České vysoké učení technické v Praze
Fakulta jaderná a fyzikálně inženýrská

Katedra fyziky
Obor: Matematické inženýrství
Zaměření: Matematická fyzika

Wienerův proces a vlastnosti jeho trajektorií
Wiener process and properties of its paths

BAKALÁŘSKÁ PRÁCE

Vypracoval: Jindřich Prokop
Vedoucí práce: Petr Veverka
Rok: 2013

Před svázáním místo téhle stránky vložíte zadání práce s podpisem děkana (bude to jediný oboustranný list ve Vaší práci) !!!!

Prohlášení

Prohlašuji, že jsem svou bakalářskou práci vypracoval samostatně a použil jsem pouze podklady uvedené v přiloženém seznamu.

V Praze dne ....

....
Jindřich Prokop

Název práce: **Wienerův proces a vlastnosti jeho trajektorií**

Autor: Jindřich Prokop

Obor: Matematické inženýrství
Druh práce: Bakalářská práce

Vedoucí práce: Petr Veverka
Katedra matematiky, Fakulta jaderná a fyzikálně inženýrská, České vysoké učení technické v Praze
Konzultant: —

Abstrakt: První část práce pokrývá základní pojmy teorie náhodných procesů. Dále je definován Wienerův proces, dokázána jeho existence a některé jeho základní vlastnosti. Konstrukci Itôva integrálu předchází důkazy některých vlastností trajektorií Wienerova procesu, zejména těch, které motivují Itôvu konstrukci. Po konstrukci integrálu následuje důkaz některých jeho vlastností. Nakonec jsou řešeny podmínky, jež je třeba klást na integrandy zahrnující řízení, aby existoval jejich Itôův integrál. Cílem práce je umožnit formulaci stochastické diferenciální rovnice a najít podmínky, které je třeba klást na členy v ní vystupující.

Klíčová slova: náhodný proces, Wienerův proces, Itôův integrál,

Title: **Wiener process and properties of its paths**

Author: Jindřich Prokop

Abstract: The basic concepts of stochastic theory are covered in the first part of the work. Wiener process is defined and further its existence is proven along with some of its properties. The construction of Itô integral is preceded by some of path properties of Wiener process, most importantly by those, which motivates Itô's construction. These properties are rigorously proven in the work. Some of the integral properties are shown after its construction. The conditions that must be imposed on the integrand containing the stochastic control in order to ensure the existence of its Itô inegral are examined at the end of the text. The aim of the work is to allow the formulation of stochastic differential equation and to examine the conditions, which must be imposed on its terms.

Key words: stochastic process, Winer process, Itô integral

Table of Contents

Introduction **7**

1 Stochastic Process **8**

1.1 Basic Definitions 8

1.2 Daniell-Kolmogorov Theorem 10

2 Wiener process **14**

2.1 Definitions 14

2.2 Existence 16

2.3 Basic Properties of Brownian Motion 18

2.4 Gaussian Property of Wiener Process 19

2.5 Markov Property of Wiener Process 19

2.6 Martingale Property of Wiener Process 20

2.7 Equivalence Tranformations 20

3 Sample Paths **22**

3.1 Strong Law of Large Numbers 22

3.2 Quadratic Variation 23

3.3 Nondifferentiability of Brownian motion 24

4 Itô Integral **27**

4.0 Integration with Respect to Brownian Motion 27

4.1 Itô Integral 29

4.2 Martingale Property if Itô Integral 35

4.3 Wiener Integral 37

5 Integration with Controlled Integrand **40**

5.1 Integration of $\sigma(X, u)$ with Respect to B 40

Conclusion **42**

List of Symbols **43**

Bibliography **44**

Introduction

The stochastic theory justifies itself purely on theoretical level. Yet its core lays in applications, be it in biology, finance, insurance or sociology. One of the fundamental notions of applied stochastic theory is stochastic differential equation. This text covers the elementary concepts of stochastic theory, such as stochastic process, and examines perhaps the most important one of them - Wiener process. Further, the Itô integral is constructed. This will allow to formulate some stochastic differential equations meaningfully.
Basic definitions are given in the first section of chapter 1; the important result from probability theory, Daniell-Kolmogorov theorem, which makes possible to prove the existence of Wiener process, is proven in the second one. The second chapter introduces the central notion - the Wiener process. In the second section the prove of its existence is carried out and in the rest of the chapter some rather elementary properties are shown. Succeeding chapter reveals the most important properties of Brownian trajectories, i.e. sample paths of Wiener process - mainly the lack of differentiability and the quadratic variance of the paths. These properties are the fundamental reason for different approach to differential and integral calculus, the failure of Lebesgue-Stieltjes integral is illustrated in the prelude of chapter 4. In section 4.1 the construction of Itô integral is carried out and again some basic properties are shown. The ultimate aim is to model stochastic systems and eventually find the control that allows us to get as close as possible to our aim, for example to maximize our profit or to destroy tumour with as little side effect as possible. Thus it is useful to analyse the conditions we have to impose on the controlled integrands in order to ensure the existence of their Itô integral and consequently to enable us to use them within stochastic differential equations. These conditions examined in chapter 5.

1 Stochastic Process

Basic concepts of stochastic theory are defined in this chapter and prove of an important result, the Daniell-Kolmogorov theorem, is given. Despite [3] as well as [4] were used, the most important reference, mainly for second section of the chapter, is [1]. These three books, with [1] being the most important one, are the source not only for this chapter, but also for two succeeding ones.

1.1 Basic Definitions

Definition 1.1.1. Stochastic Process
Let $(\Omega, \mathcal{A}, P)$ be a probability space, let $\mathcal{T} \subset \mathbb{R}$ be arbitrary nonempty set. A set $X \equiv \{X(t, \cdot) | t \in \mathcal{T}\}$, where $X(t, \cdot)$ is a random variable on $(\Omega, \mathcal{F}, P)$ for every $t \in \mathcal{T}$, is called a stochastic process.

Remark 1.1.1.

(i) We can also view stochastic process as a real-valued function on $\mathcal{T} \times \Omega$.

(ii) For notational simplicity or convenience we may write X_t or $X(t)$ instead of $X(t, \cdot)$ and also $X_t(\omega)$ instead of $X(t, \omega)$.

Definition 1.1.2. Sample Path of Stochastic Process
Having a stochastic process $X = \{X(t, \cdot) | t \in \mathcal{T}\}$ on a probability space $(\Omega, \mathcal{A}, P)$ we call $X(\cdot, \omega) \equiv X_{\cdot}(\omega)$ a sample path or a realization or a trajectory of X for any $\omega \in \Omega$.

Definition 1.1.3. Distribution and Density Function of k-th Order
Let $\{X(t, \cdot) | t \in \mathcal{T}\}$ be a stochastic process. Then the distribution function of k-th order of this stochastic process is given by

$$F(x_1, \dots, x_k, t_1, \dots, t_k) = P(X(t_1) \leqslant x_1, \dots, X(t_k) \leqslant x_k),$$

and assuming that the derivative exists, the density function is given by

$$f(x_1, \dots, x_k, t_1, \dots, t_k) = \frac{\partial^k}{\partial x_1 \dots \partial x_k} F(x_1, \dots, x_k, t_1, \dots, t_k),$$

for every $x_1, \dots, x_k \in \mathbb{R},\ t_1, \dots, t_k \in \mathcal{T}$.

Definition 1.1.4. Modification
Let $X = \{X(t, \cdot) | t \in \mathcal{T}\}$ be a stochastic process on a probability space $(\Omega, \mathcal{A}, P)$. Any other process $Y = \{Y(t, \cdot) | t \in \mathcal{T}\}$ on the same space such that $\forall t \in \mathcal{T}$ $P(\{\omega | X(t, \omega) = Y(t, \omega)\}) = 1$ is called a modification of X.

Definition 1.1.5. Filtration
Suppose we have a probability space $(\Omega, \mathcal{A}, P)$, $T \in \mathbb{R}^+$ fixed. Then a collection $\{\mathcal{F}_t | t \in [0, T]\}$ of sub-σ-algebras of $\mathcal{A}$ is a filtration if for every $s, t \in [0, T]$ such that $s \leqslant t$ we have $\mathcal{F}_s \subset \mathcal{F}_t$. A probability space with a filtration is called a stochastic basis or a filtered probability space.

Definition 1.1.6. Usual Conditions
Assume $\{\mathcal{F}_t | t \geqslant 0\}$ is a filtration on $(\Omega, \mathcal{A}, P)$, where $\mathcal{A}$ is P-complete, that is $\{A | (\exists B \in \mathcal{A})(A \subset B \; et \; P(B) = 0)\} \subset \mathcal{A}$. Then we say it satisfies the usual conditions if

1. $\mathcal{F}_0$ is P-complete, that is $\{A \in \mathcal{A} | P(A) = 0\} \subset \mathcal{F}_0$,
2. $\mathcal{F}_{t+} := \bigcap\limits_{s>t} \mathcal{F}_s = \mathcal{F}_t$.

Remark 1.1.2. The second property of last definition is called the right continuity of filtration, i.e. the filtration $\{\mathcal{F}_t\}$ satisfying 2. of definition 1.1.6 is said to be right-continuous.

Definition 1.1.7. Adapted Process
Assume we have a filtration $\{\mathcal{F}_t | t \in [0, T]\}$ on a probability space $(\Omega, \mathcal{A}, P)$. Stochastic process $X \equiv \{X(t, \cdot) | t \in [0, T]\}$ such that $\forall t \in [0, T]$ random variable $X(t)$ is $\mathcal{F}_t$-measurable is called $\{\mathcal{F}_t\}$-adapted. Sometimes a less precise notation '$\mathcal{F}_t$-adapted' is used as well, as its meaning is evident.

Remark 1.1.3. Every process generates a filtration to which it is adapted. Suppose we have a stochastic process $\{X(t, \cdot) | t \in [0, T]\}$. Then $\mathcal{F}^X(t) := \sigma(X(s) | 0 \leqslant s \leqslant t)$ is a filtration and X is obviously adapted to it. This filtration is sometimes called canonical or natural. $\mathcal{F}_t^X$ will denote the canonical filtration generated by process X throughout the text.

Definition 1.1.8. Gaussian Process
Stochastic process $\{X(t, \cdot) | t \in \mathcal{T}\}$ is called Gaussian if $\forall k \in \mathbb{N}$, $\forall t_1, \ldots, t_k \in \mathcal{T}$ random vector $(X(t_1), \ldots, X(t_k))$ has a multinormal distribution, i.e. if for every $X(t_j)$ there is a set of standard normal random variables $Z_1, \ldots, Z_m$ and $\mu, c_1, \ldots, c_m \in \mathbb{R}$ such that $X_k = \mu + \sum c_j Z_j$

Definition 1.1.9. Markov Process
Assume we have a filtration $\{\mathcal{F}_t | t \in [0, T]\}$ on a probability space $(\Omega, \mathcal{A}, P)$. Let $X \equiv \{X(t, \cdot) | t \in [0, T]\}$ be an $\mathcal{F}_t$-adapted stochastic process with $X_t \in L^1(P)$ for every t. If for every $s, t \in [0, T]$ such that $s \leqslant t$ and for every non-negative Borel-measurable function f there exists another Borel-measurable function g such that $\mathbb{E}[f(X(t)) | \mathcal{F}(s)] = g(t, X(s))$, we call the process X Markov with respect to $\{\mathcal{F}_t\}$.

Definition 1.1.10. Martingale
Assume we have a filtration $\{\mathcal{F}(t) | t \in [0, T]\}$ on a probability space $(\Omega, \mathcal{A}, P)$. Let

$\{X(t,\cdot)|t \in [0,T]\}$ be an adapted stochastic process with $X(t)$ in L^1 for every t. If $\forall s,t \in [0,T]$ such that $s \leqslant t$ we have that $\mathbb{E}[X(t)|\mathcal{F}(s)] \xlongequal{P\text{-a.s.}} X(s)$, we say that X is a martingale with respect to $\{\mathcal{F}_t\}$.

1.2 Daniell-Kolmogorov Theorem

Definition 1.2.1. n-dimensional Cylinder Set, $\mathcal{K}\left(\mathbb{R}^{[0,+\infty)}\right)$
Any set of the form $C_{t_1,\ldots,t_n,A} = \{\omega \in \mathbb{R}^{[0,+\infty)}|(\omega(t_1),\ldots,\omega(t_n)) \in A,\ t_i \geqslant 0\}$, where $A \in \mathcal{B}(\mathbb{R}^n)$ and $\mathbb{R}^{[0,+\infty)} \equiv \{f : [0,+\infty) \to \mathbb{R}\}$, is called an n-dimensional cylinder set in $\mathbb{R}^{[0,+\infty)}$. The field of all cylinder sets $\{C_{t_1,\ldots,t_n,A}|n \in \mathbb{N},\ A \in \mathcal{B}(\mathbb{R}^n)\}$ is denoted by $\mathcal{C}$, and $\sigma(\mathcal{C})$ will be denoted by $\mathcal{K}\left(\mathbb{R}^{[0,+\infty)}\right)$.

Definition 1.2.2. Consistent Family of Finite-Dimensional Distributions
Let $\mathcal{T} \subset \mathbb{R}$ be an arbitrary non-empty set and $\forall k \in \mathbb{N},\ \forall t_1,\ldots,t_k \in \mathcal{T}$ let $P_{t_1,\ldots,t_k}$ be a probability measure on $\mathbb{R}^k$ such that $\forall B_1,\ldots,B_k \in \mathcal{B}(\mathbb{R})$:

1. $P_{t_1,\ldots,t_k}(B_1 \times \ldots \times B_k) = P_{t_{\pi(1)},\ldots,t_{\pi(k)}}(B_{\pi(1)} \times \ldots \times B_{\pi(k)}) \quad \forall \pi \in S_k,$
2. $P_{t_1,\ldots,t_{k-1}}(B_1 \times \ldots \times B_{k-1}) = P_{t_1,\ldots,t_k,}(B_1 \times \ldots \times B_{k-1} \times \mathbb{R}).$

Then we call the collection $\{P_{t_1,\ldots,t_k}|k \in \mathbb{N},\ t_1,\ldots,t_k \in \mathcal{T}\}$ a consistent family of finite-dimensional distributions.

