Università degli studi di Pavia
Dipartimento di Matematica
Corso di Laurea Magistrale in Matematica

UNIVERSITÀ
DI PAVIA

Il teorema di Girsanov con applicazioni alle equazioni differenziali stocastiche
The Girsanov Theorem with Applications to Stochastic Differential Equations

Tesi di Laurea Magistrale in Matematica

Relatore (Supervisor):
Enrico Priola

Tesi di Laurea di:
Davide Paolillo
Matricola 505846

Anno Accademico 2025-2026

Abstract

The main goal of this thesis is to present Girsanov's theorem and some of its applications to Stochastic Differential Equations (SDEs) and mathematical finance. After recalling the basic tools of stochastic analysis and probability theory, we study the Wiener process and Itô stochastic integration. In particular, we discuss the reflection principle, providing a full proof, and derive Bachelier's theorem. A central part of the thesis is devoted to the construction of the Itô integral. We provide a detailed construction following the approach of N. V. Krylov [14], which is more general than some classical treatments such as [3] or [13], since it does not require progressive measurability.

The core of the thesis is to provide a complete proof of Girsanov's theorem. We present the elegant approach by N. V. Krylov [14], while including several details that are omitted in his exposition. We also clarify some points of the Krylov method. Particular attention is devoted to the exponential martingale

$$\rho_t(b) = \exp\left(\int_0^t b_s dW_s - \frac{1}{2}\int_0^t b_s^2 ds\right)$$

involved in the change of measure, to Novikov's condition, and to further sufficient criteria ensuring the martingale property of ρ_t.

We then apply Girsanov's theorem to prove weak existence and uniqueness in law for stochastic differential equations of the form

$$dX_t = b(t, X_t)\, dt + dW_t,$$

assuming that b is Borel and grows at most linearly. Finally, we discuss an application to mathematical finance and derive the Black–Scholes formula for European options.

Contents

Introduction

The main goal of this thesis is to present Girsanov's theorem and illustrate its applications to the theory of stochastic differential equations (SDEs) and mathematical finance. Girsanov's theorem is a fundamental result in stochastic calculus concerning the change of probability measures. Recall that stochastic calculus is based on a theory of stochastic integrals with respect to particular processes, such as the Wiener process or Brownian motion; this field was created by the Japanese mathematician Kiyosi Itô during World War II.

Girsanov's theorem describes how the "drift" of a stochastic process changes under a change of the underlying probability measure. It shows that, under a suitable change of probability measure, the process obtained by subtracting a possibly random drift term $\int_0^t b_s\, ds$ from a Wiener process is again a Wiener process.

Before Igor Girsanov, the foundational result in this area was the *Cameron–Martin theorem*; see [9]. In 1944, Robert Cameron and William Martin studied how the Wiener measure (the probability law of Brownian motion) transforms when the path is shifted by a deterministic function.

They proved that if one shifts a Brownian motion W_t by a function $h(t)$, the new measure is equivalent (mutually absolutely continuous) to the old one if and only if $h(t)$ is sufficiently smooth; more precisely: if it belongs to the Cameron–Martin Hilbert space. This was the first "change of measure" result, though it was limited to deterministic translations.

In 1960, the Soviet mathematician Igor Vladimirovich Girsanov published his seminal paper. He made the conceptual leap from deterministic shifts to stochastic shifts.

Roughly speaking, given a Brownian motion W_t under a probability measure $\mathbb{P}$, define a new measure $\mathbb{Q}$ via the Radon–Nikodym derivative:

$$\frac{d\mathbb{Q}}{d\mathbb{P}} = \rho_T(b) = \exp\left(\int_0^T b_s dW_s - \frac{1}{2}\int_0^T b_s^2 ds\right)$$

Then, the process defined by $\tilde{W}_t = W_t - \int_0^t b_s ds$ is a Brownian motion under the measure $\mathbb{Q}$.

One technical challenge with Girsanov's original result was ensuring that the exponential process used for the change of measure is a true martingale rather than just a local martingale. If it is not a true martingale, the new "measure" $\mathbb{Q}$ does not have a total mass of 1, and the theorem fails.

In 1972, A.A. Novikov [17] provided a sufficient condition that is now standard in the literature. The *Novikov Condition* states that if:

$$\mathbb{E}\left[\exp\left(\frac{1}{2}\int_0^T b_s^2 ds\right)\right] < \infty \quad (*)$$

then the change of measure is valid.

The core of the thesis is to provide a complete proof of Girsanov's theorem. We present the elegant approach by N.V. Krylov [14], while including several details that are omitted in his exposition. We also clarify some points of the Krylov method. This approach is different from the classical approach considered for instance in Baldi [3] and Karatzas and Shreve [13] which requires additional results from the theory of stochastic calculus. Moreover, starting from the Novikov condition (*) we consider a further sufficient condition which is more general than Novikov's condition and ensures the validity of the Girsanov theorem.

While Girsanov worked in the realm of pure probability, his theorem became important in mathematical finance starting from the Black–Scholes model (1973): [7]. Although they did not explicitly use Girsanov's name in their original derivation, the logic of "risk-neutral pricing" depends entirely on this theorem, see also [13] and [3] for more details on applications of the Girsanov theorem to finance.

The thesis is structured as follows:

Chapter 1 This is an introductory chapter, in which we set the notations used in the rest of the thesis and introduce the first theoretical bases. It collects the main preliminaries from probability theory and stochastic processes without going deep into details and proofs. We refer the reader to standard references such as Billingsley [4] and Karatzas and Shreve [3]. In particular we recall the notions and properties of probability spaces, random variables, L^p spaces, convergence, conditional expectation, stochastic processes, stopping times, and martingales.

Chapter 2 This chapter is devoted to the Wiener process and Itô stochastic integration. We introduce weak convergence, the Wiener measure on $C[0,T]$, discuss some properties of Brownian motion, including an application of the reflection principle that we use to derive Bachelier's theorem. In this part we do not provide complete proofs for all the mentioned results.

In the second part of the chapter we present a detailed description of the Itô integral, following the approach of Krylov [14], which is more general than some classical treatments, since it does not require progressive measurability. The chapter concludes with the Itô formula.

Chapter 3 We prove in detail Girsanov's theorem. We first prove some preliminary lemmas about the exponential process ρ_t associated with the change of measure in Girsanov's theorem. We then present Girsanov's theorem following the approach of N.V. Krylov. We also clarify some points of the Krylov method. Since the result requires the exponential process ρ_t to be a martingale, we then study sufficient conditions ensuring this property.

Chapter 4 This chapter is devoted to applications to the theory of stochastic differential equations with singular coefficients. We consider the equation

$$dX_t = b(t, X_t)dt + dW_t,$$

and in the first section we study weak existence and uniqueness in law assuming that b is Borel and grows at most linearly.

These results are obtained as direct consequences of Girsanov's theorem, and we present their proofs in some detail.

In the remaining sections, we discuss some further recent developments in the theory of SDEs, such as path by path uniqueness and Davie's theorem [10]. For these more recent results, we mainly refer to the existing literature without going into details.

Chapter 5 In this chapter we present an application of Girsanov's theorem to mathematical finance. We first introduce the basic notions needed to describe a financial market, including assets, portfolios, European options and discounted prices. We then show how with a change of measure based on Girsanov's theorem the discounted prices process becomes a martingale. This is the key step in pricing, since it allows one to express the value of an option as the expected discounted payoff under an equivalent martingale measure. In case of constant coefficients, this expectation is also easy to compute, in fact we conclude by deriving the Black-Scholes formula for the valuation of European options, that is an explicit computation of the expectation in the Black–Scholes model [7].

Chapter 1

Fundamentals of probability theory

In this chapter we recall some fundamental concepts from probability theory and the theory of stochastic processes that will be used throughout the thesis. Our goal is to fix notations and summarize all these basic concepts in order to prepare the reader for stochastic integration. The main references for this chapter are Billingsley [5] and Karatzas and Shreve [13].

1.1 Probability basics

The mathematical framework for probability theory is provided by measure theory. Random phenomena are modeled by assigning probabilities to suitable subsets of a given space of outcomes, in a way that is consistent with the axioms of a measure. For a detailed treatment of probability theory based on measure theory, as well as complete proofs of the results stated in this section, the reader is referred to Billingsley [5].

Let Ω be a non-empty set, called the *sample space*, whose elements represent the possible outcomes of a random experiment. In order to assign probabilities, one needs to specify a class of subsets of Ω which are called measurable sets or events, more precisely

Definition 1.1. *A* σ-algebra $\mathcal{F}$ *on* Ω *is a collection of subsets of* Ω *satisfying:*

1. $\Omega \in \mathcal{F}$,
2. *if* $A \in \mathcal{F}$, *then* $A^c \in \mathcal{F}$,
3. *if* $(A_n)_{n\in\mathbb{N}} \subset \mathcal{F}$, *then* $\bigcup_{n=1}^{\infty} A_n \in \mathcal{F}$.

From the definition it follows immediately that a $\sigma-$algebra is also closed under countable intersections.

When we say that $\mathcal{G}$ is a sub $\sigma-$ algebra of $\mathcal{F}$ we simply mean $\mathcal{G} \subset \mathcal{F}$.

The elements of $\mathcal{F}$ are called *events.* The pair $(\Omega, \mathcal{F})$ is called a measurable space. Two simple examples are worth mentioning. The smallest possible σ-algebra on Ω is the *trivial* σ-algebra $\{\emptyset, \Omega\}$. At the opposite extreme, the largest σ-algebra is $\mathcal{P}(\Omega)$, consisting of all subsets of Ω. In this case every subset is measurable. More generally, starting from a collection of subsets $\mathcal{A} \subset \mathcal{P}(\Omega)$, one can construct the smallest σ-algebra containing $\mathcal{A}$. This is denoted by

$$\sigma(\mathcal{A}),$$

and is defined as the intersection of all σ-algebras containing $\mathcal{A}$. An important example is when Ω is a topological space, endowed with a topology $\mathcal{T}$ containing all open sets. In this case one can consider $\sigma(\mathcal{T})$, i.e. the smallest $\sigma-$algebra containing all open sets of Ω. This is called the *Borel* σ-algebra of Ω and it is denoted as $\mathcal{B}(\Omega)$.

Definition 1.2. *Let $(\Omega, \mathcal{F})$ be a measurable space. A* probability measure *on $(\Omega, \mathcal{F})$ is a function $\mathbb{P} : \mathcal{F} \to [0, 1]$ such that:*

1. $\mathbb{P}(\Omega) = 1$,
2. *for every sequence of pairwise disjoint sets $(A_n)_{n\in\mathbb{N}} \subset \mathcal{F}$,*

$$\mathbb{P}\left(\bigcup_{n=1}^{\infty} A_n\right) = \sum_{n=1}^{\infty} \mathbb{P}(A_n).$$

In this setting, probabilities are assigned only to events in $\mathcal{F}$; subsets of Ω which do not belong to $\mathcal{F}$ are not measurable and therefore have no assigned probability. The second property of probability measures is called *sigma-additivity.* Probabilities are indeed sigma-additive functions taking values in $[0, 1]$ with total mass equal to 1. We recall some basic properties that follow directly from the definition. First, $\mathbb{P}(\emptyset) = 0$, and the measure is monotone, in the sense that if $A, B \in \mathcal{F}$ and $A \subseteq B$, then $\mathbb{P}(A) \leq \mathbb{P}(B)$. Moreover, for every event $A \in \mathcal{F}$ one has $\mathbb{P}(A^c) = 1 - \mathbb{P}(A)$.

Given two events $A, B \in \mathcal{F}$, the probability of their union satisfies the identity

$$\mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B) - \mathbb{P}(A \cap B),$$

which follows from σ-additivity. More generally, for any sequence of events $(A_n)_{n\in\mathbb{N}} \subset \mathcal{F}$, one has the subadditivity property

$$\mathbb{P}\left(\bigcup_{n=1}^{\infty} A_n\right) \leq \sum_{n=1}^{\infty} \mathbb{P}(A_n).$$

In particular a direct consequence of this property is that countable unions of sets having zero probability are still sets of zero probability. These properties are standard consequences of the measure-theoretic structure and will be used implicitly throughout the thesis.

If $(\Omega, \mathcal{F}, \mathbb{P})$ is a probability space we say that $A \in \mathcal{F}$ is a $\mathbb{P}$-null set if $\mathbb{P}(A) = 0$. In the next chapters we may assume that our probability spaces are *complete.* Recall that a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ is said to be complete if for any $A \subseteq B$, $B \in \mathcal{F}$, $\mathbb{P}(B) = 0$, then the set A belongs to $\mathcal{F}$. In other words if every subset of a $\mathbb{P}$-null set is an event. We now introduce the notion of random variable in a general measurable setting.

Let $(\Omega, \mathcal{F})$ be a measurable space and let $(E, \mathcal{E})$ be another measurable space. A function

$$X : \Omega \to E$$

is said to be *measurable* if for every set $A \in \mathcal{E}$ one has

$$X^{-1}(A) := \{\omega \in \Omega : X(\omega) \in A\} \in \mathcal{F}.$$

In this case, X is called a *random variable with values in* $(E, \mathcal{E})$.

In other words, the measurability condition requires that the preimage of every measurable set in E is an event in Ω.

An important case is when $E = \mathbb{R}^d$ and $\mathcal{E} = \mathcal{B}(\mathbb{R}^d)$. In this case, X is called a *Borel measurable* or *Borel* random variable. Given a random variable $X : \Omega \to (E, \mathcal{E})$ one can also define the σ-algebra generated by X as

$$\sigma(X) := \{X^{-1}(A) : A \in \mathcal{E}\},$$

that is, the smallest σ-algebra with respect to which X is measurable. More generally, given a family of random variables $(X_i)_{i \in I}$ with values in measurable spaces $(E_i, \mathcal{E}_i)$, one defines the σ-algebra generated by the family as

$$\sigma(X_i, i \in I) := \sigma\big(X_i^{-1}(A) : A \in \mathcal{E}_i,\ i \in I\big),$$

that is, the smallest σ-algebra on Ω with respect to which each X_i is measurable.

Given a random variable $X : \Omega \to E$ defined on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$, one can associate to X a probability measure μ_X on $(E, \mathcal{E})$, defined by

$$\mu_X(A) := \mathbb{P}(X \in A) = \mathbb{P}(X^{-1}(A)), \quad A \in \mathcal{E}.$$

This measure is called the *law* (or *distribution*) of X.

In the case of real-valued random variables, the law can be described more explicitly through the distribution function.

Let $X : \Omega \to \mathbb{R}$ be a real-valued random variable. The function $F_X : \mathbb{R} \to [0, 1]$ defined by

$$F_X(x) := \mathbb{P}(X \le x)$$

is called the *distribution function* of X. F_X is non-decreasing, right-continuous, and satisfies

$$\lim_{x \to -\infty} F_X(x) = 0, \qquad \lim_{x \to +\infty} F_X(x) = 1.$$

Moreover, the law of X is uniquely determined by its distribution function.

Depending on the structure of its law, one distinguishes different classes of random variables. We describe here the most common notions in the case of random variables with values in $\mathbb{R}^n$. An $\mathbb{R}^n$-valued random variable is said to be *discrete* if it takes values on a countable set, that is, if there exists a countable set $A \subseteq \mathbb{N}$ and $\{x_k\}_{k \in A} \subset \mathbb{R}^n$ such that

$$\sum_{k \in A} \mathbb{P}(X = x_k) = 1.$$

In this case, the law of X is completely described by the probabilities $\mathbb{P}(X = x_k)$.

On the other hand, X is said to be *(absolutely) continuous* if there exists a non-negative measurable function $f_X : \mathbb{R}^n \to \mathbb{R}$ such that

$$\mathbb{P}(X \in A) = \int_A f_X(x)\, dx, \qquad A \in \mathcal{B}(\mathbb{R}^n).$$

The function f_X is called the *density* of X, and in this case, if X is real–valued, the distribution function can be written as

$$F_X(x) = \int_{-\infty}^{x} f_X(y)\, dy.$$

In this case, that is, if X is a real-valued random variable with density f_X, then for every $k \in \mathbb{R}$,

$$\mathbb{P}(X = k) = \int_{\{k\}} f_X(x)\, dx = 0,$$

since singletons have Lebesgue measure zero.

A fundamental example is provided by the *Gaussian* distribution. A random variable X is said to have a Gaussian or *normal* distribution with mean $m \in \mathbb{R}$ and variance $\sigma^2 > 0$, and we write $X \sim \mathcal{N}(m, \sigma^2)$, if its density is given by

$$f_X(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(x-m)^2}{2\sigma^2}\right), \qquad x \in \mathbb{R}.$$

The Gaussian distribution plays a central role in probability theory and stochastic processes, in particular in connection with the Wiener process, which we introduce in the next chapter. For our purpose we will need the notion of multivariate normal law, that extends the one–dimensional case. A random vector $X_1, \ldots, X_n$ is said to have a multivariate normal distribution if every linear combination $a^\top X$ with $a \in \mathbb{R}^n$ is a one–dimensional random variable with normal distribution. We shall recall this notion more explicitly in Chapter 2, when actually needed.

1.1.1 Almost sure properties

Throughout the thesis, whenever a property is said to hold *almost surely*, we mean that it holds outside a set of zero probability. We will use interchangeably the notations

$$\text{a.s.} \qquad \text{and} \qquad \mathbb{P}\text{-a.s.}$$

to express this fact, where $\mathbb{P}$ is a probability measure.

In particular, if X and Y are random variables, the notation

$$X = Y \quad \text{a.s.}$$

means that

$$\mathbb{P}(\{\omega \in \Omega : X(\omega) = Y(\omega)\}) = 1.$$

1.1.2 Product spaces

We now recall the notions of product σ-algebra and product measure, which will be used later when dealing with stochastic processes as functions of both time and ω. In particular, these notions provide the natural measurable structure on spaces of the form $I \times \Omega$.

Let $(E, \mathcal{E})$ and $(F, \mathcal{F})$ be measurable spaces. The *product σ-algebra* on $E \times F$ is the σ-algebra generated by the family of measurable rectangles

$$A \times B, \qquad A \in \mathcal{E},\ B \in \mathcal{F}.$$

It is denoted by

$$\mathcal{E} \otimes \mathcal{F} := \sigma(\{A \times B :\ A \in \mathcal{E},\ B \in \mathcal{F}\}).$$

If now $(E, \mathcal{E}, \mu)$ and $(F, \mathcal{F}, \nu)$ are measure spaces, with μ and ν sigma-finite, one can define a measure on $(E \times F, \mathcal{E} \otimes \mathcal{F})$, called the *product measure*, and denoted by $\mu \times \nu$, characterized by the identity

$$(\mu \times \nu)(A \times B) = \mu(A)\nu(B)$$

for every measurable rectangle $A \times B$.

In particular, if $(\Omega, \mathcal{F}, \mathbb{P})$ is a probability space and $I \subset \mathbb{R}^+$ is an interval endowed with its Borel σ-algebra $\mathcal{B}(I)$ and Lebesgue measure dt, the product space

$$(\Omega \times I,\ \mathcal{F} \otimes \mathcal{B}(I),\ \mathbb{P} \times dt)$$

provides the natural framework for studying measurable stochastic processes. At this regard we also recall *Fubini*'s theorem that is very useful when dealing with measurable functions of two variables, in particular stochastic processes. If $f : \Omega \times I \to \mathbb{R}$ is $\mathcal{F} \otimes \mathcal{B}(I)$-measurable and integrable with respect to $\mathbb{P} \times dt$, that is,

$$\int_{\Omega \times I} |f(\omega, t)|\, (\mathbb{P} \times dt)(d\omega, dt) < \infty,$$

then, for $\mathbb{P}$-almost every ω, the function $t \mapsto f(\omega, t)$ is integrable on I, and for almost every $t \in I$, the function $\omega \mapsto f(\omega, t)$ is integrable on Ω. Moreover, the functions

$$\omega \mapsto \int_I f(\omega, t)\, dt \qquad \text{and} \qquad t \mapsto \int_\Omega f(\omega, t)\, d\mathbb{P}(\omega)$$

are defined almost everywhere, integrable, and one has

$$\int_{\Omega \times I} f(\omega, t)\, (\mathbb{P} \times dt)(d\omega, dt) = \int_\Omega \Big(\int_I f(\omega, t)\, dt \Big)\, d\mathbb{P}(\omega) = \int_I \Big(\int_\Omega f(\omega, t)\, d\mathbb{P}(\omega) \Big)\, dt.$$

In particular, under the assumptions of Fubini's theorem, one is allowed to exchange the order of integration.

1.1.3 Independence

Another key concept in probability is *independence*. Given a family of events $(A_i)_{i\in I} \subset \mathcal{F}$, we say that they are independent if for every finite collection $A_{i_1}, \dots, A_{i_n}$ one has

$$\mathbb{P}(A_{i_1} \cap \cdots \cap A_{i_n}) = \mathbb{P}(A_{i_1}) \cdots \mathbb{P}(A_{i_n}).$$

A family of sub σ-algebras $(\mathcal{F}_i)_{i\in I} \subset \mathcal{F}$ is said to be independent if for every finite collection $\mathcal{F}_{i_1}, \dots, \mathcal{F}_{i_n}$ and every choice of events $A_{i_k} \in \mathcal{F}_{i_k}$, $k = 1, \dots, n$, one has

$$\mathbb{P}(A_{i_1} \cap \cdots \cap A_{i_n}) = \mathbb{P}(A_{i_1}) \cdots \mathbb{P}(A_{i_n}).$$

This allows one to define independence of random variables. A family of random variables $(X_i)_{i\in I}$, with $X_i : \Omega \to E_i$, is said to be independent if the family of σ-algebras $(\sigma(X_i))_{i\in I}$ is independent.

1.1.4 L^p spaces and convergence of random variables

We now introduce the notion of *expectation* or *mean* of a random variable, which plays a central role in probability theory.

Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space and let $X : \Omega \to [0, +\infty]$ be a random variable. We define its expectation as the Lebesgue integral

$$\mathbb{E}[X] := \int_\Omega X \, d\mathbb{P},$$

which, for a nonnegative random variable, is always well defined as an element of $[0, +\infty]$. More generally, if X is real valued and at least one of the quantities $\mathbb{E}[X^+]$ and $\mathbb{E}[X^-]$ is finite, one defines

$$\mathbb{E}[X] := \mathbb{E}[X^+] - \mathbb{E}[X^-],$$

where

$$X^+ := \max\{X, 0\}, \qquad X^- := \max\{-X, 0\}.$$

In this case $\mathbb{E}[X]$ is well defined in the extended real line. In particular, if $\mathbb{E}[|X|] < \infty$ we say that X is *integrable*, and then both $\mathbb{E}[X^+]$ and $\mathbb{E}[X^-]$ are finite and so it is $\mathbb{E}[X]$. Some basic properties of the expectation operator are linearity and monotonicity.

All these properties follow from the fact that the expectation is defined as a Lebesgue integral with respect to the probability measure $\mathbb{P}$. For a deeper understanding of properties of the Lebesgue integral and of the following L^p spaces, we refer the reader to standard texts in analysis and measure theory, such as Rudin [18].

The expectation of a measurable function of a random variable X, can be computed by integration with respect to the law of X, more precisely we have:

Theorem 1.1. *If $X : \Omega \to E$ is a random variable with law μ_X, then for every measurable function $f : E \to \mathbb{R}$ such that $f(X)$ is integrable, one has*

$$\mathbb{E}[f(X)] = \int_E f(x) \, \mu_X(dx).$$

For $p \geq 1$, we define the space

$$L^p(\Omega, \mathcal{F}, \mathbb{P}) := \{X : \Omega \to \mathbb{R} \text{ measurable}, \mathbb{E}[|X|^p] < \infty\};$$

where elements of $L^p(\Omega, \mathcal{F}, \mathbb{P})$ are identified up to equality almost surely. For brevity, we will sometimes write $L^p(\mathcal{F}, \mathbb{P})$ in place of $L^p(\Omega, \mathcal{F}, \mathbb{P})$ when no ambiguity arises. Even simpler we will use the notation $L^p(\Omega)$ or just L^p when the probability space is implicit. $L^p(\Omega)$ is a normed vector space with the norm

$$||X||_{L^p} = (\mathbb{E}[|X^p|])^{1/p}$$

Moreover, $L^p(\Omega)$ is a *Banach* space. We also introduce the Banach space $L^\infty(\Omega, \mathcal{F}, \mathbb{P})$, defined as the set of (equivalence classes of) essentially bounded random variables, that is,

$$L^\infty(\Omega) := \{X : \Omega \to \mathbb{R} \text{ measurable} : \exists\, M > 0 \,: |X| \leq M \quad \text{a.s.}\}.$$

The corresponding norm is given by

$$||X||_\infty := \inf\{M \geq 0 : |X| \leq M \quad \text{a.s.}\}.$$

The L^p spaces satisfy the inclusions

$$L^\infty \subset L^p \subset L^q \subset L^1, \qquad 1 \leq q \leq p \leq \infty,$$

and these inclusions are continuous with

$$||X||_{L^q} \leq ||X||_{L_p}$$

for all X in the corresponding space.

These spaces play a fundamental role in probability theory. In particular, L^1 is the space of integrable random variables, while L^2 is a real Hilbert space with inner product

$$\langle X, Y \rangle := \mathbb{E}[X\,Y].$$

The quantities $||X||^p_{L^p} = \mathbb{E}[|X|^p]$ are called the p-th moments of the random variable X. In particular, for $p = 2$, the 2-th moment plays a special role.

If $X \in L^2$, the *variance* of X is defined by

$$\mathrm{Var}(X) := \mathbb{E}\big[(X - \mathbb{E}[X])^2\big],$$

and measures the dispersion of X around its mean. It can also be written as

$$\mathrm{Var}(X) = \mathbb{E}[X^2] - (\mathbb{E}[X])^2.$$

We now introduce the notion of covariance, which measures the correlation between two square-integrable random variables. Let $X, Y \in L^2(\Omega, \mathcal{F}, \mathbb{P})$. The covariance of X and Y is defined by

$$\mathrm{Cov}(X, Y) := \mathbb{E}\big[(X - \mathbb{E}[X])(Y - \mathbb{E}[Y])\big].$$

Equivalently,

$$\mathrm{Cov}(X,Y) = \mathbb{E}[XY] - \mathbb{E}[X]\,\mathbb{E}[Y].$$

In particular, the variance of a square-integrable random variable X can be written as

$$\mathrm{Var}(X) = \mathrm{Cov}(X,X).$$

Moreover, if X and Y are independent, then

$$\mathrm{Cov}(X,Y) = 0.$$

The converse does not hold in general, but it is true for Gaussian random variables.

