = 0 \end{aligned}$$

and hence

$$\begin{aligned} & P \left\{ \omega \in \Omega; \sup_{[\alpha, \beta]} \{X(t, \omega) - X(\alpha, \omega)\} \geq \lambda \right\} \\ & \geq 2P \{ \omega \in \Omega; X(\beta, \omega) - X(\alpha, \omega) \geq \lambda + 2\epsilon \} \end{aligned}$$

Again letting $\epsilon \downarrow 0$ and using the continuity of the distribution function of $X(\beta, \cdot) - X(\alpha, \cdot)$ we obtain the reverse inequality to (2.11). Thus proves the equality in the theorem. The inequality in the theorem is obtained by taking $n = 1$ and $\epsilon = \lambda$ in (2.12). \square

As we have seen, almost every sample function of a separable Brownian motion process is continuous. However almost every sample function of the process is almost everywhere non-differentiable. To prove this we need the following lemma:

Lemma 2.12. *Let X be a separable Brownian motion process on a probability space (Ω, \mathcal{B}, P) and interval $D \subset \mathbb{R}$. Then for every $t_0 \in D$,*

$$P \left\{ \omega \in \Omega; \limsup_{t \rightarrow t_0} \frac{x(t, \omega) - x(t_0, \omega)}{t - t_0} = \infty \right\} = 1.$$

Proof. The limit superior in question is \mathcal{B} -measurable. This is a consequence of the separability of X . In fact if S is a countable dense subset of D relative to which X is separable then

$$\begin{aligned} \limsup_{t \rightarrow t_0} \frac{X(t, \omega) - X(t_0, \omega)}{t - t_0} &= \lim_{n \rightarrow \infty} \sup_{0 < |t - t_0| < \frac{1}{n}} \frac{X(t, \omega) - X(t_0, \omega)}{t - t_0} \\ &= \lim_{n \rightarrow \infty} \sup_{0 < |s - t_0| < \frac{1}{n}} \frac{X(s, \omega) - X(t_0, \omega)}{s - t_0} \end{aligned}$$

which is \mathcal{B} -measurable since S is a countable set.

Let us show that

$$P \left\{ \omega \in \Omega; \limsup_{t \downarrow t_0} \frac{X(t, \omega) - X(t_0, \omega)}{t - t_0} = \infty \right\} = 1 \quad (2.13)$$

by showing that, for any $\lambda > 0$,

$$P \left\{ \omega \in \Omega; \limsup_{t \downarrow t_0} \frac{X(t, \omega) - X(t_0, \omega)}{t - t_0} \geq \lambda \right\} = 1 \quad (2.14)$$

now

$$\begin{aligned} &\left\{ \omega \in \Omega; \limsup_{t \downarrow t_0} \frac{X(t, \omega) - X(t_0, \omega)}{t - t_0} \geq \lambda \right\} \\ &= \bigcap_{n=1}^{\infty} \left\{ \omega \in \Omega; \lim_{n \rightarrow \infty} \sup_{(t_0, t_0 + \frac{1}{n}] } \frac{X(t, \omega) - X(t_0, \omega)}{t - t_0} \geq \lambda \right\} \end{aligned}$$

so that

$$\begin{aligned} &P \left\{ \omega \in \Omega; \limsup_{t \downarrow t_0} \frac{X(t, \omega) - X(t_0, \omega)}{t - t_0} \geq \lambda \right\} \\ &= \lim_{n \rightarrow \infty} P \left\{ \omega \in \Omega; \sup_{(t_0, t_0 + \frac{1}{n}] } \frac{X(t, \omega) - X(t_0, \omega)}{t - t_0} \geq \lambda \right\} \end{aligned} \quad (2.15)$$

By the reflexion principle

$$\begin{aligned}
& P \left\{ \omega \in \Omega; \sup_{(t_0, t_0 + \frac{1}{n}] } \frac{X(t, \omega) - X(t_0, \omega)}{t - t_0} \geq \lambda \right\} \\
& \geq P \left\{ \omega \in \Omega; \sup_{[t_0, t_0 + \frac{1}{n}] } \{X(t, \omega) - X(t_0, \omega)\} \geq \frac{\lambda}{n} \right\} \\
& = 2P \left\{ \omega \in \Omega; \{X(t_0 + 1/n, \omega) - X(t_0, \omega)\} \geq \frac{\lambda}{n} \right\} \\
& = 2P \left\{ \omega \in \Omega; \sqrt{n} \{X(t_0 + 1/n, \omega) - X(t_0, \omega)\} \geq \frac{\lambda}{\sqrt{n}} \right\} \tag{2.16}
\end{aligned}$$

Since the probability distribution of $X(t_0 + 1/n, \cdot) - X(t_0, \cdot)$ is $N(0, c/n)$ the probability distribution of $\sqrt{n} \{X(t_0 + 1/n, \omega) - X(t_0, \omega)\}$ is $N(0, c)$ irrespective of n . Therefore, from the continuity of the distribution function F of $N(0, c)$, we have

$$\begin{aligned}
\lim 2P \left\{ \omega \in \Omega; \sqrt{n} \{X(t_0 + 1/n, \omega) - X(t_0, \omega)\} \geq \frac{\lambda}{\sqrt{n}} \right\} \\
= 2 \{1 - F(0)\} = 1. \tag{2.17}
\end{aligned}$$

Combining (2.15), (2.16) and (2.17) we obtain (2.14), which implies (2.13). We can show similarly that

$$P \left\{ \omega \in \Omega; \limsup_{t \uparrow t_0} \frac{X(t, \omega) - X(t_0, \omega)}{t - t_0} = \infty \right\} = 1$$

which together with (2.13) yields the lemma. \square

We also need the following theorem:

Theorem 2.13. *Let X be a stochastic process on a probability space (Ω, \mathcal{B}, P) and a Lebesgue or Borel measurable set $D \subset \mathbb{R}$. If every sample function of X is continuous on D then X is separable relative to every countable dense subset $S \subset D$ and is measurable with respect to $\sigma(m_D \times \mathcal{B})$ on $\sigma(\mathcal{B}_D \times \mathcal{B})$.*

The proof of this theorem can be found on [17].

We may now handle the non-differentiability theorem.

Theorem 2.14. *Let X be a separable Brownian motion process on a probability space (Ω, \mathcal{B}, P) and an interval $D \subset \mathbb{R}$. Then for a.e. $\omega \in \Omega$ the sample function $X(\cdot, \omega)$ is a.e. non-differentiable on D , with the exception of a subset of D with Lebesgue measure 0 depending on ω .*

Proof. $X(\cdot, \omega)$ is continuous on D a.e. We may assume without loss of generality that $X(\cdot, \omega)$ is continuous on D for every $\omega \in \Omega$. Then by the theorem above, X is measurable with respect to $\sigma(\mathcal{B}_D \times \mathcal{B})$ where \mathcal{B}_D is the σ -field of Borel sets contained in D .

According to the lemma above, for every $t \in D$,

$$P \left\{ \omega \in \Omega; \limsup_{\tau \rightarrow t} \left| \frac{X(\tau, \omega) - X(t, \omega)}{\tau - t} \right| = \infty \right\} = 1$$

Let

$$D(t, \omega) = \limsup_{\tau \rightarrow t} \left| \frac{X(\tau, \omega) - X(t, \omega)}{\tau - t} \right| \quad \text{for } (t, \omega) \in D \times \Omega$$

The fact that D is measurable with respect to $\sigma(\mathcal{B}_D \times \mathcal{B})$ follows from the measurability of X with respect to $\sigma(\mathcal{B}_D \times \mathcal{B})$ and the separability of X . In fact, if S is a countable dense subset of D relative to which X is separable, then

$$\begin{aligned} D(t, \omega) &= \lim_{n \rightarrow \infty} \sup_{0 < |\tau - t| < \frac{1}{n}} \left| \frac{X(\tau, \omega) - X(t, \omega)}{\tau - t} \right| \\ &= \lim_{n \rightarrow \infty} \sup_{0 < |s - t| < \frac{1}{n}} \left| \frac{X(s, \omega) - X(t, \omega)}{s - t} \right| \end{aligned}$$

which is measurable with respect to $\sigma(\mathcal{B}_D \times \mathcal{B})$. By Fubini's theorem,

$$\int_{\Omega} \left\{ \int_D \frac{1}{D(t, \omega)} m_L(dt) \right\} P(d\omega) = \int_D \left\{ \int_{\Omega} \frac{1}{D(t, \omega)} m_L P(d\omega) \right\} (dt) = 0$$

Since $[D(t, \omega)]^{-1} \geq 0$ on $D \times \Omega$ and $\int_D [D(t, \omega)]^{-1} m_L(dt) \geq 0$ on Ω , the above implies that $\int_D [D(t, \omega)]^{-1} m_L(dt) = 0$ for a.e. $\omega \in \Omega$. Then for a.e. $\omega \in \Omega$, $[D(t, \omega)]^{-1} = 0$ for a.e. $t \in D$, i.e., $D(t, \omega) = \infty$ for a.e. $t \in D$. Therefore a.e. $\omega \in \Omega$, $X(\cdot, \omega)$ is a.e. non-differentiable on D . \square

2.2.2 Brownian motion in higher dimension

Definition 2.28. A process W_t on a probability space (Ω, \mathcal{F}, P) adapted to a filtration $(\mathcal{F}_t)_{t \in \mathbb{R}^+}$ is called a Brownian motion relative to the filtration \mathcal{F}_t if

1. Sample paths $W_t(\omega)$ are continuous functions of t for almost all ω .
2. $W_0(\omega) = 0$ for almost all ω .
3. For $0 \leq s \leq t$, the increment $W_t - W_s$ is a Gaussian random variable with zero mean and variance $t - s$.

4. For $0 \leq s \leq t$, the increment $W_t - W_s$ is independent of the σ -algebra \mathcal{F}_s .

Lemma 2.13. Let X_t , $t \in \mathbb{R}^+$, be a random process such that $X_{t_0}, X_{t_1} - X_{t_0}, \dots, X_{t_k} - X_{t_{k-1}}$ are independent random variables for every $k \geq 1$ and $0 = t_0 \leq t_1 \leq \dots \leq t_k$. Then for $0 \leq s \leq t$, the increment $X_t - X_s$ is independent of the σ -algebra \mathcal{F}_s^X .

Definition 2.29. An \mathbb{R}^d -valued process $W_t = (W_t^1, \dots, W_t^d)$ is said to be a (standard) d -dimensional Brownian motion if its components W_t^1, \dots, W_t^d are independent one-dimensional Brownian motions.

An \mathbb{R}^d -valued process $W_t = (W_t^1, \dots, W_t^d)$ is said to be a (standard) d -dimensional Brownian motion relative to a filtration \mathcal{F}_t if its components W_t^1, \dots, W_t^d are independent one-dimensional Brownian motions relative to the filtration \mathcal{F}_t .

Let X be a metric space P_α a family of probability measures on the Borel σ -algebra $\mathcal{B}(X)$. The two following concepts are widely used in probability theory.

Definition 2.30. A family of probability measures P_α on $(X, \mathcal{B}(X))$ is said to be weakly compact if from any sequence P_n , $n = 1, 2, \dots$ of measures from the family one can extract a weakly convergent subsequence P_{n_k} , $k = 1, 2, \dots$ that is $P_{n_k} \Rightarrow P$ for some probability measure P .

Definition 2.31. A family of probability measures $\{P_\alpha\}$ on $(X, \mathcal{B}(X))$ is said to be tight if for any $\epsilon > 0$ one can find a compact set $K_\epsilon \subseteq X$ such that $P(K_\epsilon) \geq 1 - \epsilon$ for each $P \in P_\alpha$.

Theorem 2.15 (Prokhorov). If a family of probability measures $\{P_\alpha\}$ on a metric space X is tight, then it is weakly compact. On a separable complete metric space the two notions are equivalent.

Definition 2.32. The space $C([0, \infty])$ is the metric space which consists of all continuous real-valued functions $\omega = \omega(t)$ on $[0, \infty)$ with the metric

$$d(\omega_1, \omega_2) = \sum_{n=1}^{\infty} \frac{1}{2^n} \min \left(\sup_{0 \leq t \leq n} |\omega_1(t) - \omega_2(t)|, 1 \right).$$

Definition 2.33. Given a finite collection of points $t_1, \dots, t_k \in \mathbb{R}^+$ and a Borel set $A \in \mathcal{B}(\mathbb{R}^k)$, we define a cylindrical subset of $C([0, \infty))$ as

$$\{\omega; (\omega(t_1), \dots, \omega(t_k)) \in A\}.$$

Denote by \mathcal{B} the minimal σ -algebra that contains all the cylindrical sets (for all choices of k, t_1, \dots, t_k and A).

Lemma 2.14. *The minimal σ -algebra \mathcal{B} that contains all the cylindrical sets is the σ -algebra of Borel sets of $C([0, \infty))$.*

Proof. Let us first show that all cylindrical sets are Borel sets. All cylindrical sets belong to the minimal σ -algebra which contains all sets of the form $B = \{\omega; \omega(t) \in A\}$, where $t \in \mathbb{R}^+$ and A is open in \mathbb{R} . But B is open in $C([0, \infty))$ since, together with any $\bar{\omega} \in B$, it contains a sufficiently small ball $B(\bar{\omega}, \epsilon) = \{\omega; d(\omega, \bar{\omega}) < \epsilon\}$. Therefore, all cylindrical sets are Borel sets. Consequently, \mathcal{B} is contained in the Borel σ -algebra.