Remark 1.2.1. Having a probability measure P on $\left(\mathbb{R}^{[0,+\infty)},\mathcal{K}\left(\mathbb{R}^{[0,+\infty)}\right)\right)$ we can define

$$P_{t_1,\ldots,t_k}(A) := P\{\omega \in \mathbb{R}^{[0,+\infty)}|(\omega(t_1),\ldots,\omega(t_k)) \in A\} \qquad \forall A \in \mathcal{B}(\mathbb{R}^k).$$

It is easily seen that:

1. $P_{t_1,\ldots,t_{k+1}}(A \times \mathbb{R}) = P\{\omega \in \mathbb{R}^{[0,+\infty)}|(\omega(t_1),\ldots,\omega(t_k)) \in A,\ \omega(t_{k+1}) \in \mathbb{R}\} =$
   $= P\{\omega \in \mathbb{R}^{[0,+\infty)}|(\omega(t_1),\ldots,\omega(t_k)) \in A\} = P_{t_1,\ldots,t_k}(A).$
2. $P_{t_1,\ldots,t_k}(B_1 \times \ldots \times B_k) = P\{\omega \in \mathbb{R}^{[0,+\infty)}|\omega(t_1) \in B_1,\ldots,\omega(t_k) \in B_k\} =$
   $= P\{\omega \in \mathbb{R}^{[0,+\infty)}|\omega(t_{\pi(1)}) \in B_{\pi(1)},\ldots,\omega(t_{\pi(k)}) \in B_{\pi(k)}\} =$
   $= P_{t_{\pi(1)},\ldots,t_{\pi(k)}}(B_{\pi(1)} \times \ldots \times B_{\pi(k)}).$

Hence $\{P_{t_1,\ldots,t_k}\}$ is consistent family of distributions. The succeeding theorem states that having a consistent family of distributions means having defined a probability measure on $\left(\mathbb{R}^{[0,+\infty)},\mathcal{K}\left(\mathbb{R}^{[0,+\infty)}\right)\right)$.

Lemma 1.2.1. Let $A \in \mathcal{B}(\mathbb{R}^n)$ and let P be a probability measure on $(\mathbb{R}^n,\mathcal{B}(\mathbb{R}^n))$. Then $\forall \varepsilon > 0$ there is $F = \overline{F}$ and $G = G^\circ$ such that $F \subset A \subset G$ and $P(G \setminus F) < \varepsilon$.

Proof. Let $\mathcal{F}$ denote the collection of all sets of desired property, that is

$$\mathcal{F} \equiv \{A \in \mathcal{B}(\mathbb{R}^n) | (\exists F = \overline{F}, G = G^\circ)(F \subset A \subset G) \ et \ P(G \setminus F) < \varepsilon)\}$$

We will first show $\mathcal{F}$ is a σ-field. Trivially $\emptyset \in \mathcal{F}$. Let $A \in \mathcal{F}$, i.e. $\forall \varepsilon > 0$ there is a $F = \overline{F}$ and $G = G^\circ$ such that $F \subset A \subset G$ and $P(G \setminus F) < \varepsilon$. Evidently $G^c = \overline{G^c}$, $F^c = (F^c)^\circ$, $G^c \subset A^c \subset F^c$, and $P(F^c \setminus G^c) = P(G \setminus F) < \varepsilon$. Therefore $A^c \in \mathcal{F}$. We now show that for $\{A_k\}_1^\infty \subset \mathcal{F}$ $\bigcup\limits_1^\infty A_k \in \mathcal{F}$. Thus suppose that $\forall \varepsilon > 0$ there is a sequence of closed sets $\{F_k\}$ and a sequence of open sets $\{G_k\}$ such that $F_k \subset A_k \subset G_k$ and $P(G_k \setminus F_k) < \frac{\varepsilon}{2^{k+1}}$. Let us set $G := \bigcup\limits_1^\infty G_k$ and $F := \bigcup\limits_1^m F_k$, where m is chosen so that $P(\bigcup\limits_1^\infty F_k \setminus \bigcup\limits_1^m F_k) < \frac{\varepsilon}{2}$. Then we have $F \subset A \subset G$, F closed, G open and

$$P(G \setminus F) \leqslant P(G \setminus \bigcup_1^\infty F_k) + P(\bigcup_1^\infty F_k \setminus F) < \sum_1^\infty \frac{\varepsilon}{2^{k+1}} + \frac{\varepsilon}{2} = \varepsilon.$$

Hence $\mathcal{F}$ is a σ-algebra.
For arbitrary $F = \overline{F}$ we define $G_k := \{x \in \mathbb{R}^n | \, \|x - y\| < \frac{1}{k}$, where $y \in F\}$. Then $G_k = G_k^\circ$, $G_{k+1} \subset G_k$ and $\bigcap\limits_1^\infty G_k = F$. Therefore $\forall \varepsilon > 0$ there is m such that $P(G_m \setminus F) < \varepsilon$, which means that $F \in \mathcal{F}$.
Since $\mathcal{B}(\mathbb{R}^n)$ is generated by closed sets and $\mathcal{F}$ contains all closed sets and is itself a σ-field we have shown that $\mathcal{F} \supset \mathcal{B}(\mathbb{R}^n)$. By the construction of $\mathcal{F}$ we have that $\mathcal{F} \subset \mathcal{B}(\mathbb{R}^n)$, hence $\mathcal{F} = \mathcal{B}(\mathbb{R}^n)$.

□

Theorem 1.2.1. Daniell-Kolmogorov
For every consistent family of finite-dimensional distributions $\{P_{t_1,\dots,t_k}\}$, there exists a probability measure P on $\left(\mathbb{R}^{[0,+\infty)}, \mathcal{K}\left(\mathbb{R}^{[0,+\infty)}\right)\right)$ such that

$$P\left(\left\{\omega \in \mathbb{R}^{[0,+\infty)} \middle| (\omega(t_1), \dots, \omega(t_k)) \in A\right\}\right) = P_{t_1,\dots,t_k}(A) \qquad \forall A \in \mathcal{B}(\mathbb{R}^k).$$

Proof. For every $C_{t_1,\dots,t_k,A} \in \mathcal{C}$ we define

$$Q(C_{t_1,\dots,t_k,A}) := P_{t_1,\dots,t_k}(A).$$

Since $\{P_{t_1,\dots,t_k}\}$ is consistent, Q is a well defined set function on $\mathcal{C}$. Assume we have $C \equiv C_{t_1,\dots,t_k,A}, D \equiv D_{\tau_1,\dots,\tau_l,B} \in \mathcal{C}$, $C \cap D = \emptyset$, then

$$\begin{aligned} C &= \{\omega | (\omega(t_1), \dots, \omega(t_k)) \in A\} = \{\omega | (\omega(t_1), \dots, \omega(t_k), \omega(\tau_1), \dots, \omega(\tau_l)) \in A \times \mathbb{R}^l\} \\ D &= \{\omega | (\omega(\tau_1), \dots, \omega(\tau_l)) \in B\} = \{\omega | (\omega(t_1), \dots, \omega(t_k), \omega(\tau_1), \dots, \omega(\tau_l)) \in \mathbb{R}^k \times B\} \\ \emptyset &= C \cap D = \{\omega | (\omega(t_1), \dots, \omega(t_k), \omega(\tau_1), \dots, \omega(\tau_l)) \in ((A \times \mathbb{R}^l) \cap (\mathbb{R}^k \times B))\} \end{aligned}$$

hence $((A \times \mathbb{R}^l) \cap (\mathbb{R}^k \times B)) = \emptyset$, which combined with the fact that $P_{t_1,\ldots,t_k,\tau_1,\ldots,\tau_l}$ is a probability measure yields

$$\begin{aligned} Q(C \cup D) &= P_{t_1,\ldots,\tau_l}\{\omega|(\omega(t_1),\ldots,\omega(t_k),\omega(\tau_1),\ldots,\omega(\tau_l)) \in (A \times \mathbb{R}^l) \cup (\mathbb{R}^k \times B)\} = \\ &= P_{t_1,\ldots,\tau_l}\{\omega|(\omega(t_1),\ldots,\omega(t_k),\omega(\tau_1),\ldots,\omega(\tau_l)) \in A \times \mathbb{R}^l\} + \\ &\quad + P_{t_1,\ldots,\tau_l}\{\omega|(\omega(t_1),\ldots,\omega(t_k),\omega(\tau_1),\ldots,\omega(\tau_l)) \in \mathbb{R}^k \times B\} = \\ &= Q(C) + Q(D). \end{aligned}$$

That is Q is finitely additive, also evidently $Q(\mathbb{R}^{[0,+\infty)}) = 1$. If we show that Q is countably additive we will know by the measure extension theorem that Q has a unique extension P to $\mathcal{K}\left(\mathbb{R}^{[0,+\infty)}\right)$ that we seek.
Let $\{B_k\}_1^\infty$ be a sequence of disjoint subsets of $\mathcal{C}$ with $B := \bigcup\limits_1^\infty B_k \in \mathcal{C}$. We want to show that $Q(B) = \sum\limits_1^\infty Q(B_k)$, or more precisely that

$$\lim_{m\to\infty} Q\left(\sum_{k=1}^m B_k\right) = \lim_{m\to\infty} \sum_{k=1}^m Q(B_k).$$

To prove it we first define the sequence $\{C_m\}_1^\infty$: $C_m := B \setminus \bigcup\limits_1^m B_k$, thus $\forall n \in \mathbb{N}$, $B = C_m \cup (\bigcup\limits_1^m B_k)$, and by finite additivity $Q(B) = Q(C_m) + \sum\limits_1^m Q(B_k)$. Hence we want to show that $Q(C_m) \xrightarrow{m\to\infty} 0$. Note that because $C_{m+1} = C_m \setminus B_{m+1}$, we have that $Q(C_{m+1}) \geqslant Q(C_m)$, therefore the limit exists. We prove that it equals 0 by contradiction. Thus suppose that $\lim\limits_{m\to\infty} Q(C_m) = \varepsilon > 0$. Since $\forall m \in \mathbb{N} \quad C_m \in \mathcal{C}$, we know C_m is of the form

$$C_m = \{\omega \in \mathbb{R}^{[0,+\infty)}|(\omega(\tau_1^{(m)}),\ldots,\omega(\tau_{k_m}^{(m)})) \in A_{k_m}\} \qquad \text{where } A_{k_m} \in \mathcal{B}(\mathbb{R}^{k_m}).$$

Furthermore we have $C_{m+1} \subset C_m$, hence we can pick $(\tau_1^{(m+1)},\ldots,\tau_{k_{m+1}}^{(m+1)})$ so that $\forall i \in \{1,\ldots,k_m\} \quad \tau_i^{(m)} = \tau_i^{(m+1)}$ and $A_{k_m+1} \subset A_{k_m} \times \mathbb{R}^{k_{m+1}-k_m}$. Therefore we can drop the upper indices of times and see that C_m is of the form

$$C_m = \{\omega \in \mathbb{R}^{[0,+\infty)}|(\omega(\tau_1),\ldots,\omega(\tau_{k_m})) \in A_{k_m}\} \qquad \text{where } A_{k_m} \in \mathcal{B}(\mathbb{R}^{k_m}).$$

To simplify the notation we now define new sequence of sets $\{D_m\}_1^\infty$:

$$D_1 := \{\omega|\omega(\tau_1) \in \mathbb{R}\} \qquad \ldots \qquad D_{k_1-1} := \{\omega|(\omega(\tau_1),\ldots,\omega(\tau_{k_1-1})) \in \mathbb{R}^{k_1-1}\}$$

$$D_{k_1} := C_1, \quad D_{k_1+1} := \{\omega|(\omega(\tau_1),\ldots,\omega(\tau_{k_1+1})) \in A_1 \times \mathbb{R}\},\ldots, D_{k_2-1} := \ldots$$

$$D_{k_2} := C_2, \ \ldots, D_{k_\ell} := C_\ell, \ldots$$

It is apparent that $\bigcap\limits_1^\infty D_m = \bigcap\limits_1^\infty C_m = \emptyset$, $D_{m+1} \subset D_m$ and D_m is of the form

$$D_m = \{\omega \in \mathbb{R}^{[0,+\infty)}|(\omega(\tau_1),\ldots,\omega(\tau_m)) \in A_m\} \qquad \text{where } A_m \in \mathcal{B}(\mathbb{R}^m).$$

According to the preceding lemma $\forall m \in \mathbb{N}$ there is $F_m = \overline{F_m}$ and $G_m = G_m^{\circ}$ such that $F_m \subset A_m \subset G_m$ and $P_{t_1,\ldots,t_m}(A_m \setminus F_m) \leqslant P_{t_1,\ldots,t_m}(G_m \setminus F_m) < \frac{\varepsilon}{2^m}$. Since F_m is closed $\forall m \in \mathbb{N}$ sets $K_m := F_m \cap \overline{B(0, r_m)} \in \mathbb{R}^m$, where r_m is large enough for $P_{t_1,\ldots,t_m}(A_m \setminus K_m) < \frac{\varepsilon}{2^m}$ to be true, are compact. Utilising K_m we define yet another sequence of sets:

$$E_m := \{\omega \in \mathbb{R}^{[0,+\infty)} | (\omega(\tau_1), \ldots, \omega(\tau_m)) \in K_m\}.$$

Hence $E_m \subset D_m$ and also $Q(D_m \setminus E_m) = P_{t_1,\ldots,t_m}(A_m \setminus K_m) < \frac{\varepsilon}{2^m}$. To ensure that $\{E_m\}$ is nonincreasing we redefine it along with $\{K_m\}$:

$$\tilde{E}_m := \bigcap_1^m E_j = \{\omega | (\omega(\tau_1), \ldots, \omega(\tau_m)) \in \tilde{K}_m\} \qquad \text{where } \tilde{K}_m = \bigcap_1^m (K_j \times \mathbb{R}^{m-j}).$$

Because

$$\begin{aligned} P_{t_1,\ldots,t_m}(\tilde{K}_m) = Q(\tilde{E}_m) &= \{\text{by finite additivity}\} = Q(D_m) - Q(D_m \setminus \tilde{E}_m) = \\ &= Q(D_m) - Q\left(\bigcup_{k=1}^m (D_m \setminus E_k)\right) \geqslant \{\text{because } D_{k+1} \subset D_k\} \geqslant \\ &\geqslant Q(D_m) - Q\left(\bigcup_{k=1}^m (D_k \setminus E_k)\right) \geqslant \varepsilon - \sum_{k=1}^m \frac{\varepsilon}{2^k} > 0, \end{aligned}$$

Thus $\tilde{K}_m \neq \emptyset$.
We thus can take some $(x_1^{(m)}, \ldots, x_m^{(m)}) \in \tilde{K}_m$ for every $m \in \mathbb{N}$. Since $\forall m \geqslant n$ $\tilde{K}_m \times \mathbb{R}^{n-m} \subset \tilde{K}_n$ we have that $(x_1^{(m)}, \ldots, x_{m-k}^{(m)}) \in \tilde{K}_{m-k}$ for any $k < m$. Specifically $\forall m,\ \{x_1^{(m)}\}_1^\infty \subset \tilde{K}_1$. Since $\tilde{K}_1$ is compact, there is a convergent subsequence $\{x_1^{(m_k)}\}$ such that $x_1^{(m_k)} \xrightarrow{k\to\infty} x_1 \in \tilde{K}_1$. $\{(x_1^{(m_k)}, x_2^{(m_k)})\} \in \tilde{K}_2$ has a convergent subsequence with limit $(x_1, x_2) \in \tilde{K}_2$ by the very same argument. Iterating this step we construct $(x_1, x_2, \ldots) \in \mathbb{R} \times \mathbb{R} \times \ldots$ such that $\forall m \in \mathbb{N}$ $(x_1, \ldots, x_m) \in \tilde{K}_m$. We now define nonempty set

$$S := \{\omega \in \mathbb{R}^{[0,+\infty)} | \forall i \in \mathbb{N}\ \omega(\tau_i) = x_i\} \subset \tilde{E}_m \subset D_m \text{ for any } m \in \mathbb{N}.$$

Hence

$$\emptyset \neq S \subset \bigcap_1^\infty D_m \Longrightarrow \bigcap_1^\infty D_m \neq \emptyset,$$

which contradicts the fact that $\bigcap_1^\infty D_m = \emptyset$ and therefore $\lim_{m\to\infty} Q(C_m) = 0$. Thus Q is countably additive. P is an extension of Q to $\mathcal{K}\left(\mathbb{R}^{[0,+\infty)}\right)$ mentioned above. $\square$

2 Wiener process

This chapter introduces one of the most important stochastic processes - the Wiener process. The basics are followed by the proof of its existence. Some useful properties of this object are proved at the end of this chapter.