Convergence in the L^p spaces is defined in terms of the corresponding norms. For $1 \leq p < \infty$, sequence (X_n) converges to X in L^p if

$$||X_n - X||_{L^p} \to 0,$$

that is,

$$\mathbb{E}[|X_n - X|^p] \to 0.$$

For $p = \infty$, convergence in L^∞ means that

$$||X_n - X||_\infty \to 0.$$

We also recall that convergence in L^p implies convergence in probability, which we recall now: A sequence of random variables (X_n) is said to converge in probability to a random variable X if for every $\varepsilon > 0$,

$$\mathbb{P}(|X_n - X| > \varepsilon) \to 0 \quad \text{as } n \to \infty.$$

We shall denote convergence in probability of a sequence $(X_n)_{n\geq 1}$ to a random variable X by

$$X_n \xrightarrow{\mathbb{P}} X.$$

Finally, we recall the notion of almost sure convergence. A sequence of random variables (X_n) is said to converge almost surely to a random variable X if

$$\mathbb{P}\big(\{\omega \in \Omega : X_n(\omega) \to X(\omega)\}\big) = 1.$$

We will often write

$$X_n \to X \quad \text{a.s..}$$

In other words, $X_n(\omega)$ converges to $X(\omega)$ for all ω outside a $\mathbb{P}-$null set. This type of convergence is also stronger than convergence in probability. Now come some known results about convergence of variables.

Theorem 1.2 (Monotone convergence)**.** *If X is a random variable and (X_n) is a sequence of random variables taking values in $[0,\infty]$ such that $X_n \to X$ almost surely and $X_n \leq X_{n+1}$ for all n, then*

$$\mathbb{E}[X_n] \to \mathbb{E}[X].$$

Theorem 1.3 (Dominated convergence)**.** *Let* (X_n) *be a sequence of random variables taking values in* $[0, \infty]$ *such that* $X_n \to X$ *almost surely, and assume that there exists an integrable random variable* Y *such that* $|X_n| \leq Y$ *almost surely for all* n*. Then,*

$$\mathbb{E}[X_n] \to \mathbb{E}[X].$$

Lemma 1.1 (Fatou)**.** *Let* (X_n) *be a sequence of non-negative random variables. Then*

$$\mathbb{E}\big[\liminf_{n\to\infty} X_n\big] \leq \liminf_{n\to\infty} \mathbb{E}[X_n].$$

These results represent the basic tools to pass to the limit under the expectation sign and will be widely used throughout all the following.

We will also make use of Scheffé's lemma, which provides a useful criterion for convergence in L^1. As we said above, convergence in L^p implies convergence in probability, while almost sure convergence implies convergence in probability as well. Scheffé's lemma gives a partial converse in the L^1 case: it allows one to pass from convergence in probability (or almost sure convergence, which is easier to check) to convergence in L^1, provided that the expectations of the absolute values converge. The precise statement is recalled below.

Lemma 1.2 (Scheffe)**.** *Let* (X_n) *be a sequence of random variables such that* $X_n \in L^1(\Omega, \mathcal{F}, P)$*, and let* $X \in L^1(\Omega, \mathcal{F}, P)$*. Suppose that:*

1. $X_n \to X$ *in probability,*
2. $E[|X_n|] \to E[|X|]$.

Then:

$$E[|X_n - X|] \to 0, \quad \textit{that is, } X_n \to X \textit{ in } L^1.$$

We conclude this part by recalling a useful convergence result, usually referred to in a more general form as *Vitali*'s theorem. The general statement is formulated in terms of uniform integrability, a notion that we shall not introduce here. For our purposes, it is enough to introduce the following result, where we use a sufficient condition for uniform integrability, that is, boundedness in L^p for some $p > 1$, which is often easier to verify in applications.

Let $(X_n)_{n\geq 1}$ be a sequence of random variables and let X be a random variable. Assume that

$$X_n \to X \qquad \text{in probability,}$$

and that there exists $p > 1$ such that

$$\sup_{n\geq 1} \mathbb{E}[|X_n|^p] < \infty.$$

Then $X \in L^1(\Omega, \mathcal{F}, \mathbb{P})$ and

$$\mathbb{E}[|X_n - X|] \to 0,$$

that is, $X_n \to X$ in L^1.

This criterion will be used as a convenient substitute for the full Vitali theorem. We refer to standard texts in probability such as Billingsley [5] and Karatzas and

Shreve [13] for the general formulation in terms of uniform integrability. We conclude by recalling some classical inequalities that will be used throughout the thesis.

Let X, Y be random variables and let $p, q > 1$ such that $\frac{1}{p} + \frac{1}{q} = 1$. *Hölder*'s inequality states that

$$\mathbb{E}[|XY|] \leq ||X||_{L_p} \, ||Y||_{L_q}.$$

As a particular case, when $p = q = 2$, one obtains the *Cauchy–Schwarz* inequality

$$\mathbb{E}[|XY|] \leq \big(\mathbb{E}[X^2]\big)^{1/2} \big(\mathbb{E}[Y^2]\big)^{1/2}.$$

Another fundamental estimate is *Chebyshev*'s inequality, which gives a precise meaning to the notion of variance as a measure of dispersion, as discussed previously. If $X \in L^2$, then for every $\varepsilon > 0$,

$$\mathbb{P}(|X - \mathbb{E}[X]| \geq \varepsilon) \leq \frac{\mathrm{Var}(X)}{\varepsilon^2}.$$

So, a small variance implies that X is concentrated around its mean. The last inequality we recall is *Jensen*'s inequality. Let X be an integrable real-valued random variable and let $\varphi : \mathbb{R} \to \mathbb{R}$ be a convex function such that $\varphi(X)$ is integrable. Then

$$\varphi(\mathbb{E}[X]) \leq \mathbb{E}[\varphi(X)].$$

1.1.5 Characteristic function

We recall that if $X = (X_1, \dots, X_n)$ is an $\mathbb{R}^n$-valued random variable, its characteristic function is

$$\varphi_X(\lambda_1, \dots, \lambda_n) := \mathbb{E}\left[\exp\left(i \sum_{j=1}^{n} \lambda_j X_j\right)\right], \qquad (\lambda_1, \dots, \lambda_n) \in \mathbb{R}^n.$$

The characteristic function φ_X is fundamental in probability because the law of X is uniquely determined by φ_X. Therefore, if two $\mathbb{R}^n$-valued random variables have the same characteristic function, then they have the same law. Moreover, we recall that if $Z \sim N(0, \sigma^2)$, then its characteristic function is given by

$$\mathbb{E}[e^{i\lambda Z}] = e^{-\frac{1}{2}\lambda^2\sigma^2}, \qquad \lambda \in \mathbb{R}.$$

Therefore, if $Y_1, \dots, Y_n$ are independent centered Gaussian random variables with variances $\sigma_1^2, \dots, \sigma_n^2$, then the characteristic function of the vector $Y = (Y_1, \dots, Y_n)$ is

$$\mathbb{E}\left[\exp\left(i \sum_{j=1}^{n} \lambda_j Y_j\right)\right] = \prod_{j=1}^{n} \mathbb{E}\left[\exp\left(i\, \lambda_j Y_j\right)\right] = \exp\left(-\frac{1}{2} \sum_{j=1}^{n} \lambda_j^2 \sigma_j^2\right).$$

It follow that, if a random vector $X = (X_1, \dots, X_n)$ has this characteristic function, then X has the same law as Y, and in particular its components are independent centered Gaussian random variables with variances $\sigma_1^2, \dots, \sigma_n^2$.

This will be used in the sequel in the proof of Girsanov's theorem.

1.1.6 Conditional expectation

Theorem 1.4 (Kolmogorov). *Let $X \in L^1(\Omega)$ be a random variable defined on the probability space $(\Omega, \mathcal{F}, \mathbb{P})$ and let $\mathcal{G} \subset F$ be a sub $\sigma-algebra$ of $\mathcal{F}$. Then there exists a (a.s.) unique random variable Z such that:*

(a) Z *is* $\mathcal{G} - measurable$,

(b) $Z \in L^1(\Omega)$,

(c) $\mathbb{E}[Z \mathbb{1}_G] = \mathbb{E}[X \mathbb{1}_G]$ *for every* $G \in \mathcal{G}$.

Definition 1.3. *With the same notation of the previous theorem Z is called a version of the conditional expectation of X given $\mathcal{G}$.*

We recall now some properties of the conditional expectation. Let $X \in L^1(\Omega)$.

1. **Linearity:** Let $Y \in L^1(\Omega)$, then
$$\mathbb{E}[aX + bY \mid \mathcal{G}] = a\,\mathbb{E}[X \mid \mathcal{G}] + b\,\mathbb{E}[Y \mid \mathcal{G}] \quad \text{a.s.,} \quad a, b \in \mathbb{R}.$$

2. **"Taking out what is measurable":**
If Y is $\mathcal{G}$-measurable and $XY \in L^1(\Omega)$, then
$$\mathbb{E}[XY \mid \mathcal{G}] = Y\,\mathbb{E}[X \mid \mathcal{G}] \quad \text{a.s.}$$

3. **Taking expectation removes conditioning:**
$$\mathbb{E}[\mathbb{E}[X \mid \mathcal{G}]] = \mathbb{E}[X] \quad \text{a.s.}$$

4. **Tower Property:**
If $\mathcal{H} \subseteq \mathcal{G}$, then
$$\mathbb{E}[\mathbb{E}[X \mid \mathcal{G}] \mid \mathcal{H}] = \mathbb{E}[X \mid \mathcal{H}] \quad \text{a.s..}$$

5. **Independence:**
If X is independent of $\mathcal{G}$ (i.e. the σ-algebra generated by X is independent of $\mathcal{G}$), then
$$\mathbb{E}[X \mid \mathcal{G}] = \mathbb{E}[X] \quad \text{a.s.}$$

6. **Monotonicity:**
If $Y \in L^1(\Omega)$ and $X \leq Y$ almost surely, then
$$\mathbb{E}[X \mid \mathcal{G}] \leq \mathbb{E}[Y \mid \mathcal{G}] \quad \text{a.s.}$$

7. **Jensen's Inequality:**
If φ is convex and $\mathbb{E}[|\varphi(X)|] < \infty$, then
$$\varphi(\mathbb{E}[X \mid \mathcal{G}]) \leq \mathbb{E}[\varphi(X) \mid \mathcal{G}] \quad \text{a.s.}$$

We now recall the notion of projection on a closed subspace H_1 of a Hilbert space H.

Remark 1.1. *The orthogonal projection operator* $\pi_{\mathcal{G}} : H \to H_1$ *maps every element* ξ *of the Hilbert space* H *to the unique element of the closed subspace* H_1 *that minimizes the* L^2 *distance from* ξ*. For further details see [8] theorem 5.2.*

Theorem 1.5. *Let* $\mathcal{G}$ *be a* σ*-algebra,* $\mathcal{G} \subset \mathcal{F}$*. Denote* $H = L^2(\mathcal{F}, \mathbb{P})$*,* $H_1 = L^2(\mathcal{G}, \mathbb{P})$*, and let* $\pi_{\mathcal{G}}$ *be the orthogonal projection operator of* H *on* H_1*. Then, for each random variable* ξ *with* $\mathbb{E}\xi^2 < \infty$*, we have* $\mathbb{E}(\xi|\mathcal{G}) = \pi_{\mathcal{G}}\xi$ *(a.s.). In particular,*

$$\mathbb{E}(\xi - \pi_{\mathcal{G}}\xi)^2 = \inf\{\mathbb{E}(\xi - \eta)^2 : \eta \textit{ is } \mathcal{G}\textit{-measurable}, \eta \in H_1\}.$$

For a proof of Theorem 1.5 see [14] page 76.

We conclude by recalling the conditional versions of the main convergence theorems. Let $\mathcal{G} \subset \mathcal{F}$ be a sub-σ-algebra.

Fatou Let (X_n) be a sequence of non-negative random variables, with $X_n \in L^1$ for all n. Then

$$\mathbb{E}\big[\liminf_{n\to\infty} X_n \big| \mathcal{G}\big] \le \liminf_{n\to\infty} \mathbb{E}[X_n \mid \mathcal{G}] \quad \text{almost surely.}$$

Monotone convergence Let (X_n) be a sequence of non-negative random variables such that $X_n \uparrow X$ almost surely, with $X \in L^1(\Omega)$. Then

$$\mathbb{E}[X_n \mid \mathcal{G}] \to \mathbb{E}[X \mid \mathcal{G}] \quad \text{almost surely.}$$

Dominated convergence Let (X_n) be a sequence of random variables such that $X_n \to X$ almost surely, and assume that there exists an integrable random variable Y such that $|X_n| \le Y$ almost surely for all n. Then

$$\mathbb{E}[X_n \mid \mathcal{G}] \to \mathbb{E}[X \mid \mathcal{G}] \quad \text{almost surely.}$$

1.2 Stochastic processes

We begin by recalling some basic definitions concerning stochastic processes, filtrations, and adaptedness. These concepts describe how random quantities evolve over time and formalize the accumulation of information in a probabilistic system. Intuitively, a stochastic process represents the random evolution of a system; the filtration encodes the information available up to a given time and the condition of adaptedness ensures that the process depends only on the past and present, not on future events.

Let $\mathcal{I} \subset [0, \infty)$ be an index set. For our purpose, it will be either $\mathcal{I} = [0, \infty)$ or $\mathcal{I} = [0, T]$ for some $T > 0$.

Definition 1.4. $(\mathcal{F}_t)_{t\in\mathcal{I}}$ *is called a filtration if for every* $t \in \mathcal{I}$ $\mathcal{F}_t$ *is a sub* $\sigma -$ *algebra of* $\mathcal{F}$ *and* $\mathcal{F}_s \subseteq \mathcal{F}_t$ *for all* $s \le t$*.*

Definition 1.5. *Let* $(\Omega, \mathcal{F}, \mathbb{P})$ *be a complete probability space, and let* $(\mathcal{F}_t)_{t\ge 0}$ *be a filtration.*

The completed filtration $(\mathcal{F}_t^{\mathbb{P}})$ *is defined, for each* $t \ge 0$*, by*

$$\mathcal{F}_t^{\mathbb{P}} := \sigma\left(\mathcal{F}_t \cup \{N \in \mathcal{F} : \ \mathbb{P}(N) = 0\}\right),$$

that is, $\mathcal{F}_t^{\mathbb{P}}$ *is the* σ*-algebra generated by* $\mathcal{F}_t$ *together with all* $\mathbb{P}$*-null sets in* $\mathcal{F}$*.*

This is a key concept, in fact we will assume completion of our filtration in order to deal with *stochastic integrals.*

Definition 1.6. *A family of random variables* $X = (X_t)_{t\in\mathcal{I}}$ *defined on a probability space* $(\Omega, \mathcal{F}, \mathbb{P})$ *taking values in a measurable space* S *is called a* stochastic process *with values in* S*.*

Definition 1.7. *The quadruple* $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\in\mathcal{I}}, \mathbb{P})$ *with* $\mathcal{F}_t$ *complete is called a stochastic basis*

Definition 1.8. *A stochastic process* X *is said to be adapted to a filtration* $(\mathcal{F}_t)_{t\in\mathcal{I}}$ *if* X_t *is* $\mathcal{F}_t -$ *measurable for all* $t \in \mathcal{I}$*.*

We will write that X is $\mathcal{F}_t-$adapted

Definition 1.9. *A process* X *(real or complex valued) is bounded if it exists a constant* $K > 0$ *such that*

$$|X_t(\omega)| \leq K, \; for\ all\, (\omega, t)\, \mathbb{P} \times dt - \; a.e.$$

Definition 1.10. *Let* $X = (X_t)_{t\in\mathcal{I}}$ *be a stochastic process and for each* $t \in \mathcal{I}$ *let* $\mathcal{F}_t = \sigma\{X_s : s \leq t\}$ *be the* $\sigma - algebra$ *generated by the random variables* $\{X_s : s \leq t\}$ *(i.e. the smallest* $\sigma - algebra$ *making these functions measurable). Then,* $\mathcal{F}_t$ *is called the* $natural\ filtration$ *of the process* X*.*

Remark 1.2. *By definition every process is adapted to its natural filtration. In what follows, we shall work with the completed natural filtration.*

Definition 1.11. *Let* $(\Omega, \mathcal{F}, \mathbb{P})$ *be a probability space and* $X = (X_t)_{t\in\mathcal{I}}$ *a stochastic process with values in a topological space. For each* $\omega \in \Omega$*, the function*

$$t \longmapsto X_t(\omega), \qquad t \in \mathcal{I},$$

is called a sample path *(or* trajectory*) of* X*. We say that* X *has* continuous *sample paths if* $t \mapsto X_t(\omega)$ *is continuous on* $\mathcal{I}$ *for* $\mathbb{P}$*–almost every* ω*. In that case we also say that the process* X *is continuous.*

Definition 1.12. *Let* $X = (X_t)_{t\in\mathcal{I}}$ *and* $Y = (Y_t)_{t\in\mathcal{I}}$ *be two stochastic processes defined on the same probability space. We say that* Y *is a* modification *of* X *if, for every* $t \in I$*,*

$$\mathbb{P}(X_t = Y_t) = 1.$$

Equivalently, this means $X_t = Y_t$ *a.s. for every* $t \in \mathcal{I}$*.*

In general, two modifications do not need to have the same paths. If instead

$$\mathbb{P}\{\omega \in \Omega : X_t(\omega) = Y_t(\omega), \quad \forall\, t \in \mathcal{I}\} = 1,$$

then the two processes are said to be *indistinguishable.*

Definition 1.13. *Let* $(\Omega, \mathcal{F}, \mathbb{P})$ *be a probability space and* $X = (X_t)_{t\in\mathcal{I}}$ *an* $\mathbb{R}^d$*–valued stochastic process. We say that* X *is* measurable *if for every* $A \in \mathcal{B}(\mathbb{R}^d)$*, the set*

$$\{(t, \omega) \in \mathcal{I} \times \Omega : \ X_t(\omega) \in A\}$$

belongs to the product σ*–algebra* $\mathcal{B}(\mathcal{I}) \otimes \mathcal{F}$*; equivalently, the mapping*

$$(t, \omega) \longmapsto X_t(\omega) : \big(\mathcal{I} \times \Omega, \ \mathcal{B}(\mathcal{I}) \otimes \mathcal{F}\big) \ \longrightarrow \ \big(\mathbb{R}^d, \ \mathcal{B}(\mathbb{R}^d)\big)$$

is measurable.

Definition 1.14. *Let* $(\mathcal{F}_t)_{t\in\mathcal{I}}$ *be a filtration on* $(\Omega, \mathcal{F}, \mathbb{P})$*. An* $\mathbb{R}^d$*–valued process* $X = (X_t)_{t\in\mathcal{I}}$ *is* progressively measurable *if, for every* $t \in \mathcal{I}$*, the mapping*

$$(s, \omega) \longmapsto X_s(\omega), \qquad (s, \omega) \in [0, t] \times \Omega,$$

is $\mathcal{B}([0, t]) \otimes \mathcal{F}_t$*–measurable.*

Remark 1.3. *Since* $\mathcal{F}_t \subseteq \mathcal{F}$*, it follows that the restriction to* $[0, n] \times \Omega$ *is* $\mathcal{B}([0, n]) \otimes \mathcal{F}$*–measurable for every* $n \in \mathbb{N}$*. For any open* $G \subset \mathbb{R}^d$ *we have*

$$\{(t, \omega) : X_t(\omega) \in G\} = \bigcup_{n=1}^{\infty} \Big(\{(t, \omega) \in [0, n] \times \Omega : X_t(\omega) \in G\}\Big),$$

a countable union of $\mathcal{B}([0, n]) \otimes \mathcal{F}$*–measurable sets; hence* $(t, \omega) \mapsto X_t(\omega)$ *is* $\mathcal{B}(\mathcal{I}) \otimes \mathcal{F}$*–measurable. Thus every progressively measurable process is (jointly) measurable.*

Proposition 1.1. *Let* $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\in\mathcal{I}}, \mathbb{P})$ *be a stochastic basis. If an* $\mathbb{R}^d$*- valued process* $X = (X_t)_{t\in\mathcal{I}}$ *is* $\mathcal{F}_t$*–adapted and has (a.s.) continuous sample paths, then* X *is progressively measurable with respect to* $(\mathcal{F}_t)_{t\in\mathcal{I}}$*. In particular,* X *is (jointly) measurable.*

The reader is referred to [13] Proposition 1.13 for the proof.

We now recall the notion of *stopping time.*

Let $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\in\mathcal{I}}, \mathbb{P})$ be a filtered probability space, where $\mathcal{I} = [0, \infty)$ or $\mathcal{I} = [0, T]$ for some $T > 0$. A map

$$\tau : \Omega \to [0, \infty]$$

is called a stopping time with respect to the filtration $(\mathcal{F}_t)$ if, for every $t \in \mathcal{I}$,

$$\{\tau \le t\} \in \mathcal{F}_t.$$

In other words, the event that the random time τ has occurred by time t must be determined by the information available up to time t.

Two basic examples are given by deterministic times and by first hitting times of suitable adapted processes. In particular, every constant time $\tau(\omega) \equiv t_0$ is a stopping time, and if $X = (X_t)_{t\in\mathcal{I}}$ is an adapted process with continuous sample paths, then the first exit time from an open set is a stopping time. We will use and recall this property later on. More precisely, we state the following lemma.

Lemma 1.3. *Let ξ_t be an $\mathcal{F}_t$-adapted continuous real-valued process, and take real numbers $a < b$. Define*

$$\tau = \inf\{t \geq 0 : \xi_t \notin (a,b)\} \quad (\inf \emptyset := \infty)$$

so that τ is the first exit time of ξ_t from (a,b). Then τ is an $\mathcal{F}_t$-stopping time.

The proof is standard and the reader can find it in Baldi [3], see Chapter 1.

Stopping times play a fundamental role in the theory of martingales and stochastic integration. We refer to standard texts such as Karatzas and Shreve [13] and Baldi [3] for proofs and further results.

1.3 Martingales

Definition 1.15. *Let $X = (X_t)_{t\in\mathcal{I}}$ be a real valued process defined on the stochastic basis $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\in\mathcal{I}}, \mathbb{P})$ with $X_t \in L^1(\Omega)$ for all $t \in \mathcal{I}$. X is a martingale (respectively supermartingale, submartingale) if:*

$$\mathbb{E}[X_t|\mathcal{F}_s] = X_s \ \ (\textit{respectively} \leq X_s\,, \geq X_s),$$

for all $s < t$. We will often write $(X_t, \mathcal{F}_t)$ is a martingale, $t \in \mathcal{I}$.

By applying the properties of the conditional expectation it is easy to see that a martingale is a process with constant mean, a supermartingale has non–increasing mean and a submartingale has non–decreasing mean.

Proposition 1.2. *Let X be a supermartingale with constant mean. Then X is a martingale*

Proof. Since X is a supermartingale, for $s < t$:

$$X_s \geq \mathbb{E}[X_t|\mathcal{F}_s],$$

And so

$$\mathbb{E}[X_s - X_t|\mathcal{F}_s] \geq 0.$$

On the other hand by property (2) of the conditional expectation and since we are assuming X having constant mean:

$$\mathbb{E}[X_s - X_t] = \mathbb{E}[\mathbb{E}[X_s - X_t|\mathcal{F}_s]] = 0.$$

This means that $\mathbb{E}[X_s - X_t|\mathcal{F}_s] = 0$ (a.s.), since it is a non negative random variable with zero mean, and it is equivalent to the definition of martingale.

□

Definition 1.16. *Let $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, \mathbb{P})$ be a stochastic basis. A real-valued adapted process $(X_t)_{t\in\mathcal{I}}$ is called a* local martingale *if there exists an increasing sequence of stopping times $(\tau_n)_{n\in\mathbb{N}}$ such that $\tau_n \uparrow \infty$ almost surely and, for every $n \in \mathbb{N}$, the stopped process $(X_{t\wedge\tau_n})_{t\geq 0}$ is a martingale.*

Chapter 2

The Wiener process and the Itô integral

In this chapter we introduce the Wiener process and the Itô stochastic integral. We begin by recalling the definition and the main properties of the Wiener process. We then prove a lemma based on the reflection principle, which will allow us to derive and extend Bachelier's theorem. We do not present complete proofs in this chapter, we will often give sketches of proofs.

The second part of the chapter is devoted to stochastic integration. In particular, we follow the approach of Krylov [14], which differs from the more classical treatments, such as those in Baldi [3] and Karatzas–Shreve [13]. In those references, the Itô integral is introduced starting from progressively measurable processes, whereas Krylov's construction does not require progressive measurability and is indeed more general. We will present this approach and recall the main properties of the resulting stochastic integral.

We now introduce the Brownian motion.

Robert Brown, an English botanist, observed (1828) that pollen grains suspended in water perform an unending chaotic motion. L. Bachelier (1900) derived the law governing the position W_t at time t of a single grain performing a one-dimensional Brownian motion starting at $a \in \mathbb{R}$ at time $t = 0$:

$$\mathbb{P}_a\{W_t \in dx\} = p(t, a, x)\, dx,$$

where

$$p(t, a, x) = \frac{1}{\sqrt{2\pi t}} e^{-(x-a)^2/(2t)}$$

is the fundamental solution of the heat equation.