To prove the converse inclusion, note that any open set is a countable union of open balls, since the space $C([0, \infty))$ is separable. We have

$$\begin{aligned} B(\bar{\omega}, \epsilon) &= \left\{ \omega; \sum_{n=1}^{\infty} \frac{1}{2^n} \min \left(\sup_{0 \leq t \leq n} |\omega(t) - \bar{\omega}(t)|, 1 \right) < \epsilon \right\} \\ &= \left\{ \omega; \sum_{n=1}^{\infty} \frac{1}{2^n} \min \left(\sup_{0 \leq t \leq n, t \in \mathbb{Q}} |\omega(t) - \bar{\omega}(t)|, 1 \right) < \epsilon \right\} \end{aligned}$$

where \mathbb{Q} is the set of rational numbers.

The function $f(\omega) = \sum_{n=1}^{\infty} \frac{1}{2^n} \min \left(\sup_{0 \leq t \leq n, t \in \mathbb{Q}} |\omega(t) - \bar{\omega}(t)|, 1 \right)$ defined on $C([0, \infty))$ is measurable with respect to the σ -algebra \mathcal{B} generated by the cylindrical sets and, therefore, $B(\bar{\omega}, \epsilon)$ belongs to \mathcal{B} . We conclude that all open sets, and therefore, all Borel sets belong to the minimal σ -algebra which contains all cylindrical sets. \square

Lemma 2.15. *A sequence of probability measures on $(C([0, \infty)), \mathcal{B})$ converges weakly if and only if it is tight and all of its finite-dimensional distributions converge weakly.*

Proof. If P_n is a sequence of probability measures converging weakly to a measure P , then it is weakly compact, and therefore tight by the Prokhorov theorem.

To prove the converse statement assume that a sequence of measures is tight, and the finite-dimensional distributions converge weakly. For each $k \geq 1$ and t_1, \dots, t_k let $\tilde{P}_n^{t_1, \dots, t_k}$ be the finite dimensional distribution of the measure P_n , and μ_{t_1, \dots, t_k} be the measure on \mathbb{R}^k such that $\tilde{P}_n^{t_1, \dots, t_k} \rightarrow \mu_{t_1, \dots, t_k}$ weakly.

Again by the Prokhorov theorem, there is a subsequence P'_m of the original sequence converging weakly to a measure P . If a different subsequence P''_n converges weakly to a measure Q , then P and Q have the same finite dimensional distributions (namely μ_{t_1, \dots, t_k})

and, therefore, must coincide. Let us demonstrate that the original sequence P_n converges to the same limit. If this is not the case there exists a bounded continuous function f on $C([0, \infty))$ and a subsequence \bar{P}_n such that $\int f d\bar{P}_n$ do not converge to $\int f dP$. Then one can find a subsequence $\bar{\bar{P}}_n$ of \bar{P}_n such that $\left| \int f d\bar{\bar{P}}_n - \int f dP \right| > \epsilon$ for some $\epsilon > 0$ and all n . On the other hand, the sequence $\bar{\bar{P}}_n$ is tight and contains a subsequence that converges to P . This leads to a contradiction, and therefore P_n converges to P . \square

Definition 2.34. A set of functions $A \subseteq C([0, \infty))$ is called equicontinuous on the interval $[0, T]$ if

$$\limsup_{\delta \rightarrow 0} m^T(\omega, \delta) = 0$$

It is called uniformly bounded on the interval $[0, T]$ if it is bounded in the $C([0, T])$ norm, that is

$$\sup_{\omega \in A} \sup_{0 \leq t \leq T} |\omega(t)| < \infty$$

Theorem 2.16 (Arzela-Ascoli). A set $A \subseteq C([0, \infty))$ has compact closure if and only if it is uniformly bounded and equicontinuous on every interval $[0, T]$.

Theorem 2.17. A sequence P_n of probability measures on $(C([0, \infty)), \mathcal{B})$ is tight if and only if the following two conditions hold:

- For any $T > 0$ and $\eta > 0$ there is a $a > 0$ such that

$$P_n \left(\left\{ \omega; \sup_{0 \leq t \leq T} |\omega(t)| > a \right\} \right) \leq \eta, \quad n \geq 1.$$

- For any $T > 0$, $\eta > 0$ and $\epsilon > 0$, there is $\delta > 0$ such that

$$P_n \left(\left\{ \omega; m^T(\omega, \delta) > \epsilon \right\} \right) \leq \eta, \quad n \geq 1.$$

Definition 2.35. A probability measure \mathcal{W} on $(C([0, \infty)), \mathcal{B})$ is called the Wiener measure if the coordinate process $W_t(\omega) = \omega(t)$ on $(C([0, \infty)), \mathcal{B}, \mathcal{W})$ is a Brownian motion relative to the filtration \mathcal{F}_t^W .

To show that \mathcal{W} is the Wiener measure, it is sufficient to show that the increments of the coordinate process $W_t - W_s$ are independent Gaussian variables with respect to \mathcal{W} , with zero mean, variance $t - s$, and $W_0 = 0$ a.s. Note that a measure which has these properties is unique.

Let ξ_1, ξ_2, \dots be a sequence of independent identically distributed (*i.i.d.*) random variables on a probability space (Ω, \mathcal{F}, P) . We assume that the expectation of each of the variables is equal to zero and the variance is equal to one. Let S_n be the partial sums, that is $S_0 = 0$ and $S_n = \sum_{i=1}^n \xi_i$ for $n \geq 1$. We define a sequence of measurable functions $X_t^n : \Omega \rightarrow C([0, \infty))$ via

$$X_t^n(\omega) = \frac{1}{\sqrt{n}} S_{[n,t]}(\omega) + (nt - [nt]) \frac{1}{\sqrt{n}} \xi_{[nt]+1}(\omega)$$

where $[t]$ stands for the integer part of t . One can think of X_t^n as a random walk with steps of order $\frac{1}{\sqrt{n}}$ and the time steps of size $\frac{1}{n}$.

Theorem 2.18 (Donsker). *The measures on $C([0, \infty))$ induced by X_t^n converge weakly to the Wiener measure.*

The proof will be based on a sequence of lemmas.

Lemma 2.16. *For $0 \leq t_1 \leq \dots \leq t_k$,*

$$\lim_{n \rightarrow \infty} (X_{t_1}^n, \dots, X_{t_k}^n) = (\eta_{t_1}, \dots, \eta_{t_k})$$

in distribution, where $(\eta_{t_1}, \dots, \eta_{t_k})$ is a Gaussian vector with zero mean and covariance matrix $E\eta_{t_i}\eta_{t_j} = t_j \wedge t_i$.

Proof. It is sufficient to demonstrate that the vector $(X_{t_1}^n, X_{t_2}^n - X_{t_1}^n, \dots, X_{t_k}^n - X_{t_{k-1}}^n)$ converges to a vector of independent Gaussian variables with variances $t_1, t_2 - t_1, \dots, t_k - t_{k-1}$. Since the term $(nt - [nt]) \frac{1}{\sqrt{n}} \xi_{[nt]+1}$ converges to zero in probability for every t , it is sufficient to establish the convergence to a Gaussian vector for

$$(V_1^n, \dots, V_k^n) = \left(\frac{1}{\sqrt{n}} S_{[nt_1]}, \frac{1}{\sqrt{n}} S_{[nt_2]} - \frac{1}{\sqrt{n}} S_{[nt_1]}, \dots, \frac{1}{\sqrt{n}} S_{[nt_k]} - \frac{1}{\sqrt{n}} S_{[nt_{k-1}]} \right)$$

Each of the components converges to a Gaussian random variable by the Central Limit Theorem for *i.i.d.* random variables. Let us write $\xi_j = \lim_{n \rightarrow \infty} V_j^n$, and let $\varphi_j(\lambda_j)$ be the characteristic function of ξ_j . Thus, $\varphi_1(\lambda_1) = e^{-\frac{t_1 \lambda_1^2}{2}}$, $\varphi_2(\lambda_2) = e^{-\frac{(t_2 - t_1) \lambda_2^2}{2}}$, etc.

In order to show that the vector (V_1^n, \dots, V_k^n) converges to a Gaussian vector, it is sufficient to consider characteristic function $\varphi^n(\alpha_1, \dots, \alpha_k) = E e^{i(\lambda_1 V_1^n + \dots + \lambda_k V_k^n)}$. Due to independence of the components of the vector (V_1^n, \dots, V_k^n) , the characteristic function $\varphi^n(\lambda_1, \dots, \lambda_k)$ is equal to the product of the characteristic functions of the components, and thus converges to $\varphi_1(\lambda_1) \cdot \dots \cdot \varphi_k(\lambda_k)$, which is the characteristic function of a Gaussian vector with independent components. \square

Lemma 2.17. *A sequence P_n of probability measures on $(C([0, \infty)))$ is tight if the following conditions hold:*

- *For any $\eta > 0$, there is $a > 0$ such that*

$$P_n(\{\omega; |\omega(0)| > a\}) \leq \eta, \quad n \geq 1.$$

- *For any $T > 0, \eta > 0$ and $\epsilon > 0$ there are $0 < \delta < 1$ and integer n_0 such that, for all $t \in [0, T]$, we have*

$$P_n \left(\left\{ \omega; \sup_{t \leq s \leq \min(t+\delta, T)} |\omega(s) - \omega(t)| > \epsilon \right\} \right) \leq \delta\eta, \quad n \geq n_0.$$

Proof. see [10] □

We now wish to apply the lemma to the sequence of measures induced by X_t^n . Since $X_0^n = 0$ a.s., we only need to verify the second assumption of the lemma. We need to show that for any $T > 0$, $\eta > 0$ and $\epsilon > 0$ there are $0 < \delta < 1$ and an integer n_0 such that for all $t \in [0, T]$ we have

$$P \left(\left\{ \sup_{t \leq s \leq \min(t+\delta, T)} |X_s^n - X_t^n| \geq \epsilon \right\} \right) \leq \delta\eta, \quad n \geq n_0.$$

Since the value of X_t^n changes linearly when t is between integer multiples of $\frac{1}{n}$, and the interval $[t, t + \delta]$ is contained inside the interval $\left[\frac{k}{n}, \frac{k+[n\delta+2]}{n} \right]$ for some integer k , it is sufficient to check that for $T > 0$, $\eta > 0$ and $\epsilon > 0$, there are $0 < \delta < 1$ and an integer n_0 such that

$$P \left(\left\{ \omega; \max_{k \leq i \leq k+[n\delta+2]} \frac{1}{\sqrt{n}} |S_i - S_k| > \frac{\epsilon}{2} \right\} \right) \leq \delta\eta, \quad n \geq n_0$$

for all k . Obviously, we can replace $\epsilon/2$ by ϵ and $[n\delta + 2]$ by $[n\delta]$. Thus, it is sufficient to show that

$$P \left(\left\{ \omega; \max_{k \leq i \leq k+[n\delta]} \frac{1}{\sqrt{n}} |S_i - S_k| > \epsilon \right\} \right) \leq \delta\eta, \quad n \geq n_0$$

Lemma 2.18. *For any $\epsilon > 0$, there is $\lambda > 1$ such that*

$$\limsup_{n \rightarrow \infty} P \left(\max_{i \leq n} |S_i| > \lambda\sqrt{n} \right) \leq \frac{\epsilon}{\lambda^2}$$

Donsker's Theorem. We have established that the sequence of measures induced by X_t^n is tight. The finite dimensional distributions converge by Lemma(2.16). Therefore, the sequence of measures induced by X_t^n converges weakly to a probability measure, which we shall denote by \mathcal{W} . By Lemma(2.16) the limiting measure \mathcal{W} satisfies the requirements of the definition of Wiener's measure. □

Chapter 3

Stochastic Integrals

We now begin the study of stochastic integrals. As we did earlier, we present two versions of the theory. The first, the simpler case, is built upon the Riemann-Stieltjes integral and could in fact be considered a generalization of this integral. The second version, which follows a more common approach, leads to the Itô formula and consequently to stochastic differential equations.

Let X be a continuous stochastic process on a probability space (Ω, \mathcal{B}, P) and an interval $D = [a, b]$. If f is a real-valued function of bounded variation, then the Riemann-Stieltjes integral $\int_a^b f(t) dX(t, \omega)$ exists as a real number for every $\omega \in \Omega$.

Let $B = B(t, \omega)$, $(t, \omega) \in D \times \Omega$ be a Brownian motion process on a probability space (Ω, \mathcal{B}, P) . And an interval $D = [a, b]$ or $[a, \infty)$ where a is finite and b may be infinite. In addition let $B(a, \cdot)$ be equal to a constant a.e. on Ω . Thus

1. For $\{t_0, t_1, \dots, t_n\} \subset D$, $t_0 \leq t_1 \leq \dots \leq t_n$, the collection of random variables $\{B(t_j, \cdot) - B(t_{j-1}, \cdot), j = 1, 2, \dots, n\}$ is an independent system.
2. For $t', t'' \in D$, $t' < t''$ the probability distribution of $B(t'', \cdot) - B(t', \cdot)$ is given by $N(0, t'' - t')$.
3. $B(a, \cdot) = c$ a.e. on Ω where $c \in \mathbb{R}$.

Consider the real Hilbert space $L_2(\Omega) = L_2(\Omega, \mathcal{B}, P)$ where an element of $L_2(\Omega)$ is an equivalence class of random variables X on (Ω, \mathcal{B}, P) with $E(X^2) < \infty$ relative to the equivalence relation of a.e. equality on Ω and where the inner product is defined by $\langle X, Y \rangle = E(XY)$ for $X, Y \in L_2(\Omega)$. In terms of this inner product and the associated Hilbert norm $\|\cdot\|$ we have

- $E(X^2) = \langle X, X \rangle = \|X\|^2$,
- $E(X) = \langle X, 1 \rangle$,
- $C(X, Y) = \langle X - E(X), Y - E(Y) \rangle$
- $V(X) = \|X - E(X)\|^2$

Consider also the real Hilbert space $L_2(D) = L_2(D, \mathcal{M}_D, m_l)$ (where \mathcal{M}_D is the σ -algebra of Lebesgue measurable sets contained in D) with inner product defined by

$$\langle f, g \rangle = \int_D f(t)g(t)m_l(dt) \quad \text{for } f, g \in L_2(D)$$

Let $S(D)$ be the collection of these elements of $L_2(D)$ which can be represented by step functions of compact support on D .