2.1 Definitions

Definition 2.1.1. Wiener process
A stochastic process $W \equiv \{W(t) | t \in \mathbb{R}_0^+\}$ such that

1) $W_0 = 0$ a. s.

2) $\forall\ t_1 < t_2 \leqslant t_3 < t_4 \in \mathbb{R}_0^+$ the increments $W(t_2) - W(t_1)$ and $W(t_4) - W(t_3)$ are independent random variables. ($W(t)$ has independent increments.)

3) $\forall t_1, t_2, s \in \mathbb{R}$ such that $t_1, t_2, t_1+s, t_2+s \in \mathbb{R}_0^+$ the increments $W(t_2+s) - W(t_1+s)$ and $W(t_2) - W(t_1)$ are equally distributed. ($W(t)$ has stationary increments.)

4) $\forall\ t_1 \leqslant t_2$ the increment $W(t_2) - W(t_1)$ is normally distributed, its expected value is 0 ($\mathbb{E}[W(t_2) - W(t_1)] = 0$) and its variance is equal to the corresponding (scaled) time difference ($\mathrm{Var}[W(t_2) - W(t_1)] = \sigma^2(t_2 - t_1)$),

is called Wiener process.

Remark 2.1.1.

i) Conditions 2) and 3) imply 4) through central limit theorem, while 4) implies 3) trivially. Hence the conditions 1), 2), 3) or 1), 2), 4) define the same object as all four together.

ii) We shall see that there are several equivalent ways of characterizing Wiener process as we prove some of its properties.

iii) Wiener process is alternatively called Brownian motion.

iv) If $\sigma^2 = 1$ we call this process standard Brownian motion.

Definition 2.1.2. Filtration for Wiener process
Assume we have a filtration $\{\mathcal{F}_t | t \geqslant 0\}$ and an adapted Wiener process $\{W_t | t \geqslant 0\}$ on a probability space $(\Omega, \mathcal{A}, P)$. Then $\{\mathcal{F}_t | t \geqslant 0\}$ is a filtration for Wiener process if and only if $\forall s, t$ such that $0 \leqslant s \leqslant t$ the increment $W_t - W_s$ is independent of $\mathcal{F}_s$.

Remark 2.1.2. A filtration generated by Wiener process itself is a filtration for this Wiener process.

It is useful to include the filtration in the definition of Wiener process.

Definition 2.1.3. $\mathcal{F}_t$-Wiener Process
Let $(\Omega, \mathcal{A}, P)$ be a probability space. An adapted process $\{W_t, \mathcal{F}_t | t \in \mathbb{R}_0^+\}$ defined on this probability space is said to be $\mathcal{F}_t$-Wiener if

1) $W_0 = 0$ P-almost surely,

2) $W(t) - W(s)$ if independent of $\mathcal{F}_s$ for every $s \leqslant t,\ s, t \in \mathbb{R}_0^+$,

3) The increments $W(t) - W(s)$ are normally distributed with mean zero and $\mathrm{Var}[W(t) - W(s)] = \sigma^2(t-s)$ for every $s \leqslant t,\ s, t \in \mathbb{R}_0^+$,

4) W is P-almost surely continuous, that is for P-almost every ω the trajectory $W(\cdot, \omega)$ is continuous.

Remark 2.1.3.

i) We have added a filtration directly into the definition of Brownian motion. But if W is a Wiener process in the sense of definition 2.1.1, it is always $\mathcal{F}_t^W$-Wiener.

ii) Every time we include a filtration $\mathcal{F}_t$ in the process $W = \{W_t, \mathcal{F}_t\}$ and say that it is a Wiener process, we actually mean that it is $\mathcal{F}_t^W$-Wiener.

iii) If $W \equiv \{W_t, \mathcal{F}_t\}$ is $\mathcal{F}_t$-Wiener, then $\frac{1}{\sigma^2} W \equiv \{\frac{1}{\sigma^2} W_t, \mathcal{F}_t\}$ is called Standard Brownian motion with respect to $\mathcal{F}_t$. It is evidently $\mathcal{F}_t$-Wiener process with normalized variance.

iv) We can always obtain general $\mathcal{F}_t$-Wiener process from standard Brownian motion by scaling it.

v) Alternatively, we may use filtration for Wiener process (definition 2.1.2) and Wiener process (definition 2.1.1) to define $\mathcal{F}_t$-Wiener Process as W on a stochastic basis $(\Omega, \mathcal{F}, \mathcal{F}_t, P)$ such that

   1) W is a Wiener process according to definition 2.1.2.

   2) $\{\mathcal{F}_t\}$ is a filtration for this Wiener process.

2.2 Existence

Suppose we have a Brownian motion $B \equiv \{B_t | t \in \mathbb{R}_0^+\}$. Given its properties, we have that cumulative distribution function of random vector $(B_{s_1}, \ldots, B_{s_n})$ is given for all $0 < s_1 < \ldots < s_n$ by

$$F_{s_1,\ldots,s_n}(x_1, \ldots, x_n) = \\ = \int_{-\infty}^{x_1} \ldots \int_{-\infty}^{x_n} p(s_1, 0, y_1) \cdot \ldots \cdot p(s_n - s_{n-1}, y_{n-1}, y_n) \mathrm{d}y_n \ldots \mathrm{d}y_1, \tag{2.1}$$

where $p(\sigma^2, x, y) = \frac{1}{\sqrt{2\pi}\sigma} \cdot e^{-\frac{(x-y)^2}{2\sigma^2}}$ for any $\sigma > 0$, $x, y \in \mathbb{R}$. On the other hand, seeing that the increments $\{B_{s_j} - B_{s_{j-1}}\}_1^n$ of a stochastic process B are independent and normally distributed with mean zero and variance $s_j - s_{j-1}$ if and only if cumulative distribution function of $(B_{s_1}, \ldots, B_{s_1})$ is given by (2.1) , we can state that $B = \{B_t, \mathcal{F}_t^B | 0 \leqslant t \in \mathbb{R}_0^+\}$ is a standard Brownian motion if and only if for every $n \in \mathbb{N}$ and any $0 < s_1 < \ldots < s_n$ the distribution of $(B_{s_1}, \ldots, B_{s_n})$ is given by (2.1) and $B_\cdot(\omega)$ is continuous for almost every $\omega \in \Omega$.
We will now construct probability measure P on $(\Omega, \mathcal{A}) \equiv \left(\mathbb{R}^{[0,+\infty)}, \mathcal{K}\left(\mathbb{R}^{[0,+\infty)}\right)\right)$ and a stochastic process B such that under P it is a Brownian motion if its sample paths are continuous almost surely and eventually show that there is a continuous modification of this process.
Thus suppose we have a stochastic process B on $\left(\mathbb{R}^{[0,+\infty)}, \mathcal{K}\left(\mathbb{R}^{[0,+\infty)}\right)\right)$. For arbitrary $n \in \mathbb{N}$ and $t_1, \ldots, t_n \geqslant 0$ distinct times let $(B_{t_1}, \ldots, B_{t_n})$ have the distribution determined by (2.1), that is $\forall A = (A_1 \times \ldots \times A_n) \in \mathcal{B}(\mathbb{R}^n)$

$$P_{t_1,\ldots,t_n}(A) := \int_{A_1} \ldots \int_{A_n} p(t_1, 0, y_1) \cdot \ldots \cdot p(t_n - t_{n-1}, y_{n-1}, y_n) \mathrm{d}y_n \ldots \mathrm{d}y_1.$$

Lemma 2.2.1. The just defined family of finite-dimensional distributions $\{P_{t_1,\ldots,t_n}\}$ is consistent.

Proof. That

$$P_{t_1,\ldots,t_n}(A_1 \times \ldots \times A_n) = P_{t_{\pi(1)},\ldots,t_{\pi(n)}}(A_{\pi(1)} \times \ldots \times A_{\pi(n)}) \qquad \forall A \in \mathcal{B}(\mathbb{R}^n)$$

for any permutation π is evident from definition of $P_{t_1,\ldots,t_n}$ as well as that

$$P_{t_1,\ldots,t_n}(A_1 \times \ldots \times A_{n-1} \times \mathbb{R}) = P_{t_1,\ldots,t_{n-1}}(A_1 \times \ldots \times A_{n-1}) \qquad \forall A \in \mathcal{B}(\mathbb{R}^n).$$

This proves the claim. □

By Daniell-Kolmogorov theorem (theorem 1.2.1) we thus have probability P on $\left(\mathbb{R}^{[0,+\infty)}, \mathcal{K}\left(\mathbb{R}^{[0,+\infty)}\right)\right)$ such that B has all needed properties but continuity of sample paths.

Theorem 2.2.1. Kolmogorov-Čentsov
Assume we have a stochastic process $X = \{X_t | 0 \leqslant t \leqslant T\}$ on a probability space $(\Omega, \mathcal{A}, P)$ such that $\exists \alpha, \beta, C > 0$ for which

$$\mathrm{E}\left[|X_t - X_s|^\alpha\right] \leqslant C|t-s|^{1+\beta} \tag{2.2}$$

holds for any $s, t \in [0, T]$ where T is arbitrary positive number.
Then there exists a continuous modification $Y = \{Y_t | 0 \leqslant t \leqslant T\}$ of X which is locally Hölder-continuous with exponent γ for any $\gamma \in (0, \frac{\beta}{\alpha})$, that is

$$P\left(\left\{\omega \,\middle|\, \sup_{\substack{0<t-s<h(\omega)\\ s,t\in[0,T]}} \frac{|Y_t(\omega) - Y_s(\omega)|}{|t-s|^\gamma} \leqslant \delta \right\}\right) = 1,$$

where h is almost surely positive random variable.

Proof. See [1], p. 53, theorem 2.8. □

Lemma 2.2.2. $B_t - B_s$ is normally distributed with mean zero and variance $t-s$ for any $0 \leqslant s < t$ implies that $\forall n \in \mathbb{N}$ there is $C_n > 0$ such that

$$\mathrm{E}\left[|B_t - B_s|^{2n}\right] = C_n |t-s|^n. \tag{2.3}$$

Proof. Because $X \equiv B_t - B_s \sim N(0, t-s)$, we have that the density of X is given by $f_X(x) = p\,(t-s, x, 0) = \frac{1}{\sqrt{2\pi(t-s)}} \cdot e^{-\frac{x^2}{2(t-s)}}$ and therefore

$$\mathrm{E}[X^{2n}] = \int_{\mathbb{R}} x^{2n} f_X(x)\mathrm{d}x = (2n-1)!!(t-s)^n.$$

Hence the sought constant C_n is $(2n-1)!!$. □

From this lemma and Kolmogorov-Čentsov theorem we conclude that for any arbitrary $T > 0$ there is a continuous modification $\tilde{W}^T$ of $\{B_t | 0 \leqslant t \leqslant T\}$ on $[0, T]$.

$$W^T := \begin{cases} \tilde{W}^T & \text{on } [0, T] \\ B & \text{on } (T, +\infty). \end{cases}$$

Thus W^T is modification of B on $[0, +\infty)$ that is continuous on $[0, T]$.
$P(\Omega_T) = 1$, where $\Omega_T := \{\omega | W_t^T(\omega) = B_t(\omega) \quad t \in [0, T] \cap \mathbb{Q}\}$, since W^T is a modification of $B = \{B_t | 0 \leqslant t < +\infty\}$ and $\Omega_T \subset \{\omega | W_t^T(\omega) = B_t(\omega),\ t \geqslant 0\}$. On $\tilde{\Omega} := \bigcap_1^\infty \Omega_T$ $\forall T_1, T_2 \in \mathbb{N}$ we have $W_t^{T_1}(\omega) = W_t^{T_2}(\omega)$ on $[0, \min\{T_1, T_2\}]$ since W^{T_1} and W^{T_2} are continuous on $[0, \min\{T_1, T_2\}]$ and coincide at every rational time in this interval. Hence the following definition of W is consistent:

$$W_t(\omega) := \begin{cases} W_t^T(\omega) & \forall \omega \in \tilde{\Omega}, \forall t \geqslant 0,\ T \geqslant t \text{ arbitrary.} \\ 0 & \forall \omega \in \Omega \setminus \tilde{\Omega}, \forall t \geqslant 0. \end{cases}$$

We have obtained continuous stochastic process W with independent increments $W_t - W_s \sim N(0, t-s)$. Therefore $\{W_t, \mathcal{F}_t^W | 0 \leqslant t < +\infty\}$ is a standard Brownian motion according to the definition 2.1.3.

2.3 Basic Properties of Brownian Motion

Some basic properties that follow from definition of Wiener process are listed here.

Theorem 2.3.1. Suppose $W = \{W_t, \mathcal{F}_t\}$ is a Wiener process. Then

i) $\mathrm{E}[W_t] = 0$,

ii) $\mathrm{E}[W_t^2] = t$,

iii) $\mathrm{E}[W_t|\mathcal{F}_s] = W_s \quad P$-almost surely, for all $s \leqslant t$,

iv) $\mathrm{Var}[W_t] = t$,

v) $\mathrm{E}[W_t W_s] = s \quad$ for all $s \leqslant t$, or equivalently $\mathrm{E}[W_t W_s] = \min\{s, t\}$,

vi) $\mathrm{Cov}(W_t, W_s) = s \quad$ for all $s \leqslant t$, equivalently $\mathrm{Cov}[W_t, W_s] = \min\{s, t\}$.