Remark 2.1. *When we write $\mathbb{P}_a\{W_t \in dx\} = p(t, a, x)\, dx$ we mean, equivalently, that the random variable W_t has distribution function*

$$F_{W_t}(x) = \mathbb{P}_a(W_t \le x) = \int_{-\infty}^{x} p(t, a, y)\, dy,$$

where $p(t,a,x)$ *is the density of the normal distribution* $\mathcal{N}(a,t)$ *with mean* a *and variance* t*. We also write* $W_t \sim \mathcal{N}(a,t)$*.*

The simplest model describing movement of a particle subject to hits by much smaller particles is the following. Let η_k, $k = 1, 2, \ldots$, be independent identically distributed random variables with $\mathbb{E}\eta_k = 0$ and $\mathbb{E}\eta_k^2 = 1$. Fix an integer n, and at times $1/n, 2/n, \ldots$ let our particle experience instant displacements by $\eta_1 n^{-1/2}$, $\eta_2 n^{-1/2}, \ldots$. At moment zero let our particle be at zero. If

$$S_k := \eta_1 + \ldots + \eta_k,$$

then at moment k/n our particle will be at the point $S_k/\sqrt{n}$ and will stay there during the time interval $[k/n, (k+1)/n)$. Since real Brownian motion has continuous paths, we replace our piecewise constant trajectory by a continuous piecewise linear one preserving its positions at times k/n. Thus we come to the process

$$\xi_t^n := \frac{S_{\lfloor nt \rfloor}}{\sqrt{n}} + (nt - \lfloor nt \rfloor)\frac{\eta_{\lfloor nt \rfloor + 1}}{\sqrt{n}}. \tag{2.1}$$

This process gives a very rough approximation of the Brownian motion. Clearly, to get the right model we have to let $n \to \infty$.

2.1 The Wiener process

2.1.1 The space $C[0,T]$, weak convergence and the Wiener measure

Let $T > 0$. We denote by

$$C[0,T] = C([0,T];\mathbb{R}) := \{x : [0,T] \to \mathbb{R} \text{ continuous}\}$$

the space of real-valued continuous functions on $[0,T]$, which is a separable Banach space with the supremum norm

$$\|x\|_\infty := \sup_{0 \le t \le T} |x(t)|.$$

We equip $C[0,T]$ with its Borel σ-algebra

$$\mathcal{B}(C[0,T])$$

i.e. the σ-algebra generated by the open sets of the metric induced by $\|\cdot\|_\infty$.

For our purpose we need to clarify what it means that a stochastic process weakly converges on $C[0,T]$. To do this, we need to introduce first the general notion of weak convergence of random variables.

Definition 2.1. *Let* (S,d) *be a metric space. A sequence of* S*-valued random variables* (X_n) *is said to converge in distribution (or weakly) to an* S*-valued random variable* X*, and we write*

$$X_n \Rightarrow X,$$

if

$$\mathbb{E}[f(X_n)] \longrightarrow \mathbb{E}[f(X)]$$

for every bounded continuous function $f : S \to \mathbb{R}$*.*

Remark 2.2. *According to [14] Theorem 1.4.2, a stochastic process with continuous sample paths on* $[0,T]$ *(see Definition 1.11) can be identified with a random variable taking values in the space* $C[0,T]$*. In fact if* $X = (X_t)_{t\in[0,T]}$ *has continuous sample paths, then for every* $\omega \in \Omega$ *the map*

$$t \mapsto X_t(\omega)$$

belongs to $C[0,T]$*. Hence* X *can be identified with the* $C[0,T]$*-valued random variable*

$$\Omega \ni \omega \mapsto X(\omega) := \big(t \mapsto X_t(\omega)\big) \in C[0,T].$$

Considering this identification we are now able to talk about weak convergence of a process and also define the notion of law of a continuous process.

Remark 2.3. *Since a stochastic process with continuous sample paths on* $[0,T]$ *can be identified with a random variable taking values in the space* $C[0,T]$ *(see the previous remark), in particular, weak convergence of continuous processes is understood as convergence in distribution of the associated* $C[0,T]$*-valued random variables.*

Definition 2.2. *Let* $(X_t)_{t\in[0,T]}$ *be a continuous process. The law of* X *is the induced probability measure on* $\big(C[0,T], \mathcal{B}(C[0,T])\big)$.

Definition 2.3. *For every* $n \geq 1$ *and every choice of times* $0 \leq t_1 < \cdots < t_n \leq T$*, the law of the random vector*

$$(X_{t_1}, \ldots, X_{t_n})$$

is called a finite-dimensional distribution of the process X*.*

For continuous processes, their law on $C[0,T]$ is determined by the finite-dimensional distributions. In particular, two continuous processes have the same law on $C[0,T]$ if their finite-dimensional distributions coincide. For more details the reader is referred to Karatzas and Shreve [13], Section 2.4 C–D.

Proposition 2.1. *Let* X *and* Y *be real-valued stochastic processes with continuous sample paths on* $[0,T]$*. Then the following are equivalent:*

(i) X *and* Y *have the same finite-dimensional distributions;*

(ii) the induced probability measures of X *and* Y *on* $\big(C[0,T], \mathcal{B}(C[0,T])\big)$ *coincide.*

Theorem 2.1 (Portmanteau)**.** *Let* (S,d) *be a separable (i.e. it contains a countable dense subset) metric space and let* X_n *and* X *be* S*-valued random variables. Then, the following are equivalent:*

1. $X_n \Rightarrow X$
2. *For every Borel set* $A \subset S$ *such that*

   $$\mathbb{P}(X \in \partial A) = 0,$$

   one has

   $$\mathbb{P}(X_n \in A) \longrightarrow \mathbb{P}(X \in A).$$

The reader can find this result in [6], see Chapter 1, Theorem 2.1.

Finally, to formally introduce the Wiener process we recall the notion of multidimensional Gaussian distributions.

Definition 2.4 (Multivariate normal distribution). *A random vector* $X = (X_1, \dots, X_n)^\top$ *in* $\mathbb{R}^n$ *is said to have a* multivariate normal distribution *with mean vector* $m = (m_1, \dots, m_n)^\top$ *and covariance matrix* $R = (R_{ij})_{1 \le i,j \le n}$*, denoted*

$$X \sim \mathcal{N}(m, R),$$

if for every $a \in \mathbb{R}^n$*, the linear combination* $a^\top X$ *has a one–dimensional normal distribution with mean* $a^\top m$ *and variance* $a^\top Ra$*.*

Definition 2.5 (Gaussian process). *Let* $(\Omega, \mathcal{F}, \mathbb{P})$ *be a probability space. A stochastic process* $X = \{X_t\}_{t \in [0,T]}$ *taking values in* $\mathbb{R}$ *is called a* Gaussian process *if for every finite collection of times* $t_1, \dots, t_n$*, the random vector*

$$(X_{t_1}, \dots, X_{t_n})$$

has a multivariate normal (Gaussian) distribution.

In this case, the finite-dimensional distributions of the process are completely determined by its mean function

$$m(t) = \mathbb{E}[X_t],$$

and its covariance function

$$R(s,t) = \mathbb{E}[(X_s - m(s))(X_t - m(t))], \qquad s, t \in [0,T].$$

Definition 2.6. *By a Wiener process or Brownian motion we mean a continuous Gaussian process*

$$W = (W_t)_{t \in [0,T]}$$

with parameters $m_t = 0$ *and* $R(s,t) = s \wedge t$, $s, t \in [0,T]$*.*

A Wiener process is a continuous Gaussian process with mean function $m(t) = 0$ and covariance function $R(s,t) = s \wedge t$. Since the finite-dimensional distributions of a Gaussian process are completely determined by its mean and covariance functions (see Definition 2.5), it follows that all Wiener processes share the same finite-dimensional distributions.

Moreover, as observed in Remark 2.2, any stochastic process with continuous sample paths on $[0,T]$ can be identified with a random variable taking values in $C[0,T]$. In particular, for continuous processes, equality of finite-dimensional distributions is equivalent to equality of the induced laws on $(C[0,T], \mathcal{B}(C[0,T]))$ (see Proposition 2.1). Consequently, all Wiener processes induce the same probability measure on $(C[0,T], \mathcal{B}(C[0,T]))$.

Definition 2.7 (Wiener measure). *The* Wiener measure *on* $(C[0,T], \mathcal{B}(C[0,T]))$ *is the probability measure induced by (equivalently, the law on* $C[0,T]$ *of) any Wiener process on* $[0,T]$*. We denote it by* $\mathbb{W}$*.*

Theorem 2.2 (Donsker). *Let $T > 0$ and let $(\xi_n)_{n\geq 1}$ be the sequence of stochastic processes defined in* (2.1) *on* $[0,T]$. *Then there exists a Wiener process $W = (W_t)_{t\in[0,T]}$ such that*

$$\xi_n \Rightarrow W \qquad \text{in } C[0,T].$$

A formal proof of this result can be found in [14] chapter 2.

We now state some known results about the Wiener process, the reader is always referred to [14] for the proofs.

There are two important basic criteria for a process to be a Wiener process.

Theorem 2.3. *A continuous process $W = (W_t)_{t\in[0,T]}$ is a Wiener process if and only if*

- *(i)* $W_0 = 0$ *(a.s.),*
- *(ii)* $W_t - W_s$ *is normal with parameters* $(0, |t-s|)$ *for every* $s,t \in [0,T]$,
- *(iii)* $W_{t_1}, W_{t_2} - W_{t_1}, \ldots, W_{t_n} - W_{t_{n-1}}$ *are independent for every* $n \geq 2$ *and* $0 \leq t_1 \leq t_2 \leq \ldots \leq t_n \leq T$.

Theorem 2.4. *A continuous process $(W_t)_{t\in[0,T]}$ is a Wiener process if and only if*

- *(i)* $W_0 = 0$ *(a.s.),*
- *(ii)* $W_t - W_s$ *is normal with parameters* $(0, |t-s|)$ *for every* $s,t \in [0,T]$,
- *(iii) for every* $n \geq 2$ *and* $0 \leq t_1 \leq t_2 \leq \ldots \leq t_n \leq T$, *the random variable* $W_{t_n} - W_{t_{n-1}}$ *is independent of* $W_{t_1}, W_{t_2}, \ldots, W_{t_{n-1}}$.

2.1.2 The reflection principle: deriving Bachelier's theorem

The following result is due to L. Bachelier (1900) [4]. We will derive it from a more general lemma based on the reflection principle for random walks.

Theorem 2.5 (Bachelier). *For every $t \in (0,1]$ we have $\max_{s\leq t} W_s \sim |W_t|$, which is to say that for every $x \geq 0$*

$$\mathbb{P}\left\{\max_{s\leq t} W_s \leq x\right\} = \frac{2}{\sqrt{2\pi t}} \int_0^x e^{-y^2/(2t)}\, dy.$$

We first present a proof of the reflection principle for random walks. This result is well known in the literature but it is not easy to find a proof with all details. Moreover, the usual proof uses a property of the Wiener process called strong Markov property; instead now we are giving a more elementary proof that does not use such property.

Lemma 2.1 (Reflection principle for random walks). *Let $(S_k)_{k\geq 0}$ be a simple symmetric random walk, i.e.*

$$S_0 = 0, \qquad S_k = \sum_{i=1}^{k} X_i, \qquad \mathbb{P}(X_i = 1) = \mathbb{P}(X_i = -1) = \tfrac{1}{2},$$

with $(X_i)_{i\geq 1}$ independent. Set $M_n := \max_{0\leq k\leq n} S_k$. Then for every $n \in \mathbb{N}$, every $m \in \mathbb{Z}$ and every $b \in \mathbb{Z}$ with $b \leq m$,

$$\mathbb{P}\big(M_n \geq m,\ S_n \leq b\big) = \mathbb{P}\big(S_n \geq 2m - b\big). \tag{2.2}$$

Proof. For $n \in \mathbb{N}$ consider the set $\Omega_n := \{-1, 1\}^n$. For $\omega = (x_1, \dots, x_n) \in \Omega_n$ define the random walk $S_k(\omega) := \sum_{i=1}^k x_i$ and $M_n(\omega) := \max_{0 \le k \le n} S_k(\omega)$.

Consider the sets

$$C_n := \{\omega \in \Omega_n : \; M_n(\omega) \ge m, \; S_n(\omega) \le b\}, \qquad D_n := \{\omega \in \Omega_n : \; S_n(\omega) \ge 2m - b\}.$$

We construct a bijection $\Phi : C_n \to D_n$.

Fix $\omega \in C_n$ and define

$$j = j(\omega) := \inf\{k \in \{0, 1, ..., n\} : S_k(\omega) \ge m\}.$$

j is the first time that the walk reaches the level m. Then $S_{j-1}(\omega) \le m - 1$ and $S_j(\omega) = m$ (because the walk has ± 1 steps). Define $\omega' = \Phi(\omega)$ by

$$x'_i := \begin{cases} x_i, & i \le j, \\ -x_i, & i > j. \end{cases}$$

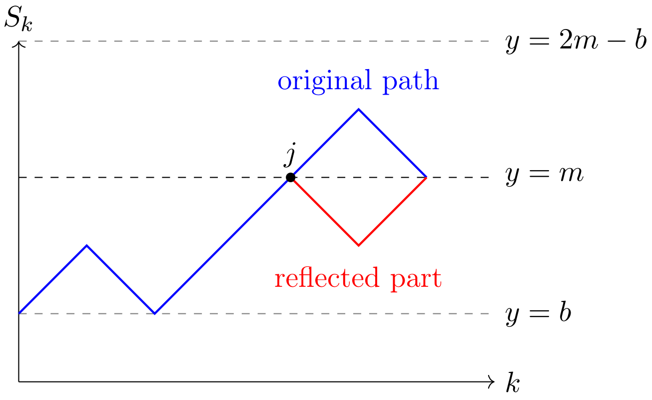
Equivalently, it is easy to see that the associated random walk satisfies

$$S_k(\omega') = \begin{cases} S_k(\omega), & k \le j, \\ 2m - S_k(\omega), & k > j. \end{cases} \tag{2.3}$$

In particular,

$$S_n(\omega') = 2m - S_n(\omega) \ge 2m - b,$$

so $\omega' \in D_n$. Hence $\Phi(C_n) \subseteq D_n$. Note that identity (2.3) means that the map Φ reflects the associated path after time j with respect to the line $y = m$, as shown in the image.



Reflection of the random walk path with respect to the line $y = m$.

Conversely, take $\omega' \in D_n$ and define $j' = j'(\omega')$ as

$$j' = \inf\{k \in \{0, 1, ..., n\} : S_k(\omega') \ge m\}$$

(j' exists because $S_n(\omega') \ge 2m - b \ge m$). Define $\omega = \Psi(\omega')$ as before by reflecting the associated random walk path after j': again ω is defined by $x_i = x'_i$ for $i \le j'$

and $x_i = -x'_i$ for $i > j'$. Then it is easy to see that $\Psi(\omega') \in C$. In fact $S_n(\omega) = 2m - S_n(\omega') \leq 2m - (2m - b) \leq b$. Obviously Ψ is the inverse map of Φ, thus Φ is a bijection between C_n and D_n.

Finally, for every $\omega \in \Omega_n$ we have $\mathbb{P}(\omega) = 2^{-n}$, and therefore

$$\mathbb{P}(C_n) = \sum_{\omega \in C_n} 2^{-n} = 2^{-n}|C_n| = 2^{-n}|D_n| = \sum_{\omega' \in D_n} 2^{-n} = \mathbb{P}(D_n),$$

which proves (2.2) □

Lemma 2.2. *Let $\{W_t\}_{t\in[0,1]}$ be a Wiener process. Then, for every $x \in \mathbb{R}$, the following identity holds:*

$$\mathbb{P}\left(\sup_{0\leq t\leq 1} W_t \geq 1,\ W_1 \leq x\right) = \begin{cases} \mathbb{P}(W_1 \geq 2 - x), & \textit{if } x \leq 1, \\ 2\mathbb{P}(W_1 \geq 1) - \mathbb{P}(W_1 \geq x), & \textit{if } x > 1. \end{cases}$$

Proof. We focus on the case $x \leq 1$ first.

Let $\{\eta_k\}_{k\geq 1}$ be a sequence of i.i.d. random variables taking values ± 1 with equal probability $1/2$. Define the symmetric random walk:

$$S_k := \sum_{j=1}^{k} \eta_j, \quad \text{and} \quad \xi_t^n := \frac{S_{\lfloor nt \rfloor}}{\sqrt{n}} + (nt - \lfloor nt \rfloor)\frac{\eta_{\lfloor nt \rfloor + 1}}{\sqrt{n}}$$

as in equation (2.1). By construction,

$$\max_{t\in[0,1]} \xi_n(t) = \max_{k\leq n} \frac{S_k}{\sqrt{n}}.$$

By Donsker's theorem, see Theorem 2.2

$$\max_{t\in[0,1]} \xi_t^n = n^{-1/2} \max_{k\leq n} S_k \longrightarrow \max_{t\in[0,1]} W_t$$

in distribution. In fact the map $F : C([0,1]) \to \mathbb{R}$, $\quad f \mapsto \max_{t\in[0,1]} f(t)$ is continuous:

$$|\max f - \max g| \leq \|f - g\|_\infty.$$

We want to study the probability:

$$\mathbb{P}\left(\max_{t\leq 1} W_t \geq 1,\ W_1 \leq x\right) = \lim_{n\to\infty} \mathbb{P}\left(\max_{k\leq n} \frac{S_k}{\sqrt{n}} \geq 1,\ \frac{S_n}{\sqrt{n}} \leq x\right). \tag{2.4}$$

This equality holds because:
By construction $\max_{t\leq 1} \xi_n(t) = \max_{k\leq n} S_k/\sqrt{n}$ and $\xi_n(1) = S_n/\sqrt{n}$, moreover since the functionals

$$f \mapsto \max_{t\leq 1} f(t) \qquad \text{and} \qquad f \mapsto f(1)$$

are continuous on $C([0,1])$ with respect to the supremum norm, thus we have

$$\left(\max_{t\leq 1} \xi_n(t),\ \xi_n(1)\right) \Rightarrow \left(\max_{t\leq 1} W_t,\ W_1\right) \quad \text{in } \mathbb{R}^2.$$

Using this fact one can prove 2.4.

Multiplying by $\sqrt{n}$:

$$\mathbb{P}\left(\max_{k\leq n} S_k \geq \sqrt{n},\ S_n \leq x\sqrt{n}\right).$$

By the reflection principle for random walks

$$\mathbb{P}\left(\max_{k\leq n} S_k \geq \sqrt{n},\ S_n \leq x\sqrt{n}\right) = \mathbb{P}(S_n \geq 2\sqrt{n} - x\sqrt{n}).$$

Dividing by $\sqrt{n}$, we get:

$$\mathbb{P}\left(\max_{k\leq n} \frac{S_k}{\sqrt{n}} \geq 1,\ \frac{S_n}{\sqrt{n}} \leq x\right) = \mathbb{P}\left(\frac{S_n}{\sqrt{n}} \geq 2 - x\right). \tag{1}$$

As $n \to \infty$, $\frac{S_n}{\sqrt{n}} \Rightarrow W_1$, thus it can be shown that

$$\lim_{n\to\infty} \mathbb{P}\left(\frac{S_n}{\sqrt{n}} \geq 2 - x\right) = \mathbb{P}(w_1 \geq 2 - x).$$

Therefore by passing to the limit $n \to \infty$ in (1),

$$\mathbb{P}\left(\max_{t\leq 1} W_t \geq 1,\ W_1 \leq x\right) = \mathbb{P}(W_1 \geq 2 - x), \quad \text{for } x \leq 1.$$

For $x \geq 1$, we use the identity

$$\mathbb{P}(A \cap B) = \mathbb{P}(A) - \mathbb{P}(A \cap B^c),$$

and write

$$\mathbb{P}(\max_{t\in[0,1]} W_t \geq 1, w_1 \leq x) = \mathbb{P}(\max_{t\in[0,1]} W_t \geq 1) - \mathbb{P}(\max_{t\in[0,1]} W_t \geq 1, W_1 > x).$$

Since $x \geq 1$,

$$\mathbb{P}\left(\max_{t\in[0,1]} W_t \geq 1,\ W_1 > x\right) = \mathbb{P}(W_1 > x).$$

To compute $\mathbb{P}(\max_{t\in[0,1]} W_t \geq 1)$, we use the previous part of the Lemma with $x = 1$, which gives

$$\mathbb{P}\left(\max_{t\in[0,1]} W_t \geq 1,\ W_1 \leq 1\right) = \mathbb{P}(W_1 \geq 1).$$

Hence

$$\mathbb{P}\left(\max_{t\in[0,1]} W_t \geq 1\right) = \mathbb{P}\left(\max_{t\in[0,1]} W_t \geq 1,\ W_1 \leq 1\right) + \mathbb{P}(\max_{t\in[0,1]} W_t \geq 1, W_1 > 1).$$

But since

$$\{W_1 > 1\} \subset \{\max_{t\in[0,1]} W_t \geq 1\},$$

the second term is equal to $\mathbb{P}(W_1 > 1)$; but since W_1 has a continuous distribution, $\mathbb{P}(W_1 > 1) = \mathbb{P}(W_1 \geq 1)$, and therefore

$$\mathbb{P}\left(\max_{t\in[0,1]} W_t \geq 1\right) = \mathbb{P}(W_1 \geq 1).$$

$\square$

Remark 2.4. *In the previous proof we have shown that*

$$\mathbb{P}\left(\sup_{0\le s\le 1} W_s \ge 1\right) = 2\,\mathbb{P}(W_1 \ge 1).$$

It can be shown that

$$\mathbb{P}\left(\sup_{0\le s\le 1} W_s = 1\right) = 0.$$

Therefore, by taking complements

$$\mathbb{P}\left(\sup_{0\le s\le 1} W_s \le 1\right) = 1 - 2\,\mathbb{P}(W_1 \ge 1),$$

Splitting

$$1 = \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-y^2/2}\,dy = \int_{-\infty}^{-1} \frac{1}{\sqrt{2\pi}} e^{-y^2/2}\,dy + \int_{-1}^{0} \frac{1}{\sqrt{2\pi}} e^{-y^2/2}\,dy + \int_{0}^{1} \frac{1}{\sqrt{2\pi}} e^{-y^2/2}\,dy + \int_{1}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-y^2/2}\,dy$$

and using the fact that the Gaussian density is even, we obtain

$$\mathbb{P}\left(\sup_{0\le s\le 1} W_s \le 1\right) = \frac{2}{\sqrt{2\pi}} \int_0^1 e^{-y^2/2}\,dy,$$

which is precisely Bachelier's formula at time 1 *and level* 1.

Our goal is to derive Bachelier's theorem in its full form, for arbitrary $t > 0$ and $x \ge 0$ and one can think about a scaling argument to derive it. The Lemma proved above, with barrier fixed at 1, only yields the formula at time 1 and level 1. This is not enough to derive the general statement by scaling alone.

In fact, the scaling argument becomes effective only once the formula is known at time 1 for an arbitrary level $a > 0$. Indeed we come to a more generalized version of the Lemma, where the barrier 1 is replaced by a generic level $a > 0$.

Lemma 2.3. *Let $\{W_t\}_{t\in[0,1]}$ be a Wiener process. Then, for every $a > 0$ and every $x \in \mathbb{R}$,*

$$\mathbb{P}\left(\sup_{0\le t\le 1} W_t \ge a,\ W_1 \le x\right) = \begin{cases} \mathbb{P}(W_1 \ge 2a - x), & \text{if } x \le a, \\ 2\mathbb{P}(W_1 \ge a) - \mathbb{P}(W_1 \ge x), & \text{if } x > a. \end{cases}$$

Proof. The proof is the same as in the previous Lemma, replacing the level 1 by the general level $a > 0$. We only indicate the point where the argument has to be modified.

Arguing exactly as before, using Donsker's theorem we are led to study, in the case $x \le a$,

$$\mathbb{P}\left(\max_{k\le n} \frac{S_k}{\sqrt{n}} \ge a,\ \frac{S_n}{\sqrt{n}} \le x\right) = \mathbb{P}\left(\max_{k\le n} S_k \ge a\sqrt{n},\ S_n \le x\sqrt{n}\right).$$

Since the random walk takes integer values, to apply reflection principle we may replace the quantities $a\sqrt{n}$ n and $x\sqrt{n}$ their integer parts. More precisely, setting

$$m_n := \lceil a\sqrt{n}\rceil, \qquad b_n := \lfloor x\sqrt{n}\rfloor,$$

we have

$$\left\{\max_{k\le n} S_k \ge a\sqrt{n}\right\} = \left\{\max_{k\le n} S_k \ge m_n\right\}, \qquad \{S_n \le x\sqrt{n}\} = \{S_n \le b_n\}.$$

Hence

$$\mathbb{P}\left(\max_{k\le n} S_k \ge a\sqrt{n},\ S_n \le x\sqrt{n}\right) = \mathbb{P}\left(\max_{k\le n} S_k \ge m_n,\ S_n \le b_n\right).$$

and therefore the reflection principle for random walks gives

$$\mathbb{P}\left(\max_{k\le n} S_k \ge m_n,\ S_n \le b_n\right) = \mathbb{P}\left(S_n \ge 2m_n - b_n\right).$$

Dividing by $\sqrt{n}$, this becomes

$$\mathbb{P}\left(\frac{S_n}{\sqrt{n}} \ge \frac{2m_n - b_n}{\sqrt{n}}\right).$$

Now

$$\frac{m_n}{\sqrt{n}} \to a, \qquad \frac{b_n}{\sqrt{n}} \to x,$$

so that

$$\frac{2m_n - b_n}{\sqrt{n}} \to 2a - x.$$

Since

$$\frac{S_n}{\sqrt{n}} \Rightarrow W_1$$

and W_1 has a continuous law, it follows that

$$\mathbb{P}\left(\sup_{0\le t\le 1} W_t \ge a,\ W_1 \le x\right) = \mathbb{P}(W_1 \ge 2a - x), \qquad x \le a.$$

The case $x > a$ is then treated exactly as in the previous Lemma: one writes

$$\mathbb{P}\left(\sup_{0\le t\le 1} W_t \ge a,\ W_1 \le x\right) = \mathbb{P}\left(\sup_{0\le t\le 1} W_t \ge a\right) - \mathbb{P}\left(\sup_{0\le t\le 1} W_t \ge a,\ W_1 > x\right),$$

uses the case $x = a$ proved above to obtain

$$\mathbb{P}\left(\sup_{0\le t\le 1} W_t \ge a\right) = 2\mathbb{P}(W_1 \ge a),$$

and finally observes that, since $x > a$,

$$\{W_1 > x\} \subset \left\{\sup_{0\le t\le 1} W_t \ge a\right\},$$

so that

$$\mathbb{P}\left(\sup_{0\le t\le 1} W_t \ge a,\ W_1 > x\right) = \mathbb{P}(W_1 > x) = \mathbb{P}(W_1 \ge x),$$

because W_1 has a continuous distribution. This gives

$$\mathbb{P}\left(\sup_{0\le t\le 1} W_t \ge a,\ W_1 \le x\right) = 2\mathbb{P}(W_1 \ge a) - \mathbb{P}(W_1 \ge x), \qquad x > a.$$

$\square$

Proof of Theorem 2.5. By applying the same argument as in Remark 2.4 to Lemma 2.3, one obtains that for every $a \ge 0$,

$$\mathbb{P}\left(\sup_{0\le s\le 1} W_s \le a\right) = \frac{2}{\sqrt{2\pi}} \int_0^a e^{-z^2/2}\, dz. \tag{2.5}$$

We now extend this identity to an arbitrary time $t > 0$ by a scaling argument.