We shall define our stochastic integral as an isometric linear transformation of $S(D)$ into $L_2(\Omega)$ first and then extend the domain of definition to $L_2(D)$.

Definition 3.1. Let $f \in S(D)$ be represented as a step function as follows

$$f = \sum_{k=1}^n c_k \chi_{J_k} \tag{3.1}$$

where $\{c_k, k = 1, 2, \dots, n\} \subset \mathbb{R}$ and the interval J_k has t_k and t_{k+1} as its endpoints and $a \leq t_1 < t_2 < \dots < t_{n+1} \leq b$. The Stochastic Integral $I(f)$ of f with respect to the Brownian motion process B is a random variable on (Ω, \mathcal{B}, P) defined by

$$I(f)(\omega) = \sum_{k=1}^n c_k \{B(t_{k+1}, \omega) - B(t_k, \omega)\} \quad \text{for } \omega \in \Omega \tag{3.2}$$

Note that $I(f)(\omega)$ is defined for every $\omega \in \Omega$ and that $I(f)$ is determined by $f \in S(D)$ but is independent of the representation of f as a step function.

Theorem 3.1. For $f, g \in S(D)$ and $\alpha, \beta \in \mathbb{R}$ the following hold:

1. $\langle I(f), 1 \rangle = E[I(f)] = 0$
2. $\langle I(f), I(g) \rangle = C(I(f), I(g)) = \langle f, g \rangle$
3. $\|I(f)\|^2 = V[I(f)] = \|f\|^2$
4. $I(f)$ is distributed according to $N(0, \|f\|^2)$

5. $I(\alpha f + \beta g)(\omega) = \alpha I(f)(\omega) + \beta I(g)(\omega)$ for every $\omega \in \Omega$.

Proof. To prove (1) note that if $f \in S(D)$ is represented as in (3.1), then by (3.2),

$$E[I(f)] = \sum_{k=1}^n c_k E[B(t_{k+1}, \cdot) - B(t_k, \cdot)] = 0$$

the term in brackets on the right-hand side equals zero due to the property of Brownian motion.

To prove (2) represent $f, g, \in S(D)$ by step functions as

$$f = \sum_{k=1}^n c'_k \chi_{J_k} \quad \text{and} \quad g = \sum_{k=1}^n c''_k \chi_{J_k} \quad (3.3)$$

where $\{c'_k, c''_k, k = 1, 2, \dots, n\} \subset \mathbb{R}$ and χ denotes the indicator function. Then

$$\begin{aligned} C(I(f), I(g)) &= E \left[\sum_{k=1}^n c'_k \{B(t_{k+1}, \cdot) - B(t_k, \cdot)\} \cdot \sum_{k=1}^n c''_k \{B(t_{k+1}, \cdot) - B(t_k, \cdot)\} \right] \\ &= E \left[\sum_{k=1}^n c'_k c''_k \{B(t_{k+1}, \cdot) - B(t_k, \cdot)\}^2 \right] + \\ &+ E \left[\sum_{k=1}^n c'_j c''_k \{B(t_{j+1}, \cdot) - B(t_j, \cdot)\} \{B(t_{k+1}, \cdot) - B(t_k, \cdot)\} \right] \\ &= \sum_{k=1}^n c'_k c''_k (t_k - t_{k+1}) = \int_D f(t)g(t)m_L(dt) = \langle f, g \rangle \end{aligned}$$

which is (2). (3) is a particular case of (2).

Since $\{B(t, \cdot), t \in D\}$ is a Gaussian system of random variables, the right-hand side of (3.2) is normally distributed. This together with (3.1) and (3.3) proves (4).

To prove (5) we use the representation (3.3) and write

$$\begin{aligned} I(\alpha f + \beta g)(\omega) &= \sum_{k=1}^n (\alpha c'_k + \beta c_k) \{B(t_{k+1}, \omega) - B(t_k, \omega)\} \\ &= \alpha \sum_{k=1}^n c'_k \{B(t_{k+1}, \omega) - B(t_k, \omega)\} + \beta \sum_{k=1}^n c_k \{B(t_{k+1}, \omega) - B(t_k, \omega)\} \\ &= \alpha I(f)(\omega) + \beta I(g)(\omega) \end{aligned}$$

□

Definition 3.2. Let $f \in L_2(D)$. From the denseness of $S(D)$ in $L_2(D)$ there exists a sequence $\{f_n, n = 1, 2, \dots\} \subset S(D)$ satisfying $\lim_{n \rightarrow \infty} \|f_n - f\| = 0$. By items (3) and (5) of the above theorem,

$$\|I(f_m) - I(f_n)\| = \|f_m - f_n\|$$

so the fact that $\{f_n, n = 1, 2, \dots\}$ is a Cauchy sequence in $L_2(D)$ implies that $\{I(f_n), n = 1, 2, \dots\}$ is a Cauchy sequence in $L_2(\Omega)$. Thus there exists $X \in L_2(\Omega)$ such that

$$\lim_{n \rightarrow \infty} \|I(f_n) - X\| = 0$$

we call X the Stochastic Integral of f with respect to the Brownian motion process B and write $I(f)$ for it. Thus

$$I(f) = s - \lim_{n \rightarrow \infty} I(f_n)$$

where $s - \lim$ denotes the limit of strong convergence (i.e. convergence in the norm of $L_2(\Omega)$).

Theorem 3.2. Let $f, g \in L_2(D)$ and $\alpha, \beta \in \mathbb{R}$. Then items (1), (2), (3) and (4) of the above theorem hold. Instead of (5) we have

$$I(\alpha f + \beta g)(\omega) = \alpha I(f)(\omega) + \beta I(g)(\omega) \quad \text{for a.e. } \omega \in \Omega$$

The collection of random variables $\{I(f), f \in L_2(D)\}$ is a Gaussian system.

Proof. (1), (2) and (3) are immediate. Let us show (2) as an example. Thus let $\{f_n, n = 1, 2, \dots\}$, $\{g_n, n = 1, 2, \dots\} \subset S(D)$ and $f = s - \lim_{n \rightarrow \infty} f_n$, $g = s - \lim_{n \rightarrow \infty} g_n$. Then

$$\begin{aligned} \langle I(f), I(g) \rangle &= \langle s - \lim_{n \rightarrow \infty} I(f_n), s - \lim_{n \rightarrow \infty} I(g_n) \rangle = \lim_{n \rightarrow \infty} \langle I(f_n), I(g_n) \rangle \\ &= \lim_{n \rightarrow \infty} \langle f_n, g_n \rangle = \langle s - \lim_{n \rightarrow \infty} f_n, s - \lim_{n \rightarrow \infty} g_n \rangle = \langle f, g \rangle \end{aligned}$$

which proves (2).

Regarding (4), since $I(f) = s - \lim_{n \rightarrow \infty} I(f_n)$, the probability distribution of $I(f_n)$ converges to that of $I(f)$. But $I(f_n)$ is normally distributed. Thus, $I(f)$ is normally distributed.

To prove (5) let f, g, f_n, g_n be as in the proof of (2). Then for $\alpha, \beta \in \mathbb{R}$ we have $\alpha f_n + \beta g_n \in S(D)$ for $n = 1, 2, \dots$ and

$$\lim_{n \rightarrow \infty} \|(\alpha f_n + \beta g_n) - (\alpha f + \beta g)\| = 0$$

so that

$$\begin{aligned} I(\alpha f + \beta g) &= s - \lim_{n \rightarrow \infty} I(\alpha f_n + \beta g_n) = s - \lim_{n \rightarrow \infty} \alpha I(f_n) + s - \lim_{n \rightarrow \infty} \beta I(g_n) \\ &= \alpha I(f) + \beta I(g). \end{aligned}$$

Thus (5) holds.

Since $\{I(f), f \in S(D)\}$ is a collection of linear combinations of numbers of the Gaussian System $\{B(t, \cdot), t \in D\}$, it is a Gaussian System. Then since $\{I(f), f \in L_2(D)\}$ is a collection of strong limits of sequences of numbers of the Gaussian system $\{I(f), f \in S(D)\}$ it is a Gaussian System. \square

The following theorem relates our stochastic integral to the Riemann-Stieltjes integral.

Theorem 3.3. *If the Brownian motion process B is continuous on $[a, b]$ and f is of bounded variation on $[a, b]$ then*

$$I(f)(\omega) = \int_a^b f(t) dB(t, \omega) \quad \text{for a.e. } \omega \in \Omega$$

Proof. Consider first the case where f is real valued and monotone increasing on $[a, b]$. Let

$$M = f(b) - f(a)$$

and

$$D_{n,k} = \left\{ t \in [a, b]; f(a) + \frac{k}{n}M \leq f(t) < f(a) + \frac{k+1}{n}M \right\}$$

for $k = 0, 1, 2, \dots, n-1$.

Since f is monotone, each $D_{n,k}$ is either an interval or a point or \emptyset . If $D_{n,k}$ is a point adjoin it to $D_{n,k-1}$ or $D_{n,k+1}$. In this way we have a decomposition of $[a, b]$ into finitely many intervals. If necessary, decompose the intervals further so that the lengths of the resulting intervals $J_{n,k}$, $k = 1, 2, \dots, p(n)$, with endpoints $t_{n,k}$ and $t_{n,k+1}$ do not exceed $\frac{b-a}{n}$. Let

$$f_n = \sum_{k=1}^{p(n)} f(t_{n,k}) \chi_{J_{n,k}} \quad \text{for } n = 1, 2, \dots$$

Then $\{f_n, n = 1, 2, \dots\} \subset S(D)$ and

$$|f_n(t) - f(t)| \leq 2\frac{M}{n} \quad \text{for } t \in [a, b] \quad \text{and } n = 1, 2, \dots$$

so that

$$\lim_{n \rightarrow \infty} \|f_n - f\| = \lim_{n \rightarrow \infty} \int_{[a,b]} |f_n(t) - f(t)|^2 m_L(dt) \leq \lim_{n \rightarrow \infty} \frac{4M^2}{n^2} (b-a) = 0.$$

Thus, according to the above definition

$$I(f) = s - \lim_{n \rightarrow \infty} I(f_n)$$

and hence there exists a subsequence $\{f_{n_m}, m = 1, 2, \dots\}$ such that

$$I(f)(\omega) = \lim_{m \rightarrow \infty} I(f_{n_m})(\omega) \quad \text{for } a.e. \omega \in \Omega$$

Now, according to the definition of $I(t)$:

$$I(f_{n_m})(\omega) = \sum_{k=1}^{p(n)} f(t_{n_m, k}) \{B(t_{n_m, k+1}, \omega) - (B_{n_m, k}, \omega)\}$$

which approximates the Riemann-Stieltjes integral $\int_a^b f(t) dB(t, \omega)$ which exists for every $\omega \in \Omega$. Thus

$$\lim_{m \rightarrow \infty} I(f_{n_m})(\omega) = \int_a^b f(t) dB(t, \omega) \quad \text{for every } \omega \in \Omega$$

and hence

$$I(f)(\omega) = \int_a^b f(t) dB(t, \omega) \quad \text{for } a.e. \omega \in \Omega$$

This proves the theorem for the case where f is monotone increasing. If f is of bounded variation we express it as the difference of two monotone increasing functions and apply (5) of the preceding theorem. \square

Theorem 3.4. *The stochastic integral I transforms $L_2(D)$ one-to-one into $L_2(\Omega)$. The image R of $L_2(D)$ under the transformation I is a closed linear subspace of $L_2(\Omega)$ and I is an isomorphism between $L_2(D)$ and R as Hilbert spaces.*

Proof. Since I transforms $f \in L_2(D)$ into a random variable whose probability distribution is given by $N(0, \|f\|^2)$, $0 \in L_2(D)$ is the only element of $L_2(D)$ that is transformed into $0 \in L_2(\Omega)$. This proves the one-to-one property of the transformation I .

From the linearity of I , R is a linear subspace of $L_2(\Omega)$. To show that R is closed we show that if $\{X_n, n = 1, 2, \dots\} \subset R$, $X \in L_2(\Omega)$, and $X = s - \lim_{n \rightarrow \infty} X_n$ then $X \in R$. Now for $X_n \in R$ there exists $f_n \in L_2(D)$ such that $X_n = I(f_n)$. Since the transformation I preserves metric, so does the transformation I^{-1} of R onto $L_2(D)$. Then the fact that $\{X_n, n = 1, 2, \dots\}$ is a Cauchy sequence in R implies that $\{f_n, n = 1, 2, \dots\}$ is a Cauchy sequence in $L_2(D)$ and hence the existence of $f \in L_2(D)$ such that $\lim_{n \rightarrow \infty} \|f_n - f\| = 0$.

Since the transformation of I of $L_2(D)$ into $L_2(\Omega)$ preserves metric we have $\lim_{n \rightarrow \infty} \|I(f_n) - I(f)\| = 0$, *i.e.*

$$X = s - \lim_{n \rightarrow \infty} X_n = s - \lim_{n \rightarrow \infty} I(f_n) = I(f) \in R.$$

This proves the closedness of R .

As a closed linear subspace of the Hilbert space $L_2(\Omega)$, R is itself a Hilbert space. The fact that I is an isomorphism between $L_2(D)$ and R as Hilbert spaces follows from theorem 3.2. \square

Corollary 3.1. *For $\{f_\alpha, \alpha \in A\} \subset L_2(D)$ the following hold:*

1. $\{f_\alpha, \alpha \in A\}$ is an orthonormal system in $L_2(D)$ if and only if $I(f_\alpha)\alpha \in A$ is an orthonormal system in R , and the former is complete if and only if the latter is.
2. $\{f_\alpha, \alpha \in A\}$ is an orthonormal system in $L_2(D)$ if and only if $I(f_\alpha), \alpha \in A$ is an independent system of random variables.