Proof.

i) Since W_0 is almost surely zero we have $E[W_t] = E[W_t - W_0] = 0$ by definition.

ii) For the same reason we have $\mathrm{E}[W_t^2] = \mathrm{E}[(W_t - W_0)^2] = t - 0 = t$.

iii) Because $W_t - W_s$ is independent of $\mathcal{F}_s$ by definition

$$\begin{aligned}\mathrm{E}[W_t|\mathcal{F}_s] &= \mathrm{E}[W_t - W_s + W_s|\mathcal{F}_s] = \mathrm{E}[W_t - W_s|\mathcal{F}_s] + \mathrm{E}[W_s|\mathcal{F}_s] \\ &\overset{P\text{-a.s.}}{=\!=\!=} \mathrm{E}[W_t - W_s] + W_s = W_s.\end{aligned}$$

iv) $\mathrm{Var}[W_t] = \mathrm{E}[W_t^2] - \mathrm{E}^2[W_t] = t - 0 = t$.

v) Since $W_t - W_s$ and $W_s - W_0$ are independent we have

$$\begin{aligned}\mathrm{E}[W_t W_s] &= \mathrm{E}[W_s(W_t - W_s + W_s)] = \mathrm{E}[W_s(W_t - W_s)] + \mathrm{E}[W_s^2] \\ &= \mathrm{E}[(W_s - W_0)(W_t - W_s)] + s = s.\end{aligned}$$

vi) Since $\mathrm{Cov}(W_s, W_t) = \mathrm{E}\left[(W_s - \mathrm{E}\left[W_s\right])(W_t - \mathrm{E}\left[W_t\right])\right]$, this point follows from i) and v).

□

2.4 Gaussian Property of Wiener Process

Theorem 2.4.1. Wiener process is Gaussian.

Proof. Suppose we have a Wiener process $\{W(t)|t \geqslant 0\}$. Let $n \in \mathbb{N}$, $\{t_i\}_1^n \subset \mathbb{R}_0^+$ and $t_1 < \ldots < t_n$. Then we have

$$W(t_k) = \sum_{j=1}^{k}(W(t_j) - W(t_{j-1})) \qquad \text{where } t_0 = 0.$$

$(W(t_j) - W(t_{j-1}))$ are increments of Brownian motion, hence they are independent and normally distributed. Therefore a random vector $(W(t_1), \ldots, W(t_n))$ has a multinormal distribution. ☐

Remark 2.4.1. Seeing that Wiener process is Gaussian with zero mean and covariance function $\rho(s,t) = \mathrm{E}[(W(s) - \mathrm{E}[W(s)])(W(t) - \mathrm{E}[W(t)])] = \mathrm{E}[W(s)W(t)] = = \mathrm{E}[W^2(\min\{s,t\})] = \sigma^2\min\{s,t\}$ and knowing that one-dimensional Gaussian process is completely determined by its mean and covariance we have another alternative definition of Wiener process. Any Gaussian process with mean $\mathrm{E}[W_t] = 0$ and covariance $\rho(s,t) = \sigma^2\min\{s,t\}$ for all $s,t \geqslant 0$ and with almost surely continuous paths is a Wiener process.

2.5 Markov Property of Wiener Process

Theorem 2.5.1. Brownian motion is a Markov process.

Proof. Let $\{\mathcal{F}_t|t \geqslant 0\}$ be a filtration for Brownian motion $\{W_t|t \geqslant 0\}$, let $0 \leqslant s \leqslant t$ and let f be a bounded Borel-measurable function.

$$\mathrm{E}[f(W_t)|\mathcal{F}_s] = \mathrm{E}[f(W_t - W_s + W_s)|\mathcal{F}_s].$$

Define $g(x,y) := f(x+y)$, thus we can rewrite

$$\mathrm{E}[f(W_t)|\mathcal{F}_s] = \mathrm{E}[g(W_t - W_s, W_s)|\mathcal{F}_s].$$

Since f is bounded by assumption, by [4], p. 130, ex. 7.6 we have that

$$g(W_t - W_s, W_s) = \lim \sum \varphi_j(W_t - W_s)\psi_j(W_s),$$

hence by linearity and continuity of expectation

$$\mathrm{E}[f(W_t)|\mathcal{F}_s] = \lim \sum \mathrm{E}\left[\varphi_j(W_t - W_s)\psi_j(W_s)|\mathcal{F}_s\right]$$

by adaptedness of W

$$= \lim \sum \psi_j(W_s) \mathrm{E}\left[\varphi_j(W_t - W_s)|\mathcal{F}_s\right]$$

by $\mathcal{F}_s$-independence of $W_t - W_s$

$$= \lim \sum \psi_j(W_s) \mathrm{E}\left[\varphi_j(W_t - W_s)\right]$$

by law of $W_t - W_s$

$$\begin{aligned} &= \lim \sum \psi_j(W_s) \frac{1}{\sqrt{2\pi(t-s)}} \int \varphi(x) \cdot e^{\frac{-x^2}{2(t-s)}} \mathrm{d}x \\ &= \frac{1}{\sqrt{2\pi(t-s)}} \int g(x, W_s) \cdot e^{\frac{-x^2}{2(t-s)}} \mathrm{d}x =: h(t, W_s). \end{aligned}$$

It is easily seen that h is Borel measurable. □

2.6 Martingale Property of Wiener Process

Theorem 2.6.1. Wiener process is a martingale.

Proof. Let $\{\mathcal{F}(t)|t \geqslant 0\}$ be a filtration, let $s, t \in [0, T]$ be arbitrary, but fixed numbers such that $s \leqslant t$. By linearity of (conditional) expectation and independence of increments we have

$$\begin{aligned} \mathrm{E}[W(t)|\mathcal{F}(s)] &= \mathrm{E}[W(t) + W(s) - W(s)|\mathcal{F}(s)] \\ &= \mathrm{E}[W(t) - W(s)|\mathcal{F}(s)] + \mathrm{E}[W(s)|\mathcal{F}(s)] \\ &\overset{P\text{-a.s.}}{=\!=} \underbrace{\mathrm{E}[W(t) - W(s)]}_{0} + W(s) = W(s). \end{aligned}$$

□

2.7 Equivalence Tranformations

Theorem 2.7.1. Equivalence Transformations
Assume $W \equiv \{W_t, \mathcal{F}_t\}$ is a standard Brownian motion. Processes S, I, R, $-W$ obtained by following transformations are standard Brownian motions.

i) Scaling
$S \equiv \{S_t, \mathcal{F}_{ct}|t \in [0, +\infty)\}$ where $S_t := \frac{W_{ct}}{\sqrt{c}}$ for any $c > 0$.

ii) Time inversion
$I \equiv \{I_t, \mathcal{F}_t^I | t \in [0, +\infty)\}$ where

$$I_t := \begin{cases} tW_{\frac{1}{t}} & 0 < t < +\infty \\ 0 & t = 0. \end{cases}$$

iii) Time reversal
For arbitrary fixed $T > 0$ we define $R_t := W_T - W_{T-t}$ on $[0, T]$ and $R \equiv \{R_t, \mathcal{F}_t^R | t \in [0, T]\}$.

iv) Mirror reflection
$-W \equiv \{-W_t, \mathcal{F}_t | t \in [0, +\infty)\}$.

Proof. All of the processes are evidently Gaussian, since the defining property must hold for every time t and we are transforming the process by the means of scaling, inverting, reverting or reflecting t. Hence the P-a.s. continuity, zero mean, and proper covariance property remains to be shown in order to obtain the result.

i) Scaling continuous function does not interfere with continuity, $\mathrm{E}[S_t] = 0$ evidently, and
$\mathrm{Cov}(S_t, S_s) = \mathrm{Cov}(\frac{1}{\sqrt{c}} W_{ct}, \frac{1}{\sqrt{c}} W_{cs}) = \frac{1}{c}\mathrm{Cov}(W_{ct}, W_{st}) = \frac{\min\{cs,ct\}}{c} = \min\{s, t\}$.

ii) Continuity on $(0, +\infty)$ is again evident, but in this we need to show that the process is continuous at the origin. But this is also evident using the strong law of large numbers (theorem 3.1.1):

$$\lim_{t\to 0} I_t = \lim_{t\to 0} tW_{\frac{1}{t}} = \lim_{s\to +\infty} \frac{W_s}{s} = 0.$$

The expectation is again null trivially and thus for covariance we have

$$\begin{aligned} \mathrm{Cov}(I_s, I_t) &= \mathrm{E}[I_s I_t] = \mathrm{E}\left[sW_{\frac{1}{s}} tW_{\frac{1}{t}}\right] = st\mathrm{E}\left[W_{\frac{1}{s}} W_{\frac{1}{t}}\right] = st \cdot \min\left\{\frac{1}{s}, \frac{1}{t}\right\} \\ &= \min\{s, t\}. \end{aligned}$$

iii)

$$\mathrm{E}[R_t] = \mathrm{E}[W_T - W_{T-t}] = 0,$$

$$\begin{aligned} \mathrm{Cov}(R_t, R_s) &= \mathrm{E}[R_s R_t] = \mathrm{E}[(W_T - W_{T-t})(W_T - W_{T-s})] \\ &= \mathrm{E}\left[W_T^2 - W_T W_{T-t} - W_{T-s} W_T + W_{T-s} W_{T-t}\right] \\ &= T - (T - t) - (T - s) + \min\{T - s, T - t\} = \min\{s, t\}. \end{aligned}$$

iv) This is evident.

□

REMARK 2.7.1. Unlike ii) and iii), i) and iv) are standard Brownian motions relative to the transformations of the original filtration of W, but it is evident that they are adapted to the corresponding transformations.

3 Sample Paths

Properties of the sample paths describe the behaviour of the process itself to a high degree. Thus it is important to examine what they are. Since the ultimate aim is to formulate and solve stochastic differential equations, it is useful to know that the paths of Wiener process are nondifferentiable and moreover that their quadratic variation is nonzero. The proofs of these properties are given in this chapter. As aforementioned in the preamble of chapter one, the main reference is [1].
$W = \{W_t, \mathcal{F}_t | 0 \leqslant t < +\infty\}$ will denote standard Wiener process on a probability space $(\Omega, \mathcal{F}, P)$ throughout this chapter.

3.1 Strong Law of Large Numbers

Theorem 3.1.1. Strong Law of Large Numbers

$$\lim_{t \to +\infty} \frac{W_t}{t} = 0 \qquad P\text{-a.s.}$$

Proof. By Doob's maximal inequality ([1], p. 4, iv) of theorem 1.3.8) we have that

$$\mathrm{E}\left[\sup_{t \in [\sigma,\tau]} \left(\frac{W_t}{t}\right)^2\right] \leqslant \left(\frac{2}{2-1}\right)^2 \cdot \mathrm{E}\left[\left(\frac{W_\tau}{\tau}\right)^2\right] = \frac{4}{\tau^2}\mathrm{E}\left[W_\tau^2\right] = \frac{4}{\tau} \,.$$

This combined with Čebyšev's inequality gives

$$P\left(\sup_{t \in [\sigma,\tau]} \frac{|W_t|}{t} \geqslant \varepsilon\right) \leqslant \frac{\mathrm{E}\left[\sup_{[\sigma,\tau]} \left(\frac{W_t}{t}\right)^2\right]}{\varepsilon^2} \leqslant \frac{4}{\varepsilon^2 \tau}.$$

Now by setting $\sigma := 2^n$ and $\tau := 2^{n+1}$ we obtain

$$p_n \equiv P\left(\sup_{2^n \leqslant t \leqslant 2^{n+1}} \frac{|W_t|}{t} \geqslant \varepsilon\right) \leqslant \frac{2}{\varepsilon^2} 2^{-n} \qquad \text{all } \varepsilon > 0, \text{ all } n \in \mathbb{N},$$

thus

$$\sum_{n=1}^{\infty} p_n \leqslant \frac{2}{\varepsilon^2} \sum_{n=1}^{\infty} \frac{1}{2^n} < +\infty.$$

It now follows from Borel-Cantelli lemma that

$$P\left(\left\{\sup_{2^n \leqslant t \leqslant 2^{n+1}} \frac{|W_t|}{t} \geqslant \varepsilon\right\} \quad \text{i. o.}\right) = 0,$$

which readily shows the result. □

3.2 Quadratic Variation

Definition 3.2.1. Partition of Interval
Let $[a, b] \subset \mathbb{R}$, then the partition of $[a, b]$ is a sequence of numbers $\sigma \equiv \{t_i\}_0^n \subset [a, b]$ such that $t_0 = a,\ t_n = b$ and $t_0 < t_1 < \ldots < t_n$. We define the norm of the partition σ to be

$$\|\sigma\| := \max\{t_i - t_{i-1} | t_i, t_{i-1} \in \sigma\}.$$

The sequence of partitions $\{\sigma_n\}_1^\infty$ is called normal if $\|\sigma_n\| \xrightarrow{n\to\infty} 0$.

Definition 3.2.2. Quadratic Variation
Assume a stochastic process $X = \{X_t | t \in T\}$ on $(\Omega, \mathcal{F}, P)$ with $[0, t] \subset T$ and a partition $\sigma = (t_0, \ldots, t_m)$ of $[0, t]$. Then $V_\sigma^{(2)}(X)$ given by

$$V_\sigma^{(2)}(X) = \sum_{k=1}^{m} \left|X_{t_k} - X_{t_{k-1}}\right|^2,$$

is called quadratic variation of X over σ. Further assume $\{\sigma_n\}_1^\infty$ is a normal sequence of partitions. Then the P-limit of $V_{\sigma_n}^{(2)}(X)$ is called quadratic variation of X up to t and it is denoted by $\langle X \rangle_t$, provided it exists.

Theorem 3.2.1. Quadratic Variation of Brownian Motion
Let $\{\sigma_n\}_1^\infty$ be a sequence of partitions of $[0, t]$ and let $V_t^{(2)}(\sigma_n)$ denote the quadratic variation of W over σ_n, then

i) $\lim\limits_{n\to\infty} \|\sigma_n\| = 0 \implies V_t^{(2)}(\sigma_n) \xrightarrow[n\to\infty]{L^2} t$, and

ii) $\sum\limits_{n=1}^{\infty} \|\sigma_n\| < +\infty \implies V_t^{(2)}(\sigma_n) \xrightarrow[n\to\infty]{\text{a. s.}} t$.

Specifically $\langle W \rangle_t = t$.