Fix $t > 0$ and define, for $u \in [0,1]$,

$$\widetilde{W}_u := \frac{1}{\sqrt{t}} W_{tu}.$$

By applying Theorem 2.3 it is easy to check that $(\widetilde{W}_u)_{u\in[0,1]}$ is again a Wiener process on $[0,1]$. In fact, $\widetilde{W}_0 = 0$ almost surely, its sample paths are continuous, and for $0 \le u_1 \le \cdots \le u_n \le 1$ the increments

$$\widetilde{W}_{u_k} - \widetilde{W}_{u_{k-1}} = \frac{1}{\sqrt{t}}\big(W_{tu_k} - W_{tu_{k-1}}\big)$$

are independent because the corresponding increments of W are independent. Moreover, for $0 \le v \le u \le 1$,

$$\widetilde{W}_u - \widetilde{W}_v = \frac{1}{\sqrt{t}}(W_{tu} - W_{tv}) \sim N(0, u - v),$$

since $W_{tu} - W_{tv} \sim N(0, t(u - v))$.

We want to study the probability of the event

$$\{\sup_{0\le s\le t} W_s \le x\}.$$

Let $u := \frac{s}{t}$. We have

$$\left\{\sup_{0\le s\le t} W_s \le x\right\} = \left\{\sup_{0\le u\le 1} W_{tu} \le x\right\} = \left\{\sup_{0\le u\le 1} \widetilde{W}_u \le \frac{x}{\sqrt{t}}\right\}$$

Therefore, applying identity (2.5) to the process $\widetilde{W}$ with $a = \frac{x}{\sqrt{t}}$ we get, for every $x \ge 0$,

$$\mathbb{P}\left(\sup_{s\le t} W_s \le x\right) = \mathbb{P}\left(\sup_{0\le u\le 1} \widetilde{W}_u \le \frac{x}{\sqrt{t}}\right) = \frac{2}{\sqrt{2\pi}} \int_0^{x/\sqrt{t}} e^{-z^2/2}\, dz.$$

Finally, with the change of variable $y = \sqrt{t}\,z$, we obtain

$$\int_0^{x/\sqrt{t}} e^{-z^2/2}\,dz = \frac{1}{\sqrt{t}} \int_0^x e^{-y^2/(2t)}\,dy,$$

so that

$$\mathbb{P}\left(\sup_{0\le s\le t} W_s \le x\right) = \frac{2}{\sqrt{2\pi t}} \int_0^x e^{-y^2/(2t)}\,dy.$$

Since $W_t \sim N(0,t)$, the right-hand side is exactly $\mathbb{P}(|W_t| \le x)$, and this proves that

$$\sup_{0\le s\le t} W_s \sim |W_t|.$$

□

In analogy with Definition 2.6 we can define:

Definition 2.8. *$W = (W_t)_{t\ge0}$ is a Wiener process if for all $T > 0$ $(W_t)_{t\in[0,T]}$ is a Wiener process according to Definition 2.6.*

2.2 Itô's stochastic integral

Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a complete probability space, $(\mathcal{F}_t)_{t\ge0}$ a filtration and $(W_t)_{t\ge0}$ a Wiener process defined on the stochastic basis $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\ge0}, \mathbb{P})$.

Definition 2.9. *Let $f_t = f_t(\omega)$ be a real valued process defined on $\Omega \times [0,\infty)$. We say that $f \in H_0$ if there exist $0 \le t_0 \le t_1 \le ...t_n < \infty$ such that the f_{t_i} are $\mathcal{F}_{t_i}-$measurable, $Ef_{t_i}^2 < \infty$ and*

$$f_t = \begin{cases} f_{t_i} \text{ if } t \in [t_i, t_{i+1}],\ i \le n \\ 0 \text{ if } t \ge t_n \end{cases}$$

A process of class H_0 is called *elementary*. Elementary processes are the foundation for defining Itô integrals, just like the simple functions are the starting point for defining Lebesgue integral.

Definition 2.10. *By H we denote the set of all real-valued $\mathcal{F}_t$-adapted functions $f_t(\omega)$ which are $\mathcal{F} \otimes \mathcal{B}(0,\infty)$-measurable and satisfy*

$$\mathbb{E}\left(\int_0^\infty f_t^2\,dt\right) < \infty.$$

Remark 2.5. *Obviously*

$$H_0 \subset H$$

For $f \in H_0$, we define its stochastic integral as

Definition 2.11.

$$I(f) = If = \sum_{j=0}^{n-1} f_{t_j}(W_{t_j+1} - W_{t_j})$$

Remark 2.6. *Note that $I(f)$ is independent of the partition if $f \in H_0$. Also, the notation If makes sense, as I is a linear operator on H_0.*

For the proof of the following lemma, the reader is referred to [14] chapter 6, section 1.

Lemma 2.4. *Let $f \in H_0$, then $\mathbb{E}If = 0$ and*

$$\mathbb{E}(If)^2 = \mathbb{E}\int_0^\infty f_t^2 dt \quad \textit{(Isometry formula)} \tag{2.6}$$

Remark 2.7. *Identity* (2.6) *tells us that the operator I acting from H_0 as a subspace of $L^2(\mathcal{F} \otimes \mathcal{B}[0,\infty), \mathbb{P} \times dt)$ into $L^2(\mathcal{F}, \mathbb{P})$ is an isometry.*

Lemma 2.5. *For $f \in H_0$ define $I_s f := I(\mathbb{1}_{[0,s)} f)$.*
Then $(I_s f, \mathcal{F}_s)$ is a martingale $(s \geq 0)$

Proof. Without loss of generality we can suppose that $s \in \{t_o, \dots, t_n\}$, since the values of f depend just on $\{t_0, ..., t_n\}$. If $s = t_k$, then

$$\mathbb{1}_{[0,s)} f = \sum f_{t_i} \mathbb{1}_{[t_i, t_{i+1}]}(t), \quad I_s f = \sum_{i=0}^{k-1} f_{t_i}(W_{t_{i+1}} - W_{t_i}).$$

It follows that $I_s f$ is $\mathcal{F}_s-$ measurable. Furthermore, if $t \leq s$, $t, s \in \{t_0, ..., t_n\}$, $t = t_r$, $s = t_k$, then

$$\mathbb{E}[I_s f - I_t f | \mathcal{F}_t] = \sum_{i=r}^{k-1} \mathbb{E}[f_{t_i} \mathbb{E}[W_{t_{i+1}} - W_{t_i} | \mathcal{F}_{t_i}] | \mathcal{F}_t] = 0.$$

□

Remark 2.8 (Extension to $\overline{H_0}$)**.** *Lemma 2.4 allows us to follow a familiar pattern. Namely, consider H_0 as a subset of $L^2(\mathcal{F} \otimes \mathcal{B}[0,\infty), \mu)$, where $\mu(d\omega, dt) = \mathbb{P}(d\omega)dt$. On H_0 we have defined the operator I which maps H_0 isometrically to a subset of $L^2(\mathcal{F}, \mathbb{P})$. Moreover, the operator I admits a unique extension to an isometric operator acting from $\overline{H_0}$ into $L^2(\mathcal{F}, \mathbb{P})$. We keep the same notation I for the extension, and for a function $f \in \overline{H_0}$ we define its Itô stochastic integral by the formula*

$$\int_0^\infty f_t \, dW_t = If.$$

Remark 2.9. *In particular as the extension is still an isometry identity* (2.6) *is still valid for $f, g \in \overline{H_0}$*

While initially we had $H_0 \subset H$, now we are showing that $H \subseteq \overline{H_0}$,

Theorem 2.6.

$$H \subseteq \overline{H_0},$$

where $\overline{H_0}$ is the closure of H_0 in $L^2(\Omega \times [0,\infty))$.

Since this result is not widely known and typically appears in the literature with additional assumptions, we provide a full proof.

Proof. It suffices to prove that $f \in \overline{H_0}$ for $f \in H$ such that $f_t(\omega) = 0$ for $t \geq T$, where T is a constant. In fact

$$\mathbb{E} \int_0^\infty |f_t - f_t \mathbb{1}_{t \leq n}|^2 dt = \mathbb{E} \int_n^\infty f_t^2 dt \to 0$$

as $n \to \infty$ for the dominated convergence theorem: in fact if $f_n = \int_n^{+\infty} f_t^2 \, dt$, then $f_n \to 0$ $\omega - a.s.$, and $|f_n| \leq \int_0^\infty f_t^2 \, dt \in L^1(\Omega)$, since $f \in L^2(\Omega \times [0, \infty))$. So, if $f_t \mathbb{1}_{t \leq n} \in L^2(\Omega \times [0, \infty))$, then $f_t \in L^2(\Omega \times [0, \infty))$ due to the completeness of the space $L^2(\Omega \times [0, \infty))$.

Therefore we fix $f \in H$ and assume that $f_t = 0$ for $t \geq T$. It is convenient to assume that f_t is defined for negative t as well, and $f_t = 0$ for $t \leq 0$. Now we recall that it is known from integration theory that every $L^2(\mathbb{R})$ function is continuous in $L^2(\mathbb{R})$. A proof of this property can be found in [18] chapter 9. More precisely, if $h \in L^2([0, T])$ and $h(t) = 0$ outside $[0, T]$, then

$$\lim_{a \to 0} \int_{-T}^T |h_{t+a} - h_t|^2 dt = 0. \tag{2.7}$$

(2.7) implies

$$\lim_{a \to 0} \mathbb{E} \int_{-T}^T |f_{t+a} - f_t|^2 dt = 0. \tag{2.8}$$

We prove now equation (2.8) using the dominated convergence theorem. For a.e. ω the function $t \mapsto f_t(\omega) \in L^2([-T, T])$, so

$$\lim_{a \to 0} \int_{-T}^T |f_{t+a} - f_t|^2 dt = 0$$

follows from equation (2.7). Moreover,

$$\int_{-T}^T |f_{t+a} - f_t|^2 \, dt \leq 2 \left(\int_{-T}^T f_{t+a}^2 \, dt + \int_{-T}^T f_t^2 \, dt \right) \leq 4 \int_0^T f_t^2 \, dt$$

Where in the last inequality we used the fact that f vanish outside $[0, T]$ and changed variable $s = t + a$:

$$\int_{-T}^T f_{t+a}^2 \, dt = \int_{-T+a}^{T+a} f_s^2 \, ds \leq \int_{-T}^{T+a} f_s^2 ds = \int_0^T f_s^2 \, ds.$$

In conclusion $\int_{-T}^T |f_{t+a} - f_t|^2 \leq 4 \int_0^T f_t^2 \, dt \in L^1(\Omega)$, since $f \in L^2(\Omega \times [-T, T])$ and (2.8) is proved by dominated convergence. Now let $n \in \mathbb{N}$ and define the dyadic partition of $[0, T]$ by

$$t_j^n := j \, 2^{-n} \quad for \; j = 0, 1, ..., [T 2^n]$$

$$\rho_n(t) = t_j^n \quad for \; t \in [t_j^n, t_{j+1}^n).$$

In other words $\rho_n(t)$ is the largest dyadic point less than or equal to t. Let $s \in [0,1]$, then by changing variables $t+s=u,\ v=t$ we obtain

$$\int_0^1 \mathbb{E}(\int_0^T |f_{\rho_n(t+s)-s} - f_t|^2 dt)ds = \int_0^{T+1} (\mathbb{E}\int_{u-1}^u |f_{\rho_n(u)-u+v} - f_v|^2 dv)du.$$

Let $u \in [0, T+1]$, then $\rho_n(u) \to u$ uniformly as $n \to \infty$, since $|u - \rho_n(u)| \leq 2^{-n}$, so the last expectation tends to zero by (2.8). It follows that there is a sequence $n(k) \to \infty$ such that

$$\lim_{k\to\infty} \mathbb{E}\int_0^T |f_{\rho_{n(k)}(t+s)-s} - f_t|^2\, dt = 0 \tag{2.9}$$

for almost every $s \in [0,1]$. Fix any s for which (2.9) holds and denote $f_t^k = \rho_{n(k)}(t+s)-s$. Then by using the inequality $|a|^2 \leq 2|b|^2 + 2|a-b|^2$ with $a = f_t^k$, $b = f_t$

$$\mathbb{E}\int_0^T |f_t^k|^2\, dt \leq 2\mathbb{E}\int_0^T |f_t|^2\, dt + 2\mathbb{E}\int_0^T |f_t^k - f_t|^2\, dt$$

shows that $f_t^k \in L^2(\Omega \times [0,T])$ at least for large k (i.e. $f_t^k \in H_0$) and the theorem is proved.

□

Remark 2.10. *H can be identified with the subspace of $L^2(\Omega \times [0,\infty))$*

$$\widetilde{H} = \{X \in L^2(\Omega \times [0,\infty) : \exists\, \tilde{X} \in H,\ X = \tilde{X} \quad \mathbb{P}\times dt \quad a.e.\}$$

More precisely, since $L^2(\Omega \times [0,\infty))$ consists of equivalence classes, and adaptation is a pointwise property, the space H is not itself a subspace of $L^2(\Omega \times [0,\infty))$.

However, $\widetilde{H}$ is a well-defined subspace, and the canonical projection

$$H \longrightarrow \widetilde{H}, \quad f \mapsto [f]$$

induces an isomorphism between $\tilde{H}$ and the quotient space

$$H/\sim,$$

where $f \sim g$ if and only if $f = g$ $\mathbb{P}\times dt$-a.e.

In this sense, we identify H with $\widetilde{H}$ as a subspace of L^2

Theorem 2.7. *Let $T > 0$. Define*

$$\tilde{H}_T := \left\{ f \in L^2([0,T]\times\Omega) \ :\ \exists\, \tilde{f} \in H \text{ such that } f = \tilde{f} \quad a.e. \text{ on } [0,T]\times\Omega \right\}.$$

Then $\widetilde{H}_T$ is a closed subspace of $L^2([0,T]\times\Omega)$.

Proof. Let $f_n \in \widetilde{H}_T$ such that $f_n \to f$ in $L^2([0,T)\times\Omega)$. We have to show that $f \in \widetilde{H}_T$. Since $f_n \in \widetilde{H}_T$ for all n, then there exists a representative $\tilde{f}_n \in H$ for all n. Moreover:

$$||\tilde{f}_n - f||_{L^2([0,T]\times\Omega)} = ||f_n - f||_{L^2([0,T]\times\Omega)} \longrightarrow 0,$$

so by Fubini's theorem:

$$\int_0^T \mathbb{E}|\tilde{f}_n(t) - f(t)|^2\, dt \longrightarrow 0,$$

implying that there exist n_k, $A \in \mathcal{B}([0,T])$, with $leb(A) = T$, such that for all $t \in A$

$$\mathbb{E}|\tilde{f}_{n_k}(t) - f(t)|^2 \longrightarrow 0.$$

Fix $t \in A$, then there exists a subsequence $n_{k_j}(t) = n_j(t)$ such that

$$\tilde{f}_{n_j(t)}(t) \longrightarrow \tilde{f}(t) \quad a.s.$$

We write

$$\tilde{f}_{n_j(t)}(t,\omega) \to f(t,\omega) \quad \forall\, \omega \in \Omega_t \in \mathcal{F},\ \mathbb{P}(\Omega_t) = 1.$$

Since $(\mathcal{F}_t)$ is complete, then $\Omega_t \in \mathcal{F}_t$, and so $f(t)\mathbb{1}_{\Omega_t}$ is $\mathcal{F}_t$-adapted, since

$$\tilde{f}_{n_j}(t)\mathbb{1}_{\Omega_t} \longrightarrow f(t)\mathbb{1}_{\Omega_t} \quad \forall\, \omega \in \Omega.$$

On the other hand $f(t)\mathbb{1}_{\Omega_t{}^c}$ is $\mathcal{F}_t$-measurable. In fact if $B \in \mathcal{B}(\mathbb{R})$ and $0 \notin B$, then $\{f(t)\mathbb{1}^C_{\Omega_t} \in B\} = \{f(t) \in B\} \cap \Omega_t^C \in \mathcal{F}_t$. If $0 \in B$, then $\{f(t)\mathbb{1}^C_{\Omega_t} \in B\} = (\{f(t) \in B\} \cap \Omega_t^C) \cup \Omega_t \in \mathcal{F}_t$.

So $f(t) = f(t)\mathbb{1}_{\Omega_t} + f(t)\mathbb{1}_{\Omega_t{}^c}$ is $\mathcal{F}_t$-measurable for all $t \in A$. We define

$$\tilde{f}(t,\omega) = \mathbb{1}_{A\times\Omega} f(t,\omega).$$

We have that $\tilde{f} = f$ $dt \times \mathbb{P}$ a.e. and $\tilde{f} \in H$, in fact: for $t \in A$, then $\tilde{f}(t,\cdot)$ is $\mathcal{F}_t$ adapted, if $t \notin A$, then $\tilde{f} = 0$ is $\mathcal{F}_t$-adapted. Moreover, since $f \in L^2([0,T]\times\Omega)$, it is measurable with respect to $\mathcal{B}([0,T]) \otimes \mathcal{F}$, and $\mathbb{1}_A$ is measurable. Therefore, $\tilde{f} = \mathbb{1}_A f$ is also $\mathcal{B}([0,T]) \otimes \mathcal{F}$-measurable.

□

Corollary 2.1. *$\widetilde{H}$ is closed in $L^2([0,\infty)\times\Omega)$*

Proof. Let $f_n \in \tilde{H}$ such that $f_n \to f$ in $L^2([0,\infty)\times\Omega)$, then $f_n \to f$ in $L^2([0,T)\times\Omega)$ for all $T > 0$. Let $T > 0$, for the previous lemma there exists $\tilde{f}^{(T)} \in H$ such that $\tilde{f}^{(T)} = f$ for all $(t,\omega) \in A_T \in \mathcal{B}([0,T] \otimes \mathcal{F})$, $\mu((A_T)^C) = 0$. Define

$$\tilde{f}(t,\omega) = \sum_{T=1}^{\infty} \mathbb{1}_{[T-1,T)}(t) \cdot \tilde{f}^{(T)}(t,\omega),$$

$$A = \bigcup_{T\geq 1} A_T.$$

Since

$$\mu\left(\left(\bigcup_{T=1}^{\infty} A_T\right)^c\right) = \mu\left(\bigcap_{T=1}^{\infty} A_T^c\right) = 0,$$

then A is a subset of full measure of $[0,\infty)\times\Omega$ and of course $\tilde{f} = f$ for all $(t,\omega) \in A$. Moreover, for each T the functions $\mathbb{1}_{[T-1,T)} \cdot \tilde{f}^{(T)}$ are jointly measurable and so $\tilde{f}$ is jointly measurable as countable sum of jointly measurable functions. Finally, for each $t > 0$ there exist a unique T such that $t \in [T-1,T)$, so $\tilde{f}(t,\cdot) = \tilde{f}^{(T)}(t,\cdot)$ that is $\mathcal{F}_t$-measurable because $\tilde{f}^{(T)} \in H$; so $\tilde{f}$ is $\mathcal{F}_t$-adapted and $\tilde{f} \in H$, hence $f \in \tilde{H}$.

□

Remark 2.11. *Identifying H with $\widetilde{H}$ we can conclude that*

$$\overline{H_0} = H,$$

since $H_0 \subseteq H = \widetilde{H} \subseteq \overline{H_0}$ by Theorem 2.6 and $\widetilde{H}$ is closed in $L^2(\Omega \times [0,\infty))$, implying that $\overline{H_0} \subseteq \widetilde{H}$.

Remark 2.12. *The space H can also be identified with $L^2(\mathcal{P},\mu)$, where μ is the product measure $\mathbb{P} \times Leb$ and $\mathcal{P} = \sigma\{A \times (s,t], 0 \le s \le t < \infty, A \in \mathcal{F}_s\}$ is called the σ-algebra of predictable sets. More precisely the space $L^2(\mathcal{P},\mu)$ is the space of square integrable $f(t,\omega)$ functions which are measurable with respect to the μ-completion of the σ-algebra $\mathcal{P}$. The reader is referred to [14] chapter 2, section 7 for more details.*

Definition 2.12. *For $f \in \overline{H_0}$ and $s \ge 0$ define*

$$\int_0^s f_t \, dW_t = \int_0^\infty \mathbb{1}_{[0,s)}(t) \, f_t \, dW_t.$$

Theorem 2.8. *Let $f \in \overline{H_0}$. Then the process $\int_0^s f_t \, dW_t$ admits a continuous modification (see Definition 1.12).*

The reader can found the proof of this result in [14] chapter 6, section 1.

2.2.1 Properties of the Stochastic integral on H

Theorem 2.9. *Let $f, g \in H$, $a, b \in \mathbb{R}$. Then:*

(i) (linearity) With probability one for all t at once

$$\int_0^t (af_s + bg_s) \, dW_s = a \int_0^t f_s \, dW_s + b \int_0^t g_s \, dW_s; \tag{1}$$

(ii) $\mathbb{E}\left(\int_0^\infty f_t \, dW_t\right) = 0$;

(iii) the process $\int_0^t f_s \, dW_s$ is a martingale relative to $\mathcal{F}_t$.

We give a sketch of proof.

Proof. (i) Take $t \in \mathbb{Q}_+$, then equation (1) holds almost surely for $f, g \in H_0$ and we can extend it for $f, g \in H$ via approximation. So for each rational t, it exists a set Ω_t, $P(\Omega_t) = 1$, such that for all $\omega \in \Omega_t$ (1) holds. Define

$$\Omega' = \bigcap_{t \in \mathbb{Q}_+} \Omega_t.$$

Ω' is still a set of full probability, then for all $\omega \in \Omega'$ (1) holds for all $t \in \mathbb{Q}_+$. Furthermore, both sides of (1) are *continuous* in t (see Theorem 2.8), indeed indeed they coincide (a.s.) for all $t \ge 0$ since they coincide on a dense subset.

(ii) Take $f^n \in H_0$ such that $f^n \to f$ in $L_2(\Omega \times [0,\infty))$. Then use Cauchy's inequality and Lemma 2.4 to find that

$$|\mathbb{E}If| = |\mathbb{E}I(f-f^n)| \le \mathbb{E}\left[(I(f-f^n))^2\right]^{1/2} = \left[\mathbb{E}\left(\int_0^\infty (f_t - f_t^n)^2\,dt\right)\right]^{1/2} \to 0.$$

Where in the last identity we used the Isometry formula for processes in H, see Remarks 2.8 and 2.9

(iii) Take the same sequence f^n as above and remember that Lemma 2.5 allows us to write

$$\mathbb{E}\left\{\int_0^t f_s^n\,dW_s \Big| \mathcal{F}_r\right\} = \int_0^r f_s^n\,dW_s \quad \text{(a.s.)} \quad \forall\, 0 \le r \le t. \tag{3}$$

Furthermore,

$$\mathbb{E}\left(\int_0^t f_s^n\,dW_s - \int_0^t f_s\,dW_s\right)^2 = \mathbb{E}\left(\int_0^t (f_s^n - f_s)^2\,ds\right) \to 0,$$

and if $\mathcal{G} \subset \mathcal{F}$, then evaluating conditional expectation is a continuous operator (as a projection operator) defined on the space $L^2(\mathcal{F},\mathbb{P})$ with values in the space $L^2(\mathcal{G},\mathbb{P})$ (Theorem 1.5), and this continuity allows as to pass to the limit in L^2 in (3). Hence upon passing to the limit in (3) in the mean-square sense, we get an equality which shows that

(a) $\int_0^r f_s\,dW_s$ is $\mathcal{F}_r^{\mathbb{P}}$-measurable as a function almost surely equal to an $\mathcal{F}_r$-measurable $\mathbb{E}(\cdot|\mathcal{F}_r)$;

(b) the martingale equality holds.

□

Lemma 2.6. *Let τ be an $\mathcal{F}_t$-stopping time. Then $\mathbb{1}_{t<\tau}$ and $\mathbb{1}_{t\le\tau}$ are $\mathcal{F}_t$-adapted.*

Proof. For all $t \ge 0$ $\{\tau > t\} = \{\tau \le t\}^C \in \mathcal{F}_t$, meaning $\mathbb{1}_{t<\tau}$ is $\mathcal{F}_t$-adapted. $\{\tau \ge t\} = \{\tau < t\}^C$ and

$$\{\tau < t\} = \bigcup_{n\ge1}\{\tau \le t - \frac{1}{n}\}$$

and $\{\tau \le t - \frac{1}{n}\} \in \mathcal{F}_t$ for all n, so $\mathbb{1}_{t\le\tau}$ is $\mathcal{F}_t$-adapted.

□

Lemma 2.7. *Let $f = f_t(\omega)$ be nonnegative, $\mathcal{F}_t$-adapted, and $\mathcal{F}\otimes\mathcal{B}[0,\infty)$-measurable. Then, for any $t \ge 0$, $\int_0^t f_s\,ds$ is $\mathcal{F}_t$-measurable.*

Proof. If the assertion holds for $f \wedge n$ in place of f, then by letting $n \to \infty$ and using the monotone convergence theorem we get the result for our f. It follows that without losing generality we may assume that f is bounded. Furthermore, we can cut the function f in t by taking $t \ge 0$ and setting $f_s = 0$ for $s \ge t$. Then we see that it suffices to prove our assertion for $f \in H$.