Proof. (1) is implied by the isomorphism of $L_2(D)$ and R as Hilbert spaces. (2) is from the fact that $\{f_\alpha, \alpha \in A\}$ is an orthogonal system in $L_2(D)$ if and only if $\{I(f_\alpha)\alpha \in A\}$ is an orthogonal system in $L_2(\Omega)$ and the following corollary \square

Corollary 3.2. *Let $\{X_\alpha, \alpha \in A\}$ be a Gaussian system of random variables on a probability space (Ω, \mathcal{B}, P) with $E(X_\alpha) = 0$ for every $\alpha \in A$. Then $\{X_\alpha, \alpha \in A\}$ is an independent system of random variables if and only if it is an orthogonal system in the Hilbert space $L_2(\Omega)$.*

Proof. Since the independence of an infinite system of random variables is defined as the independence of every finite subsystem, it is sufficient to prove the corollary for a finite Gaussian system $\{X_j, j = 1, 2, \dots\}$ with $E(X_j) = 0$ for $j = 1, 2, \dots, n$. Then since $\langle X_j, X_l \rangle = C(X_j, X_l)$, the system is an independent system if and only if it is an orthogonal system. \square

We now proceed to our second version of stochastic integrals.

We shall now define stochastic integrals of the form $\int_{[0,t]} X dM$ where M is a right continuous local L^2 Martingale.

Definition 3.3. *The family of subsets of $\mathbb{R}_+ \times \Omega$ containing all sets of the form $0 \times F_0$ and $(s, t] \times F$, where $F_0 \in \mathcal{F}_0$ and $F \in \mathcal{F}_s$ for $s < t$ in \mathbb{R}_+ , is called the class of predictable rectangles and is denoted by \mathcal{R} .*

Definition 3.4. *The σ -field \mathcal{P} of subsets of $\mathbb{R}_+ \times \Omega$ generated by \mathcal{R} is called the predictable σ -field and sets in \mathcal{P} are called predictable sets.*

Definition 3.5. A function $X : \mathbb{R}_+ \times \Omega \rightarrow \mathbb{R}$ is called predictable if X is \mathcal{P} -measurable.

Definition 3.6. For optional times η and τ , the set

$$[\eta, \tau] = \{(t, \omega) \in \mathbb{R}_+ \times \Omega : \eta(\omega) \leq t < \tau(\omega)\}$$

is called a stochastic interval.

Lemma 3.1. Stochastic intervals of the form $[0, \tau]$ and $(\eta, \tau]$ are predictable.

Proposition 3.1. Every optional time is predictable if and only if every local martingale adapted to $\{\mathcal{F}_t\}$ has a continuous version.

Suppose that $Z = \{Z_t, t \in \mathbb{R}_+\}$ is a real-valued process adapted to the standard filtration $\{\mathcal{F}_t, t \in \mathbb{R}_+\}$ and $Z_t \in L^1$ for each $t \in \mathbb{R}_+$.

We define a set function λ_Z in R by

$$\begin{aligned} \lambda_Z((s, t] \times F) &= E(I_F(Z_t - Z_s)) \quad \text{for } F \in \mathcal{F}_s \text{ and } s < t \text{ in } \mathbb{R}_+ \\ \lambda_Z(0 \times F_0) &= 0 \quad \text{for } F_0 \in \mathcal{F}_0. \end{aligned} \quad (3.4)$$

We extend λ_Z to be a finitely additive set function on the ring \mathcal{A} generated by R by defining

$$\lambda_Z(A) = \sum_{j=1}^n \lambda_Z(R_j)$$

for any $A = \bigcup_{j=1}^n R_j$, where $\{R_j, 1 \leq j \leq n\}$ is a finite collection of disjoint sets in \mathcal{R} . The value of $\lambda_Z(A)$ is the same for all representations of A as a finite disjoint union of sets in \mathcal{R} . We call λ_Z a content if $\lambda_Z \geq 0$ on \mathcal{R} and hence on \mathcal{A} .

It is clear that if Z is a martingale then $\lambda_Z \equiv 0$, and if Z is a sub-martingale then $\lambda_Z \geq 0$. In particular suppose $M = \{M_t, t \in \mathbb{R}_+\}$ is a L^2 -Martingale, then $(M)^2 = \{(M_t)^2, t \in \mathbb{R}_+\}$ is a sub-martingale and hence $\lambda_{(M)^2} \geq 0$. More explicitly, for $F \in \mathcal{F}_s$ and $s < t_1$

$$\lambda_{(M)^2}((s, t] \times F) = E \{I_F(M_t - M_s)^2\} \quad (3.5)$$

we use μ_M to denote the unique measure on \mathcal{P} which extends $\lambda_{(M)^2}$. This is called the Doléans measure of M .

First we define the stochastic integral $\int X dM$ when X is an \mathcal{R} -simple process and show that the map $X \rightarrow \int X dM$ is an isometry from a subspace of \mathcal{L}^2 into L^2 . This isometry is the key to the extension of the definition to all X in \mathcal{L}^2 .

When X is the indicator function of a predictable rectangle, the integral $\int X dM$ is defined as follows. For $s < t$ in \mathbb{R}_+ and $F \in \mathcal{F}_s$,

$$\int I_{(s,t] \times F} dM \equiv I_F(M_t - M_s)$$

and for $F_0 \in \mathcal{F}_0$,

$$\int I_{0 \times F_0} dM \equiv 0$$

Let ε denote the class of all functions $X : \mathbb{R}_+ \times \Omega \rightarrow \mathbb{R}$ that are finite linear combinations of indicator functions of predictable rectangles. Such a function will be called an \mathcal{R} -simple process.

Thus, $X \in \varepsilon$ can be expressed in the form

$$X = \sum_{j=1}^n c_j I_{(s_j, t_j] \times F_j} + c_0 I_{0 \times F_0}$$

where $c_j \in \mathbb{R}$, $F_j \in \mathcal{F}_{s_j}$, $s_j < t_j$ in \mathbb{R}_+ for $1 \leq j \leq n$, $n \in \mathbb{N}$, $c_0 \in \mathbb{R}$ and $F_0 \in \mathcal{F}_0$.

The integral $\int X dM$ for $X \in \varepsilon$ is defined by linearity. Thus, for X of the above form we have

$$\int X dM \equiv \sum_{j=1}^n c_j I_{F_j}(M_{t_j} - M_{s_j})$$

Since $I_{\mathcal{R}} \in \mathcal{L}^2$ for any predictable rectangle \mathcal{R} , it follows that ε is a subspace of \mathcal{L}^2 ; and since $M_t \in L^2$ for each t , $\int X dM$ is in L^2 for each $X \in \varepsilon$. The following theorem shows that the linear map $X \rightarrow \int X dM$ is an isometry from $\varepsilon \subset \mathcal{L}^2$ onto its image in L^2 .

Theorem 3.5. *For $X \in \varepsilon$ we have the isometry*

$$E \left\{ \left(\int X dM \right)^2 \right\} = \int_{\mathbb{R}_+ \times \Omega} (X)^2 d\mu_M$$

Proof. See [2] □

If we regard \mathcal{L}^2 and L^2 as Hilbert spaces, then the map $X \rightarrow \int X dM$ is a linear isometry from the dense subspace ε of \mathcal{L}^2 into L^2 . For $X \in \mathcal{L}^2$, we define $\int X dM$ as the image of X under this isometry.

We denote the space of all $X \in \mathcal{P}$ such that $I_{[0,t]} X \in \mathcal{L}^2$ for each $t \in \mathbb{R}_+$ by $\Lambda^2(\mathcal{P}, M)$ and for each t let $Y_t = \int I_{[0,t]} X dM$. Then $Y = Y_t, t \in \mathbb{R}_+$ is a zero mean L^2 -Martingale and there is a version of Y with all paths right continuous.

Theorem 3.6. *Suppose the hypothesis of the theorem above hold and M has continuous paths. Then there is a version of Y with continuous paths.*

Theorem 3.7. Let $X \in \Lambda^2(\mathcal{P}, M)$ and let Y denote the right continuous stochastic integral process $\left\{ \int_{[0,t]} X dM, t \in \mathbb{R}_+ \right\}$. Then the following properties hold:

1. For $s < t$ in \mathbb{R}_+ and any random variable $Z \in b\mathcal{F}_s$, we have $I_{(s,t]}Z \in \mathcal{P}$, $I_{(s,t]}ZX \in \Lambda^2(\mathcal{P}, M)$, and a.s.

$$\int I_{(s,t]}ZX dM = Z \int_{(s,t]} X dM$$

2. The measure μ_Y associated with the right continuous L^2 -Martingale Y has density $(X)^2$ with respect to μ_M , i.e., for any $A \in \mathcal{P}$,

$$\mu_Y(A) = \int_A (X)^2 d\mu_M$$

3. For any bounded optional time τ ,

$$Y_\tau \equiv \int_{[0,\tau]} X dM = \int I[0,t]X dM \quad a.s.$$

The proofs of the last two theorems can be found in [2]

3.1 The Itô Formula

We now study one of the most important results in stochastic integration. Before defining this formula and explaining the intuition behind it we need a few definitions and results.

Definition 3.7. For $t \in \mathbb{R}_+$, a partition π_t of $[0, t]$ is a finite ordered subset $\pi_t = \{t_0, t_1, \dots, t_k\}$ of $[0, t]$ such that $0 = t_0 < t_1 < \dots < t_k = t$. We denote the mesh of π_t by

$$\delta\pi_t \equiv \max \{|t_{j+1} - t_j|, j = 0, 1, \dots, k-1\}.$$

If $\{\pi_t^n, n \in \mathbb{N}\}$ is a sequence of partitions of $[0, t]$, then for each n , let t_j^n , $j = 0, 1, \dots, k_n$ denote the numbers of π_t^n .

Theorem 3.8. Let $t \in \mathbb{R}_+$ and $\{\pi_t^n, n \in \mathbb{N}\}$ be a sequence of partitions of $[0, t]$ such that $\lim_{n \rightarrow \infty} \delta\pi_t^n = 0$. Suppose M is a continuous local martingale and for each n let

$$S_t^n = \sum_j (M_{t_{j+1}^n} - M_{t_j^n})^2$$

where the sum is over all j such that both t_j and t_{j+1} are in π_t^n . Then

1. if M is bounded, $\{S_t^n, n \in \mathbb{N}\}$ converges in L^2 to

$$[M]_t \equiv (M_t)^2 - (M_0)^2 - 2 \int_0^t M dM \quad (3.6)$$

2. $\{S_t^n, n \in \mathbb{N}\}$ converges in probability to $[M]_t$. We call $[M]_t$, defined by (3.6), the quadratic variation of M at time t , and $[M] = \{[M]_t, t \in \mathbb{R}_+\}$ the quadratic variation process associated with M .

In the case where M is a Brownian motion B in \mathbb{R} , the quadratic variation process $[B]$ is indistinguishable from $\{t, t \in \mathbb{R}_+\}$.

We are now able to proceed to one of the most important results in the theory of stochastic integrals, the rule of change of variables, known as the Itô formula. Essentially it states that if M is a continuous local Martingale and f is a twice continuously differentiable real-valued function on \mathbb{R} , then

$$f(M_t) - f(M_0) = \int_0^t f'(M_s) dM_s + \frac{1}{2} \int_0^t f''(M_s) d[M]_s.$$

Comparing this with the Fundamental Theorem of Calculus we see that here we have an additional term involving the quadratic variation process.

When $f(x) = x^2$, this equation reduces to the definition of the quadratic variation process.

Theorem 3.9 (The Itô Formula). *Let M be a continuous local Martingale and V be a continuous process which is locally of bounded variation. Let f be a continuous real-valued function defined on \mathbb{R}^2 such that the partial derivatives $\frac{\partial f}{\partial x}(x, y)$, $\frac{\partial^2 f}{\partial x^2}(x, y)$ and $\frac{\partial f}{\partial y}(x, y)$, exist and are continuous for all (x, y) in \mathbb{R}^2 . Then a.s., we have for each t*

$$\begin{aligned} f(M_t, V_t) - f(M_0, V_0) &= \int_0^t \frac{\partial f}{\partial y}(M_s, V_s) dM_s + \\ &+ \int_0^t \frac{\partial f}{\partial y}(M_s, V_s) dV_s + \\ &+ \frac{1}{2} \int_0^t \frac{\partial^2 f}{\partial x^2}(M_s, V_s) d[M]_s. \end{aligned} \quad (3.7)$$

A suggestive way to write (3.7) is by using differentials:

$$df(M_t, V_t) = \frac{\partial f}{\partial x}(M_t, V_t) dM_t + \frac{\partial f}{\partial y}(M_t, V_t) dV_t + \frac{1}{2} \frac{\partial^2 f}{\partial x^2}(M_t, V_t) d[M]_t \quad (3.8)$$

Proof. Since both sides of (3.7) are continuous processes, it suffices to prove for each t that (3.7) holds a.s. Let $\{\pi_t^n, n \in \mathbb{N}\}$ be a sequence of partitions of $[0, t]$ such that $\lim_{n \rightarrow \infty} \delta \pi_t^n = 0$. We omit the superscript from t_j^n . The left-hand side of (3.7) may be written as the telescopic sum

$$\begin{aligned} f(M_t, V_t) - f(M_0, V_0) &= \sum_j \{f(M_{t_{j+1}}, V_{t_{j+1}}) \\ &\quad - f(M_{t_{j+1}}, V_{t_j}) + f(M_{t_{j+1}}, V_{t_j}) \\ &\quad - f(M_{t_j}, V_{t_j})\} \end{aligned} \tag{3.9}$$

By Taylor's theorem the right side of (3.9) may be written as:

$$\begin{aligned} \sum_j \left\{ \left(\left(\frac{\partial f}{\partial y} \right) (M_{t_j}, V_{t_j}) + \epsilon_j^1 \right) (V_{t_{j+1}} - V_{t_j}) \right. \\ \left. + \frac{\partial f}{\partial x} (M_{t_j}, V_{t_j}) (M_{t_{j+1}} - M_{t_j}) \right. \\ \left. \frac{1}{2} \left(\frac{\partial^2 f}{\partial x^2} (M_{t_j}, V_{t_j}) + \epsilon_j^2 \right) (M_{t_{j+1}} - M_{t_j})^2 \right\} \end{aligned} \tag{3.10}$$

where

$$\epsilon_j^1 = \frac{\partial f}{\partial y} (M_{t_{j+1}}, V_{\tau_j}) - \frac{\partial f}{\partial y} (M_{t_j}, V_{t_j})$$

and

$$\epsilon_j^2 = \frac{\partial^2 f}{\partial x^2} (M_{\eta_j}, V_{t_j}) - \frac{\partial^2 f}{\partial x^2} (M_{t_j}, V_{t_j})$$

for some random times τ_j and η_j in $[t_j, t_{j+1}]$.