Proof. Assume $\sigma = (t_0, \ldots, t_m)$ partition of $[0, t]$ and assume $\sigma_n = t_0^{(n)}, \ldots, t_{m_n}^{(n)}$.

i) We are interested in expectation of squared term $V_t^{(2)}(\sigma_n) - t$ as n tends to infinity. Let us first rewrite such term:

$$V_t^{(2)}(\sigma) - t = \sum_{k=1}^{m} \left((W_{t_k} - W_{t_{k-1}})^2 - (t_k - t_{k-1})\right),$$

$\left\{W_{t_k} - W_{t_{k-1}}\right\}_1^m$ are independent by definition of Brownian motion, thus the "cross" terms in following sum vanish:

$$\begin{aligned}
\mathrm{E}\left[\left(V_t^{(2)}(\sigma)-t\right)^2\right] &= \mathrm{E}\left[\left(\sum_{k=1}^{m}\left((W_{t_k}-W_{t_{k-1}})^2-(t_k-t_{k-1})\right)\right)^2\right] \\
&= \sum_{k=1}^{m}\mathrm{E}\left[\left((W_{t_k}-W_{t_{k-1}})^2-(t_k-t_{k-1})\right)^2\right] \\
&= \sum_{k=1}^{m}(t_k-t_{k-1})^2\mathrm{E}\left[\left(\left(\frac{W_{t_k}-W_{t_{k-1}}}{\sqrt{t_k-t_{k-1}}}\right)^2-1\right)^2\right]
\end{aligned}$$

(By properties of Wiener process we have $N\equiv\frac{W_{t_k}-W_{t_{k-1}}}{\sqrt{t_k-t_{k-1}}}\sim N(0,1)$.)

$$\begin{aligned}
&= \sum_{k=1}^{m}(t_k-t_{k-1})\underbrace{(t_k-t_{k-1})}_{\leqslant\|\sigma\|}\mathrm{E}\left[\left(N^2-1\right)^2\right] \\
&\leqslant \|\sigma\|\sum_{k=1}^{m}(t_k-t_{k-1})\mathrm{E}\left[\left(N^2-1\right)^2\right] = t\|\sigma\|\mathrm{E}\left[(N^2-1)^2\right].
\end{aligned}$$

Hence

$$\mathrm{E}\left[\left(V_t^{(2)}(\sigma_n)-t\right)^2\right]\xrightarrow{n\to\infty}0,$$

since $\|\sigma\|\xrightarrow{n\to\infty}0$, which yields the result.

ii) By Čebyšev's inequality we have that

$$\begin{aligned}
p_n\equiv P\left(\left|V_t^{(2)}(\sigma_n)-t\right|>\varepsilon\right) &\leqslant \frac{\mathrm{E}\left[\left(V_t^{(2)}(\sigma_n)-t\right)^2\right]}{\varepsilon^2} \\
&\leqslant \frac{t\mathrm{E}\left[\left(N^2-1\right)^2\right]\|\sigma_n\|}{\varepsilon^2}\equiv\frac{c\|\sigma_n\|}{\varepsilon^2},
\end{aligned}$$

and by assumption

$$\sum_{n=1}^{\infty}p_n\leqslant\sum_{n=1}^{\infty}\frac{c\|\sigma_n\|}{\varepsilon^2}<+\infty.$$

It now follows from Borel-Cantelli lemma that

$$P\left(\left|V_t^{(2)}(\sigma_n)-t\right|>\varepsilon\ \text{ i. o.}\right)=0\quad\text{for arbitrary, fixed }\varepsilon>0.$$

□

3.3 Nondifferentiability of Brownian motion

Definition 3.3.1. Dini Derivatives
Assume a continuous function $f:\mathbb{R}_0^+\to\mathbb{R}$, then we define the right upper $D^+f(t)$

and right lower $D_+f(t)$ Dini derivative of f at t by

$$D^+f(t) := \limsup_{h\downarrow 0} \frac{f(t+h)-f(t)}{h} \quad \text{and} \quad D_+f(t) := \liminf_{h\downarrow 0} \frac{f(t+h)-f(t)}{h}.$$

f is differentiable at t from the right, if the upper $D^+f(t)$ and lower $D_+f(t)$ Dini derivative of f at t agree. The left Dini derivatives are defined analogously.

Theorem 3.3.1. Nondifferentiability of Brownian motion
There exists an event $F \in \mathcal{F}$, such that $P(F) = 1$ and

$$F \subset \left\{\omega \in \Omega \,\middle|\, \forall t \in [0,+\infty)\ D^+W_t(\omega) = +\infty \text{ or } D_+W_t(\omega) = -\infty\right\}.$$

Proof. Let us first show that the claim holds if we replace $\forall t \in [0,+\infty)$ by $\forall t \in [0,1]$. Since $\mathbb{R}_0^+$ is a countable union of such intervals the general result will follow.
For $j,k \in \mathbb{N}$ arbitrary, fixed, we define

$$A_{jk} := \bigcup_{t\in[0,1]} \bigcap_{h\in\left[0,\frac{1}{k}\right]} \left\{\omega \,\middle|\, |W_{t+h}(\omega) - W_t(\omega)| \leqslant jh\right\}.$$

Thus

$$C_{[0,1]} := \left\{\omega \,\middle|\, \exists t \in [0,1]\ -\infty < D_+W_t(\omega) \leqslant D^+W_t(\omega) < +\infty\right\} = \bigcup_{j=1}^{\infty}\bigcup_{k=1}^{\infty} A_{jk}.$$

For fixed j,k, assume arbitrary $\omega \in A_{jk}$ fixed, that is $\exists t \in [0,1]$, such that $\forall h \in [0,\frac{1}{k}]$ we have $|W_{t+h}(\omega) - W_t(\omega)| \leqslant jh$. Thus for $n \in \mathbb{N},\ n \geqslant 4k$ we have:

1. $\exists i \in \mathbb{N},\ i \leqslant n$ such that $\frac{i-1}{n} \leqslant t \leqslant \frac{i}{n}$,
2. for $\iota \in \{0,1,2,3\}$ we have $\frac{i+\iota}{n} - t \leqslant \frac{\iota+1}{n} \leqslant \frac{1}{k}$, this follows from 1.
3. From 2. and assumption that $\omega \in A_{jk}$ it follows that

$$\begin{aligned}
\left|W_{\frac{i}{n}}(\omega) - W_t(\omega)\right| &\leqslant j\left(\frac{i}{n} - t\right) \leqslant \frac{j}{n},\\
\left|W_{\frac{i+1}{n}}(\omega) - W_t(\omega)\right| &\leqslant j\left(\frac{i+1}{n} - t\right) \leqslant \frac{2j}{n},\\
\left|W_{\frac{i+2}{n}}(\omega) - W_t(\omega)\right| &\leqslant j\left(\frac{i+2}{n} - t\right) \leqslant \frac{3j}{n},\\
\left|W_{\frac{i+3}{n}}(\omega) - W_t(\omega)\right| &\leqslant j\left(\frac{i+3}{n} - t\right) \leqslant \frac{4j}{n}.
\end{aligned}$$

4. Therefore by triangle inequality

$$\begin{aligned}
\left|W_{\frac{i+1}{n}}(\omega) - W_{\frac{i}{n}}(\omega)\right| &\leqslant \frac{3j}{n},\\
\left|W_{\frac{i+2}{n}}(\omega) - W_{\frac{i+1}{n}}(\omega)\right| &\leqslant \frac{5j}{n},\\
\left|W_{\frac{i+3}{n}}(\omega) - W_{\frac{i+2}{n}}(\omega)\right| &\leqslant \frac{7j}{n}.
\end{aligned}$$

Let us define $F_{jk}^{(n,i)} := \bigcap_{\iota=1}^{3} \left\{\omega \,\middle|\, \left|W_{\frac{i+\iota}{n}}(\omega) - W_{\frac{i+\iota-1}{n}}(\omega)\right| \leqslant j\frac{2\iota+1}{n}\right\}$.
Since $N \equiv \sqrt{n}\left(W_{\frac{i+\iota}{n}} - W_{\frac{i+\iota-1}{n}}\right) \sim N(0,1)$ we rewrite $F_{jk}^{(n,i)}$ as

$$F_{jk}^{(n,i)} = \bigcap_{\iota=1}^{3} \left\{\omega \,\middle|\, \sqrt{n}\left|W_{\frac{i+\iota}{n}}(\omega) - W_{\frac{i+\iota-1}{n}}(\omega)\right| \leqslant j\frac{2\iota+1}{\sqrt{n}}\right\} = \bigcap_{\iota=1}^{3} \left\{\omega \,\middle|\, |N(\omega)| \leqslant j\frac{2\iota+1}{\sqrt{n}}\right\}$$

and use the fact that for standard normal random variable N, $P\left(|N| \leqslant \varepsilon\right) \leqslant \varepsilon$. It follows by the independence of the increments above that

$$P\left(F_{jk}^{(n,i)}\right) \leqslant \frac{3j}{\sqrt{n}} \cdot \frac{5j}{\sqrt{n}} \cdot \frac{7j}{\sqrt{n}} = \frac{105j^3}{\sqrt{n}^3}.$$

To obtain an event that contains A_{jk} and has zero probability we finally define

$$F_{jk} := \bigcap_{n=4k}^{\infty} \bigcup_{i=1}^{\infty} F_{jk}^{(n,i)} \qquad \Longrightarrow \qquad P(F_{jk}) \leqslant \inf_{n \geqslant 4k} P\left(\bigcup_{i=1}^{\infty} F_{jk}^{(n,i)}\right) = 0.$$

Recall that $C_{[0,1]} = \bigcup_{j=1}^{\infty} \bigcup_{k=1}^{\infty} A_{jk}$, hence the event

$$F_{[0,1]} := \bigcup_{j=1}^{\infty} \bigcup_{k=1}^{\infty} F_{jk},$$

is the event that contains $C_{[0,1]}$ and has probability zero.
We could find $F_{[\ell,\ell+1]}$ to every $C_{[\ell,\ell+1]}$ analogously, the sought F would then be $\left(\bigcup_{\ell=0}^{\infty} F_{[\ell,\ell+1]}\right)^C$. The proof is now complete.

$\square$

4 Itô Integral

From the lack of differentiability of Wiener process stems the need for special approach to defining the integral with respect to it. Such approach is illustrated - the Itô integral is constructed in the first part of this chapter. Some properties of the integral succeed the construction. The main reference for this chapter is [2], the less rigorous, yet somewhat more intuitive source is [4].

4.0 Integration with Respect to Brownian Motion

It is interesting and also, for various reasons, useful to integrate other processes with respect to Brownian motion. Consider for example the following. Suppose we are interested in trading certain commodity and suppose we would like to model the price of this commodity with the Brownian motion B. Assume that in some period, say a month, we buy or sell certain number of contracts, let t_j denote j-th day and let f_{t_j} denote the number of contracts bought the j-th day. We might be interested in the gain from trading this commodity after a month. The gain is easily obtained as

$$\sum_{j=1}^{\text{last day}} f_{t_j} \left(B_{t_{j+1}} - B_{t_j} \right).$$

This is a simple example of integral of f with respect to B and it is easily seen why such object should be of interest.

f in the example above was somewhat special though. It took on one value at t_j and held it up to t_{j+1}. That is every day morning we bought some contracts and focused on something completely different until the next morning. The aim of this chapter is to extend this concept to more general f. $B = \{B_t, \mathcal{F}_t\}$ will henceforth be a Brownian motion on $(\Omega, \mathcal{F}, P)$. Further assume that f is a process defined on the same probability space and adapted to $\{\mathcal{F}_t\}$. We would like to assign some meaning to what looks like Lebesgue-Stieltjes integral

$$\int_a^b f_t(\omega) \mathrm{d}B_t(\omega).$$

It is evident that moving straightforward with defining such object through the Lebesgue integral as

$$\int_a^b f_t(\omega) \frac{\mathrm{d}B_t}{\mathrm{d}t}(\omega) \mathrm{d}t$$

will simply not work since $\frac{\mathrm{d}B_t}{\mathrm{d}t}$ has almost surely no meaning anywhere. Hence because of nondifferentiability of Brownian motion we must take different approach. In the example above we have seen that if the integrand has certain properties $\int_a^b f_t(\omega)\mathrm{d}B_t(\omega)$ may be written as a sum. Assume we would like to approximate $\int_0^T B_t(\omega)\mathrm{d}B_t(\omega)$ somehow. Define

$$\begin{aligned}\varphi_0^{(n)}(t,\omega) &= \sum_{j\geqslant 0} B_{\frac{j}{2^n}}(\omega)\mathbb{1}_{\left[\frac{j}{2^n},\frac{j+1}{2^n}\right)}(t)\\ \varphi_1^{(n)}(t,\omega) &= \sum_{j\geqslant 0} B_{\frac{j+1}{2^n}}(\omega)\mathbb{1}_{\left[\frac{j}{2^n},\frac{j+1}{2^n}\right)}(t),\end{aligned}$$

and let us examine $\int_0^T \varphi_\iota^{(n)}(t,\omega)\mathrm{d}B_t(\omega)$ as $n\to\infty$ in a sense indicated by the example, that is

$$\int_0^T \varphi_\iota^{(n)}(t,\omega)\mathrm{d}B_t(\omega) = \sum_{j\geqslant 0} B_{\frac{j+\iota}{2^n}}(B_{\frac{j+1}{2^n}} - B_{\frac{j}{2^n}}) \qquad \iota = 0,1.$$

For the expectation of the first integral we have

$$\begin{aligned}\mathrm{E}\left[\int_0^T \varphi_0^{(n)}(t,\omega)\mathrm{d}B_t(\omega)\right] &= \sum \mathrm{E}\left[B_{\frac{j}{2^n}}(B_{\frac{j+1}{2^n}} - B_{\frac{j}{2^n}})\right]\\ &= \sum \mathrm{E}\left[(B_{\frac{j}{2^n}} - B_0)(B_{\frac{j+1}{2^n}} - B_{\frac{j}{2^n}})\right] = 0\end{aligned}$$

while for the second

$$\begin{aligned}\mathrm{E}\left[\int_0^T \varphi_1^{(n)}(t,\omega)\mathrm{d}B_t(\omega)\right] &= \sum \mathrm{E}\left[B_{\frac{j+1}{2^n}}(B_{\frac{j+1}{2^n}} - B_{\frac{j}{2^n}})\right]\\ &= \sum \mathrm{E}\left[(B_{\frac{j+1}{2^n}} - B_{\frac{j}{2^n}} + B_{\frac{j}{2^n}})(B_{\frac{j+1}{2^n}} - B_{\frac{j}{2^n}})\right]\\ &= \sum \left(\mathrm{E}\left[(B_{\frac{j+1}{2^n}} - B_{\frac{j}{2^n}})^2\right] + \mathrm{E}\left[B_{\frac{j}{2^n}}(B_{\frac{j+1}{2^n}} - B_{\frac{j}{2^n}})\right]\right)\\ &= \sum \left(\frac{j+1}{2^n} - \frac{j}{2^n}\right) = T.\end{aligned}$$

Thus the expectations of these integrals do not agree neither do they depend on n and they will not be any closer as n will grow. It is as if the Riemann upper and lower integral would not agree. We can see that the difference in this case is caused by the fact that the Brownian quadratic variation is not null. This indeed is a reason why ordinary calculus fails here.

We can overcome this difficulty by simply choosing which point in the intervals we will use to obtain integral that we want, or more precisely by choosing different points to obtain different integrals. The first choice in the example above leads to the Itô integral which will be constructed in the following section.

4.1 Itô Integral

As we have seen it is relatively easy to define the integral for 'simple' process. The idea is to use these 'simple' integrands to define the integral for some larger class of processes. The class of processes for which the Itô integral will be defined is given in the following definition.