In that case, we conclude that there exist $f^n \in H_0$ such that

$$\mathbb{E}\left|\int_0^t f_s\,ds - \int_0^t f_s^n\,ds\right| \le \mathbb{E}\left(\int_0^t |f_s - f_s^n|\,ds\right) \le \sqrt{t}\left(\mathbb{E}\int_0^\infty |f_s - f_s^n|^2\,ds\right)^{1/2} \to 0.$$

Furthermore, $\int_0^t f_s^n\,ds$ is obviously written as a sum in which all terms are $\mathcal{F}_t$-measurable. The mean-square limit of $\mathcal{F}_t$-measurable variables is $\mathcal{F}_t$-measurable, and the lemma is proved.

□

Theorem 2.10. *Let $f \in H$, and let $N, c > 0$, and $T \le \infty$ be constants. Then*

$$\mathbb{P}\left\{\sup_{t\le T}\left|\int_0^t f_s\,dW_s\right| \ge c\right\} \le \mathbb{P}\left\{\int_0^T f_s^2\,ds \ge N\right\} + \frac{1}{c^2}E\left(N \wedge \int_0^T f_s^2\,ds\right).$$

Proof. We use the standard way of stopping stochastic integrals such as

$$\xi_t = \int_0^t f_s\,dW_s$$

by using their "brackets", defined as

$$\langle\xi\rangle_t := \int_0^t f_s^2\,ds.$$

We will see that this is exactly the definition of quadratic variation of a Itô process (see Definition 2.15). Let $\tau = \inf\{t \ge 0 : \langle\xi\rangle_t \ge N\}$, so that τ is the first exit time of $\langle\xi\rangle_t$ from $(-1, N)$. By Lemma1.3 and Lemma 2.7 we have that τ is a stopping time. Furthermore,

$$\{\omega : \tau < T\} \subset \{\omega : \langle\xi\rangle_T \ge N\}$$

and on the set $\{\omega : \tau \ge T\}$ we have $\mathbb{1}_{s<\tau} f_s = f_s$ if $s < T$. Note that the process

$$M_t := \left(\int_0^t \mathbb{1}_{s<\tau} f_s\,dW_s\right)^2$$

is a non-negative submartingale. This follows from the fact that $\int_0^t \mathbf{1}_{s<\tau} f_s\,dw_s$ is a martingale (by Theorem 2.9) with zero mean, and the square of a martingale is a submartingale.

We now apply the Doob maximal inequality for submartingales, which states that if $\{X_t\}$ is a non-negative submartingale, then for any $c > 0$,

$$\mathbb{P}\left(\sup_{t\le T} X_t \ge c\right) \le \frac{\mathbb{E}[X_T]}{c}.$$

Applying this to M_t with c^2 in place of c, we obtain

$$\mathbb{P}\left(\sup_{t\le T}\left|\int_0^t \mathbb{1}_{s<\tau} f_s\,dw_s\right| \ge c\right) = \mathbb{P}\left(\sup_{t\le T} M_t \ge c^2\right) \le \frac{1}{c^2}\mathbb{E}\left[\int_0^T \mathbb{1}_{s<\tau} f_s^2\,ds\right].$$

Therefore, upon denoting

$$A = \left\{ \omega : \sup_{t \le T} |\xi_t| \ge c \right\},$$

by the Doob-Kolmogorov inequality for continuous submartingales we get

$$\mathbb{P}(A, \tau \ge T) = \mathbb{P}\left\{ \tau \ge T, \sup_{t \le T} \left| \int_0^t \mathbb{1}_{s<\tau} f_s \, dw_s \right| \ge c \right\}$$

$$\le \mathbb{P}\left\{ \sup_{t \le T} \left| \int_0^t \mathbb{1}_{s<\tau} f_s \, dw_s \right|^2 \ge c^2 \right\} \le \frac{1}{c^2} \mathbb{E}\left(\int_0^T \mathbb{1}_{s<\tau} f_s \, dw_s \right)^2$$

$$= \frac{1}{c^2} \mathbb{E}\left(\int_0^{T \wedge \tau} \mathbb{1}_{s<\tau} f_s^2 \, ds \right) = \frac{1}{c^2} \mathbb{E}\left(\int_0^T \mathbb{1}_{s<\tau} f_s^2 \, ds \wedge \int_0^\tau f_s^2 \, ds \right)$$

$$\le \frac{1}{c^2} \mathbb{E}\left(N \wedge \int_0^T f_s^2 \, ds \right).$$

Where in the last inequality we have used the fact that, if $\tau < \infty$, then obviously $\langle \xi \rangle_\tau = N$, and if $\tau = \infty$, then $\langle \xi \rangle_\tau \le N$. Hence

$$\mathbb{P}(A) = \mathbb{P}(A, \tau < T) + \mathbb{P}(A, \tau \ge T) \le \mathbb{P}(\tau < T) + \mathbb{P}(A, \tau \ge T)$$

$$\le \mathbb{P}\left\{ \int_0^T f_s^2 \, ds \ge N \right\} + \frac{1}{c^2} \mathbb{E}\left(N \wedge \int_0^T f_s^2 \, ds \right).$$

The theorem is proved. □

2.3 Defining the Itô integral if $\int_0^T f_s^2 ds < \infty$

Definition 2.13. *Denote by $\mathcal{S}$ the set of all $\mathcal{F}_t$-adapted, $\mathcal{F} \otimes \mathcal{B}(0,\infty)$-measurable processes f_t such that*

$$\int_0^T f_s^2 \, ds < \infty \quad \text{(a.s.)} \quad \forall T < \infty.$$

Our task here is to define $\int_0^t f_t \, dW_t$ for $f \in \mathcal{S}$. We give just a sketch of proof.

Define

$$\tau(n) = \inf\left\{ t \ge 0 : \int_0^t f_s^2 \, ds \ge n \right\}.$$

$\tau(n)$ are stopping times by Lemma 1.3) and

$$\int_0^{\tau(n)} f_s^2 \, ds \le n. \quad (1)$$

Furthermore, obviously $\tau(n) \uparrow \infty$ (a.s.) as $n \to \infty$. Finally, notice that $\mathbb{1}_{s<\tau(n)} f_s \in H$. Indeed, the fact that this process is $\mathcal{F}_t$-adapted follows from Lemma 2.6. Also

$$\mathbb{E}\left(\int_0^\infty \mathbb{1}_{s<\tau(n)} f_s^2 \, ds \right) = \mathbb{E}\left(\int_0^{\tau(n)} f_s^2 \, ds \right) \le n < \infty.$$

It follows from the above that the stochastic integrals

$$\xi_t(n) := \int_0^t \mathbb{1}_{s<\tau(n)} f_s \, dW_s \tag{2.10}$$

are well defined. We define

$$\int_0^{t\wedge\tau(n)} f_s \, dw_s := \int_0^t \mathbb{1}_{s<\tau(n)} f_s \, dw_s.$$

Lemma 2.8. *Let $f \in \mathcal{S}$. Then there exists a set $\Omega' \subset \Omega$ such that $\mathbb{P}(\Omega') = 1$ and, for every $\omega \in \Omega'$, $m \geq n$, and $t \in [0, \tau(n,\omega))$, we have $\xi_t(n) = \xi_t(m)$.*

Corollary 2.2. *If $f \in \mathcal{S}$, then with probability one the sequence $\xi_t(n)$ (see* (2.10)*) converges uniformly on each finite time interval.*

The reader can find a proof of the two in [14] chapter 6.

Definition 2.14. *Let $f \in \mathcal{S}$. For those ω for which the sequence $\xi_t(n)$ (see* (2.10)*) converges uniformly on each finite time interval, we define*

$$\int_0^t f_s \, dw_s = \lim_{n\to\infty} \int_0^t \mathbb{1}_{s<\tau(n)} f_s \, dw_s.$$

For all other ω we define $\int_0^t f_s \, dw_s = 0$.

Of course, one has to check that Definition 2.14 does not lead to anything new if $f \in H$. Observe that if $f \in H$, then by Fatou's theorem and the dominated convergence theorem

$$\mathbb{E}\left|\lim_{n\to\infty} \int_0^t \mathbb{1}_{s<\tau(n)} f_s \, dw_s - \int_0^t f_s \, dw_s\right|^2 \leq \lim_{n\to\infty} \mathbb{E}\left(\int_0^t (1 - \mathbb{1}_{s<\tau(n)}) f_s^2 \, ds\right) = 0.$$

Therefore both definitions give the same result almost surely for any given t.

We recall a useful result about stopping times.

Lemma 2.9. *Let τ be a stopping time with respect to the filtration $(\mathcal{F}_t)_t$, with $\tau \leq T$. Then, if $f \in S$, it also holds that $(f_t \mathbb{1}_{\{t<\tau\}})_t \in S$) and*

$$\int_0^\tau f_s \, dW_s = \int_0^T f_s \mathbb{1}_{\{s<\tau\}} \, dW_s.$$

The reader can consult [3] (Lemma 6.19) for a proof. Now come some properties of the stochastic integral on $\mathcal{S}$.

Theorem 2.11. *Let $f, f^n, g \in \mathcal{S}$, and let $\delta, \varepsilon > 0$, $T \in [0,\infty)$ be constants. Then:*

(i) the stochastic integral $\int_0^t f_s \, dw_s$ is continuous in t and $\mathcal{F}_t$-adapted;

(ii) we have

$$\mathbb{P}\left(\sup_{t\le T}\left|\int_0^t f_s\,dw_s-\int_0^t g_s\,dw_s\right|\ge\varepsilon\right)\le\mathbb{P}\left(\int_0^T|f_s-g_s|^2\,ds\ge\delta\right)$$

$$+\frac{1}{\varepsilon^2}\mathbb{E}\delta\wedge\int_0^T(f_s-g_s)^2\,ds\le\mathbb{P}\left(\int_0^T|f_s-g_s|^2\,ds\ge\delta\right)+\frac{\delta}{\varepsilon^2};\qquad(2)$$

(iii) we have

$$\int_0^T|f_s^n-f_s|^2\,ds\xrightarrow{\mathbb{P}}0\quad\Rightarrow\quad\sup_{t\le T}\left|\int_0^t f_s^n\,dw_s-\int_0^t f_s\,dw_s\right|\xrightarrow{\mathbb{P}}0.$$

Proof. (i) The continuity of $\int_0^t f_s\,ds$ follows from Definition 2.14, in which

$$\int_0^t \mathbb{1}_{s<\tau(n)}f_s\,ds$$

are continuous and $\mathcal{F}_t$-adapted (even $\mathcal{F}_t$-martingales). Their limit is also $\mathcal{F}_t$-adapted.

To prove (ii), first notice that all expressions in (2) are monotone and right-continuous in ε and δ. Therefore, it suffices to prove (2) only at points of their continuity. Also notice that the second inequality in (2) is obvious since $\delta\wedge\cdot\le\delta$.

Now fix appropriate ε, and δ and define

$$\tau(n)=\inf\left\{t\ge0:\int_0^t f_s^2\,ds\ge n\right\},\quad\sigma(n)=\inf\left\{t\ge0:\int_0^t g_s^2\,ds\ge n\right\},$$

$$f_s^n=\mathbb{1}_{s<\tau(n)}f_s,\quad g_s^n=\mathbb{1}_{s<\sigma(n)}g_s.$$

Since f^n and g^n belong to H, inequality (2) holds with f^n,g^n in place of f,g due to the linearity of the stochastic integral on H and Theorem 2.10. Furthermore, almost surely, as $n\to\infty$,

$$\sup_{t\le T}\left|\int_0^t f_s^n\,dw_s-\int_0^t g_s^n\,dw_s\right|\to\sup_{t\le T}\left|\int_0^t f_s\,dw_s-\int_0^t g_s\,dw_s\right|,$$

$$\int_0^T|f_s^n-g_s^n|^2\,ds\to\int_0^T|f_s-g_s|^2\,ds.$$

These convergences of random variables imply convergence of the corresponding distribution functions at all points of their continuity. Adding to this tool the dominated convergence theorem, we get (2) from its version for f^n,g^n.

To prove (iii) it suffices to take $g=f^n$ in (2) and let first $n\to\infty$ and then $\delta\downarrow0$. The theorem is proved.

□

2.4 Itô formula

Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a complete probability space carrying a Wiener process $(W_t, \mathcal{F}_t)$ and a continuous $\mathcal{F}_t$-adapted process ξ_t. Assume that we are also given a process σ_t and a process b_t such that $\sigma \in \mathcal{S}$ and b is jointly measurable in (ω, t), $\mathcal{F}_t$-adapted, and

$$\int_0^T |b_s|\, ds < \infty \quad \text{(a.s.) for any } T < \infty.$$

Then we write

$$d\xi_t = \sigma_t\, dW_t + b_t\, dt$$

if and only if (a.s.) for all t

$$\xi_t = \xi_0 + \int_0^t \sigma_s\, dW_s + \int_0^t b_s\, ds. \tag{1}$$

In that case one says that ξ_t has *stochastic differential* equal to $\sigma_t\, dw_t + b_t\, dt$ and ξ is said to be a *Itô process*

Definition 2.15. *Let X_t be a Itô process satisfying (1), then we define its quadratic variation as*

$$\langle X \rangle_t = \int_0^t \sigma_s^2\, ds$$

Remark 2.13. *The definition of quadratic variation of a Itô process coincides with the more general concept of quadratic variation of a martingale which we do not treat here and it is discussed in details by Karatzas and Shreve in [13] section 1.5. In particular it can be shown that $\langle X \rangle_t$ is the limit in probability of the sum*

$$\sum_{i=0}^{k-1} (X_{t_{i+1}} - X_{t_i})^2$$

along a partition $\pi = \{0 = t_0 < t_1 < \cdots < t_k = t\}$ of $[0, t]$, as the mesh of the partition tends to zero (we will not need this fact in the sequel). Note that by definition, $\langle X \rangle_t$ is an Itô process, whose stochastic differential is given by

$$\mathrm{d}\langle X \rangle_t = \sigma_t^2\, \mathrm{d}t.$$

We can therefore define the integral with respect to $\langle X \rangle$ by setting

$$\int_0^t Y_s\, \mathrm{d}\langle X \rangle_s := \int_0^t Y_s\, \sigma_s^2\, \mathrm{d}s, \tag{6.10}$$

for any process Y for which this makes sense.

We say that a function $F = F(t, x) : \mathbb{R}^+ \times \mathbb{R} \to \mathbb{R}$ is of class $C^{1,2}$ if it is continuously differentiable once in t and twice in x, that is, if the partial derivatives

$$\frac{\partial F}{\partial t}(t, x), \quad \frac{\partial F}{\partial x}(t, x), \quad \frac{\partial^2 F}{\partial x^2}(t, x)$$

exist and are continuous functions of $(t, x) \in \mathbb{R}^+ \times \mathbb{R}$.

We now present Itô's formula in its rigorous integral form, followed by its commonly used differential version.

Itô's formula is the stochastic analogue of the classical chain rule and one of the most fundamental results in stochastic calculus. It provides the correct way to differentiate functions of stochastic processes, playing a key role in both theory and applications.

Theorem 2.12 (Itô formula: integral form). *Let $X = \{X_t\}_{t\geq 0}$ be an Itô process, with*

$$\mathrm{d}X_t = \sigma_t \,\mathrm{d}W_t + b_t dt,$$

and let $F = F(t,x) : \mathbb{R}^+ \times \mathbb{R} \to \mathbb{R}$ be a function of class $C^{1,2}$. Then, for almost every ω and for all $t \geq 0$, we have:

$$F(t,X_t) - F(0,X_0) = \int_0^t \frac{\partial F}{\partial s}(s,X_s)\,\mathrm{d}s + \int_0^t \frac{\partial F}{\partial x}(s,X_s)\,\mathrm{d}X_s + \frac{1}{2}\int_0^t \frac{\partial^2 F}{\partial x^2}(s,X_s)\,\mathrm{d}\langle X\rangle_s.$$

While the integral form provides the precise mathematical statement, Itô formula can be expressed also in differential form, which is a useful symbolic representation, especially in applied contexts.

Theorem 2.13 (Itô formula: differential form). *Let X_t be an Itô process:*

$$dX_t = b_t\,dt + \sigma_t\,dW_t,$$

and let $F = F(t,x) \in C^{1,2}$. Then:

$$dF(t,X_t) = \frac{\partial F}{\partial t}(t,X_t)\,dt + \frac{\partial F}{\partial x}(t,X_t)\,dX_t + \frac{1}{2}\frac{\partial^2 F}{\partial x^2}(t,X_t)\,d\langle X\rangle_t.$$

We will need the general multidimensional version of Itô formula, the one–dimensional version stated above is the special case with $m = 1$.

Theorem 2.14 (Multidimensional Itô formula). *Let X_t^i, $i = 1, \ldots, m$, be stochastic processes admitting stochastic differentials of the form*

$$dX_t^i = F_i(t)\,dt + G_i(t)\,dB_t, \qquad i = 1, \ldots, m,$$

where $(B_t)_{t\geq 0}$ is a Wiener process and the coefficient processes F_i and G_i are adapted and satisfy suitable integrability conditions.

Let $f : \mathbb{R}_+ \times \mathbb{R}^m \to \mathbb{R}$ be a function of class $C^{1,2}$. Then, setting $X_t = (X_t^1, \ldots, X_t^m)$, the process $f(t,X_t)$ satisfies

$$\begin{aligned} df(t,X_t) = {} & \frac{\partial f}{\partial t}(t,X_t)\,dt + \sum_{i=1}^m \frac{\partial f}{\partial x_i}(t,X_t)\,dX_t^i \\ & + \frac{1}{2}\sum_{i,j=1}^m \frac{\partial^2 f}{\partial x_i \partial x_j}(t,X_t)\,G_i(t)\,G_j(t)\,dt. \end{aligned}$$

The notion of quadratic variation introduced above allows us to state another useful characterization theorem for a Wiener process, that is the following one by P.Levy which we state for a one dimensional process. The reader can find this result in [13] in the more general d-dimensional case.

Theorem 2.15 (Lévy)**.** *Let* $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, \mathbb{P})$ *be a filtered probability space and let* $X = (X_t)_{t\geq 0}$ *be a continuous, adapted process with* $X_0 = 0$*. Assume that*

- X *is a continuous local martingale with respect to* $(\mathcal{F}_t)$*, and*
- *its quadratic variation satisfies*

$$\langle X \rangle_t = t, \qquad t \geq 0 \quad a.s.$$

Then X *is an* $(\mathcal{F}_t)$*–Wiener process. Conversely, a one-dimensional Wiener process satisfies these properties.*

Chapter 3

Girsanov's theorem

In this chapter, we present a complete proof of Girsanov's theorem, following the approach of Krylov [14] and adding a lot of details he is omitting. This approach is different with respect to more classical treatments such as those found in Baldi [3] or Karatzas and Shreve [13] which require more results from stochastic calculus. Krylov uses an argument based on complex analysis.

We consider a Wiener's process $(W_t, \mathcal{F}_t)$ defined on a complete probability space $(\Omega, \mathcal{F}, \mathbb{P})$ and assume that the $\mathcal{F}_t$ are complete, that is $\mathcal{F}_t = \mathcal{F}_t^{\mathbb{P}}$ (see Definition 1.5)

Lemma 3.1. *Let $b \in S$ (see Definition 2.13) be a real valued process. Denote*

$$\rho_t(b) = \exp(\int_0^t b_s\, dW_s - \frac{1}{2}\int_0^t b_s^2\, ds).$$

Then

(i) $d\rho_t = b_t \rho_t dW_t$

(ii) ρ_t *is a supermartingale*

(iii) if b is a bounded process (see Definition 1.9), then ρ_t is a martingale and, in particular $\mathbb{E}\rho_t = 1$

(iv) if $T \in [0, \infty)$ and $\mathbb{E}\rho_T = 1$, then ρ_t is a martingale for $t \in [0, T]$ and also for any sequence of bounded processes $b^n \in S$ such that $\int_0^T |b_s^n - b_s|^2 ds \longrightarrow 0$ (a.s.), we have:

$$\mathbb{E}|\rho_T(b^n) - \rho_T(b)| \longrightarrow 0$$

Proof. Assertion (i) follows from Itô formula. In fact if one defines

$$X_t = \int_0^t b_s\, dW_s - \frac{1}{2}\int_0^t b_s^2\, ds,$$

then $q_t(b) = f(X_t)$ where $f(x) = e^x$ (non-depending on time), applying Itô formula we obtain:

$$d\rho_t = \frac{\partial}{\partial x} f(X_t)\, dX_t + \frac{1}{2}\frac{\partial}{\partial x^2} f(X_t)\, d\langle X\rangle_t = e^{X_t}(b_t W_t - \frac{1}{2}{b_t}^2 dt) + \frac{1}{2} e^{X_t} {b_t}^2 dt = \rho_t b_t dW_t.$$

To prove (ii) define $\tau_n := \inf\{t \geq 0 : \int_0^t |b_s|^2 \geq n\}$.
τ_n is a stopping time by Lemma 1.3 and $\mathbb{1}_{\{t<\tau_n\}} b_t \rho_t \in H$, and so $\int_0^t \mathbb{1}_{\{s<\tau_n\}} b_s \rho_s \, dW_s$ is a martingale by Theorem 2.9. Using (i) we can write $\rho_t = \rho_0 + \int_0^t b_s \rho_s dW_s$, and so

$$\rho_{t\wedge\tau_n} = 1 + \int_0^{t\wedge\tau_n} b_s \rho_s \, dW_s = 1 + \int_0^t \mathbb{1}_{\{s<\tau_n\}} b_s \rho_s \, dW_s$$

is a martingale. The last identity is true since if $\tau_n > t$, then $\tau_n \wedge t = t$ and $\mathbb{1}_{s<\tau_n} = 1$ on $[0, t]$, so both sides coincide. On the other hand if $\tau_n \leq t$, then $\tau_n \wedge t = \tau_n$ and the equality is obtained by Lemma 2.9.

Hence, for $t_2 \geq t_1$

$$\mathbb{E}[\rho_{t_2\wedge\tau_n} | \mathcal{F}_{t_1}] = \rho_{t_1\wedge\tau_n} (a.s.).$$

As $n \longrightarrow \infty$, we have $\tau_n \longrightarrow \infty$. In fact if we suppose that

$$\lim_n \tau_n < \infty$$

on a set A, $\mathbb{P}(A) > 0$, then it would exist $T < \infty$ such that for all $\omega \in A$ $\exists$ $N = N(\omega)$ such that

$$\tau_n(\omega) < T \quad \forall\, n \geq N.$$

Hence, for all $n \geq N$:

$$\int_0^T b_s^2 \, ds \geq \int_0^{\tau_n} b_s^2 \, ds \geq n,$$

implying

$$\int_0^T b_s^2 \, ds = \infty\,,$$

but this is not possible since $b \in S$ (see Definition 2.13). So that $t_i \wedge \tau_n \longrightarrow t_i$ and by Fatou's theorem (a.s.)

$$\mathbb{E}[\rho_{t_2} | \mathcal{F}_{t_1}] \leq \rho_{t_1}.$$

. This proves (ii) and implies:

$$\mathbb{E}\rho_t = \mathbb{E}exp(\int_0^t b_s dW_s - \frac{1}{2}\int_0^t b_s^2 ds) \leq \mathbb{E}\rho_0 = 1. \tag{3.1}$$

To prove (iii) notice that

$$\rho_s^2(b) = \rho_s(2b) \exp \int_0^s |b_r|^2 dr,$$

since

$$\rho_s^2(b) = \exp \int_0^s 2b_r \, dW_r - \int_0^s {b_r}^2 dr,$$

and

$$\rho_s(2b) = \exp(\int_0^s 2b_r\, dW_r - \frac{1}{2}\int_0^s 4b_r^2 dr) = \exp(\int_0^s 2b_r\, dW_r - \int_0^s 2b_r^2 dr).$$

Hence, if we suppose $|b_s| \le K$, then $b_s\rho_s \in H$, in fact:

$$\mathbb{E}\int_0^t |b_s|^2\rho_s^2 ds \le K^2\mathbb{E}\int_0^t \rho_s^2 ds = K^2\int_0^t (\mathbb{E}\rho_s(2b)\exp\int_0^s |b_r|^2 dr)ds$$

$$\le K^2\int_0^t \mathbb{E}[\rho_s(2b)e^{K^2 s}]ds \le K^2\int_0^t e^{K^2 s}ds < \infty.$$

Hence

$$\int_0^t b_s\rho_s dW_s \quad \text{and } \rho_t = 1 + \int_0^t b_s\rho_s dW_s$$

are martingales and (iii) is proved.
To prove (iv) suppose $\mathbb{E}\rho_T = 1$, this means that ρ_t is a supermartingale with constant mean (for $t \le T$), so ρ_t is a martingale by Proposition 1.2. Notice that $\mathbb{E}\rho_T(b^n) = 1$ for all n by (iii), $\mathbb{E}\rho_T(b) = 1$ by the assumption, and $\rho_T(b^n) \longrightarrow \rho_T(b)$ in probability by Theorem 2.11.

This implies that $\mathbb{E}|\rho_T(b^n) - \rho_T(b)| \longrightarrow 0$ by Scheffe's theorem (see Lemma 1.2). ☐

An immediate consequence of inequality (2.1) is the following.

Corollary 3.1. *Let $b \in S$. Let t be a time such that $\int_0^t |b_s|^2 ds \le C_t$ (i.e. $\int_0^t |b_s|^2 ds$ is a bounded variable), then*

$$\mathbb{E}\left[\exp\int_0^t b_s\, dW_s\right] \le e^{C_t}.$$

Proof. Fix $t > 0$ and let $C_t > 0$ be the constant such that $\int_0^t |b_s|^2 ds \le C_t$ (a.s.).