For each ω , the functions $(r, s) \rightarrow \frac{\partial f}{\partial y} (M_r, V_s)(\omega)$ and $(r, s) \rightarrow \frac{\partial^2 f}{\partial x^2} (M_r, V_s)(\omega)$ are uniformly continuous on $[0, t]^2$, and hence $\sup_j |\epsilon_j^1(\omega)|$ and $\sup_j |\epsilon_j^2(\omega)|$ tend to zero as $n \rightarrow \infty$. (Note that ϵ_j^1 and ϵ_j^2 depend on n although the notation does not specifically indicate this because the indexes n on the t_j^n have been omitted).

From this property of $\epsilon_j^1(\omega)$, the continuity of $s \rightarrow \frac{\partial f}{\partial y} (M_s, V_s)(\omega)$, and since $s \rightarrow V_s(\omega)$ is of bounded variation on $[0, t]$ for almost every ω , it follows that

$$\sum_j \left(\frac{\partial f}{\partial y} (M_{t_j}, V_{t_j}) + \epsilon_j^1 \right) (V_{t_{j+1}} - V_{t_j}) \rightarrow \int_0^t \frac{\partial f}{\partial y} (M_s, V_s) dV_s$$

almost surely as $n \rightarrow \infty$. From the above property of $\epsilon_j^2(\omega)$, and since $\sum_j (M_{t_{j+1}} - M_{t_j})^2 \rightarrow [M]_t$ in probability as $n \rightarrow \infty$ by theorem (3.8), it follows that $\sum_j \epsilon_j^2 (M_{t_{j+1}} - M_{t_j})^2 \rightarrow 0$ in probability as $n \rightarrow \infty$.

From this property of $\epsilon_j^1(\omega)$, the continuity of $s \rightarrow \frac{\partial f}{\partial y}(M_s, V_s)(\omega)$, and since $s \rightarrow V_s(\omega)$ is of bounded variation on $[0, t]$ for almost every ω , it follows that

$$\sum_j \left(\frac{\partial f}{\partial y}(M_{t_j}, V_{t_j}) + \epsilon_j^1 \right) (V_{t_{j+1}} - V_{t_j}) \rightarrow \int_0^t \frac{\partial f}{\partial y}(M_s, V_s) dV_s$$

almost surely as $n \rightarrow \infty$. From the above property of $\epsilon_j^2(\omega)$, and since $\sum_j (M_{t_{j+1}} - M_{t_j})^2 \rightarrow [M]_t$ in probability as $n \rightarrow \infty$ by theorem (3.8), it follows that $\sum_j \epsilon_j^2 (M_{t_{j+1}} - M_{t_j})^2 \rightarrow 0$ in probability as $n \rightarrow \infty$.

The proof will be completed in two steps. First we prove that when M and V are bounded, the terms in (3.10) involving the x -partial derivative of f converge in probability to the appropriate terms in (3.7). Then we extend (3.7) to the general case by using a localizing sequence for M and V .

Suppose that M and V are bounded. Then $\frac{\partial f}{\partial x}$ and $\frac{\partial^2 f}{\partial x^2}$ are bounded on the range of (M, V) and μ_M is a finite measure on \mathcal{P} . For each n , the process X^n defined by

$$X^n = \sum_j \frac{\partial f}{\partial x}(M_{t_j}, V_{t_j}) I_{(t_j, t_{j+1}]} + \frac{\partial f}{\partial x}(M_0, V_0) I_0$$

is predictable and $\{X^n\}$ converges pointwise to $I_{[0, t]} \frac{\partial f}{\partial x}(M, V)$ on $\mathbb{R}_+ \times \Omega$ and hence by bounded convergence in \mathcal{L}^2 . Then by the isometry,

$$\sum_j \frac{\partial f}{\partial x}(M_{t_j}, V_{t_j}) (M_{t_{j+1}} - M_{t_j}) = \int X_s^n dM_s \rightarrow \int_0^t \frac{\partial f}{\partial x}(M_s, V_s) dM_s$$

in L^2 as $n \rightarrow \infty$. Also, with $Y = \frac{\partial^2 f}{\partial x^2}(M, V)$ we have

$$\sum_j \frac{\partial^2 f}{\partial x^2}(M_{t_j}, V_{t_j}) (M_{t_{j+1}} - M_{t_j})^2 \rightarrow \int_0^t \frac{\partial^2 f}{\partial x^2}(M_s, V_s) d[M]_s \quad (3.11)$$

in probability as $n \rightarrow \infty$. It follows that the expression in (3.10) converges in probability to the right side of (3.7) and hence (3.7) holds a.s. Thus we have proved the theorem when M and V are bounded.

To extend to the general case, for each n let $\tau_n = \inf t \geq 0 : |M_t| \vee |V_t| > n$. Then $M = M_{\cdot \wedge \tau_n} I_{\tau_n} > 0$ and $V^n = V_{\cdot \wedge \tau_n}$ are bounded, and it follows from the above that (3.7) holds a.s. with $t \wedge \tau_n$ in place of t . Then (3.7) follows on letting $n \rightarrow \infty$. \square

Applications:

- Let B denote the Brownian motion in \mathbb{R} and let $f(x) = x^2$. With $M = B$ and $f(x, y) = f(x)$, (3.8) becomes

$$d(B_t)^2 = 2B_t dB_t + dt$$

Formally this suggests $(dB_t)^2 = dt$. For general M the appropriate formalism is $(dM_t)^2 = d[M]_t$. Heuristically this explains the presence of the additional term in the Itô formula.

- A direct application of the theorem yields the following representation for the stochastic integral $\int_0^t V_s dM_s$, where M is a continuous local martingale and V is a continuous process that is locally of bounded variation:

$$\int_a^b V_s dM_s = M_b V_b - M_a V_a - \int_a^b M_s dV_s$$

since the paths of M are continuous and V is locally of bounded variation, the integral on the right is defined pathwise as a Riemann-Stieltjes integral, and by integration by parts, the right member defines the Riemann-Stieltjes of V with respect to M . It follows that $\int_0^t V_s dM_s$ defines the same random variable whether the integral is regarded as a stochastic integral or a Riemann-Stieltjes integral.

Chapter 4

Semi-groups and Generators

Suppose that (Ω, \mathcal{F}) is a measure space. Let $B(\Omega, \mathcal{F})$ denote the space of all bounded \mathcal{F} -measurable functions from Ω to \mathbb{R} , suppose that $P(t, x, A)$ is a transition function satisfying the conditions:

1. For fixed t and x , $P(t, x, A)$ is a measure on the σ -algebra \mathcal{F} ;
2. For fixed t and A , $P(t, x, A)$ is a \mathcal{F} -measurable function of x ;
3. $P(0, x, \Omega) \leq 1$;
4. $P(0, x, \Omega \setminus x) = 0$;
5. $P(s + t, x, A) = \int_{\Omega} P(s, x, dy)P(t, y, A)$, $s, t \geq 0$

We say that the transition function $P(t, x, A)$ is normal if $P(+0, x, E) = 1$ for all $x \in E$. We call it conservative if $P(t, x, E) = 1$ for all $t \geq 0, x \in E$.

Definition 4.1. A family of bounded linear operators is called a semi-group if for all $s \geq 0, t \geq 0$,

$$T_{s+t} = T_s T_t$$

A semi-group T_t is called a contraction semi-group if for all $t \geq 0, \|T_t\| \leq 1$, i.e. if for all $t \geq 0, f \in L$,

$$\|T_t f\| \leq \|f\|.$$

The infinitesimal operator of a semi-group T_t is defined by the formula

$$Af = s - \lim_{h \downarrow 0} \frac{T_h f - f}{h}$$

It can be shown (see [3] page 51) that the operator T_t defined by

$$T_t f(x) = \int_{\Omega} P(t, x, dy) f(y), \quad f \in B(\Omega, \mathcal{F})$$

forms a contraction semi-group on $B(\Omega, \mathcal{F})$ (sometimes the terminology sub-Markov semi-group is used).

Suppose now that Ω is a locally compact space with countable basis (LCCB) and $\mathcal{F} = \mathcal{F}(E)$ then Ω is δ -compact and Polish. We write:

- $C(\Omega)$ for the space of all (real-valued) continuous functions on Ω ;
- $C_b(\Omega)$ for the space of bounded continuous functions on Ω ;
- $\hat{C}_0(\Omega)$ for the space of (bounded) continuous functions on Ω which can vanish at infinity;
- $C_c(\Omega)$ for the space of continuous functions on Ω with compact support.

It is normal to say T_t is a Feller semi-group in case

$$T_t : C_b(\Omega) \rightarrow C_b(\Omega), \quad \forall t \geq 0.$$

and a strongly Feller semi-group in case

$$T_t : B(t, \mathcal{F}) \rightarrow C_b(\Omega), \quad \forall t \geq 0.$$

In [16] Williams suggests the name Feller-Dynkin semi-group in case T_t is a strongly continuous normal contraction semi-group on $C_0(\Omega)$.

If Ω is not compact, we can adjoin to Ω a point ζ so that $\Omega_\zeta = \Omega \cup \zeta$ is compact metrisable, in fact ζ is the point at infinity in the one point compactification. If Ω is compact adjoin an isolated point ζ to Ω .

One extends the transition function P_t to a conservative transition function on $(\Omega_\zeta, \mathcal{F}(\Omega_\zeta))$ by

$$P^\zeta(t, x, \zeta) = 1 - P_t(x, \Omega), \quad x \in \Omega, t \geq 0;$$

$$P^\zeta(t, \zeta, \cdot) = \epsilon_\zeta$$

$$P^\zeta(t, \cdot, \cdot) = P(t, \cdot, \cdot) \quad \text{on} \quad \Omega \times \mathcal{F}$$

One says that a map $f : [0, \infty) \rightarrow \mathbb{R}$ is a Skorokhod map if f right-continuous and has limits at the left. These functions are also known by the name of $c\tilde{A} \, d\tilde{A} \, g$. The collection of these functions forms a Skorokhod space.

The basic standard result is that if f is a Skorokhod map, Ω is LCCB, $\mathcal{F} = \mathcal{F}(\Omega)$ and T_t is a Feller-Dynkin semi-group on $\hat{C}_0(\Omega)$ then there exists a strong Markov, Ω_ζ -valued stochastic process X with Skorokhod trajectories whose transition function gives rise to T_t .

Let us observe that if T_t is a Feller-Dynkin semi-group then $t \rightarrow T_t f(x)$ is right-continuous and $T_t f(x)$ is $\mathcal{F} \times \mathcal{F}[0, \infty)$ -measurable in (x, t) considered as a map from $\Omega \times [0, \infty)$ to \mathbb{R} .

Taking the Laplace transform of the semi-group we define the associated resolvent ¹ $R_\lambda, \lambda > 0$, by

$$R_\lambda f(x) = \int_0^\infty e^{-\lambda t} T_t f(x) dt. \quad (4.1)$$

Standard applications of the monotone class theorem and Fubini's theorem show that $\{R_\lambda, \lambda > 0\}$ is a sub-Markovian resolvent on \hat{C}_0 :

1. $R_\lambda : \hat{C}_0 \rightarrow \hat{C}_0$;
2. $0 \leq f \leq 1 \Rightarrow 0 \leq \lambda R_\lambda f \leq 1$;
3. The resolvent equation holds:

$$R_\lambda - R_\mu + (\lambda - \mu)R_\lambda R_\mu = 0, \quad \forall \lambda, \mu > 0;$$

4. $f_n \downarrow 0 \Rightarrow R_\lambda f_n \downarrow 0, \quad \forall \lambda > 0$.

In the conservative case $T_t 1 = 1, \quad t \Leftrightarrow \lambda R_\lambda 1 = 1, \quad \forall \lambda$. R_λ is often called the λ -potential associated with the semi-group T_t .

Note that the range $R = R_\lambda \hat{C}_0$ of R_λ is independent of λ (from item 3 above). Now define $B_0 = \overline{R_\lambda \hat{C}_0(\Omega)}$. For more on this see [16] page 115.

In fact we shall demonstrate that under conditions 1 to 4 above, $B_0 = \hat{C}_0(\Omega)$. Suppose we have a family $\{R_\lambda, \lambda > 0\}$ satisfying:

$$\|\lambda R_\lambda f\| \leq 1, \quad \forall \lambda > 0$$

¹For a definition of the resolvent see [18] page 209 or [3] page 24

and item 3 above. Define B_0 as above. Then it is essentially a special case of the Hille-Yosida theorem (see [16] pages 109-114) that there exists a unique strongly continuous semi-group $\{U_t, t \geq 0\}$ such that

$$\int_0^\infty e^{-\lambda t} U_t f dt = R_\lambda f, \quad \forall \lambda > 0, \forall f \in B_0$$

One can define one operator G by

$$(\lambda - G)R_\lambda f = f, \quad \forall \lambda > 0, \quad \forall f \in B_0, \quad D(G) = D = R_\mu B_0$$

which is closed in the operator theoretic sense.