Definition 4.1.1. $L^2_{ad}([a,b]\times\Omega)$

$$L^2_{ad}([a,b]\times\Omega) := \left\{ f : [a,b]\times\Omega \to \mathbb{R} \;\middle|\; \begin{array}{ll} \text{i)} & f \text{ is } \mathcal{F}_t\text{-adapted stochastic process.} \\ \text{ii)} & f \text{ is } \mathcal{B}(\mathbb{R}_0^+)\otimes\mathcal{F}\text{-measurable.} \\ \text{iii)} & \int_a^b \mathrm{E}\left[f^2(t,\omega)\right]\mathrm{d}t < +\infty \end{array} \right\}.$$

Some rather technical assumptions are usually omitted within the construction of Itô integral. These assumptions are, though somewhat partially, illustrated now. In the following we assume that

1. $\mathcal{F}$ is P-complete, that is $\mathcal{F}$ contains every P-negligible set,
2. $\{\mathcal{F}_t\}$ satisfies the usual conditions (definition 1.1.6),
3. During the construction of the integral we use progressively measurable modification of f, while we denote it simply by f. This will be explained further.

Definition 4.1.2. Progressively Measurable Stochastic Process
Let X be a stochastic process on a stochastic basis $(\Omega,\mathcal{F},\mathcal{F}_t,P)$. We say that X is progressively measurable with respect to $\{\mathcal{F}_t\}$ if for every t the mapping $(s,\omega)\mapsto X_s(\omega) : [0,t]\times\Omega\to\mathbb{R}$, is $\mathcal{B}([0,t])\otimes\mathcal{F}_t$-measurable.

Theorem 4.1.1. Existence of Progressively Measurable Modification
Assume $X \equiv \{X_t | t \geqslant 0\}$ on $(\Omega,\mathcal{F},\mathcal{F}_t,P)$ is $\mathcal{F}_t$-adapted and $\mathcal{B}(\mathbb{R}_0^+)\otimes\mathcal{F}$-measurable. Then it has a progressively measurable modification (with respect to $\{\mathcal{F}_t\}$).

Proof. See [5]. □

By this theorem for every $f \in L^2_{ad}([a,b]\times\Omega)$ there exists progressively measurable modification. By definition of modification and definition of $L^2_{ad}([a,b]\times\Omega)$, this modification is also contained in $L^2_{ad}([a,b]\times\Omega)$. In fact if we factorize $L^2_{ad}([a,b]\times\Omega)$ as is customary for L^p spaces, both these processes will be contained in the same L^2 class. Since the Itô integral is constructed as L^2-limit, integrals of arbitrary functions from the same L^2 class will agree. Though, existence of progressively measurable modification is important as it will allow us to apply Fubini's theorem.

As has already been mentioned, we are unable to define the integral for the whole $L^2_{ad}([a,b] \times \Omega)$ directly. Instead we must resort to a certain subclass - the class of step processes.

Definition 4.1.3. Step Stochastic Process
Let $f \in L^2_{ad}([a,b] \times \Omega)$ and let $\sigma = \{t_i\}_0^n$ be a partition of $[a,b]$ such that

$$f(t,\omega) = \sum_{i=1}^{n} c_{i-1}(\omega) \mathbb{1}_{[t_{i-1},t_i)}(t),$$

where c_i is $\mathcal{F}_{t_i}$-measurable random variable with $\mathrm{E}\left[c_i^2\right] < +\infty$. Then we say that f is a step stochastic process. The class of all step processes on $L^2_{ad}([a,b] \times \Omega)$ will be denoted by $\mathcal{I}\left(L^2_{ad}([a,b] \times \Omega)\right)$.

Definition 4.1.4. Itô Integral of Step Process
Assume $f \in \mathcal{I}\left(L^2_{ad}([a,b] \times \Omega)\right)$ as in preceding definition, then the Itô integral of f over $[a,b]$ is a random variable on $(\Omega, \mathcal{F}, P)$ given by

$$I(f) \equiv \left(\int_a^b f(t) \mathrm{d}B_t\right)(\omega) := \sum_{j=1}^{n} c_{j-1}(\omega) \left(B_{t_j} - B_{t_{j-1}}\right)(\omega).$$

Remark 4.1.1. *I is linear mapping on $\mathcal{I}\left(L^2_{ad}([a,b] \times \Omega)\right)$. To see this let f be step with respect to σ_f, let g be step with respect to σ_g. Then both f and g are step with respect to the common refinement $\sigma = \{t_i\}_0^n$ of σ_f and σ_g, thus $I(\alpha f + g)$ is merely a finite sum of terms $\alpha c_j + k_j$. All the terms of the finite sum commute.*

Theorem 4.1.2. Itô Isometry for Step Processes
For every step process we have

$$\mathrm{E}\left[\left(\int_a^b f(t) \mathrm{d}B_t\right)^2 (\omega)\right] = \mathrm{E}\left[\int_a^b f^2(t,\omega) \mathrm{d}t\right].$$

Proof. Let

$$f(t,\omega) = \sum_{i=1}^{n} c_{i-1}(\omega) \mathbb{1}_{[t_{i-1},t_i)}(t).$$

Assume without loss of generality that $j \geqslant i$. Then by properties of conditional expectation (specifically partial averaging property), by $\mathcal{F}_{t_k}$-measurability of c_k, and by independence and variance of Brownian increments we obtain

$$\begin{aligned}
\mathrm{E}\Big[c_i c_j \underbrace{\left(B_{t_{i+1}} - B_{t_i}\right)}_{\equiv \Delta_i} \underbrace{\left(B_{t_{j+1}} - B_{t_j}\right)}_{\equiv \Delta_j}\Big] &= \mathrm{E}\left[\mathrm{E}\left[c_i c_j \Delta_i \Delta_j \left| \mathcal{F}_{t_j}\right.\right]\right] \\
&= \mathrm{E}\left[c_i c_j \mathrm{E}\left[\Delta_i \Delta_j \left| \mathcal{F}_{t_j}\right.\right]\right] \\
&= \begin{cases} \mathrm{E}\left[c_i c_j \Delta_i \mathrm{E}\left[\Delta_j \left| \mathcal{F}_{t_j}\right.\right]\right] & i < j \\ \mathrm{E}\left[c\right]_j^2 \mathrm{E}\left[\Delta_j^2 \left| \mathcal{F}_{t_j}\right.\right] & i = j \end{cases} \\
&= \mathrm{E}\left[c_j^2\right] \left(t_{j+1} - t_j\right) \delta_{ij}.
\end{aligned}$$

Thus

$$\mathrm{E}\left[\left(\int_a^b f\mathrm{d}B\right)^2\right] = \sum_{i,j}\mathrm{E}\left[c_ic_j\Delta_i\Delta_j\right] = \sum_j \mathrm{E}\left[c_j^2\Delta_j^2\right]$$
$$= \sum_j \mathrm{E}\left[c_j^2\right](t_{j+1}-t_j) = \mathrm{E}\left[\int_a^b f^2\mathrm{d}t\right].$$

□

REMARK 4.1.2. It readily follows from the proof that $\mathrm{E}\left[I(f)\right] = 0$.

We will now show that step functions from $L^2_{ad}([a,b]\times\Omega)$ are in some sense dense in $L^2_{ad}([a,b]\times\Omega)$.

Lemma 4.1.1. Bounded functions in $L^2_{ad}([a,b]\times\Omega)$ are L^2-dense in $L^2_{ad}([a,b]\times\Omega)$, that is for every f there is a sequence $\{f_n\}$ of bounded functions such that

$$\lim_{n\to\infty}\int_a^b \mathrm{E}\left[|f(t)-f_n(t)|^2\right]\mathrm{d}t = 0.$$

Proof. Let f be arbitrary function from $L^2_{ad}([a,b]\times\Omega)$. Define

$$f_n(t,\omega) := \begin{cases} f(t,\omega) & |f(t,\omega)| \leqslant n \\ 0 & |f(t,\omega)| > n. \end{cases}$$

By Lebesgue dominated convergence theorem we obtain

$$\int_a^b \mathrm{E}\left[|f(t)-f_n(t)|^2\right]\mathrm{d}t \xrightarrow{n\to\infty} 0.$$

□

Theorem 4.1.3. Functions of $\mathcal{I}\left(L^2_{ad}([a,b]\times\Omega)\right)$ are L^2-dense in bounded functions of $L^2_{ad}([a,b]\times\Omega)$.

Proof. Let f be arbitrary bounded function from $L^2_{ad}([a,b]\times\Omega)$.

i) Let us first assume $\mathrm{E}\left[f(s)f(t)\right]$ is a continuous function of (s,t).
Let $\{\sigma_n\}$ be a normal sequence of partitions of $[a,b]$, such that $\sigma_n = \{t_i^{(n)}\}_0^n$. For every $n\in\mathbb{N}$ we define

$$f_n(t,\omega) := \sum f(t_{i-1}^{(n)},\omega)\mathbb{1}_{\left(t_{i-1}^{(n)},t_i^{(n)}\right]}(t).$$

It is evident that $\{f_n\}_1^\infty \subset \mathcal{I}\left(L^2_{ad}([a,b]\times\Omega)\right)$. Moreover by continuity of $\mathrm{E}\left[f(s)f(t)\right]$ we have

$$\lim_{n\to\infty}\mathrm{E}\left[|f(t)-f_n(t)|^2\right] = 0.$$

Furthermore

$$\mathrm{E}\left[|f(t)-f_n(t)|^2\right] \leqslant 2\left(\mathrm{E}\left[|f(t)|^2\right]+\mathrm{E}\left[|f_n(t)|^2\right]\right) \leqslant 4\sup_{[a,b]}\mathrm{E}\left[|f(t)|^2\right],$$

hence by Lebesgue dominated convergence

$$\lim_{n\to\infty}\int_a^b \mathrm{E}\left[|f(t)-f_n(t)|^2\right]\mathrm{d}t = 0.$$

ii) We now set

$$g_n(t,\omega) := \int_0^{n(t-a)} e^{-\tau} f\left(t-\frac{\tau}{n},\omega\right)\mathrm{d}\tau,$$

and we show that $\mathrm{E}\left[g_n(t)g_n(s)\right]$ is continuous in (t,s) (and we can thus apply i) to g_n for every n), and that L^2-$\lim g_n = f$. We first rewrite g_n:

$$\begin{aligned} g_n(t,\omega) &= \int_0^{n(t-a)} e^{-\tau} f\left(t-\frac{\tau}{n},\omega\right)\mathrm{d}\tau \\ &= \left\{u := t-\frac{\tau}{n}\right\} = \int_a^t n e^{-n(t-u)} f(u,\omega)\mathrm{d}u. \end{aligned}$$

It is now clear that g_n is $\mathcal{F}_t$-adapted by Fubini's theorem. This was the reason for requiring f to be progressively measurable. Since integral as a function of the upper bound is continuous and e^{-nt} is continuous as well, we can see that

$$\lim_{t\to s}\mathrm{E}\left[g_n(t)g_n(s)\right] = \mathrm{E}\left[g_n^2(s)\right].$$

It now remains to show that $\int_a^b \mathrm{E}\left[|f(t)-g_n(t)|^2\right]\mathrm{d}t \xrightarrow{n\to\infty} 0$. We may write

$$f(t,\omega) = \int_0^{+\infty} e^{-\tau} f(t,\omega)\mathrm{d}\tau,$$

and if we set $f(t,\omega) := 0$ for every $t < a$

$$g_n(t,\omega) = \int_0^{n(t-a)} e^{-\tau} f\left(t-\frac{\tau}{n},\omega\right)\mathrm{d}\tau = \int_0^{+\infty} e^{-\tau} f\left(t-\frac{\tau}{n},\omega\right)\mathrm{d}\tau,$$

hence

$$f(t,\omega)-g_n(t,\omega) = \int_0^{+\infty} e^{-\tau}\left(f(t,\omega)-f\left(t-\frac{\tau}{n},\omega\right)\right)\mathrm{d}\tau.$$

Since $e^{-\tau}\mathrm{d}\tau$ is a probability measure on $[0,+\infty)$ we can apply Schwarz inequality to obtain

$$|f(t,\omega)-g_n(t,\omega)|^2 \leqslant \int_0^{+\infty}\left|\left(f(t,\omega)-f\left(t-\frac{\tau}{n},\omega\right)\right)\right|^2 e^{-\tau}\mathrm{d}\tau.$$

We can use this upper bound and Fubini's theorem to obtain

$$\begin{aligned}\int_a^b \mathrm{E}\left[|f(t)-g_n(t)|^2\right]\mathrm{d}t &\leqslant \int_a^b\int_0^{+\infty} e^{-\tau}\mathrm{E}\left[\left|\left(f(t,\omega)-f\left(t-\frac{\tau}{n},\omega\right)\right)\right|^2\right]\mathrm{d}\tau\mathrm{d}t\\ &= \int_0^{+\infty} e^{-\tau}\left(\int_a^b \mathrm{E}\left[\left|\left(f(t,\omega)-f\left(t-\frac{\tau}{n},\omega\right)\right)\right|^2\right]\mathrm{d}t\right)\mathrm{d}\tau\\ &= \int_0^{+\infty} e^{-\tau}\mathrm{E}\left[\int_a^b \left|\left(f(t,\omega)-f\left(t-\frac{\tau}{n},\omega\right)\right)\right|^2\mathrm{d}t\right]\mathrm{d}\tau.\end{aligned}$$

Since f is bounded by assumption we have

$$\lim_{n\to\infty}\int_a^b \left|\left(f(t,\omega)-f\left(t-\frac{\tau}{n},\omega\right)\right)\right|^2\mathrm{d}t \overset{\text{a.s.}}{=\!=} 0.$$

To each g_n we can apply i) to obtain a sequence of step processes $\{f_k^{(n)}\}_k$ from $\mathcal{I}\left(L^2_{ad}([a,b]\times\Omega)\right)$ which converges to g_n in L^2. Thus for every $n\in\mathbb{N}$ we can pick some index k_n such that

$$\int_a^b \mathrm{E}\left[|g_n(t)-f_{k_n}^{(n)}(t)|^2\right]\mathrm{d}t < \frac{1}{n}.$$

We finally set $f_n := f_{n_k}^{(n)}$ and observe that $\{f_n\}$ is the sought sequence, that is

$$\int_a^b \mathrm{E}\left[|f(t)-f_n(t)|^2\right]\mathrm{d}t \xrightarrow{n\to\infty} 0.$$

□

Corollary 4.1.1. By lemma 4.1.1 and theorem 4.1.3 $\mathcal{I}\left(L^2_{ad}([a,b]\times\Omega)\right)$ are L^2-dense in $L^2_{ad}([a,b]\times\Omega)$.