Note that

$$\exp\left(\int_0^t b_s\, dW_s\right) = \rho_t(b)\cdot\exp\left(\frac{1}{2}\int_0^t |b_s|^2\, ds\right).$$

Therefore,

$$\mathbb{E}\left[\exp\left(\int_0^t b_s\, dW_s\right)\right] = \mathbb{E}\left[\rho_t(b)\cdot\exp\left(\frac{1}{2}\int_0^t |b_s|^2\, ds\right)\right].$$

Applying the Cauchy-Schwarz inequality,

$$\mathbb{E}\left[\rho_t(b)\cdot\exp\left(\frac{1}{2}\int_0^t |b_s|^2\, ds\right)\right] \le \left(\mathbb{E}[\rho_t^2(b)]\right)^{1/2}\cdot\left(\mathbb{E}\left[\exp\left(\int_0^t |b_s|^2\, ds\right)\right]\right)^{1/2} =$$

$$= (\mathbb{E}\left[\rho_t(2b)\exp\int_0^t |b_s|^2 ds\right])^{\frac{1}{2}}\cdot(\mathbb{E}\left[\exp\int_0^t |b_s|^2 ds\right])^{\frac{1}{2}}$$

$$\le e^{\frac{C_t}{2}}(\mathbb{E}\rho_t(2b))^{\frac{1}{2}}(\mathbb{E}e^{C_t})^{\frac{1}{2}} \le e^{C_t}.$$

Where in the last inequality we used the fact that $\rho_t(2b)$ is a supermartingale, thus its expected value is less or equal than 1, since $\rho_0(2b) = 1$ ☐

Definition 3.1. *Let Z_t be a complex valued process. We say that Z_t is a martingale if its real part and imaginary part are martingales.*

Definition 3.2. *Let $Z_t \in S$ be a complex valued process (i.e., $\Re(Z) \in S$ and $\Im(Z) \in S$). We define its stochastic integral as*

$$\int_0^t Z_s \, dW_s := \int_0^t \Re(Z_s) \, dW_s + i \int_0^t \Im(Z_s) \, dW_s.$$

Before presenting the proof of Girsanov's theorem, we state a remarkable auxiliary result due to Krylov. This lemma plays a central role in the approach adopted by Krylov for the proof of Girsanov's theorem and differs significantly from the classical treatments found in Baldi [3] or Karatzas and Shreve [13].

Lemma 3.2. *If b_t is a bounded complex-valued process of class S, then $\rho_t(b)$ is a (complex-valued) martingale and, in particular, $\mathbb{E}\rho_t(b) = 1$ for any t.*

Proof. Take $t_2 > t_1 \geq 0$ and $A \in \mathcal{F}_{t_1}$.To prove the lemma it suffices to prove that if f_t, g_t are bounded real valued process of class S, then for all complex z

$$\mathbb{E}\left[\mathbb{1}_A \exp(\int_0^{t_2} (f_s + z g_s) dW_s - \frac{1}{2}\int_0^{t_2} (f_s + z g_s)^2 ds)\right]$$
$$= \mathbb{E}\left[\mathbb{1}_A \exp(\int_0^{t_1} (f_s + z g_s) dW_s - \frac{1}{2}\int_0^{t_1} (f_s + z g_s)^2 ds)\right].$$

Observe that the identity above holds for real z by Lemma 3.1 (iii). Therefore, to prove the lemma we just need to show that both sides are holomorphic functions of z, since by the identity theorem of holomorphic functions, see [18] chapter 10, we know that if two holomorphic functions coincide on a line (in our case the real line), then they coincide on the whole complex plane. To prove that, we will use *Morera's theorem*, showing that both sides are continuous and their integrals along closed piecewise constant paths vanish.

We divide the proof in 2 steps.

Step 1 We consider a simplified case in order to illuminate the main idea and clarify the structure of the argument. Let $z \in \mathbb{C}$ and let g_s be bounded real valued processes of class S.

We prove using Morera's theorem that the function

$$z \mapsto \mathbb{E}\left[\mathbb{1}_A e^{z \int_0^1 g_s \, dW_s}\right]$$

is holomorphic. We prove now that $z \mapsto \mathbb{E}\left[\mathbb{1}_A e^{z \int_0^1 g_s \, dW_s}\right]$ is continuous. Since the exponential function is continuous we have that if $z_n \to z$, then

$$\mathbb{1}_A \exp(z_n \int_0^1 g_s \, dW_s) \to \mathbb{1}_A \exp(z \int_0^1 g_s \, dW_s) \quad \text{(a.s.)}.$$

Moreover, if $z_n \to z$, then there exists $R > 0$ such that $|\Re(z_n)| \leq |z_n| \leq R$ for all n. Then :

$$\mathbb{E}|\mathbb{1}_A \exp(z_n \int_0^1 g_s\, dW_s)|^2 \leq \mathbb{E}\left[\exp(2\Re(z_n)\int_0^1 g_s\, dW_s)\right] \leq \mathbb{E}\left[\exp(\int_0^1 2Rg_s\, dW_s)\right].$$

Now note that $\int_0^1 4R^2 g_s^2\, ds \leq 4R^2\, \|g\|_\infty^2 = C$, indeed by Corollary 3.1:

$$\mathbb{E}\left[\exp(\int_0^1 2Rg_s\, dW_s)\right] \leq e^C,$$

in other words we have shown that the sequence $\mathbb{1}_A \exp(z_n \int_0^1 g_s\, dW_s)$ is bounded in $L^2(\Omega)$. We can conclude by Vitali's theorem that the function $z \mapsto \mathbb{E}\left[\mathbb{1}_A e^{z\int_0^1 g_s\, dW_s}\right]$ is continuous.

Denote for semplicity

$$F(z) := \mathbb{E}\left[\mathbb{1}_A e^{z\int_0^1 g_s\, dW_s}\right] := \mathbb{E}\left[f(z,\omega)\right],$$

and

$$X = \int_0^1 g_s\, dW_s$$

We would need to show that the integral of $F(z)$ along every closed piecewise constant curve vanish. To do so we will need to use Fubini's theorem, so first of all let's see that we can properly use Fubini.

Let $\gamma : [a,b] \to \mathbb{C}$ be a closed piecewise regular path. Then

$$\mathbb{E}\int_a^b |\mathbb{1}_A e^{\gamma(t)X(\omega)}\gamma\prime(t)|dt \leq$$

$$\leq \mathbb{E}\int_a^b e^{\Re(\gamma(t))X(\omega)}\left|\gamma\prime(t)\right| dt \leq \mathbb{E}\int_a^b e^{RX(\omega)}R'\, dt \leq R'(b-a)\mathbb{E}[e^{RX(\omega)}] < \infty$$

being the last expectation finite by Corollary 3.1. In conclusion we use Fubini's theorem to find

$$\int_\gamma \mathbb{E}\left[f(z,\omega)\right]\, dz = \mathbb{E}\int_\gamma f(z,\omega)\, dz = 0$$

for *Cauchy*'s theorem (see [1] chapter 4), since $z \mapsto f(z,\omega) = \mathbb{1}_A \exp(zX(\omega))$ is holomorphic for all ω a.s.

Step 2 We treat the general case.

Let $z \in \mathbb{C}$ and let g_s, f_s be bounded real valued processes of class S. Fix $t_j > 0$, $j = 1, 2$. Then, as we did in the previous step we need to show that

$$F(z) := \mathbb{E}\left[\exp(\int_0^{t_j}(f_s + zg_s)\, dW_s) - \frac{1}{2}\int_0^{t_j}(f_s + zg_s)^2\, ds\right] =: \mathbb{E}\left[f(z,\omega)\right]$$

is continuous and that its line integral over a closed piecewise linear path vanish. Of course the function $z \mapsto f(z,\omega)$ is continuous (ω $a.s.$) as it is the exponential function composed to other continuous functions of z. So if $z_n \to z$, then $f(z_n,\omega) \to f(z,\omega)$. Now we are looking for a bound in L^2 as we did in the previous step. Let $M > 0$ such that $|f_s|, |g_s| \le M$ $\mathbb{P} \times dt$ $a.s.$ and let $|z_n| \le R$. Note that

$$\int_0^{t_j} 4(f_s + \Re(z_n)g_s)^2\, ds \le 4(M^2 + R^2M^2 + 2RM^2)t_j =: C \tag{3.2}$$

and

$$|f(z_n,\omega)|^2 = \exp(2\Re \int_0^{t_j} (f_s + z_n g_s)\, dW_s - \Re \int_0^{t_j} (f_s + z_n g_s)^2\, ds) =$$

$$= \exp(\int_0^{t_j} 2(f_s + \Re(z_n)g_s)\, dW_s) \cdot \exp(-\Re \int_0^{t_j} (f_s + z_n g_s)^2\, ds).$$

We can control directly the second exponential

$$\exp(-\Re \int_0^{t_j} (f_s + z_n g_s)^2\, ds) \le \exp(- \int_0^{t_j} (f_s^2 + \Re({z_n}^2)\, {g_s}^2 + 2g_s\, f_s\, \Re(z_n)\, ds)$$

$$\le e^{-t_j(M^2+R^2M^2+2RM^2)=e^{-C/4}}$$

Where we have used the inequality $\Re(z^2) \le R^2$ if $|z| \le R$. So, by Corollary 3.1 and inequality (3.2):

$$\mathbb{E}|f(z_n,\omega)|^2 \le \mathbb{E}\left[\exp(\int_0^{t_j} 2(f_s + \Re(z_n)g_s)\, dW_s) \cdot e^{-t_j(M^2+R^2M^2+M^2)}\right] \le$$

$$\le e^{-C/4}\, e^{C}.$$

We have shown that $f(z_n,\omega)$ is bounded in L^2, so applying Vitali's theorem we conclude that $z \mapsto F(z)$ is a continuous map. Now we need, again, to justify the use of Fubini's theorem. We need to show that if $\gamma : [a,b] \to \mathbb{C}$ is a closed piecewise linear curve, then

$$\int_a^b \mathbb{E}|f(\gamma(t),\omega)| \cdot |\gamma'(t)|\, dt < \infty$$

. Observe that

$$\int_0^{t_j} (f_s + \Re(z(t))g_s)^2\, ds \le C/4 \tag{3.3}$$

So

$$\mathbb{E}|f(\gamma(t),\omega)| = \mathbb{E}\left[\exp(\Re \int_0^{t_j} (f_s + \gamma(t)g_s)\, dW_s - \frac{1}{2}\Re \int_0^{t_j} (f_s + \gamma(t)g_s)^2\, ds)\right] \le$$

$$\leq e^{-C/8}\,\mathbb{E}\left[\exp\int_0^{t_j}\left(f_s+\Re(\gamma(t))g_s\right)dW_s\right]\leq e^{-C/8}e^{C/4}.$$

The last inequality follows from Corollary 3.1 and inequality (3.3). In conclusion

$$\int_a^b \mathbb{E}|f(\gamma(t),\omega)|\,|\gamma\prime(t)|\,dt\leq e^{C/8}R'(b-a)<\infty.$$

Now we use Fubini's theorem and conclude the proof as in step 1, after observing that for all ω (a.s.) the map $z\mapsto f(z,\omega)$ is holomorphic.

□

Lemma 3.3. *Let $Z_t\in S$ be a complex valued process, $T\in[0,\infty)$, then*

$$|e^{\int_0^T Z_s\,dW_s}|=e^{\int_0^T \Re(Z_s)\,dW_s},$$

and

$$|e^{\int_0^T Z_s\,ds}|=e^{\int_0^T \Re(Z_s)\,ds}.$$

We give the proof of this lemma in the case of stochastic integrals.

Proof. Let's write Z_s as $Z_s=X_s+iY_s$. By properties of the complex exponential function

$$|e^{\int_0^T Z_s\,dW_s}|=e^{\Re\int_0^T Z_s\,dW_s}.$$

We just need to remark that $\Re\int_0^T Z_s\,dW_s=\int_0^T\Re(Z_s)\,dWs$, see Definition 3.2.

□

Before proving Girsanov's theorem we quickly recall some of the most important notions about absolutely continuous measures. Let $(\Omega,\mathcal{F},\mathbb{P})$ be a probability space. We would like to define another probability measure $\tilde{\mathbb{P}}$ on the same space $(\Omega,\mathcal{F})$. One easy way to do that is to start with a random variable f defined on $(\Omega,\mathcal{F})$ such that:

$$f\geq 0\;\mathbb{P}-\text{a.s.}\quad\text{and}\quad\mathbb{E}[f]=1,$$

where $\mathbb{E}[\cdot]$ denotes the expectation operator with respect to P.
Then, we can obtain a probability measure $\tilde{\mathbb{P}}$ on $(\Omega,\mathcal{F},\mathbb{P})$ by setting for all $A\in\mathcal{F}$:

$$\tilde{\mathbb{P}}(A)=\mathbb{E}[f\mathbb{1}_A].$$

$\tilde{\mathbb{P}}$ is a probability measure on $(\Omega,\mathcal{F})$ because

$$\tilde{\mathbb{P}}(\Omega)=\mathbb{E}[f]=1.$$

Furthermore $\tilde{\mathbb{P}}(A)=0$ whenever $\mathbb{P}(A)=0$.
In this case we say that $\tilde{\mathbb{P}}$ is absolutely continuous with respect to $\mathbb{P}$, and denote it $\tilde{\mathbb{P}}<<\mathbb{P}$. In fact, the Radon-Nikodym theorem states that the above construction of $\tilde{\mathbb{P}}$ is the only way to obtain probability measures which are absolutely continuous with respect to $\mathbb{P}$, namely: If $\tilde{\mathbb{P}}$ is a probability measure on $(\Omega,\mathcal{F})$ and $\tilde{\mathbb{P}}<<\mathbb{P}$, then there exists a unique $(\mathbb{P}-\text{a.s.})$ random variable f on $(\Omega,\mathcal{F})$ such that $f\geq 0$ and $\mathbb{E}[f]=1$, for which $\tilde{\mathbb{P}}(A)=\mathbb{E}[f\mathbb{1}_A]$.
The random variable f is called the *Radon-Nikodym derivative* or *density* of $\tilde{\mathbb{P}}$ with respect to $\mathbb{P}$ and is often denoted as $f=\frac{d\tilde{\mathbb{P}}}{d\mathbb{P}}$.

Theorem 3.1 (Girsanov). *Let $T \in [0,\infty)$ and let $b \in S$ be a real valued process such that $\mathbb{E}\rho_T(b) = 1$. On the measurable space $(\Omega, \mathcal{F})$ introduce the measure $\tilde{\mathbb{P}}$ defined by:*

$$\tilde{\mathbb{P}}(d\omega) = \rho_T(b)(\omega)\mathbb{P}(d\omega).$$

Then $(\Omega, \mathcal{F}, \tilde{\mathbb{P}})$ is a probability space and $\xi_t = W_t - \int_0^t b_s\, ds$ is a Wiener process under $\tilde{\mathbb{P}}$ for $t \leq T$.

Proof. $(\Omega, \mathcal{F}, \tilde{\mathbb{P}})$ is a probability space since

$$\tilde{\mathbb{P}}(\Omega) = \int_\Omega \tilde{\mathbb{P}}(d\omega) = \int_\Omega \rho_T(b)(\omega)\mathbb{P}(d\omega) = \mathbb{E}\rho_T(b) = 1.$$

Consider $\xi_t = W_t - \int_0^t b_s\, ds$. Notice that $\xi_0 = 0$ and that ξ_t is continuous. Therefore, by Theorem 2.3, to show that ξ_t is a Wiener process we just need to show that the increments $(\xi_{t_n} - \xi_{t_{n-1}}, ..., \xi_{t_1} - \xi_{t_0})$ are independent and distributed as normal variables of mean 0 and variance equal to Δt_i. In other words, it suffices to show that the distribution of the vector $(\xi_{t_n} - \xi_{t_{n-1}}, ..., \xi_{t_1} - \xi_{t_0})$ with respect to $(\Omega, \mathcal{F}, \tilde{\mathbb{P}})$ is the same of the distribution of the vector of the increments of W_t with respect to $(\Omega, \mathcal{F}, \mathbb{P})$. To show this, we use the characteristic function, see Section 1.1.5.

Let $0 \leq t_0 \leq t_1 \leq \ldots \leq t_n = T$. Let $\lambda_j \in \mathbb{R}$, $j = 0, \ldots, n-1$, and define $\lambda_s = i\lambda_j$ on $[t_j, t_{j+1}]$, $j = 0, \ldots, n-1$. Also denote by $\tilde{\mathbb{E}}$ the expectation sign relative to $\tilde{\mathbb{P}}$. We need to show that

$$\tilde{\mathbb{E}}\Big[\exp\Big(i\sum_{j=0}^{n-1}\lambda_j(\xi_{t_{j+1}} - \xi_{t_j})\Big)\Big] = \exp(-\frac{1}{2}\sum_{j=0}^{n-1}|\lambda_j|^2(t_{j+1} - t_j)). \tag{3.4}$$

Step (1) : We prove equation (3.4) for a *bounded* process $b \in S$, so suppose b is bounded.

$$\tilde{\mathbb{E}}\Big[\exp\big(i\sum_{j=0}^{n-1}\lambda_j(\xi_{t_{j+1}} - \xi_{t_j})\big)\Big] = \mathbb{E}\Big[\exp\big(\sum_{j=0}^{n-1} i\lambda_j(\xi_{t_{j+1}} - \xi_{t_j})\big)\,\rho_T(b)\Big]$$

Observe that this quantity is integrable, in fact:

$$\big|\exp(\sum_{j=0}^{n-1} i\lambda_j(\xi_{t_{j+1}} - \xi_{t_j}))\,\rho_T(b)\big| = \Big|\prod_j \exp(i\lambda_j(\xi_{t_{j+1}} - \xi_{t_j}))\rho_T(b)\Big| \leq \rho_T(b) \in L^1.$$

$$\mathbb{E}\big[\exp(\sum_{j=0}^{n-1} i\lambda_j(\xi_{t_{j+1}} - \xi_{t_j}))\,\rho_T(b)\big] = \mathbb{E}\Big[\exp(\sum_{j=0}^{n-1} i\lambda_j(W_{tj+1} - \int_0^{t_{j+1}} b_s ds - W_{t_j} + \int_0^{t_j} b_s ds))\,\rho_T(b)\Big]$$

$$= \mathbb{E}\Big[\exp(\int_0^T \lambda_s dW_s - \sum_{j=0}^{n-1} i\lambda_j \int_{t_j}^{tj+1} b_s ds)\rho_T(b)\Big] = \mathbb{E}[\exp(\int_0^T \lambda_s dW_s - \int_0^T \lambda_s b_s ds)\rho_T(b)] =$$

$$= \mathbb{E}\Big[\exp\Big(\int_0^T \lambda_s dW_s - \int_0^T \lambda_s b_s ds\Big)\exp(\int_0^T b_s dW_s - \frac{1}{2}\int_0^T b_s^2 ds)\Big] =$$

$$= \mathbb{E}[\exp(\int_0^T (\lambda_s + b_s)dW_s - \int_0^T (\lambda_s b_s + \frac{1}{2}b_s^2)ds] =$$

$$= \mathbb{E}[\rho_T(\lambda + b)e^{\frac{1}{2}\int_0^T \lambda_s^2 ds}] = e^{\frac{1}{2}\int_0^T \lambda_s^2 ds};$$

where in the last equality we used the fact that if b is bounded, then $\lambda + b$ is bounded, and so $\mathbb{E}\rho_T(\lambda + b) = 1$ for Lemma 3.2.
Also notice that λ_s is purely imaginary, and so $\lambda_s^2 = -|\lambda_s|^2$.
Finally, we have

$$\tilde{\mathbb{E}}\Big[\exp(i\sum_{j=0}^{n-1}\lambda_j(\xi_{t_{j+1}} - \xi_{t_j}))\Big] = e^{\frac{1}{2}\int_0^T \lambda_s^2 ds} = \exp(\frac{1}{2}\sum_{j=0}^{n-1}(i\lambda_j)^2(t_{j+1} - t_j)) =$$

$$= \exp(-\frac{1}{2}\sum_{j=0}^{n-1}|\lambda_j|^2(t_{j+1} - t_j)).$$

We proved (3.4) in the case of a bounded b.

Step (2) : If b is not bounded take a sequence of bounded $b^n \in S$ such that (a.s.) $\int_0^T |b_s^n - b_s|^2 ds \longrightarrow 0$. Then, we can do the same proof as above and pass to the limit for $n \to \infty$. In fact

$$\mathbb{E}\rho_T(\lambda + b) = \lim_{n\to\infty} \mathbb{E}\rho_T(\lambda + b_n). \tag{3.5}$$

We prove now equality (3.5).

$$\Lambda_n = \mathbb{E}|\rho_T(\lambda + b^n) - \rho_T(\lambda + b)| =$$

$$= \mathbb{E}\Big[\Big|\exp(\int_0^T (\lambda_s + b_s^n)dW_s - \frac{1}{2}\int_0^T (\lambda_s^2 + (b_s^n)^2 + 2\lambda_s b_s^n)ds) -$$

$$- \exp(\int_0^T (\lambda_s + b_s)dWs - \frac{1}{2}\int_0^T (\lambda_s^2 + b_s^2 + 2\lambda_s b_s)ds)\Big|\Big] =$$

$$= e^{\frac{1}{2}\int_0^T |\lambda_s|^2 ds}\mathbb{E}\Big[\big|e^{\int_0^T \lambda_s\, dW_s}\big| \cdot \big|\rho_T(b^n)e^{-\int_0^T \lambda_s b_s^n ds} - \rho_T(b)e^{-\int_0^T \lambda_s b_s ds}\big|\Big]$$

Note that $\left|e^{\int_0^T \lambda_s dW_s}\right| = 1$ and $\left|e^{\int_0^T \lambda_s b_s^n ds}\right| = 1$ by Lemma 3.3, since $\Re(\lambda_s) = 0$ and $\Re(\lambda_s b_s^n) = 0$

$$\Lambda_n = e^{\frac{1}{2}\int_0^T |\lambda_s|^2 ds}\ \mathbb{E}\Big[\Big|\rho_T(b^n)e^{-\int_0^T \lambda_s b_s^n ds} + \rho_T(b)e^{-\int_0^T \lambda_s b_s^n ds} - \rho_T(b)e^{-\int_0^T \lambda_s b_s^n ds} - \rho_T(b)e^{-\int_0^T \lambda_s b_s ds}\Big|\Big] \le$$

$$\le e^{\frac{1}{2}\int_0^T |\lambda_s^2| ds}\ \mathbb{E}\Big[\big|\rho_T(b^n) - \rho_T(b)\big|\Big] + \mathbb{E}\Big[\big|e^{\int_0^T \lambda_s b_s^n ds} - e^{\int_0^T \lambda_s b_s ds}\big|\rho_T(b)\Big] \longrightarrow 0.$$

In fact:

- $\mathbb{E}|\rho_T(b^n) - \rho_T(b)| \to 0$ for Lemma 3.1 (iv).

- $\mathbb{E}|e^{\int_0^T \lambda_s b_s^n ds} - e^{\int_0^T \lambda_s b_s ds}|\, \rho_T(b)) \to 0$ by the dominated convergence theorem;
  in particular, for Lemma 3.3 again $|e^{\int_0^T \lambda_s b_s^n ds}| \leq 1$ and
  $$\int_0^T |\lambda_s b_s^n - \lambda_s b_s| ds \leq \max_j |\lambda_j| \int_0^T |b_s^n - b_s| ds \longrightarrow 0 \ (a.s.).$$
  Finally,
  $$e^{\int_0^T \lambda_s b_s^n ds} \to e^{\int_0^T \lambda_s b_s ds} \ (a.s.)$$
  by the continuity of the (complex) exponential function.

□

Let us show the application of Girsanov's theorem to finding
$$\mathbb{P}\left(\max_{t\leq 1}(W_t + t) \geq 1\right),$$
where W_t is a Wiener process defined on $(\Omega, \mathcal{F}, \mathbb{P})$. Let $b = -1$ and on $(\Omega, \mathcal{F})$ introduce the new measure
$$\tilde{\mathbb{P}}(d\omega) = e^{-w_1 - 1/2}\mathbb{P}(d\omega).$$

By Girsanov's theorem, $\tilde{W}_t := W_t + t$ is a Wiener process for $t \in [0, 1]$.

$$\mathbb{P}\Big(\max_{t\leq 1}(W_t + t) \geq 1\Big) = \mathbb{E}\Big[\mathbb{1}_{\{\max_{t\leq 1}(W_t+t)\geq 1\}}\Big] = \tilde{\mathbb{E}}\Big[\mathbb{1}_{\{\max_{t\leq 1}(W_t+t)\geq 1\}}\frac{d\mathbb{P}}{d\tilde{\mathbb{P}}}\Big].$$
$$= \tilde{\mathbb{E}}\left[\mathbb{1}_{\max_{t\leq 1}(W_t+t)\geq 1} \cdot e^{W_1 + \frac{1}{2}}\right] = \int_\Omega \mathbb{1}_{\max_{t\leq 1}(W_t+t)\geq 1} e^{W_1+\frac{1}{2}}\tilde{\mathbb{P}}(d\omega)$$
$$= \int_\Omega \mathbb{1}_{\max_{t\leq 1}\tilde{W}_t \geq 1} e^{\tilde{W}_1 - 1/2}\tilde{\mathbb{P}}(dw) = \mathbb{E}\left[\mathbb{1}_{\max_{t\leq 1} W_t \geq 1} e^{W_1 - 1/2}\right].$$

Where in the last identity we used the fact that $\tilde{W}_t$ is a Wiener process under $\tilde{\mathbb{P}}$ and W_t is a Wiener process under $\mathbb{P}$.

Then we can evaluate the last expectation using Lemma 2.2 which we recall now:
$$\mathbb{P}\left(\max_{0\leq t\leq 1} W_t \geq 1,\ W_1 \leq x\right) = \begin{cases} \mathbb{P}(W_1 \geq 2 - x) & \text{if } x \leq 1 \\ 2\mathbb{P}(W_1 \geq 1) - P(W_1 \geq x) & \text{if } x \geq 1 \end{cases} \tag{1}$$
and the following result proved in [14] Sec. 6.8. If there exists $f \geq 0, \quad f \in L^1(\mathbb{R})$ such that $\mathbb{P}(\xi \leq 1, \eta \leq b) = \int_{-\infty}^b f(x)\, dx$ for every b, then:
$$\mathbb{E}[g(\eta) \cdot \mathbb{1}_{\{\xi \leq 1\}}] = \int_\mathbb{R} g(x) f(x)\, dx.$$

We apply this to:
$$\mathbb{E}\left[\mathbb{1}_{\{\max_{t\leq 1} W_t \geq 1\}} \cdot e^{W_1 - \frac{1}{2}}\right]$$

with:

$$\begin{aligned}\xi &= \max_{t\leq 1} W_t\\ \eta &= W_1\\ g(x) &= e^{x-\frac{1}{2}}\end{aligned}$$

Now we differentiate the joint distribution function (1) to find the density $f(x)$.