The task is then to make precise the idea that

$$U_t = \exp(tG) \tag{4.2}$$

A proof is sketched in [16] and the details can be found in [18]. One calls G the infinitesimal generator.

As a part of the Hille-Yosida theorem one knows that if $f \in D(G)$ then

$$Gf = \lim_{t \downarrow 0} \frac{U_t f - f}{t} \tag{4.3}$$

Suppose the items 1 to 4 above hold, then $B_0 = C_0(\Omega)$. Indeed, if not, there exists by the Hahn-Banach theorem a non-trivial linear functional H on $\hat{C}_0(\Omega)$ which annihilates B_0 . This functional is represented by a Riesz measure μ so that

$$\int \lambda R_\lambda f(x) \mu(dx) = \langle \lambda R_\lambda, H \rangle = 0, \quad \forall f \in \hat{C}_0(\Omega), \quad \lambda > 0.$$

Noting that $\lambda R_\lambda f(x) = \int_0^\infty e^{-s} T_{s/\lambda} f(x) ds \rightarrow f(x)$, $\lambda \rightarrow \infty$, it follows by the demonstrated convergence theorem that

$$\langle f, H \rangle = \int f(x) \mu(dx) = 0, \quad \forall f \in \hat{C}_0(\Omega)$$

or $H = 0$, a contradiction.

We need a few definitions from the theory of stochastic processes. We say a point $x \in \Omega$ is absorbed (with respect to the process X) if

$$P_X(x(t, \omega) = x, \forall t) = 1$$

or equivalently if

$$P(t, x, x) = 1, \quad \forall t.$$

Given a Markov process $(X_t, \tau, \mathcal{M}_t, P_X)$ given on state space Ω . A is a subset of Ω_ζ we define two functions

$$D_A(\omega) = \inf \{t \geq 0, X_t(\omega) \in A\};$$

$$T_A(\omega) = \inf \{t > 0, X_t(\omega) \in A\}$$

Where the infimum of the empty set is understood to be $+\infty$. We call D_A the first entry time of A and T_A the first hitting time of A (D_A is also referred to as the first exit time from the set $\Omega \setminus A$). The question as to when D_A, T_A are stopping times is difficult and its resolution involves Choquet capacity theory. Given a Markov process with state space (E, ϵ^*) (ϵ^* being the σ -algebra of universally measurable sets) which is right continuous, quasi-left continuous and such that $\mathcal{M}_t \supset \Omega_t$ one can say the following: D_A and T_A are $\{\mathcal{F}_t\}$ stopping times for all Borel (or more generally, analytic) subsets A of Ω_ζ .

Let Ω be LCCB and f an ϵ^* -measurable function. Let $\tau(U)$ be the first exit time from the open set U_x which is a neighborhood of x . The set of all ϵ^* -measurable functions for which the limit

$$\mathcal{A}f(x) = \lim_{U_x \downarrow x} \frac{E_x f(x_{\tau(U)}) - f(x)}{E_x(x_{\tau(U)})} \quad (4.4)$$

where E_x is the expectation with respect to P^X exists and is finite, will be denoted by $D_{\mathcal{A}}(x)$. We write $f \in D_{\mathcal{A}}$ if $f \in D_{\mathcal{A}}(x)$ for all $x \in \Omega$. For such f we shall call the operator defined by (4.4) the characteristic operator of the process (clearly with respect to the topology) because the following result holds. See the discussion about these operators on [3] p. 143 ff.

Proposition 4.1. *Let X be a Feller-Dynkin Markov process on a LCCB space Ω . Suppose that x_0 is non-absorbing. Then there is a neighborhood U of x_0 such that if $\tau(U)$ is the first exit time from U then $\sup_{x \in U} E_x \tau(U) < \infty$.*

Proof. Since x_0 is non-absorbing for some $t_0 > 0$, it is the case that $P(t_0, x_0, x_0) < 1$. Let d be a metric giving the topology of Ω . Then we can find a closed ball $B(x_0, E)$ around x_0 such that $P(t_0, x_0, B(x_0, \epsilon)^c) > 0$. Since $B(x_0, \epsilon)^c$ is open and the process is Feller-Dynkin $P(t_0, x_0, B(x_0, \epsilon)^c)$ is lower semi-continuous function of x such that $P(t_0, x, B(x_0, \epsilon)^c) \geq \epsilon_0 > 0$ for $x \in U$. We suppose that $U \subset B(x_0, \epsilon)$ so that

$$P(t, x, U^c) \geq \epsilon_0 \quad \text{for all } x \in U. \quad (4.5)$$

Let $\tau(U)$ be the first exit time of the process from the set U . Then $P_x(\tau(U) > t_0) \leq 1 - \epsilon_0$ for $x \in U$. Moreover,

$$\begin{aligned}
P_x(\tau(U) > nt_0) &= P_x(x(t) \in U, 0 \leq t \leq nt_0) \\
&= P_x \{x(t) \in U, 0 \leq t \leq (n-1)t_0; x(t) \in U, (n-1)t_0 \leq t \leq nt\} \\
&= \int_{\tau(U) > (n-1)t_0} P_{x((n-1)t_0)}(\tau(U) > t_0) dP_x \\
&\leq (1 - \epsilon_0) P_x(\tau(U) > (n-1)t_0).
\end{aligned} \tag{4.6}$$

This implies that $P_x(\tau(U) > nt_0) \leq (1 - \epsilon_0)^n$ so that

$$E_x(U) \leq \frac{\epsilon_0 t_0}{1 - \epsilon_0} \quad \text{for } x \in U.$$

□

The question which arises is the connection between the infinitesimal generator (4.3) and the characteristic operator (4.4). For the Feller-Dynkin process there is a simple reply:

Theorem 4.1 (Dynkin). $G = \mathcal{A}$.

The proof of this theorem, while not complicated, relies on many intermediate steps and definitions, therefore the reader is referred to [16] page 131.

We want now to close this circle of ideas and return to the theory of diffusion processes in \mathbb{R}^n (more generally in manifolds generalizing the Wiener process). By a Feller-Dynkin diffusion on \mathbb{R}^n is meant a Feller-Dynkin process with continuous paths: $t \rightarrow x_t(\omega)$ on $[0, \tau(\omega))$ and such that the domain $D(G)$ contains $C_0^\infty(\mathbb{R}^n)$ (with the usual notation of analysis). For such diffusions we see that $H = G \in C_0^\infty(\mathbb{R}^n)$ satisfies the following conditions:

1. H is linear from C_0^∞ to \hat{C}_0 ;
2. H is local;
3. H satisfies the maximum principle

Of these properties the first is trivial and the third is simply Dynkin's maximum principle. From the definition of \mathcal{A} and the fact that X has continuous paths we conclude that \mathcal{A} is local and the second property follows as $H \subset \mathcal{A}$.

One can readily characterize the infinitesimal generator of a Feller-Dynkin diffusion in \mathbb{R}^n . Indeed we have the following theorem.

Theorem 4.2 (Dynkin). *H is a second order elliptic operator of the form*

$$Hf(x) = \frac{1}{2} \sum_{ij} a_{ij}(x) \partial_i \partial_j f(x) + \sum_i b_i(x) \partial_i f(x) - c(x) f(x)$$

where ∂_i denotes $\frac{\partial}{\partial x_i}$ and

1. $a_{ij}(x)$, $b_i(x)$ and $c(x)$ are continuous $i, j = 1, \dots, n$;
2. $(a_{ij}(x))$ is a non-negative definite symmetric matrix;
3. $c(x) \geq 0$, $\forall x \in \mathcal{R}^n$

Proof. One observes from the locality of H and the fact that the local maximum principle holds: $f \in D(H)$ has a local maximum at x and $f(x) \geq 0$ then $Hf \leq 0$. For $x \in \mathbb{R}^n$, there exists $\phi \in C_0^\infty(\mathbb{R}^n)$ such that $\phi = 1$ in a neighborhood of x . For such a ϕ define $c(x) = -G\phi(x)$ independently of ϕ . Further, set $b_i(x) = G\phi_i(x)$, where $\phi_i \in C_0^\infty(\mathbb{R}^n)$ and $\phi_i(y) = y_i - x_i$ near x and $a_{ij}(x) = G(\phi_i \phi_j)(x)$.

The functions c, b_i and a_{ij} are continuous. For $(\lambda_1, \dots, \lambda_n) \in \mathbb{R}^n$ the function h defined by $h(y) = -(\sum \lambda_i \phi_i(y))^2$ has a local maximum at x , so that

$$\sum_{ij} a_{ij}(x) \lambda_i \lambda_j = -Hh(x) \geq 0$$

so that $(a_{ij}(x))$ is non-negative definite.

If $f \in C_0^\infty(\mathbb{R}^n)$ the Taylor's formula gives

$$f(y) = \psi(y) + o(|y - x|^2)$$

where $\psi(y) = f(y)\phi(y) + \sum \partial_i f(x) \partial_i(y) + \frac{1}{2} \sum_{ij} \partial_i \partial_j f(x) \phi_i(y) \phi_j(y)$. Note that

$$H\psi(x) = -c(x)f(x) + \sum_i b_i(x) \partial_i f(x) + \frac{1}{2} \sum_{ij} a_{ij}(x) \partial_i \partial_j f(x).$$

For $\epsilon > 0$, the functions in $C_0^\infty(\mathbb{R}^n)$ defined near x by

$$y \rightarrow \pm(f(y) - \psi(y) \mp \epsilon|y - x|^2)$$

have a local maxima at x . Hence,

$$\pm(Hf(x) - H\psi(x) \mp \frac{1}{2}\epsilon \sum_{ij} a_{ij}) \leq 0$$

from which it follows that $Hf(x) = H\psi(x)$. □

Chapter 5

Stochastic Differential Equations and Diffusions

The connection between stochastic differential equations of Itô and a limited class of homogeneous Feller-Dynkin diffusions is described. The regularity assumptions made here are quite strong and results under much weaker conditions have been established by Stroock and Varadhan.

Consider the differential generator

$$H = \frac{1}{2} \sum_{ij} a_{ij}(x) \partial_i \partial_j + \sum_i b_i(x) \partial_i - c(x), \quad x \in \mathbb{R}^n \quad (5.1)$$

We suppose that for positive constants γ, c, M, K we have:

1. $\max_{ij} \sup_{x \in \mathbb{R}^n} \|a_{ij}(\cdot)\|_{C^2(\mathbb{R}^n)} \leq M$;
2. $a(x) = (a_{ij}(x)) \geq \gamma I, \forall x \in \mathbb{R}^n$;
3. $\max_i \sup_{x \in \mathbb{R}^n} |b_i(x)| \leq M$;
4. $|b_i(x) - b_i(y)| \leq K|x - y|, i = 1, \dots, n, \forall x, y \in \mathbb{R}^n$;
5. $c(x) \geq 0, |c(x) - c(y)| \leq K|x - y|^\alpha, 0 < \alpha \leq 1$;
6. $c(x) \leq M$.

Items 1 and 4 can be weakened to a Hölder condition in constructions of fundamental solutions. We have assumed the above form to compress the exposition.

Let Γ be a smooth contour lying in the half-plane $\Re z > 0$ of the complex plane and containing all the eigenvalues of $a(x)$. Then $\sigma(x) = \frac{1}{2\pi} \int_{\Gamma} \sqrt{z}(a(x) - zI)^{-1} dz$ can be defined via spectral calculus and it can be shown that $\sigma^2(x) = a(x)$ (see [6]). Further, we have

$$|\sigma(x) - \sigma(y)| \leq c|x - y|, \quad \forall x, y \in \mathbb{R}^n \quad (5.2)$$

$$\sup_{x \in \mathbb{R}^n} |\sigma(x)|, \quad \sigma \geq 0$$

with the same constant.

Associated with the generator (5.1), satisfying conditions 1 to 6, is the diffusion matrix σ satisfying (5.2), and drift b satisfying items 3 and 4. We considered consequently the related stochastic differential

$$d\xi(t) = b(\xi(t)) dt + \sigma(\xi(t))d\omega(t) \quad (5.3a)$$

$$\xi(0) = \xi_0 \quad (5.3b)$$

where ξ_0 is any n -dimensional random vector independent of

$$\sigma \{ \omega(t), 0 \leq t \leq T, T > 0 \}; \quad E(|\xi_0|^2) < \infty \quad (5.4)$$

Observe that if H satisfies items 1 to 6 then the parabolic equation

$$\frac{\partial u}{\partial t} = Hu \quad (5.5)$$

(Kolmogorov's backward equation) has a unique fundamental solution P in the sense that: on the set $t > 0, x, y \in \mathbb{R}^n, P(t, x, y) \in C(\mathbb{R}_+, \mathbb{R}^n, \mathbb{R}^n)$. $P(t, x, y)$ is C^2 in x and satisfies (5.5). Also for any continuous bounded function $f(x), x \in \mathbb{R}^n$,

$$\lim_{t \downarrow 0} \int_{\mathbb{R}^n} P(t, x, y) f(y) dy = f(x). \quad (5.6)$$

Moreover,

$$0 < P(t, x, y) \leq M_0 t^{-\frac{n}{2}} e^{-\theta \frac{|x-y|^2}{t}}, \quad M_0, \theta > 0, t > 0. \quad (5.7)$$

These properties can be obtained from any text on parabolic equations, see for example [5].