Proof. By lemma 4.1.1 for every $f\in L^2_{ad}([a,b]\times\Omega)$ there is a sequence of bounded functions $\{g_n\}\subset L^2_{ad}([a,b]\times\Omega)$ that converges to f in L^2. By theorem 4.1.3 for every g_n there is a sequence $\{f_k^{(n)}\}\subset\mathcal{I}\left(L^2_{ad}([a,b]\times\Omega)\right)$ converging to g_n in L^2. Thus we can pick the indices in the same manner as in the proof of the preceding theorem to construct a sequence $\{f_n\}$ such that

$$\int_a^b \mathrm{E}\left[|f(t)-f_n(t)|^2\right]\mathrm{d}t \xrightarrow{n\to\infty} 0.$$

□

We will utilise the fact that step processes are dense in the whole $L^2_{ad}([a,b]\times\Omega)$ to define the Itô integral for the whole class $L^2_{ad}([a,b]\times\Omega)$.

Definition 4.1.5. Itô Integral
Let $f \in L^2_{ad}([a,b] \times \Omega)$ then the Itô integral of f on $[a,b]$ is given by

$$I(f) \equiv \int_a^b f(t)\mathrm{d}B_t := L^2\text{-}\lim I(f_n),$$

where f_n is arbitrary step process converging to f in L^2.

REMARK 4.1.3.

i) The Itô integral is well defined. To observe this we use the linearity of I on step processes (remark 4.1.1) and the isometry for step processes (theorem 4.1.2) to obtain

$$\mathrm{E}\left[|I(f_n) - I(f_m)|^2\right] = \mathrm{E}\left[|I(f_n - f_m)|^2\right] = \mathrm{E}\left[\int_a^b |f_n(t) - f_m(t)|^2 \mathrm{d}t\right].$$

Since $\{f_n\}$ is Cauchy we conclude that $\{I(f_n)\}$ is $L^2(\Omega)$-Cauchy. Thus $\lim I(f_n)$ exists by completeness of $L^2(\Omega)$. Furthermore assume $\{g_n\}, \{f_n\}$ are two sequences satisfying the conditions of the preceding definition. We want to show that $L^2(\Omega)\text{-}\lim I(g_n) = L^2(\Omega)\text{-}\lim I(f_n)$. By the same argument

$$\mathrm{E}\left[|I(g_n) - I(f_n)|^2\right] = \mathrm{E}\left[|I(g_n - f_n)|^2\right] = \mathrm{E}\left[\int_a^b |g_n(t) - f_n(t)|^2 \mathrm{d}t\right] \xrightarrow{n\to\infty} 0,$$

which concludes that the limit is independent of the choice of $\{f_n\}$.

ii) Itô integral is linear, that is $I(\alpha f + g) = \alpha I(f) + I(g)$. Assume $\{f_n\}$, resp. $\{g_n\}$ is a sequence from $\mathcal{I}\left(L^2_{ad}([a,b] \times \Omega)\right)$ converging to f, resp. g in L^2. Then $\{\alpha f_n + g_n\}$ is a step process (with respect to the common refinement of some partitions σ_f, resp. σ_g such that f, resp. g is step with respect to σ_f, resp. σ_g) as well. The result easily follows from linearity of integral for step processes.

iii) For every $f \in L^2_{ad}([a,b] \times \Omega)$ and every $[c,d] \subset [a,b]$ we have

$$\int_c^d \mathrm{E}\left[|f(t)|^2\right] \mathrm{d}t \leqslant \int_a^b \mathrm{E}\left[|f(t)|^2\right] \mathrm{d}t,$$

thus $f \in L^2_{ad}([c,d] \times \Omega)$. Therefore the Itô integral of f on $[c,d]$ is defined. It is evident that the Itô integral is additive in its bounds, that is

$$\int_a^c f(t)\mathrm{d}B_t + \int_c^d f(t)\mathrm{d}B_t = \int_a^d f(t)\mathrm{d}B_t.$$

iv) $\int_a^t f(\tau)\mathrm{d}B_\tau$ is constructed as a limit of $\mathcal{F}_t$-measurable random variables. Thus it is $\mathcal{F}_t$-measurable as well.

v) The result of theorem 4.1.2 remains true for the whole $L^2_{ad}([a,b] \times \Omega)$. This will be shown in the following theorem.

Theorem 4.1.4. Isometry of Itô Integral
$I : L^2_{ad}([a,b] \times \Omega) \to L^2(\Omega)$ is an isometry.

Proof. Let $f \in L^2_{ad}([a,b] \times \Omega)$, let $f_n \xrightarrow[n\to\infty]{L^2} f$. We have

$$\|I(f)\|_{L^2(\Omega)} \equiv \mathrm{E}\left[|I(f)|^2\right] \overset{\text{definition 4.1.5}}{=\!=\!=} \lim \mathrm{E}\left[|I(f_n)|^2\right]$$
$$\overset{\text{theorem 4.1.2}}{=\!=\!=} \lim \int_a^b \mathrm{E}\left[|f_n(t)|^2\right] \mathrm{d}t = \int_a^b \mathrm{E}\left[|f(t)|^2\right] \mathrm{d}t \equiv \|f\|_{L^2_{ad}([a,b]\times\Omega)},$$

which shows the result. □

REMARK 4.1.4. This only is a corollary of the fact that I preserves the inner product, that is $\forall f, g \in L^2_{ad}([a,b] \times \Omega)$

$$\mathrm{E}\left[I(f)I(g)\right] = \int_a^b \mathrm{E}\left[f(t)g(t)\right] \mathrm{d}t.$$

This is easily concluded using the same argument as in the proof of theorem 4.1.2 and definition 4.1.5.

4.2 Martingale Property if Itô Integral

Assume $f \in L^2_{ad}([a,b] \times \Omega)$. By point iii) of remark 4.1.3 we can construct stochastic process X on $[a,b]$, by setting

$$X_t := \int_a^t f(\tau)\mathrm{d}B_\tau \quad t \in [a,b]. \tag{4.1}$$

Moreover X_t is $\mathcal{F}_t$-measurable, that is $\{X_t\}$ is $\{\mathcal{F}_t\}$-adapted. One of the advantages of Itô integral is that $X \equiv \{X_t, \mathcal{F}_t | t \in [a,b]\}$ is a martingale.

Theorem 4.2.1. Martingale Property of Itô Integral
X defined above is a martingale (with respect to $\{\mathcal{F}_t\}$).

Proof. Let $a \leqslant s < t \leqslant b$. We want to show that $\mathrm{E}\left[X_t | \mathcal{F}_s\right] \overset{\text{a.s.}}{=\!=} X_s$. By remark 4.1.3, iii) we have

$$X_t = \int_a^t f(\tau)\mathrm{d}B_\tau = \left(\int_a^s + \int_s^t\right) f(\tau)\mathrm{d}B_\tau = X_s + \int_s^t f(\tau)\mathrm{d}B_\tau.$$

Thus by adaptedness of X we obtain $\mathrm{E}\left[X_t | \mathcal{F}_s\right] \overset{\text{a.s.}}{=\!=} X_s + \mathrm{E}\left[\int_s^t f(\tau)\mathrm{d}B_\tau \middle| \mathcal{F}_s\right]$ and it remains to show that $\mathrm{E}\left[\int_s^t f(\tau)\mathrm{d}B_\tau \middle| \mathcal{F}_s\right] \overset{\text{a.s.}}{=\!=} 0$. We will again begin with a step process and then generalize the result to the whole $L^2_{ad}([a,b] \times \Omega)$. Henceforth we will write f instead of $f|_{[s,t]}$.

a) Assume $f \in \mathcal{I}(L^2_{ad}([a,b]\times\Omega))$, $f = \sum_1^n c_{k-1}\mathbb{1}_{(t_{k-1},t_k]}$ with $\{t_k\}_1^n$ being the partition of $[s,t]$. Then

$$\mathrm{E}\left[\int_s^t f(\tau)\mathrm{d}B_\tau \,\middle|\, \mathcal{F}_s\right] = \mathrm{E}\left[\sum c_{k-1}(B_{t_k} - B_{t_{k-1}}) \,\middle|\, \mathcal{F}_s\right] \overset{\mathcal{F}_s \subset \mathcal{F}_{t_{k-1}}}{=\!=\!=} $$
$$= \mathrm{E}\left[\mathrm{E}\left[\sum c_{k-1}(B_{t_k} - B_{t_{k-1}}) \,\middle|\, \mathcal{F}_{t_{k-1}}\right] \,\middle|\, \mathcal{F}_s\right]$$

by $\mathcal{F}_{t_{k-1}}$-measurability of $c_{t_{k-1}}$

$$= \mathrm{E}\left[c_{k-1}\sum \mathrm{E}\left[B_{t_k} - B_{t_{k-1}} | \mathcal{F}_{t_{k-1}}\right] \,\middle|\, \mathcal{F}_s\right] = 0,$$

by independence of Brownian increments.

b) Let f be arbitrary process from $L^2_{ad}([a,b]\times\Omega)$, let $\{f_n\} \subset \mathcal{I}(L^2_{ad}([a,b]\times\Omega))$ such that $\int_s^t \mathrm{E}\left[|f(\tau) - f_n(\tau)|^2\right]\mathrm{d}\tau \xrightarrow{n\to\infty} 0$. By linearity of conditional expectation and Itô integral we have

$$\mathrm{E}\left[\int_s^t f(\tau)\mathrm{d}B_\tau \,\middle|\, \mathcal{F}_s\right] - \mathrm{E}\left[\int_s^t f_n(\tau)\mathrm{d}B_\tau \,\middle|\, \mathcal{F}_s\right] = \mathrm{E}\left[\int_s^t (f(\tau) - f_n(\tau))\,\mathrm{d}B_\tau \,\middle|\, \mathcal{F}_s\right].$$

Since the second term on the left-hand side is almost surely zero, it suffice to show that

$$= \lim \mathrm{E}\left[\int_s^t (f(\tau) - f_n(\tau))\,\mathrm{d}B_\tau \,\middle|\, \mathcal{F}_s\right] \overset{\text{a.s.}}{=\!=} 0,$$

or equivalently that

$$\mathrm{E}\left[\left(\mathrm{E}\left[\int_s^t (f(\tau) - f_n(\tau))\,\mathrm{d}B_\tau \,\middle|\, \mathcal{F}_s\right]\right)^2\right] \xrightarrow{n\to\infty} 0.$$

We use the special case of conditional Jensen's inequality $\mathrm{E}^2[Y] \leqslant \mathrm{E}\left[Y^2\right]$ and the partial averaging property of conditional expectation to obtain

$$\mathrm{E}\left[\left(\mathrm{E}\left[\int_s^t (f(\tau) - f_n(\tau))\,\mathrm{d}B_\tau \,\middle|\, \mathcal{F}_s\right]\right)^2\right] \leqslant \mathrm{E}\left[\mathrm{E}\left[\left|\int_s^t (f(\tau) - f_n(\tau))\,\mathrm{d}B_\tau\right|^2 \,\middle|\, \mathcal{F}_s\right]\right]$$
$$= \mathrm{E}\left[\left|\int_s^t (f(\tau) - f_n(\tau))\,\mathrm{d}B_\tau\right|^2\right]$$

by isometry of Itô integral (theorem 4.1.4)

$$= \int_s^t \mathrm{E}\left[|f(\tau) - f_n(\tau)|^2\right]\mathrm{d}\tau \xrightarrow{n\to\infty} 0.$$

□

4.3 Wiener Integral

Assume we have a deterministic function $f : \mathbb{R}_0^+ \to \mathbb{R}$ and recall that B is a standart Brownian motion. We now take interest in

$$\int_a^b f(t)\mathrm{d}B_t \qquad (b > a \geqslant 0).$$

Since we have already defined the Itô integral we can use it in obvious manner to help us define the notion above. Let us consider process g such that $g(t,\omega) = f(t)$ for every $\omega \in \Omega$ and every $t \geqslant 0$. Then, given that $f \in L^2([a,b])$, we see that $g \in L^2_{ad}([a,b] \times \Omega)$. The measurability of $g(t,\cdot)$ for every t is evident - the preimage of every set is either Ω or $\emptyset$. Also the adaptedness follows easily from the definition:

$$\mathrm{E}\left[g(t)|\mathcal{F}_t\right] = \mathrm{E}\left[f(t)|\mathcal{F}_t\right] = f(t) = g(t).$$

Thus we can use the Itô integral of g to define the stochastic integral of deterministic function of f. We have thus defined the integral $\int_a^b f(t)\mathrm{d}B_t$ for every $f \in L^2([a,b])$. There is another approach - we might have defined this integral the very same way we defined Itô integral, that is use step functions from $L^2([a,b])$, density of step functions in $L^2([a,b])$ and a limit transition. This would lead to the same result though. The just defined integral is thus the special case of Itô integral and it is called Wiener integral. We will use the same notation - the Wiener integral of f over $[a,b]$ will be denoted by $\int_a^b f(t)\mathrm{d}B_t$ and if the domain will be clear it will be denoted by $I(f)$ as well. Of course, the results valid for Itô remain valid for Wiener integral. Yet there is more we can say about Wiener integral - since the integrand is deterministic we 'should know more'. This indeed is true as for deterministic function f we know the law of $\int_a^b f(t)\mathrm{d}B_t$ precisely. This is stated in the following theorem.

Theorem 4.3.1. Law of Wiener Integral
Let $f \in L^2([a,b])$, then the Wiener integral of f on $[a,b]$ is a centered normal random variable with variance $\int_a^b f^2(t)\mathrm{d}t$, that is $\int_a^b f(t)\mathrm{d}B_t \sim N(0, \int_a^b f^2(t)\mathrm{d}t)$.

Proof. We use the same approach as before, that is we will prove the claim for the step functions and then use the limit transition.

i) Assume $f(t) = \sum\limits_{i=1}^{n} c_i \mathbb{1}_{[t_{i-1},t_i)}(t)$, where $\{t_i\}_1^n$ is a partition of $[a,b]$. Note that c_k are now constant numbers. Since

$$I(f) = \sum c_i (B_{t_i} - B_{t_{i-1}}),$$

$I(f)$ is the linear combination of standart normal random variables (the Brownian increments). It is the general result for every Itô integral, that its expectation is zero. Thus it remains to show that the variance of $I(f)$ indeed has the claimed value. This readily follows from the properties of Brownian increments:

$$\begin{aligned}
Var[I(f)] = \mathrm{E}\left[I^2(f)\right] &= \mathrm{E}\left[\sum_i\sum_j c_i c_j (B_{t_i} - B_{t_{i-1}})(B_{t_j} - B_{t_{j-1}})\right] \\
&= \sum_i\sum_j c_i c_j \mathrm{E}\left[(B_{t_i} - B_{t_{i-1}})(B_{t_j} - B_{t_{j-1}})\right] \\
&= \sum_i c_i^2 \mathrm{E}\left[(B_{t_i} - B_{t_{i-1}})^2\right] = \sum_i c_i^2 (t_i - t_{i-1}) = \int_a^b f^2(t)\mathrm{d}t.
\end{aligned}$$

ii) For general f from $L^2([a,b])$ it suffice to use that if $X_n \sim N(\mu_n, \sigma_n)$ converge to X in $L^2(\Omega)$, then $X \sim N(\lim \mu_n, \lim \sigma_n)$.