If $x < 1$:

$$\mathbb{P}(\max W_t \geq 1,\ W_1 \leq x) = \mathbb{P}(W_1 \geq 2-x) = \int_{2-x}^{\infty} \phi(y)\, dy\,, \quad \phi(y) = \frac{1}{\sqrt{2\pi}} e^{-y^2/2}$$

Then:

$$\frac{d}{dx}\mathbb{P}(W_1 \geq 2-x) = \phi(2-x)$$

If $x \geq 1$:

$$\mathbb{P}(\max W_t \geq 1,\ W_1 \leq x) = 2\mathbb{P}(W_1 \geq 1) - \mathbb{P}(W_1 \geq x) = \text{const} - \int_x^{\infty} \phi(y)\, dy$$

Then:

$$\frac{d}{dx} = \phi(x).$$

In conclusion

$$\begin{aligned}\mathbb{E}\left[\mathbb{1}_{\{\max W_t \geq 1\}} \cdot e^{W_1-\frac{1}{2}}\right] &= \int_{-\infty}^{1} e^{x-\frac{1}{2}}\phi(2-x)\, dx + \int_1^{\infty} e^{x-\frac{1}{2}}\phi(x)\, dx\\ &= \int_1^{\infty} e^{x-1/2} \cdot \frac{1}{\sqrt{2\pi}} e^{-x^2/2}\, dx + \int_{-\infty}^{1} e^{x-1/2} \cdot \frac{1}{\sqrt{2\pi}} e^{-(2-x)^2/2}\, dx\\ &= \frac{1}{\sqrt{2\pi e}} \int_1^{\infty} \left(e^x + e^{2-x}\right) e^{-x^2/2}\, dx\end{aligned}$$

3.1 Novikov condition

In this section b is a process of class S. After having proved the Girsanov theorem we are interested in finding sufficient conditions that ensure $\rho_t = \rho_t(b)$ to be a martingale.

Lemma 3.4. *If $\mathbb{E}\sup_{t\leq T} \rho_t < \infty$, then ρ_t is a martingale*

Proof. We proceed like in the proof of Lemma 3.1 and define

$$\tau_n := \inf\{t \in [0,T] : \int_0^t |b_s|^2|\rho_s|^2 ds \geq n\}.$$

Thus, $\int_0^t \mathbb{1}_{s<\tau_n} b_s\rho_s\, dW_s$ and $\rho_{t\wedge\tau_n}$ are martingales. So, for $t_2 > t_1$ we have

$$\mathbb{E}[\rho_{t_2\wedge\tau_n}|\mathcal{F}_{t_1}] = \rho_{t_1\wedge\tau_n}. \tag{3.6}$$

Passing to the limit for $n \to \infty$ in (3.6) we come to the thesis. In fact the right hand side tends to ρ_{t_1}, the left member tends to $\mathbb{E}[\rho_{t_2}|\mathcal{F}_1]$ by the dominated convergence theorem, since $\rho_{t_2\wedge\tau_n} \leq \sup_{t\leq T} \rho_t \in L^1$ by our hypothesis. □

Lemma 3.5. *If there exist $p > 1$, $N < \infty$ such that $\mathbb{E}\rho_\tau^p \leq N$ for every stopping time $\tau \leq T$, then $(\rho_t, \mathcal{F}_t)$ is a martingale for $t \in [0,T]$.*

Proof. By the previous Lemma it is sufficient to prove that

$$\mathbb{E} \sup_{t \leq T} \rho_t < \infty.$$

Fix $a > 0$. We define the stopping time

$$\hat{\tau}_a := \inf\{t \geq 0 : \rho_t \geq a\} \in [0, \infty].$$

Set $\tau := \hat{\tau}_a \wedge T$, so $\tau \leq T$ is a stopping time. On the event $\{\hat{\tau}_a \leq T\}$ we have $\tau = \hat{\tau}_a$ and hence $\rho_\tau = \rho_{\hat{\tau}_a} \geq a$, therefore:

$$\rho_\tau^p \ \geq \ a^p \, \mathbf{1}_{\{\hat{\tau}_a \leq T\}} \quad \Longrightarrow \quad a^p \, \mathbb{P}(\hat{\tau}_a \leq T) \ \leq \ \mathbb{E}[\rho_\tau^p] \ \leq \ N.$$

Since $\{\hat{\tau}_a \leq T\} = \{\sup_{t \leq T} \rho_t \geq a\}$, we obtain

$$\mathbb{P}\left(\sup_{t \leq T} \rho_t \geq a \right) \ \leq \ \frac{N}{a^p}, \qquad a > 0. \tag{3.7}$$

Now we apply a result from measure theory (see [18] theorem 8.16), in particular for a non-negative random variable X the following identity holds:

$$\mathbb{E}[X] = \int_0^\infty \mathbb{P}(X \geq a) \, da$$

We apply this to $X = \sup_{t \leq T} \rho_t$, then we split the integral and use (3.7) for $a \geq 1$:

$$\mathbb{E}\left[\sup_{t \leq T} \rho_t \right] \ = \ \int_0^1 \mathbb{P}\left(\sup_{t \leq T} \rho_t \geq a \right) \, da \ + \ \int_1^\infty \mathbb{P}\left(\sup_{t \leq T} \rho_t \geq a \right) \, da \ \leq \ 1 + \int_1^\infty \frac{N}{a^p} \, da.$$

Because $p > 1$, the improper converges and equals to $N/(p-1)$, hence

$$\mathbb{E}\left[\sup_{t \leq T} \rho_t \right] \ \leq \ 1 + \frac{N}{p-1} \ < \ \infty.$$

This proves the claim. ☐

We now state a classical result due to Novikov: a simple criterion ensuring that the exponential martingale associated with b is in fact a martingale.

Theorem 3.2 (Novikov)**.** *Let $b \in S$. If*

$$\mathbb{E} \exp\left(\frac{1}{2} \int_0^T |b_t|^2 \, dt \right) < \infty,$$

then $(\rho_t)_{t \in [0,T]}$ is a martingale on $[0,T]$

We prove now a sufficient condition that is stronger than Novikov's (it implies Novikov's).

Proposition 3.1. *If*

$$\mathbb{E}\left[\exp\left(\int_0^T c\,|b_t|^2\,dt\right)\right] < \infty \tag{3.8}$$

for a constant $c > \frac{1}{2}$*,then* $\mathbb{E}\big[\rho_T(b)\big] = 1$.

Proof. Let $c > \frac{1}{2}$ such that (3.8) holds. Let τ be a stopping time, $\tau \le T$. Let $p > 1$, let r and r' be conjugates exponents (i.e. $1/r + 1/r' = 1$), then a simple calculation gives

$$\rho_\tau^p = \exp(p\int_0^\tau b_s\,dW_s)\cdot\exp(\frac{-p}{2}\int_0^\tau b_s\,ds) = \rho_\tau^{1/r}(pr\cdot b)\cdot\exp(\frac{1}{2}(p^2r-p)\int_0^\tau b_s^2\,ds)$$

Using Holder's inequality we get

$$\mathbb{E}[\rho_\tau^p] \le (\mathbb{E}[\rho_\tau(pr\cdot b)])^{1/r}\cdot(\mathbb{E}[\exp(\frac{1}{2}r'(p^2r-p)\int_0^\tau b_s^2\,ds)])^{1/r'}$$

Now note $\int_0^\tau b_s^2\,ds \le \int_0^T b_s^2\,ds$ and that $\mathbb{E}[\rho_\tau(pr\cdot b) \le 1]$ since S is closed under multiplication by constants, thus $b \in S$ implies $pr\cdot b \in S$, and therefore the process $\rho(pr\cdot b)$ is a a supermartingale (see Lemma 3.1 (ii)). Thus,

$$\mathbb{E}[\rho_\tau^p] \le (\mathbb{E}[\exp(\frac{1}{2}r'(p^2r-p)\int_0^\tau b_s^2\,ds])^{1/r'}.$$

Consider the function

$$k(p,r) = \frac{1}{2}r'(p^2r-p).$$

Note that if $p \to 1^+$, then $k(1,r) = \frac{1}{2}r'(r-1) = \frac{1}{2}r$.

Fix $r \in (1, 2c)$ and let $r' = \frac{r}{r-1}$ be the conjugate exponent so that $k(1,r) = \frac{1}{2}r < c$. By continuity of the function $k(\cdot, r)$ there exist $p > 1$ sufficiently close to 1 such that $k(p,r) < c$. With that choice we have

$$\sup_{\tau\le T}\mathbb{E}[\rho_\tau^p] \le (\mathbb{E}[\exp(c\int_0^T b_s^2\,ds)])^{1/r'} =: N,$$

then we conclude by Lemma 3.5 $\square$

Now we are going to improve Proposition 3.1 with the weaker condition

$$\lim_{\varepsilon\downarrow 0}\ \varepsilon\,\ln\mathbb{E}\left[\exp\left(\frac{1-\varepsilon}{2}\int_0^T |b_t|^2\,dt\right)\right] = 0. \tag{3.9}$$

We learn condition (3.9) by the article of Krylov [15]. It is an improvement of Novikov. In particular if Novikov condition holds, then (3.9) holds, since if $X := \int_0^T |b_t|^2\,dt \ge 0$. For any $\varepsilon \in (0, 1]$,

$$1 \ \le\ \mathbb{E}e^{\frac{1-\varepsilon}{2}X} \ \le\ \mathbb{E}e^{\frac{1}{2}X} = K < \infty,$$

since $\frac{1-\varepsilon}{2}X \ge 0$ and $x \mapsto e^{\alpha x}$ is increasing for $\alpha \ge 0$. Taking logs and multiplying by ε gives

$$0 \ \le\ \varepsilon\ln\mathbb{E}e^{\frac{1-\varepsilon}{2}X} \ \le\ \varepsilon\ln K \xrightarrow[\varepsilon\downarrow 0]{} 0.$$

Remark 3.1. *Condition* (3.9) *immediately implies that there exists* $\varepsilon_0 \in (0,1)$ *such that*

$$\sup_{0<\epsilon\leq\epsilon_0} \mathbb{E}e^{\frac{1-\epsilon}{2}X} < \infty.$$

Proposition 3.2. *Assume that 3.9 holds, then* $\mathbb{E}[\rho_T(b)] = 1$.

Proof. Let $\varepsilon \in (0,\varepsilon_0)$ (see the previous Remark). By Proposition 3.1 applied to $(1-\varepsilon)b$ (choose any $c_\varepsilon \in \big(\frac{1}{2}, \frac{1}{2(1-\varepsilon)}\big]$ so that $c_\varepsilon(1-\varepsilon)^2 \leq \frac{1-\varepsilon}{2}$ and $\mathbb{E}e^{c_\varepsilon \int_0^T |(1-\varepsilon)b_s|^2\,ds} \leq \mathbb{E}e^{\frac{1-\varepsilon}{2}\int_0^T |b_s|^2 ds} < \infty$), we have

$$\mathbb{E}\big[\rho_T((1-\varepsilon)b)\big] = 1.$$

Note that

$$\rho_T((1-\varepsilon)b) = \exp(\int_0^T (1-\varepsilon)b_s\,dW_s - \frac{1}{2}\int_0^T (1-\varepsilon)^2|b_s|^2\,ds)$$

and

$$\rho_T(b)^{1-\varepsilon} = \exp(\int_0^T (1-\varepsilon)b_s\,dW_s - \frac{1}{2}\int_0^T (1-\varepsilon)|b_s|^2\,ds)$$

Thus

$$\rho_T((1-\varepsilon)b) = \rho_T(b)^{\,1-\varepsilon}\exp\Big(\tfrac{1}{2}\,\varepsilon(1-\varepsilon)\int_0^T |b_t|^2\,dt\Big).$$

We apply the Hölder inequality with conjugate exponents $p = \frac{1}{1-\varepsilon}$ and $q = \frac{1}{\varepsilon}$:

$$1 = \mathbb{E}\big[\rho_T((1-\varepsilon)b)\big] \leq \big(\mathbb{E}[\rho_T(b)]\big)^{1-\varepsilon}\Big(\mathbb{E}\exp\Big(\frac{1-\varepsilon}{2}\int_0^T |b_t|^2\,dt\Big)\Big)^{\varepsilon}.$$

Taking logarithms gives

$$0 \leq (1-\varepsilon)\,\ln\mathbb{E}[\rho_T(b)] \;+\; \varepsilon\,\ln\mathbb{E}\exp\Big(\frac{1-\varepsilon}{2}\int_0^T |b_t|^2\,dt\Big).$$

Hence

$$\ln\mathbb{E}[\rho_T(b)] \;\geq\; -\frac{\varepsilon}{1-\varepsilon}\,\ln\mathbb{E}\exp\Big(\frac{1-\varepsilon}{2}\int_0^T |b_t|^2\,dt\Big).$$

Letting $\varepsilon \downarrow 0$ and using (3.9) yields $\ln\mathbb{E}[\rho_T(b)] \geq 0$, i.e. $\mathbb{E}[\rho_T(b)] \geq 1$. Since $\rho_t(b)$ is a supermartingale, $\mathbb{E}[\rho_T(b)] \leq \mathbb{E}[\rho_0(b)] = 1$. Thus $\mathbb{E}[\rho_T(b)] = 1$. □

The results above highlight the central role of exponential integrability in ensuring that the exponential martingale ρ is martingale. Novikov's condition is the classical sufficient criterion. Stronger assumptions like the one in Proposition 3.1 automatically imply Novikov, while Krylov's condition relaxes it and extends its applicability.

3.1.1 A consequence of Novikov's theorem

As a direct consequence of Novikov's condition we give the proof of another result. This is another sufficient condition for $\rho_t(b)$ to be a martingale but the integrability assumption is weakened as follows.

Corollary 3.2. *Let $b \in S$. If there exists a sequence of real numbers $\{t_n\}_{n=0}^{\infty}$ with*

$$0 = t_0 < t_1 < t_2 < \cdots < t_n \uparrow T,$$

such that for all $n \geq 1$,

$$\mathbb{E}\left[\exp\left(\frac{1}{2}\int_{t_{n-1}}^{t_n} |b_s|^2\, ds\right)\right] < \infty. \tag{3.10}$$

Then

$$\rho_t(b) \;=\; \exp\left(\int_0^t b_s\, dW_s \;-\; \frac{1}{2}\int_0^t b_s^2\, ds\right), \quad t \in [0,T].$$

is a martingale.

Proof. For $n \geq 1$, let $b_t^n := b_t \mathbb{1}_{[t_{n-1},t_n]}$. Then by Novikov condition, the exponential martingale

$$\rho_t(b^n) \;=\; \exp\left(\int_{t_{n-1}}^t b_s\, dW_s - \frac{1}{2}\int_{t_{n-1}}^t |b_s|^2\, ds\right), \qquad t_{n-1} \leq t \leq t_n,$$

is a martingale. In particular,

$$\mathbb{E}\big[\rho_{t_n}(b^n) \,\big|\, \mathcal{F}_{t_{n-1}}\big] = \rho_{t_{n-1}}(b^n) = 1, \qquad n \geq 1.$$

Note that for all $n \geq 1$:

$$\rho_{t_n}(b) = \rho_{t_{n-1}}(b) \cdot \rho_{t_n}(b^n).$$

But then,

$$\mathbb{E}[\rho_{t_n}(b)] = \mathbb{E}[\rho_{t_{n-1}}(b) \cdot \rho_{t_n}(b^n)] = \mathbb{E}\Big[\rho_{t_{n-1}}(b)\mathbb{E}[\rho_{t_n}(b^n) \mid \mathcal{F}_{t_{n-1}}]\Big] = \mathbb{E}[\rho_{t_{n-1}}(b)].$$

In the second equality we used properties of the conditional expectation and the fact that ρ is an adapted process. By induction on n, since $\mathbb{E}[\rho_{t_0}] = \mathbb{E}[\rho_0] = 1$ and $\mathbb{E}[\rho_t(b)]$ is nonincreasing in t, this shows that

$$\mathbb{E}[\rho_{t_n}(b)] = 1 \qquad \text{for all } n \geq 1.$$

Since $\mathbb{E}[\rho_t(b)]$ is nonincreasing in t and $\lim_{n\to\infty} t_n = \infty$, we conclude that

$$\mathbb{E}[\rho_t(b)] = 1, \qquad \forall\, t \in [0,T],$$

which proves the result. □

Corollary 3.3. *Let W be a one dimensional Wiener process defined on a stochastic basis $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, \mathbb{P})$. Let $x : [0,\infty) \times \mathbb{R} \to \mathbb{R}$ be a Borel measurable function and, for each $T > 0$, suppose there exists $K_T > 0$ such that*

$$|x(t,\xi)| \leq K_T(1 + |\xi|), \qquad \text{for all } \ 0 \leq t \leq T, \ \xi \in \mathbb{R}. \tag{3.11}$$

Define the process

$$b_t(\omega) := x(t, W_t(\omega)), \qquad t \geq 0,$$

Then $\rho_t(b)_{t\leq T}$ is a martingale.

We follow here the argument of Karatzas and Shreve [13] adding some details, the proof is based on the previous corollary and Doob's L^p inequality, which we recall now: Let $(M_t)_{0\leq t\leq T}$ be a nonnegative continuous submartingale and let $p > 1$. Then

$$\mathbb{E}\left[\left(\sup_{0\leq t\leq T} M_t\right)^p\right] \leq \left(\frac{p}{p-1}\right)^p \mathbb{E}[M_T^p].$$

Proof. Fix $T > 0$. First of all, observe that $b \in S$, in fact it is $\mathcal{F}_t$-adapted since for fixed t, W_t is $\mathcal{F}_t$–measurable and $x(t,\cdot)$ is Borel; hence $b_t = x(t, W_t)$ is $\mathcal{F}_t$–measurable. The process b is also jointly-measurable: since W has continuous paths, by Proposition 1.1 the mapping $(\omega, t) \mapsto W_t(\omega)$ is $\mathcal{F} \otimes \mathcal{B}([0,\infty))$–measurable. Therefore the map

$$\Phi : \Omega \times [0,\infty) \to [0,\infty) \times \mathbb{R}, \qquad \Phi(\omega, t) := (t, W_t(\omega)),$$

is $\mathcal{F} \otimes \mathcal{B}([0,\infty))$ to $\mathcal{B}([0,\infty)) \otimes \mathcal{B}(\mathbb{R})$ measurable. Composing with the Borel map x gives $b(\omega,t) = x \circ \Phi(\omega,t)$ $\mathcal{F} \otimes \mathcal{B}([0,\infty))$–measurable

Moreover, by (3.11) and the continuity of W,

$$\int_0^T b_s^2 \, ds \leq T K_T^2 \big(1 + (W_T^*)^2\big) < \infty \quad \text{a.s.}$$

where

$$W_T^* := \max_{0\leq s\leq T} |W_s|.$$

We would like to construct a sequence

$$0 = t_0 < t_1 < \cdots < t_{n(T)} = T$$

such that condition (3.10) holds.

From (3.11), for $0 \leq t_{n-1} < t_n \leq T$ we have

$$\int_{t_{n-1}}^{t_n} |b_s|^2 \, ds \ \leq \ (t_n - t_{n-1}) K_T^2 \big(1 + W_T^*\big)^2.$$

The process

$$Y_t := \exp\big((t_n - t_{n-1}) K_T^2 \big(1 + |W_t|^2\big)/4\big)\,, \qquad 0 \leq t \leq T,$$

is a nonnegative submartingale with continuous paths. In fact

$$\phi(x) = \exp\big(\tfrac{1}{4}(t_n - t_{n-1}) K_T^2 (1 + x)\big)\,, \qquad x \geq 0,$$

is convex and increasing, and $|W_t|^2$ is a submartingale by Jensen's inequality and since W_t is a martingale. By Jensen inequality, if M_t is a submartingale and ϕ is convex and increasing, then $\phi(M_t)$ is again a submartingale.

Then by Doob's L^2-inequality we obtain

$$\mathbb{E}\exp\Bigg(\tfrac{1}{2}(t_n - t_{n-1})K_T^2\big(1+(W_T^*)^2\big)\Bigg) = \mathbb{E}\Bigg(\max_{0\le t\le T} Y_t^2\Bigg) \;\le\; 4\mathbb{E}[Y_T^2].$$

The right-hand side is finite provided that

$$t_n - t_{n-1} \;\le\; \frac{1}{TK_T^2}.$$

In fact, since $W_T \sim \mathcal{N}(0,T)$, for $\lambda \in \mathbb{R}$ we compute

$$\mathbb{E}\Big[e^{\lambda W_T^2}\Big] = \frac{1}{\sqrt{2\pi T}}\int_{\mathbb{R}} \exp\Big(\lambda x^2 - \frac{x^2}{2T}\Big)\, dx = \frac{1}{\sqrt{2\pi T}}\int_{\mathbb{R}} \exp\Big(-\Big(\frac{1}{2T}-\lambda\Big)x^2\Big)\, dx.$$

The last integral is finite if and only if $\frac{1}{2T} - \lambda > 0$, i.e. $\lambda < \frac{1}{2T}$. This covers our case with $\lambda = \frac{1}{2}(t_n - t_{n-1})K_T^2$, indeed the condition $t_n - t_{n-1} \le \frac{1}{TK_T^2}$.

Thus we can construct the partition $\{t_0, t_1, \dots, t_{n(T)}\}$ satisfying the assumption of Corollary 3.2. This completes the proof. □

Chapter 4

Applications to the theory of SDEs

In this chapter we present some applications, including recent developments, of Girsanov's theorem to the theory of stochastic differential equations. For simplicity, we restrict our attention to the one-dimensional setting, although the multidimensional case is equally important and widely treated in the literature. For a deeper treatment of the theory of SDEs, and the multidimensional case, we refer the reader to Karatzas and Shreve [13].

The first section, called *Weak existence and uniqueness by Girsanov's theorem*, is more detailed and we present all proofs in detail. Its purpose is to show how Girsanov's theorem provides an effective tool for constructing weak solutions and studying their uniqueness in law.

In the next two sections, called *Pathwise uniqueness* and *Path-by-path uniqueness*, we do not present complete proofs, but rather describe the main ideas and state some important recent results, referring to the literature for the technical details. This choice is motivated both by the length and complexity of the arguments and by the fact that these topics include more recent developments, especially in the case of path-by-path uniqueness and, in particular, Davie's theorem. For these we refer again to Karatzas and Shreve [13] and to the original articles.

Let's introduce the topic by recalling a known fact from the theory of ODEs. From the theory of ODEs, in fact, it is known that the equation $dx_t = b(t, x_t)dt$ in general does not have a solution for any bounded Borel b. We show this with a particular example.

Consider the ordinary differential equation

$$\frac{dx}{dt} = b(x), \quad x(0) = 0 \tag{4.1}$$

where the function $b(x)$ is defined as

$$b(x) = \begin{cases} 1, & x \leq 0 \\ -1, & x > 0 \end{cases} \tag{4.2}$$

We are proving that there are no solutions, that is it does not exist a continuous function $x(t)$ such that

$$x(t) = \int_0^t b(x(s))\, ds, \quad \forall\, t \in [0, T].$$

Proof. Suppose that a solution $x(t)$, $t \in [0,T]$ exists.

1. If there exists $s \in [0,T]$ such that $x(s) > 0$, then by continuity there exists $[a,b]$, with $0 < a < s < b < T$, such that $x(t) > 0$ for all $t \in [a,b]$.
   So, for $t \in [a,b]$:

   $$x(t) = \int_0^a b(x(s))\,ds + \int_a^t b(x(s))\,ds = C_a - (t-a).$$

   Then $x(a) = C_a > 0$, so by continuity there exist $\delta > 0$, such that $x(t) > 0$ for $t \in [a-\delta, a+\delta]$. Then, for $t \in [a-\delta, a]$:

   $$x(t) = \int_0^{a-\delta} b(x(s))\,ds + \int_{a-\delta}^t b(x(s))\,ds = x(a-\delta) - (t-a+\delta).$$

   But $x(a) = C_a$, so $x(a-\delta) - \delta = C_a$, and thus $x(t) = C_a - t + a$, $t \in [a-\delta, a]$.
   Proceeding we come to

   $$a_{min} = \inf\{\delta > 0 : x(t) = C_a + a - t,\, t \in [\delta, b]\}.$$

   (a) If $a_{min} = 0$ we come to a contradiction, in fact $x(t) = C_a - t + a$ for $t \in [0,b]$, so $x(0) = C_a + a > 0$ which contradicts the initial condition $x(0) = 0$.

   (b) If $a_{min} > 0$ we come again to a contradiction, in fact: There exists $\delta_n \to a_{min}^+$ such that, for all n, $x(t) = C_a - t + a$ for $t \in [\delta_n, b]$, then by continuity

   $$x(a_{min}) = \lim_{\delta_n \to a_{min}^+} x(\delta_n) = C_a - a_{min} + a.$$

   So $x(a_{min}) = C_a + a - a_{min} > 0$. Thus there exist $\epsilon > 0$ such that $x(t) > 0$ for $t \in [a_{min} - \epsilon, a_{min}]$. Then, for $t \in [a_{min} - \epsilon, a_{min}]$:

   $$x(t) = \int_0^{a_{min}-\epsilon} b(x(s))\,ds + \int_{a_{min}-\epsilon}^t b(x(s))\,ds = x(a_{min}-\epsilon) - t + a_{min} - \epsilon. \tag{1}$$

   But $x(a_{min}) = C_a - a_{min} + a$, and so $x(a_{min}) = x(a_{min} - \epsilon) - \epsilon = C_a - a_{min} + a$.
   Substituting what we obtained in (1) we find that for all $t \in [a_{min} - \epsilon, a_{min}]$:

   $$x(t) = C_a - a_{min} + a - t + a_{min} = C_a + a - t,$$

   which contradicts the definition of a_{min}.