We define a semi-group \tilde{T}_t on \hat{C}_0 by

$$\tilde{T}_t f(x) = \int_{\mathbb{R}^n} P(t, x, y) f(y) dy, \quad t \in \hat{C}_0 \quad (5.8)$$

Note that for $t > 0$ and $|x| > r$

$$\begin{aligned} |\tilde{T}_t f(x)| &\leq \int_{|y| \geq r} P(t, x, y) |f(y)| dy + \int_{|y| < r} P(t, x, y) |f(y)| dy \\ &\leq Kt^{-\frac{n}{2}} \int_{\mathbb{R}^n} e^{-\theta \frac{|y|^2}{2t}} dy \sup_{|y| \geq r} |f(y)| \\ &\quad + Kt^{-\frac{n}{2}} \|f\|_\infty e^{-\theta \frac{(|x|-r)^2}{2r}} \int_{|y| \leq r} dy \end{aligned}$$

hence, $\lim_{|x| \rightarrow \infty} \sup_{0 \leq t \leq r} |\tilde{T}_t f(x)| = 0$ or $\tilde{T}_t \hat{C}_0 \subseteq \hat{C}_0$. Also, by (5.6) $\lim_{t \downarrow 0} \|\tilde{T}_t f - f\|_\infty = 0$, $f \in \hat{C}_0$. The semi-group property is well known, it follows that \tilde{T}_t is a Feller-Dynkin semi-group and generates a Feller-Dynkin process \tilde{X} . It is readily shown that the infinitesimal generator of \tilde{T}_t contains $C_0^\infty(\mathbb{R}^n)$ and the stochastic regularity conditions hold

$$\lim_{y \rightarrow \infty} \sup_{t \leq n} P(t, y, \Gamma) = 0$$

and

$$\lim_{t \downarrow 0} \sup_{x \in \Gamma} P(t, x, \overline{B(x, \epsilon)^c}) = 0.$$

The latter imply that the process has continuous paths. We conclude that \tilde{X} is a Feller-Dynkin diffusion $\tilde{X} = (\tilde{x}(t, \omega), \tilde{\tau}(\omega), \tilde{M}_t, \tilde{P}_x)$.

Let us return to the stochastic differential equation (5.3a). Supposing that $\xi(0) = x$ a.e.; $x \in \mathbb{R}^n$. One seeks a solution in $M_\omega^2[0, T]$, $T > 0$ via the method of successive approximations:

$$\xi_{n+1} = \xi_0 + \int_0^t b(\xi_n) ds + \int_0^t \sigma(\xi_n) d\omega$$

making the inductive hypothesis that $\xi_n \in M_\omega^2[0, T]$ and

$$E(|\xi_{k+1}(t) - \xi_k(t)|^2) \leq \frac{(Mt)^{k+1}}{(k+1)!}, \quad 0 \leq k \leq n-1, \quad M > 0.$$

Using a Martingale inequality one shows that

$$E\left(\sup_{0 \leq t \leq T} |\xi_{n+1}(t) - \xi_n(t)|^2\right) \leq \frac{c(MT)^n}{n!}$$

and the Borel-Cantelli lemma to show that

$$\sup_{0 \leq t \leq T} |\xi_{n+1}(t) - \xi_n(t)|^2 \leq \frac{1}{2^n} \quad n \geq n_0(\omega) \quad a.e. \quad \omega$$

One concludes that $\xi_k = \xi_0 + \sum_{n=0}^{k-1} (\xi_{n+1}(t) - \xi_n(t))$ converges uniformly and passes to the limit in the stochastic integral. This gives a solution in $M_\omega^2[0, T]$ which is continuous and path-wise unique.

Let $\xi_x(t)$ be the solution corresponding to $\xi(0) = x$ a.e. $x \in \mathbb{R}^n$. For any Borel set Γ in \mathbb{R}^n , $t \geq 0$, we set

$$P(t, x, \Gamma) = P(\xi_x(t) \in \Gamma) \quad (5.9)$$

It is not difficult to show the following result:

Theorem 5.1. *If σ satisfies (5.2); b satisfies conditions 3 and 4; and if ξ_0 is independent of $\sigma(\omega(t), t \geq 0)$, $E(|\xi_0|^2) < \infty$ then the unique solution of (5.3a) and (5.3b) satisfies*

$$\begin{aligned} P(\xi(t) \in \Gamma | \mathcal{F}_s) &= P(\xi(t) \in \Gamma | \xi(s)) \\ &= P(t-s, \xi_s, \Gamma) \quad a.e. \end{aligned} \quad (5.10)$$

and $P(t, x, \Gamma)$ is a transition probability function.

There is of course an associated Feller-Dynkin Markov process which can be constructed as follows. Let $W = C([0, \infty), \mathbb{R}^n)$ and introduce the σ -algebra

$$\mathcal{F} = \sigma \{ \omega : \omega(u) \in A \subseteq \mathbb{R}^n, \text{ Borel}, u \geq 0 \}$$

and sub σ -algebras $\mathcal{M}_t = \sigma \{ \omega : \omega(u) \in A \subseteq \mathbb{R}^n \text{ Borel}; 0 \leq u \leq t \}$ in W . Define a continuous process X by

$$X(t) = x(t, \omega) = \omega(t), \quad \omega \in W.$$

and set

$$P_X(\omega \in B) = P(\omega : \xi_x(\cdots, \omega) \in B), \quad \forall B \in \mathcal{M}_t \quad (5.11)$$

P_X is a probability measure on \mathcal{M}_t .

Note that

$$P_X(X(t+h) \in \Gamma | \mathcal{M}_t) = P(h, X(t), \Gamma) \quad a.e. \quad (5.12)$$

In fact,

$$\begin{aligned} P(\xi_x(t+h) \in \Gamma | \sigma(\xi_x(\lambda), \lambda \leq t)) \\ = P(h, \xi_x, \Gamma) \end{aligned}$$

so that for $\forall s \leq t_1 < t_2 < \cdots < t_n \leq t$ and Borel sets A_1, \cdots, A_n we have

$$\begin{aligned} P(\xi_x(t+h) \in \Gamma, \xi_x(t_1) \in A_1, \cdots, \xi_x(t_n) \in A_n) \\ = \int_{\cap_i \xi_x(t_i) \in A_i} P(h, \xi_x(t), \Gamma) dP \\ = \int I_{A_1}(\xi_x(t_1)) \cdots I_{A_n}(\xi_x(t_n)) P(h, \xi_x(t), \Gamma) dP \end{aligned}$$

and using (5.10) and (5.11) we conclude that

$$\begin{aligned}
& P_X(X(t+h) \in \Gamma, X(t_1) \in A_1, \dots, X(t_n) \in A_n) \\
&= \int I_{A_1}(X(t_1)) \cdots I_{A_n}(X(t_n)) P(h, X(t), \Gamma) dP_X \\
&= \int_{\cap_i X(t_i) \in A_i} P(h, X(t), \Gamma) dP_X
\end{aligned} \tag{5.13}$$

this implies (5.12).

The semi-group which corresponds to the process X is given by

$$T_t f(x) = E_x f(X(t)) = E f(\xi_x(t))$$

Taking (5.3a) in the integrated form we have

$$\xi_x(t) = x + \int_0^t b(\xi_x(u)) du + \int_0^t \sigma(\xi_x(u)) d\omega(u). \tag{5.14}$$

Clearly, $\left| \int_0^t b(\xi_x(u)) du \right| \leq Mt$.

Hence, for $|x| > 4Mt$ we have

$$\begin{aligned}
\left\{ |\xi_x(t)| \leq \frac{|x|}{2} \right\} &\subseteq \left\{ |\xi_x(t) - x| > \frac{|x|}{2} \right\} \\
&\subseteq \left\{ \left| \int_0^t \sigma(\xi_x(u)) d\omega(u) \right| > \frac{|x|}{4} \right\}
\end{aligned}$$

Now $\xi_x(t) \in M_\omega^2[0, T]$, so that by the Martingale inequality we have:

$$\begin{aligned}
P \left\{ \left| \int_0^t \sigma(\xi_x(u)) d\omega(u) \right| \geq \frac{|x|}{4} \right\} &\leq \frac{256}{x^2} \int_0^t |E \sigma(\xi_x(u))|^2 du \\
&\leq \frac{256}{x^2} M^2 t^2
\end{aligned}$$

Hence, for any $|x| > 4Mt$

$$P(|\xi_x(t)| \leq \frac{|x|}{2}) \leq \frac{256}{x^2} M^2 t$$

and as $|x| \rightarrow \infty$, $|\xi_x(t)| \rightarrow \infty$ a.e. P , so that, for $f \in C_0$, $f(\xi_x(t)) \rightarrow 0$ a.e. P . We conclude that $T_t \hat{C}_0 \subseteq \hat{C}_0$. Also $\|T_t f - f\|_{\hat{C}_0} \rightarrow 0$ as $t \downarrow 0$ is evident. We conclude that X is a Feller-Dynkin process with continuous paths (and strong Markov).

We now establish the connection between this process X and H .

Theorem 5.2. *Suppose that $c \equiv 0$. The process $X = (x_t, \infty, \mathcal{M}_t, P_X)$ constructed from the stochastic differential equation (5.3a) is a Feller-Dynkin diffusion with infinitesimal generator H .*

Proof. It will be convenient to denote by $F^*(y)$ the fixed translate $F(x+y)$, $x \in \mathbb{R}^n$. We set

$$\begin{aligned} y_t^s &= \xi_x(t) - \xi_x(s) \\ y_t &= y_t^0 = \xi_x(t) - x \end{aligned}$$

From (5.3a) it follows that

$$y_t^s = \int_s^t b^*(y_u) du + \int_s^t \sigma^*(y_u) d\omega(u)$$

Let $f \in C^2(\mathbb{R}^n)$. Applying the Ito formula to f^* we obtain

$$f^*(y_t) - f^*(y_0) = \int_0^t b^*(y_u) du + \int_0^t h^*(y_u) d\omega(u) \quad (5.15)$$

with

$$g(y) = Hf(y); \quad h = \sum_{i,l} \partial_i f \sigma_{i,l}$$

We note that $F^*(y_t) = F(\xi_x(t))$. Therefore (5.15) can be rewritten

$$f(\xi_x(t)) - f(x) = \int_0^t g(\xi_x(u)) du + \int_0^t h(\xi_x(u)) d\omega(u). \quad (5.16)$$

If $f \in C_0^2(\mathbb{R}^n)$ then $g, h \in C_0(\mathbb{R}^n)$ and

$$Ef(\xi_x(t)) - f(x) = \int_0^t Eg(\xi_x(u)) du$$

Therefore, from (5.16) we obtain

$$Ef(\xi_x(t)) - f(x) = \int_0^t Eg(\xi_x(u)) du$$

or

$$T_t f(x) - f(x) = \int_0^t T_u g(u) du \quad (5.17)$$

We conclude that

$$\|T_t f - f\| \leq \|g\|_\infty t. \quad (5.18)$$

Approximating $f \in C_0(\mathbb{R}^n)$ by $f_n \in C_0^2(\mathbb{R}^n)$ and passing to the limit in (5.18) we conclude that (5.18) holds for all $f \in C_0(\mathbb{R}^n)$ and in particular for g .

From (5.17) we conclude that

$$\begin{aligned} \left| \frac{T_t f - f}{t} - g(x) \right| &\leq \frac{1}{t} \int_0^t |T_u g(x) - g(x)| du \\ &\leq \sup_{0 \leq u \leq t} |T_u g - g| \rightarrow 0 \quad \text{by} \quad (5.18) \end{aligned}$$

Hence, $f \in D(G)$ and $Gf = g$.

Assume now that $f \in D(x)$. Then there exists a function $F \in C_0^2(\mathbb{R}^n)$ such that $f(y) = F(y)$ in some neighborhood U_x of x . In view of the local character of \mathcal{A} and G this implies that $f \in D_{\mathcal{A}}(x)$ and

$$\mathcal{A}f(x) = \mathcal{A}F(x) = GF(x) = HF(x) = Hf(x)$$

Hence, X is a Feller-Dynkin diffusion and H its generator.

It follows that we may identify the transition function of the Feller-Dynkin process determined by (5.3a) with the fundamental solutions of (5.5) ($c \equiv 0$). \square

Chapter 6

The Dirichlet Problem

It is well known that the classical Dirichlet problem for a bounded $\Omega \subseteq \mathbb{R}^n$, $n \geq 2$, in its usual form is not always well posed.

Given $f \in C(\partial\Omega)$, find $u \in C(\bar{\Omega})$ such that

$$\Delta u = 0 \quad \text{in} \quad \Omega \tag{6.1}$$

and

$$\lim_{\substack{x \rightarrow y \\ x \in \Omega}} u(x) = f(y), \quad \forall y \in \partial\Omega \tag{6.2}$$

Gauss thought that he had solved the problem by using the "Dirichlet Principle" but his reasoning was invalid. Even ignoring the counter-example of Zaremba of 1909, there exists the much more significant observation of H. Lebesgue in 1913 on the so called Lebesgue's thorn in dimension $n > 2$.

Consider a $n \geq 3$ dimensional spherical surface and push a sharp thorn into its side. Think of D as a container being heated by the tip of the thorn, the sides ∂D being held at temperature $f \in C(\partial D)$, $0 \leq f \leq 1$. Analytically

$$D = \left\{ (x, y, z) : x^2 + y^2 + z^2 \leq 1, x^2 + y^2 > \exp\left(-\frac{1}{2z}\right), z > 0 \right\}. \tag{6.3}$$

Look at an extreme case in which $f = 0$ except near the tip of the thorn at which $f = 1$. As $t \uparrow +\infty$ the temperature inside of D should converge to the solution of the Dirichlet problem with $u = f$ on D .

However, heuristically the heat radiated by the thorn is proportional to its surface area and if the thorn is sharp enough the interior will be cold or $\lim_{\substack{x \rightarrow 0 \\ x \in D}} u(x) < 1 = f(0)$.

One can give a rigorous analytical argument justifying these observations.