□

Example 4.3.1. From [4], p. 37, ex. 1
For arbitrary $t \in \mathbb{R}^+$ we have

$$\int_0^t s \, \mathrm{d}B_s = tB_t - \int_0^t B_s \mathrm{d}s \qquad \text{almost surely.} \tag{4.2}$$

Proof. We first discuss the term on the left-hand side and the second term on the right-hand side separately. Let $\sigma_n \equiv \left\{t_i^{(n)}\right\}_0^n$ be a normal sequence of equidistant partitions of $[0,t]$, we shall drop the upper indices of times to simplify the notation ($\sigma_n = \{t_i\}_0^n$). This shall not cause any confusion as the times from different partitions will not be mixed. Since f given by $f(s) = s$ on $[0,t]$ is in $L^2_{ad}([a,b] \times \Omega)$ and since $\sum_{i=1}^n t_{i-1}\mathbb{1}_{(t_{i-1},t_i]} \xrightarrow[n\to\infty]{L^2} f$, we obtain by definition of Itô integral

$$\int_0^t s \, \mathrm{d}B_s = \lim_{n\to\infty} \int_0^t \sum_{i=1}^n t_{i-1}\mathbb{1}_{(t_{i-1},t_i]}(s)\mathrm{d}B_s = \lim \sum_1^n t_{i-1}\left(B_{t_i} - B_{t_{i-1}}\right). \tag{4.3}$$

Almost every sample path of B is continuous by definition of Wiener process. Hence $B(\omega)$ is continuous on $[0,t]$ for almost every $\omega \in \Omega$ and is thus Riemann-integrable on $[0,t]$. Therefore $\int_0^t B_s \mathrm{d}s$ can be defined path-by-path for almost every ω. By definition of Riemann integral we write

$$\int_0^t B_s \mathrm{d}s = \lim_{n\to\infty} \sum_1^n B_{\tilde{t}_i}(t_i - t_{i-1}), \qquad \tilde{t}_i \in [t_{i-1}, t_i] \text{ arbitrary.} \tag{4.4}$$

Using (4.3) and (4.4) we now rewrite the original equation (4.2) as

$$\begin{aligned}
tB_t &= \lim_{n\to\infty} \sum_1^n t_{i-1}\left(B_{t_i} - B_{t_{i-1}}\right) + \lim_{n\to\infty} \sum_1^n B_{\tilde{t}_i}(t_i - t_{i-1}) \\
&= \lim \left(\sum_1^n t_{i-1}\left(B_{t_i} - B_{t_{i-1}}\right) + \sum_1^n B_{\tilde{t}_i}(t_i - t_{i-1})\right).
\end{aligned}$$

To prove the claim it suffice to show that the term on the right-hand side indeed is equal to the left-hand side. We shall drop the limit and rewrite the sums as follows:

$$\sum_1^n t_{i-1}\left(B_{t_i} - B_{t_{i-1}}\right) + B_{\tilde{t}_i}(t_i - t_{i-1}) = \sum_1^n t_{i-1}B_{t_i} - t_{i-1}B_{t_{i-1}} + B_{\tilde{t}_i}t_i - B_{\tilde{t}_i}t_{i-1} \equiv R.$$

By construction of Riemann integral $\tilde{t}_i$ may be chosen freely within $[t_{i-1}, t_i]$. By setting $\tilde{t}_i := t_i$ we finally obtain

$$\begin{aligned} R &= \sum_1^n t_{i-1}B_{t_i} - t_{i-1}B_{t_{i-1}} + B_{t_i}t_i - B_{t_i}t_{i-1} = \sum_1^n -t_{i-1}B_{t_{i-1}} + B_{t_i}t_i \\ &= t_1B_{t_1} - \underbrace{t_0B_{t_0}}_{=0 \text{ a. s.}} + t_2B_{t_2} - t_1B_{t_1} + \ldots + t_nB_{t_n} - t_{n-1}B_{t_{n-1}} \\ &\overset{\text{a.s.}}{=\!=} tB_t. \end{aligned}$$

This readily yields the result. ☐

5 Integration with Controlled Integrand

It is useful to know under which circumstances the integral $\int_a^b \sigma(X_t, u_t)\mathrm{d}B_t$ is well defined as these are exactly the circumstances under which the term $\sigma(X_t, u_t)\mathrm{d}B_t$ in stochastic differential equation makes sense. The said conditions are given for several special cases of σ in this chapter.

5.1 Integration of $\sigma(X, u)$ with Respect to B

Assume $\sigma : \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ is Borel measurable function. We are interested in conditions we must impose on the integrand $\sigma(X, u)$ in order to ensure the existence of Itô integral $\int_a^b \sigma(X_t, u_t)\mathrm{d}B_t$. From the construction of the integral in the preceding chapter, it is evident, that if $\sigma(X, u) \in L^2_{ad}([a, b] \times \Omega)$, then the integral is well defined. The aim is to simplify this condition in some special cases as well as in case of general Borel measurable σ. Recall the definition of $L^2_{ad}([a, b] \times \Omega)$ (definition 4.1.1). Let us examine the $\{\mathcal{F}_t\}$-adaptedness of the integrand first. Since σ is Borel measurable by assumption, the adaptedness of X and u is sufficient to ensure the adaptedness of the whole integrand. This is rather natural condition as the control $u_t(\omega)$ depends on observations of the system up to t. Obviously, the same result applies to the measurability with respect to $\mathcal{B}(\mathbb{R}) \otimes \mathcal{F}$.
The second concern is the finiteness of L^2 norm of the integrand. Here the assumption $X, u \in L^2$ will simply not suffice. Consider for example the following. Let $X \in L^2([a, b])$ such that $X \notin L^1([a, b])$, let $\sigma : (x, y) \to \sqrt{|x|}$ and let u be arbitrary process from $L^2_{ad}([a, b] \times \Omega)$. Then $\sigma(X, u) \notin L^2_{ad}([a, b] \times \Omega)$ since the L^2 norm is infinite. Therefore in the general case, we have to assume that X, u are $\{\mathcal{F}_t\}$-adapted processes on $\Omega \times [a, b]$ with $\int_a^b \mathrm{E}\left[\sigma^2(X_t, u_t)\right] \mathrm{d}t < +\infty$. Some special cases of σ will now be discussed.

i) $\sigma(x, y) = \sigma(x)$
The very same example as in general case may be applied here. Thus the assumption of X itself being in L^2 is insufficient and the whole $Y \equiv \sigma(X)$ must be in L^2. In fact in this case the problem is reduced to the original question from the beginning of the preceding chapter, that is for which processes it is reasonable to construct the Itô integral, since Y is a process on $[a, b] \times \Omega$.
This case seemingly does not involve any control, but it is not entirely true. If the control u only depends on the process X, that is if $u = \sigma(X)$, then this case shows the assumptions we have to make in order to work with this controll, more precisely to give a good meaning to the stochastic differential equation

containing the term $u_t \mathrm{d}W_t = \sigma(X_t)\mathrm{d}W_t$. Such control is called a closed loop or feedback control.

ii) $\sigma(x, y) = \alpha x + \beta y$, where $\alpha, \beta \in \mathbb{R}$
By property of L^2 space being a vector space we conclude that $X, u \in L^2$ is enough for $\int_a^b (\alpha X + \beta u)\mathrm{d}B$ to exist.

iii) $\sigma(x, y) = \alpha xy$, where $\alpha, \beta \in \mathbb{R}$
This will be dissolved in four subcases:

a) X or u is bounded.
Let $X, u \in L^2_{ad}([a, b] \times \Omega)$ and let $X_t(\omega) \leqslant K \in \mathbb{R}$ for all t and ω without any loss of generality. Then

$$\int_a^b \mathrm{E}\left[\alpha^2 X_t^2 u_t^2\right] \mathrm{d}t \leqslant \alpha^2 K \int_a^b \mathrm{E}\left[u_t^2\right] \mathrm{d}t < +\infty,$$

which yields $\alpha X u \in L^2_{ad}([a, b] \times \Omega)$.

b) Neither X nor u is bounded.
Let $X = u \in L^2_{ad}([a, b] \times \Omega)$ such that

$$X_t(\omega) = \begin{cases} \frac{1}{\sqrt[4]{t-a}} Z & t \in (a, b] \\ \text{arbitrary} & t = a, \end{cases}$$

where Z is standard normal random variable. Then

$$\int_a^b \mathrm{E}\left[\alpha^2 X_t^2 u_t^2\right] \mathrm{d}t = \alpha^2 \int_a^b \frac{1}{t-a} \underbrace{\mathrm{E}\left[Z^4\right]}_{3} \mathrm{d}t = 3\alpha^2 \left[\ln(t-a)\right]_a^b$$

$$= 3\alpha^2 \left(\ln(b-a) - \lim_{\tau \downarrow 0} \ln(\tau) \right) = +\infty,$$

hence the sole assumption of X, u having finite L^2 norm is too weak for $\alpha X u \in L^2$ to hold in this case.

c) $X, u \in L^4$, that is $\int_a^b \mathrm{E}\left[X^4\right] \mathrm{d}t$, $\int_a^b \mathrm{E}\left[u^4\right] \mathrm{d}t < +\infty$.
Using the higher order of integrability and Cauchy-Schwarz inequality we obtain

$$\int_a^b \mathrm{E}\left[\alpha^2 X_t^2 u_t^2\right] \mathrm{d}t \leqslant \alpha^2 \sqrt{\int_a^b \mathrm{E}\left[X_t^4\right] \mathrm{d}t} \sqrt{\int_a^b \mathrm{E}\left[u_t^4\right] \mathrm{d}t} < +\infty$$

which means that under these assumptions $I(\alpha X u)$ exists.

d) More generally if there is $p, q > 1$ such that $\frac{1}{p} + \frac{1}{q} = 1$ and $\int_a^b \mathrm{E}\left[X_t^{2p}\right] \mathrm{d}t$, $\int_a^b \mathrm{E}\left[X_t^{2q}\right] \mathrm{d}t < +\infty$, we can use Hölder inequality to see that these assumptions are sufficient to ensure the existence of the integral as well:

$$\int_a^b \mathrm{E}\left[\alpha^2 X_t^2 u_t^2\right] \mathrm{d}t \leqslant \alpha^2 \left(\int_a^b \mathrm{E}\left[X_t^{2p}\right] \mathrm{d}t \right)^{\frac{1}{p}} \left(\int_a^b \mathrm{E}\left[u_t^{2q}\right] \mathrm{d}t \right)^{\frac{1}{q}} < +\infty.$$

Note that the inequality has been applied to $X^2 u^2$, rather than to Xu itself.

Conclusion

Groundwork for formulation of stochastic system was laid in the work. The first part covers the basic concepts needed for construction of the stochastic integral. The first chapter introduces the elementary notions of stochastic theory and the proof of Daniell-Kolmogorov theorem is given. This allows proving the existence of Wiener process defined in the second chapter, the proof is carried out there as well. Properties of sample paths important for understanding the problems of stochastic calculus are shown in the third chapter. These properties motivate the construction of Itô integral, which is carried out in the fourth chapter. The Itô integral is the central notion of the work as it allows to formulate (and possibly solve) stochastic differential equations describing the stochastic systems we want to model.

Despite the fact that the description of the system is important issue, our aim is to manipulate the system in our favour. Thus the stochastic control is of concern. The last part of the work specifies the conditions we must impose on the integrand (which corresponds to the controlled term in stochastic differential equation) in order to ensure the existence of its Itô integral. The general conditions following directly from the construction of Itô integral are simplified or transformed in different, and perhaps more easily verified, conditions.

This is the starting point for formulating a problem of finding the 'optimal' control of some stochastic system. The conditions extracted here specify which terms may be contained in stochastic differential equations. The next step is to formulate a stochastic differential equation involving stochastic control and examine the solutions. This may have be done with various restrictions on the control which consequently allows to model numerous situations in a way, which reflects the reality with less distortion. For example, the control applied to the system might not be continuous. This corresponds to the situation in which the control is applied by someone at discrete times. Such problems are of great importance in applications.

List of Symbols

$\mathrm{E}[X]$	Expectation of X.
$\int_a^b f(t)\mathrm{d}B_t,\ \int_a^b f\mathrm{d}B$	Itô integral of f over $[a,b]$, page 34.
$\int_a^b f(t)\mathrm{d}B_t,\ \int_a^b f\mathrm{d}B$	Wiener integral of f over $[a,b]$, page 37.
$\langle X\rangle_t$	Quadratic variation of process X, page 23.
$\mathbb{R}^+$	The set of strictly positive real numbers.
$\mathbb{R}^{[0,+\infty)}$	Space of real-valued functions on $[0,+\infty)$, page 10.
$\mathbb{R}_0^+$	The set of nonnegative real numbers.
$\mathcal{A},\mathcal{F}$	σ-algebra, σ-field.
$\mathcal{A}\otimes\mathcal{F}$	Product σ-algebra.
$\mathcal{B}(X)$	Borel σ-algebra on X.
$\mathcal{I}(X)$	The set of step functions from $\{X\}$.
$\mathcal{K}\left(\mathbb{R}^{[0,+\infty)}\right)$	Kolomogorov σ-algebra on $\mathbb{R}^{[0,+\infty)}$, page 10.
$\{\mathcal{F}_t^X\}$	Canonical filtration generated by process X, page 9.
D^+ and D_+	Upper and lower Dini derivatives, page 24.
$L^2_{ad}([a,b]\times\Omega)$	The space of adapted, quadratically integrable processes, page 29.
$N(\mu,\sigma^2)$	Normal (Gaussian) distribution with mean μ nd variance σ^2.
P-a. s.	Almost surely with respect to probability measure P.
S_n	The set of all permutations on $\{1,\ldots,n\}$.
$W,\ B$	Wiener process, Brownian motion.
$(\Omega,\mathcal{F},P)$	Abstract probability space.
a. s.	Almost surely.

Bibliography

[1] Ioannis Karatzas and Steven E. Shreve. *Brownian Motion and Stochastic Calculus*. Springer-Verlag, 1991.

[2] Hui-Hsiung Kuo. *Introduction to Stochastic Integration*. Springer-Verlag, 2006.

[3] Mario Lefebvre. *Applied Stochastic Processes*. Springer-Verlag, 2006.

[4] Bernt Øksendal. *Stochastic Differential Equations: An Introduction with Applications*. Springer-Verlag, 2003.

[5] Martin Ondreját and Jan Seidler. On existence of progressively measurable modifications. To appear.

[6] Steven E. Shreve. *Stochastic Calculus for Finance II: Continuous Time Models*. Springer-Verlag, 2004.