We say that a distribution u in Ω is a sub-solution of $\lambda - \Delta$ if $(\Delta - \lambda)u$ is a positive Radon measure in Ω . We say u is a super-solution in case $-u$ is a sub-solution. When $\lambda = 0$ we refer to these distributions as subharmonic and superharmonic respectively.

If a subharmonic distribution u is in fact a upper semi-continuous function in Ω then it is a subharmonic function in the classical sense that it satisfies the addition condition: for any w subharmonic in a subdomain $\Omega' \subset \Omega$ the difference $u - w$ is either constant or fails to take a maximum in Ω' . Let x_0 be a point in $\partial\Omega$. A continuous function β in Ω will be called a barrier at x_0 (for $\lambda - \Delta$) if β is a sub-solution of $\lambda - \Delta$ and satisfies:

$$\begin{aligned} \beta &< 0 \quad \text{in} \quad \Omega \\ \overline{\lim}_{\substack{y \rightarrow x \\ x \in \Omega}} \beta(y) &< 0, \quad x \in \partial\Omega, \quad x \neq x_0 \\ \lim_{\substack{x \rightarrow x_0 \\ x \in \Omega}} \beta(x) &= 0 \end{aligned}$$

We say $x_0 \in \partial\Omega$ is regular if a barrier exists at x_0 . It can be shown that this is equivalent to the existence of a function $\beta \in H'(\Omega)$ such that

$$\begin{aligned} (\lambda - \Delta)\beta &= -1 \quad \text{in} \quad \Omega; \\ \beta &< 0 \quad \text{in} \quad \Omega \\ \forall x \in \partial\Omega, \quad x \neq x_0 \quad \overline{\lim}_{\substack{y \rightarrow x \\ y \in \Omega}} \beta(y) &< 0 \\ \lim_{\substack{x \rightarrow x_0 \\ x \in \Omega}} \beta(x) &= 0 \end{aligned}$$

See for example [15].

Let us recall that O. Perron in 1923 with a remarkable simple and direct method resolved the classical Dirichlet problem for boundary functions f in the following way. Let $V(f)$ be the class of all subharmonic functions v in Ω satisfying:

$$\overline{\lim}_{x \rightarrow y} v(x) \leq f(y), \quad \forall y \in \partial\Omega.$$

Then

$$u(x) = \sup_{v \in V(f)} v(x) \tag{6.4}$$

is either harmonic or $\equiv \pm\infty$ in Ω . This uses the Poisson integral and Harnack's principle. Then we have

Theorem 6.1 (Perron 1923). *If f is bounded on $\partial\Omega$ and $x_0 \in \partial\Omega$ is regular then*

$$\underline{\lim}_{x \rightarrow x_0} f(x) \leq \underline{\lim}_{x \rightarrow x_0} u(x) \leq \overline{\lim}_{x \rightarrow x_0} u(x) \leq \overline{\lim}_{x \rightarrow x_0} f(x)$$

Clearly, if all points of $\partial\Omega$ are regular and f is continuous on $\partial\Omega$ then the classical Dirichlet problem for the Laplacian (6.1) has a solution.

Wiener in 1924 gave a test for regularity couched in terms of potential theory and capacity. He did not point any connection however with his own published theory of Brownian motion. This connection was later observed by Kakutani and Doob. Let us briefly remark on the notion of capacity.

We say a domain $D \subset \mathbb{R}^n$ ($n \geq 2$) is Greenian if there exists a function $G(x, y) = K + H$ on $D \times D$, $< +\infty$ off the diagonal and such that:

$$\begin{aligned} G &\geq 0; \\ G(x, y) &= G(y, x); \\ K &= -(2\pi)^{-1} \log |x - y|; \\ K &= \frac{\Gamma\left(\frac{n}{2} - 1\right)}{4\pi^{n/2}} |x - y|^{2-n}; \\ \Delta_x H &= \Delta_y H = 0. \end{aligned}$$

Let B be a compact subset of D . Associated G , B be a Radon measure $\mu \geq 0$ is the Green potential $U^\mu = \int_B G(x, y) d\mu(y)$.

The capacity potential V_k is the greatest Green potential U^μ of positive Radon measures supported on K such that $U^\mu \leq 1$ in D . The capacity $C(B)$ is the supremum of the masses $\mu(B)$ as μ ranges over the set of all positive Radon measures supported by B whose Green's potential does not exceed 1 in D . $D = \mathbb{R}^n$, $n \geq 2$, for example is Greenian and each dimensional domain is Greenian.

One can define the inner capacity of arbitrary sets A by

$$C(\Omega) = \sup_{\substack{K \subset A \\ K \text{ compact}}} C(K)$$

and similarly the outer capacity of A as the infimum of inner capacity of open sets containing A . One says that a set is capacitable if the inner and outer capacities agree. Choquet's results show that the analytic sets generated by the compact subsets are capacitable. With these somewhat summary observations Wiener's test can be stated.

Wiener's Test

$x_0 \in \partial\Omega$ is regular if given λ , $0 < \lambda < 1$, and setting $B_k = \left\{ y \in \Omega^c : \lambda^{k+1} < |y - x_0| \leq \lambda^k \right\}$ then

$$\sum_k \lambda^{k(2-n)} C(B_k) = \infty \tag{6.5}$$

See [5] or [8] for a proof.

Let τ_D be the hitting time of a Borel set D with respect to Brownian motion. Doob gave a probabilistic reformulation of Wiener's test in terms of Brownian motion.

Theorem 6.2 (Doob). $x_0 \in \partial\Omega$ is regular for the Dirichlet problem for the Laplacian if

$$P_{x_0}(\tau_{\Omega^c} = 0) = 1 \tag{6.6}$$

See [11] for a proof.

Let us now return to the Feller-Dynkin generator H of the preceding chapter, which we assume satisfies conditions 1-6 (see the definitions on the preceding chapter), and let Ω be some bounded domain.

In a natural extension of the ideas of sub (-super) harmonic functions for the Laplacian one says that f is H -harmonic if $f \in C^2(\Omega)$ and $Hf = 0$, $x \in \Omega$ and that f is superharmonic if

- f is lower semi-continuous on Ω ;
- $f \neq \infty$;
- if U is open with compact closure and h is H -harmonic on U and continuous on \bar{U} and satisfies

$$h(x) \leq f(x), x \in \partial U$$

then $h(x) \leq f(x)$ for all $x \in U$.

There is a related idea of being superharmonic for the associated Feller-Dynkin diffusion X . In fact, it can be shown that functions that are X -superharmonic and not identically $+\infty$ are H -superharmonic.

For a Feller-Dynkin diffusion X we say a point $x_0 \in \partial\Omega$ has a barrier f : if there exists a neighborhood U of x_0 such that f is positive and superharmonic in and satisfies $\lim_{\substack{x \rightarrow x_0 \\ x \in \Omega}} f(x) = 0$.

Generalizing (6.6) we say that $x_0 \in \partial\Omega$ is regular for X if

$$P_X(\tau_{\Omega^c} = 0) = 1 \tag{6.7}$$

A necessary and sufficient condition that $x_0 \in \partial\Omega$ be a regular point for a Feller-Dynkin diffusion X with differential H is that there exists a barrier at x_0 .

The Dirichlet problem for H can be formulated as:

$$Hu = 0 \quad \text{in} \quad \Omega \tag{6.8}$$

$$\lim_{\substack{x \rightarrow y \\ x \in \Omega}} u(x) = f(y), \quad y \in \partial\Omega^r$$

for some subset $\partial\Omega^r \subset \partial\Omega$.

The theory for the Laplacian generalizes in the following form. Let $X = (\xi_t, \tau, \mathcal{M}_t, P_X)$ be the Feller-Dynkin diffusion associated with H , τ the first exit time from Ω .

Theorem 6.3 (Doob-Dynkin). $u(x) = E_x f(\xi_\tau)$, $x \in \Omega$ is H -harmonic on Ω and

$$u(x) = E_x f(\xi_\tau), \quad x \in \Omega \text{ } H\text{-harmonic on } \Omega \tag{6.9a}$$

$$\lim_{\substack{x \rightarrow x_0 \\ x \in \Omega}} u(x) = f(x_0) \tag{6.9b}$$

if x_0 is regular for X and f is continuous at $x_0 \in \partial\Omega$.

Note that if $y \in \partial\Omega$ there exists a closed ball K , centered at x_0 , $K \cap \Omega = \emptyset$, $K \cap \bar{\Omega} = y$

$$w(x, y) = k (|x_0 - y|^{-\rho} - |x - y|^{-r\rho})$$

is a barrier for H if k and ρ are sufficiently large. It follows that if $\partial\Omega$ is C^2 then the barriers exist at all points $y \in \partial\Omega$ and so $\partial\Omega$ is regular for X . We prove a version of (6.9a) and (6.9b) in this restricted case using Ito's formula.

Theorem 6.4. *Let H be a Feller-Dynkin differential generator satisfying items 1 to 6 (see preceding chapter) and X the process associated with $H_0 = H + c$. Let $\partial\Omega$ be C^2 and f continuous on C^2 . Then the Dirichlet problem for H (with $\partial\Omega^r = \partial\Omega$) has solution*

$$u(x) = E_x(f(\xi_\tau)) \exp\left(\int_0^\tau c(\xi(u)) du\right)$$

where τ is the exit time from Ω .

Proof. Note that under these conditions it is well known from the theory of elliptic differential equations that a solution exists so that we are merely giving an explicit representation. (6.9a) and (6.9b) say much more.

It is not difficult to extend the integrated version of Ito's formula from intervals $[0, t]$ to $[0, \tau]$ with stopping times τ . Consider the associated stochastic differential equation:

$$d\xi(t) = b(\xi(t))dt + \sigma(\xi(t))d\omega(t). \tag{6.10}$$

Let V_ϵ be the closed ϵ - neighborhood of $\partial\Omega$ and $\Omega_\epsilon = \partial \setminus V_\epsilon$. Let v be a function in $C^2(\mathbb{R}^n)$ which coincides with the classical solutions of the Dirichlet problem u in $\Omega_{\epsilon/2}$ and let $\tau = \tau_\epsilon \wedge T$ where τ_ϵ is the hitting time of V_ϵ . Then an extension of Ito's formula states that

$$\begin{aligned} & E_x(V(\xi(\tau))) \exp\left(\int_0^\tau c(\xi(u)) du\right) - v(x) \\ &= E_x \int_0^\tau \left[v(\xi(u)) \exp\left(\int_0^t c(\xi(u)) du\right) \right] dt \end{aligned} \quad (6.11)$$

Let $x \in \Omega_\epsilon$. Then $v(\xi(t)) = u(\xi(t))$ for all t , $0 \leq t \leq \tau$. Hence (6.11) holds with $v = u$. Taking $\epsilon \rightarrow 0$ and using the Lebesgue dominated convergence theorem we have

$$u(x) = E_x u(\xi(\tau \wedge T)) \exp\left(\int_0^{\tau \wedge T} C(\xi(u)) du\right) \quad (6.12)$$

Finally, observe that $E_x \tau < \infty$. Introduce $h(x) = -Ae^{\lambda x}$. If A, λ are sufficiently large $\sum_{ij} a_{ij} \partial_i \partial_j h + \sum_i b_i \partial_i h \leq -1$ in Ω . By Ito's formula

$$E_x h(\xi(\tau \wedge T)) - h(x) \leq -E_x(\tau \wedge T)$$

Since $|h(x)| \leq K$ in Ω , we have

$$E_x(\tau \wedge T) \leq 2K$$

and taking $T \uparrow \infty$ and using the monotone convergence theorem

$$E_x \tau \leq 2K$$

It follows that we can take $T \uparrow +\infty$ in (6.12) obtaining the result. □

Bibliography

- [1] Krishna Athreya and Soumendra Lahiri. *Measure Theory and Probability Theory*. Springer, 2006.
- [2] K.L. Chung and R.J. Williams. *Introduction to Stochastic Integration*. Birkhauser, 1990.
- [3] E.B. Dynkin. *Markov Processes Vol. I*. Springer, 1965.
- [4] Albert Einstein. Zur theorie des brownshen bewegung. *Ann. Phys.*, 19(4):371–381, 1966.
- [5] Avner Fridman. *Partial Differential Equations of Parabolic Type*. Prentice-Hall, 1964.
- [6] Avner Friedman. *Stochastic Differential Equations and Applications. Vol.1*. Academic Press, 1975.
- [7] Nobuyuki Ikeda and Shinzo Watanabe. *Stochastic Differential Equations and Diffusion Processes*. North-Holland/Kodansha, 1981.
- [8] K. Ito and H.P. McKean. *Diffusion Processes and their Sample Paths*. Springer, 1965.
- [9] Ioannis Karatzas and Steven E. Shreve. *Brownian Motion and Stochastic Calculus*. Springer, 1988.
- [10] Leonid Korolov and Yakov Sinai. *Theory of Probability and Random Processes*. Springer, 2013.
- [11] S.C. Port and C.J. Stone. *Brownian Motion and Classical Potential Theory*. Academic Press, 1978.

- [12] A.N. Shiryaev. *Probability*. Springer, 1996.
- [13] Elias M. Stein and Rami Shakarchi. *Real Analysis: Measure Theory, Integration, and Hilbert Spaces*. Princeton University Press, 2005.
- [14] Mark Thompson. Differential operators in infinite dimensional spaces. *Proc. of the 18th Brazilian Seminar in Analysis*, 1983.
- [15] Francois Trèves. *Basic Partial Differential Equations*. Academic Press, 1975.
- [16] D. Williams. *Diffusions, Markov Processes and Martingales*. Wiley, 1979.
- [17] J. Yeh. *Stochastic Processes and the Wiener Integral*. Marcel Dekker, 1973.
- [18] Kosaku Yosida. *Functional Analysis*. Springer, 1966.