2. If there exists $s \in [0,T]$ such that $x(s) < 0$ we proceed in the same way we did in (1).

In other words we have shown that equation (4.1) does not admit a solution.

□

4.1 Weak existence and uniqueness by Girsanov's theorem

We now apply Girsanov's theorem to show that the stochastic differential equation with bounded Borel drift $b : [0, \infty) \times \mathbb{R} \to \mathbb{R}$

$$dY_t = b(t, Y_t + W_t)\, dt, \quad Y_0 = x \in \mathbb{R}. \tag{4.3}$$

admits a "weak" solution. This is equivalent to

$$X_t = x + \int_0^t b(s, X_s)\, ds \; + W_t \tag{4.4}$$

The precise formulation is given in Theorem 4.1 below. We give now the definition of weak solution.

Definition 4.1. *Let $b : [0, \infty) \times \mathbb{R} \to \mathbb{R}$ be a Borel-measurable function. Let us fix $T > 0$ and $x \in \mathbb{R}$. A* weak solution *of equation* (4.4) *is a triple (X, W), $(\Omega, \mathcal{F}, \mathbb{P})$, $\{\mathcal{F}_t\}$, where*

(i) $(\Omega, \mathcal{F}, \mathbb{P})$ is a probability space, and $\{\mathcal{F}_t\}$ is a filtration of sub-σ-algebras of $\mathcal{F}$;

(ii) $X = \{X_t, \mathcal{F}_t;\ 0 \le t \le T\}$ is a continuous, adapted $\mathbb{R}$-valued process, $W = \{W_t, \mathcal{F}_t;\ 0 \le t < \infty\}$ is a Wiener process;

(iii) $\mathbb{P}[\int_0^t |b(s, X_s)|\, ds < \infty] = 1$ for all $0 \le t \le T$;

(iv) $X_t = x + \int_0^t b(s, X_s)\, ds \; + W_t$ holds almost surely.

The same terminology will be used for analogous equations, replacing (*iii*) and (*iv*) with the corresponding integrability condition and integral equation.

This kind of solution is called weak because we do not fix the probability space and the Wiener process W in advance. We present a first result. It will be generalized in Proposition 4.1.

Theorem 4.1. *Let $b(t, x)$ be an $\mathbb{R}$-valued Borel bounded function on $(0, \infty) \times \mathbb{R}$. Then, for every $x \in \mathbb{R}$, there exists a stochastic basis*

$$(\Omega, \mathcal{F}, (\mathcal{F}_t)_{0 \le t \le T}, \mathbb{P}),$$

a continuous $(\mathcal{F}_t)$-adapted process $\xi = (\xi_t)_{0 \le t \le T}$, and a Wiener process $W = (W_t)_{0 \le t \le T}$ with respect to $(\mathcal{F}_t)$, such that

$$\xi_t = x + \int_0^t b(s, \xi_s)\, ds + W_t, \qquad 0 \le t \le T. \tag{4.5}$$

In particular, equation (4.4) *admits a weak solution.*

Proof. Let $(\Omega, \mathcal{F}, \tilde{\mathbb{P}})$ be a complete probability space and let ξ_t be a Wiener process defined on $(\Omega, \mathcal{F}, \tilde{\mathbb{P}})$. Define, for $x \in \mathbb{R}$:

$$W_t = \xi_t - \int_0^t b(s, \xi_s)\, ds\, - x$$

and on $(\Omega, \mathcal{F})$ introduce a new measure $\mathbb{P}$ by the formula

$$\mathbb{P}(d\omega) = \exp\left(\int_0^T b(s, \xi_s)\, d\xi_s - \frac{1}{2}\int_0^T |b(s, \xi_s)|^2\, ds\right)\tilde{\mathbb{P}}(d\omega).$$

Take

$$\mathcal{F}_t = \sigma(\xi_s : s \le t).$$

Then, by Lemma 3.1 (note that b is bounded) and Girsanov's theorem $(\Omega, \mathcal{F}, \mathbb{P})$ is a probability space, W_t is a Wiener process on $(\Omega, \mathcal{F}, \mathbb{P})$ for $t \in [0, T]$ with respect to $\mathcal{F}_t$, and, by definition, ξ_t solves (4.5). Since b is bounded, we also have

$$\mathbb{P}\Big[\int_0^t |b(s, \xi_s)|\, ds < \infty\Big] = 1, \qquad 0 \le t \le T.$$

In conclusion, we have proved that (ξ, W), $(\Omega, \mathcal{F}, \mathbb{P})$, $(\mathcal{F}_t)$ is a weak solution of equation (4.5). □

The solution given by Girsanov's theorem is a weak solution. Girsanov gives: a stochastic basis $(\Omega, \mathcal{F}, \mathbb{P}, \mathcal{F}_t)$, a Wiener process (W_t) on that stochastic basis, and a (continuous) process ξ such that under $\mathbb{P}$

$$\xi_t = x + \int_0^t b(s, \xi_s)\, ds + W_t \quad a.s.$$

We proved that if b is bounded and Borel measurable, then Girsanov's theorem gives the existence of a weak solution to equation (4.4). More generally the existence of a solution in the weak sense can be proved also if b grows at most linearly.

Proposition 4.1. *Consider the stochastic differential equation*

$$dX_t = b(t, X_t)\, dt + dW_t; \quad 0 \le t \le T, \quad X_0 = x. \tag{4.6}$$

where T is a fixed positive number, $x \in \mathbb{R}$, W is a Wiener process and $b(t, x)$ is a Borel measurable, $\mathbb{R}$-valued function defined on $[0, T] \times \mathbb{R}$ satisfying

$$|b(t, x)| \le K\,(1 + |x|), \qquad 0 \le t \le T,\ x \in \mathbb{R};$$

for some $K > 0$. Then, for any $x \in \mathbb{R}$ equation (4.6) has a weak solution.

Proof. Let $(B_t)_{t\ge 0}$ be a Wiener process on $(\Omega, \mathcal{F}, \mathbb{P})$. For any $x \in \mathbb{R}$ we define a family

$$X^x = \{x + B_t,\, 0 \le t \le T\}.$$

$X^x - x$ is a Wiener process on $(\Omega, \mathcal{F}, \mathbb{P})$. As we did in the proof of Corollary 3.3, we have $b(s, X_s) \in S$. Thus, by Corollary 3.3, the process

$$\rho_t^x = \exp\left(\int_0^t b(s, X_s^x)\, dX_s \; - \; \frac{1}{2} \int_0^t |b(s, X_s^x)|^2 ds \right), \qquad 0 \le t \le T,$$

is a martingale.

Thus, by Girsanov's theorem, under the probability measure $\mathbb{Q}^x$ defined by

$$\frac{d\mathbb{Q}^x}{d\mathbb{P}} = \rho_T(b), \qquad 0 \le t \le T,$$

the process

$$W_t^x := X_t^x - x - \int_0^t b(s, X_s^x)\, ds, \qquad 0 \le t \le T$$

is a Wiener process with $\mathbb{Q}^x[W_0^x = 0] = 1$.

Rearranging gives

$$X_t^x = x + \int_0^t b(s, X^x)\, ds + W_t^x, \qquad 0 \le t \le T.$$

Finally, with the filtration

$$\mathcal{F}_t := \sigma(X_s^x, \; s \le t),$$

and since X^x is an $\mathcal{F}_t-$adapted, continuous process and W^x is a Wiener process; all the conditions of Definition 4.1 are satisfied the triple

$$(X^x, W^x), \quad (\Omega, \mathcal{F}, \mathbb{Q}^x), \quad \{\mathcal{F}_t\}$$

is a weak solution to (4.6) (equivalently of its integral form). ☐

Remark 4.1. *Under the hypothesis of Proposition 4.1; if*

$$(X, W), \; (\Omega, \mathcal{F}, \mathbb{P}), \; (\mathcal{F}_t)$$

is a weak solution of

$$dX_t = b(t, X_t)\, dt + dW_t, \qquad X_0 = x,$$

then the process Y *defined by*

$$Y_t := X_t - W_t$$

satisfies

$$Y_t = x + \int_0^t b(s, Y_s + W_s)\, ds,$$

and therefore

$$(Y, W), \; (\Omega, \mathcal{F}, \mathbb{P}), \; (\mathcal{F}_t)$$

is a weak solution of equation (4.3)

Once existence of a weak solution has been established, it is natural to ask whether such a solution is unique. For weak solutions there are two important notion of uniqueness. We treat now the weaker notion of uniqueness associated with weak solutions, that is called *uniqueness in the sense of probability law*, or uniqueness in law, or just weak uniqueness.

Recall what we mean when we talk about the law of a continuous process, that is, the induced probability measure on the path space $C[0,T] = C([0,T],\mathbb{R})$ (see Section 2.1.1)

Definition 4.2. *We say that* uniqueness in the sense of probability law *holds for equation* (4.3) *if for any* $x \in \mathbb{R}$*, for any two weak solutions*

$$(X,W),(\Omega,\mathcal{F},\mathbb{P}),\{\mathcal{F}_t\}, \quad \textit{and} \quad (\tilde{X},\tilde{W}),(\tilde{\Omega},\tilde{\mathcal{F}},\tilde{\mathbb{P}}),\{\tilde{\mathcal{F}}_t\}$$

with

$$X_0 = \tilde{X}_0 = x \quad a.s.$$

the two processes X *and* $\tilde{X}$ *have the same law on* $C[0,T]$

Girsanov's theorem is also helpful in the study of uniqueness in law of weak solutions.

Proposition 4.2 (Uniqueness in law for $dX_t = b(t,X_t)\,dt + dW_t$)**.** *Let* $b : [0,T]\times\mathbb{R} \to \mathbb{R}$ *be Borel and bounded. Let* $x \in \mathbb{R}$ *and suppose* $(X^{(i)}, W^{(i)}, \mathbb{P}^{(i)})$*,* $i = 1,2$*, are weak solutions of*

$$dX_t = b(t,X_t)\,dt + dW_t, \qquad X_0 = x \quad a.s.$$

Then $X^{(1)}$ *and* $X^{(2)}$ *have the same law.*

Proof. Let $(X,W),(\Omega,\mathcal{F},\mathbb{P}),(\mathcal{F}_t)$ be a weak solution of

$$dX_t = b(t,X_t)\,dt + dW_t, \qquad X_0 = x \quad \text{a.s.}$$

Set

$$L_t := W_t + \int_0^t b(s,X_s)\,ds, \qquad 0 \le t \le T.$$

Since

$$X_t = x + \int_0^t b(s,X_s)\,ds + W_t,$$

we have

$$X_t = x + L_t, \qquad 0 \le t \le T.$$

Define

$$Z_t := \exp\Big(-\int_0^t b(s,X_s)\,dW_s - \frac{1}{2}\int_0^t |b(s,X_s)|^2\,ds \Big), \qquad 0 \le t \le T.$$

Since b is bounded, $(Z_t)_{0\le t\le T}$ is a martingale by Lemma 3.1. We define a probability measure $\mathbb{Q}$ on $(\Omega,\mathcal{F})$ by

$$\frac{d\mathbb{Q}}{d\mathbb{P}} = Z_T.$$

By Girsanov's theorem, the process $L = (L_t)_{0 \leq t \leq T}$ is a Wiener process under $\mathbb{Q}$, with $\mathbb{Q}[L_0 = 0] = 1$. Hence $X = x + L$ is, under $\mathbb{Q}$, a Brownian motion starting from x.

Now observe that $Z_T > 0$ $\mathbb{P}$-a.s., so $\mathbb{P} \ll \mathbb{Q}$. Moreover, $\frac{d\mathbb{P}}{d\mathbb{Q}} = Z_T^{-1}$,

$$Z_T^{-1} = \exp\left(\int_0^T b(s, X_s)\, dW_s + \frac{1}{2} \int_0^T |b(s, X_s)|^2 \, ds \right) \cdot \mathbf{1}_{\{Z_T > 0\}}.$$

Note that $\mathbb{E}^{\mathbb{Q}}[Z_T^{-1}] = 1$. Using

$$dX_s = b(s, X_s)\, ds + dW_s,$$

we obtain

$$\int_0^T b(s, X_s)\, dW_s = \int_0^T b(s, X_s)\, dX_s - \int_0^T |b(s, X_s)|^2 \, ds,$$

and therefore

$$Z_T^{-1} = \exp\left(\int_0^T b(s, X_s)\, dX_s - \frac{1}{2} \int_0^T |b(s, X_s)|^2 \, ds \right) \cdot \mathbb{1}_{\{Z_T > 0\}}.$$

In particular, Z_T^{-1} is a measurable functional of the path of X up to time T. Let us write

$$Z_T^{-1} = F(X)$$

for a suitable Borel measurable functional F on $C([0, T])$ (recall that we are seeing X as a random variable with values in $C[0, T]$, see Remark 2.2).

Let $A \in \mathcal{B}(C[0, T])$. Then

$$\mathbb{P}(X \in A) = \mathbb{E}^{\mathbb{P}}[\mathbb{1}_{\{X \in A\}}] = \mathbb{E}^{\mathbb{Q}}[\mathbb{1}_{\{X \in A\}}\, Z_T^{-1}] =$$

$$= \mathbb{E}^{\mathbb{Q}}[\mathbb{1}_A(X)\, F(X)] = \mathbb{E}^{\mu_x}[\mathbb{1}_A\, F]\,.$$

Where μ_x is the law of X under $\mathbb{Q}$ and the last equality follows from Theorem 1.1. The last expectation is done in the path space $C[0, T]$. Since under $\mathbb{Q}$ the process X is a Wiener process starting from x, its law under $\mathbb{Q}$ is the law of a Wiener process starting from x. Hence the right-hand side depends only on the initial point x, on the fixed functional F, and on A, but not on the particular weak solution chosen.

Therefore the law of X is uniquely determined. In particular, $X^{(1)}$ and $X^{(2)}$ have the same law. □

Uniqueness in the sense of probability law for the shifted equation follows immediately.

4.2 Pathwise uniqueness

Another form of uniqueness that can be related to weak solution is pathwise uniqueness.

Definition 4.3. *Suppose that whenever* $(X, W), (\Omega, \mathcal{F}, \mathbb{P}), \{\mathcal{F}_t\}$ *and* $(\tilde{X}, W), (\Omega, \mathcal{F}, \mathbb{P}), \{\mathcal{F}_t\}$ *are weak solutions to equation* (4.3) *with common Wiener process* W *on a common probability space* $(\Omega, \mathcal{F}, \mathbb{P})$ *and with common deterministic initial value, i.e.,*

$$\mathbb{P}[X_0 = \tilde{X}_0 = x] = 1,$$

the two processes X *and* $\tilde{X}$ *are indistinguishable:*

$$\mathbb{P}[X_t = \tilde{X}_t, \ \forall\, 0 \leq t \leq T] = 1.$$

We say then that pathwise uniqueness *holds for equation* (4.3).

This theorem highlights how uniqueness in law is a weaker property than pathwise uniqueness. The reader can find it proved in [13] Section 5.3.

Proposition 4.3. *Pathwise uniqueness implies uniqueness is the sense of probability law*

Conversely, uniqueness in the sense of probability law does not imply pathwise uniqueness. The following example illustrates this point.

Example 1 (H.Tanaka, 1974)**.** *This classical example, illustrating uniqueness in law without pathwise uniqueness, can be found for instance in [13] chapter 5. We provide additional details for completeness.*

Consider the one-dimensional equation

$$X_t = \int_0^t \operatorname{sgn}(X_s)\, dW_s, \qquad 0 \leq t \leq T, \tag{4.7}$$

where

$$\operatorname{sgn}(x) = \begin{cases} 1, & x > 0, \\ -1, & x \leq 0. \end{cases}$$

If $(X, W), (\Omega, \mathcal{F}, \mathbb{P}), \{\mathcal{F}_t\}$ *is a weak solution, then the process* $X = \{X_t, \mathcal{F}_t; 0 \leq t \leq T\}$ *is a continuous, square-integrable martingale with quadratic variation process*

$$\langle X \rangle_t = \int_0^t \operatorname{sgn}^2(X_s)\, ds = t.$$

Therefore, X *is a Brownian motion (see Theorem 2.15), and uniqueness in the sense of probability law holds. On the other hand,* $(-X, W), (\Omega, \mathcal{F}, \mathbb{P}), \{\mathcal{F}_t\}$ *is also a weak solution, since*

$$-X_t = -\int_0^t \operatorname{sgn}(X_s)\, dW_s = \int_0^t \operatorname{sgn}((-X)_s)\, dW_s,$$

So pathwise uniqueness *can not hold for the equation.*

4.2.1 Strong solutions

We saw that Girsanov's theorem can give us weak solutions, now we want to explore the notion of *strong solution.* In order to develop the concept of strong solution, we choose a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ as well as a one-dimensional Wiener process $W = \{W_t\}_{t\in[0,T]}$ on it.

Definition 4.4 (Strong solution)**.** *Let $b : [0, \infty) \times \mathbb{R} \to \mathbb{R}$ be a Borel-measurable function, and let $T > 0$. Fix a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ and a Wiener process $W = \{W_t\}_{t\in[0,T]}$ defined on $(\Omega, \mathcal{F}, \mathbb{P})$. Let $x \in \mathbb{R}$. A process $X = \{X_t : 0 \le t \le T\}$ is called a* strong solution *of*

$$dX_t = b(t, X_t)\, dt + dW_t, \qquad X_0 = x, \tag{4.8}$$

if X has continuous sample paths and the following properties hold:

(i) X is adapted to the filtration $\mathcal{F}_t^W = \sigma(W_s,\, s \le t)$;

(ii) $\mathbb{P}[X_0 = x] = 1$;

(iii) for every $t \in [0, T]$,

$$\mathbb{P}\left[\int_0^t |b(s, X_s)|\, ds < \infty\right] = 1;$$

(iv) the integral form of (4.8) *holds almost surely for all $t \in [0, T]$:*

$$X_t \;=\; x \;+\; \int_0^t b(s, X_s)\, ds \;+\; W_t. \tag{4.9}$$

We conclude this section by recalling the Yamada–Watanabe principle, which clarifies the relation between weak and strong solutions. We only state the result and refer to the literature for the proof and for a more precise formulation; see in particular Karatzas and Shreve [13].

Theorem 4.2 (Yamada–Watanabe)**.** *Assume that for every initial condition $x \in \mathbb{R}$ the SDE* (4.8) *admits a weak solution, and that pathwise uniqueness holds for every initial condition $x \in \mathbb{R}$. Then, for every $x \in \mathbb{R}$, the SDE admits a strong solution.*

Thus, under these assumptions, weak existence together with pathwise uniqueness give strong existence.

It is important to stress that the condition on the initial date x is essential. In general, it is not enough to know weak existence and pathwise uniqueness only for one fixed initial condition x; this does not by itself imply strong existence for that particular x.

4.3 Path-by-path uniqueness for additive-noise SDEs

Up to now we have considered two notions of existence and uniqueness for the additive-noise SDE

$$dX_t = b(t, X_t)\, dt + dW_t, \qquad X_0 = x, \tag{4.10}$$

under some assumptions on b. These types of uniqueness do not address whether, for a *fixed* realization of the driving Brownian motion, the solution path is uniquely determined. As already posed by Krylov (see [2] p.29), one may ask:

> *Does the SDE* (4.10) *with bounded (Borel) drift b admit a* ***unique*** *solution for* ***almost every*** *realization of the driving Brownian path?*

This is the *path-by-path* uniqueness problem. In what follows we introduce the notions needed to state and discuss results at this level, and we outline the Davie's theorem, which shows that SDEs with *bounded* measurable drifts indeed enjoy this uniqueness property. In the following formulation for the notion of path-by-path uniqueness, we follow the approach introduced by Flandoli in [11]. This requires recalling the notion of Wiener measure.

4.3.1 Set-up and path by path equation

Fix $T > 0$. Let $\mathcal{C}_T = C([0,T];\mathbb{R})$ be the space of real valued continuous functions on $[0,T]$. Let $b : [0,T] \times \mathbb{R} \to \mathbb{R}$ be a measurable function. For $x_0 \in \mathbb{R}$ and a *fixed* path $\gamma \in \mathcal{C}_T$, consider the deterministic, pathwise integral equation

$$x_t \;=\; x_0 \;+\; \int_0^t b(s, x_s)\, ds \;+\; \gamma_t, \qquad t \in [0,T]. \tag{4.11}$$

We denote by

$$C(\gamma, x_0) \;:=\; \Big\{ x \in C([0,T];\mathbb{R}) :\; x \text{ satisfies } (4.11) \Big\}$$

the (possibly empty) set of path by path solutions associated with (γ, x_0).

Definition 4.5 (Path-by-path uniqueness)**.** *We recall that $\mathbb{W}$ denotes the Wiener measure (see Definition 2.7). We say that* (4.10) *enjoys path-by-path uniqueness if for $\mathbb{W} - a.e.\, \gamma \in \mathcal{C}_T$ the set $C(\gamma, x_0)$ is a singleton.*

Indeed we are interested in knowing when $C(\gamma, x_0)$ is a singleton. A first basic result about this singleton property is the following.

Proposition 4.4. *Assume b is globally Lipschitz in x, uniformly in t, and that there exists $C > 0$ satisfying the condition*

$$|b(t,x)| \le C(1 + |x|), \quad \forall\, t \in [0,T],\ \forall\, x \in \mathbb{R}.$$

Then for every $\gamma \in \mathcal{C}_T$ and every $x_0 \in \mathbb{R}$, the set $C(\gamma, x_0)$ is a singleton.

Proof. Define $y_t := x_t - \gamma_t$. Then (4.11) is equivalent to

$$y_t \;=\; x_0 + \int_0^t b(s, y_s + \gamma_s)\, ds. \tag{4.12}$$

By Cauchy–Lipschitz theorem, this ODE has a unique continuous solution y; setting $x := y + \gamma$ yields the unique pathwise solution. □

Under certain assumptions, $C(\gamma, x_0)$ is non-empty, this happens for instance when b is continuous with linear growth conditions in x uniformly in t, by Peano's theorem for the auxiliary Eq. (4.12)

With the following proposition we now introduce a consequence of the theory of strong solutions, see [20, 21] for more details.

Proposition 4.5. *Under the same assumptions of Proposition 4.1, namely b Borel and satisfies the linear growth condition, given $x_0 \in \mathbb{R}$, for $\mathbb{W} - a.e.\, \gamma \in \mathcal{C}_T$, the set $C(\gamma, x_0)$ is non-empty.*

Proposition 4.5 admits important extensions to less regular drift, due to Krylov and Röckner; we do not enter into the details and refer the reader to [16]. However, it must be clear that this does not answer our question about the singleton property of $C(\gamma, x_0)$.

4.3.2 Davie's theorem

We now state the first result on the singleton property of $C(\gamma, x_0)$ or path-by-path uniqueness. The proof is highly nontrivial; see [10] and the overview in [11]). Further remarks and refinements can be found in [19].

Theorem 4.3 (Davie). *Let $b \in L^\infty([0, T] \times \mathbb{R})$. Then, for every $x_0 \in \mathbb{R}$,*

$$C(\gamma, x_0) \text{ is a singleton for } \mathbb{W} - a.e.\, \gamma \in \mathcal{C}_T.$$

Chapter 5

An application to mathematical finance

5.1 Modeling of Financial Markets

As discussed in the first chapter, stochastic processes provide a natural framework for modeling phenomena that evolve randomly over time. An important class of applications arises in the context of financial markets, where quantities such as stock prices, bond values, and exchange rates are subject to uncertainty and are influenced by the continuous arrival of information. For this part our main reference is Baldi [3].

In this section, we introduce a probabilistic model for the time evolution of financial assets prices and establish the basic setting for the valuation of derivative securities. In particular, we begin by describing the price dynamics of the underlying asset through a stochastic differential equation. We then introduce, in Section 5.1.1, the class of derivative securities that will be considered, with particular attention to European options and their payoff. In Section 5.1.2 we present the financial market in which these claims are traded, together with the notion of portfolio and self–financing strategy. This will allow us, in the following sections, to pass to discounted prices, apply Girsanov's theorem, and derive the valuation formula under the risk–neutral measure.

We denote by S_t, for $t \geq 0$, the price at time t of a financial or monetary quantity. More precisely, S_t may represent the value of a stock, a bond, or the exchange rate of a currency with respect to another.

We first observe that the price of a financial asset must remain non–negative and, under normal market conditions, strictly positive. Moreover, price variations are naturally considered in relative rather than absolute terms: when it is said that between time s and time t there has been an increase in percentage of $p\%$, this means $\frac{S_t}{S_s} = 1 + \frac{p}{100.}$. Thus, it is more convenient to model the logarithm of the price that the price itself. For this reason, it is natural to describe the evolution of the asset price S_t by means of a stochastic differential equation of the form

$$\frac{dS_t}{S_t} = b(S_t, t)\, dt + \sigma(S_t, t)\, dB_t, \tag{5.1}$$

where $(B_t)_{t\geq 0}$ is a Brownian motion defined on a suitable filtered probability space, and the functions b and σ represent the drift and diffusion coefficients, respectively. These functions are assumed to be measurable and to satisfy appropriate regularity conditions such as boundedness and local Lipschitz continuity. Throughout this work we shall mainly consider the case in which the coefficients b and σ are constant. In this case, the equation becomes

$$\frac{dS_t}{S_t} = b\,dt + \sigma\,dB_t,$$

and by Itô formula, one can see that its solution with initial condition $S_s = x$ is given by

$$S_t^{s,x} = x \exp\biggl(\Bigl(b - \frac{\sigma^2}{2}\Bigr)(t-s) + \sigma(B_t - B_s)\biggr).$$

Introducing the logarithmic variable $\xi_t = \log S_t$ and applying Itô formula, one obtains an additive stochastic differential equation for ξ_t.

5.1.1 Derivative Securities

A derivative is a financial contract whose value depends on the value of another financial quantity, called the underlying asset. One of the most common derivatives is an option. Now we focus on European options.

A European call option is a contract that grants its holder the right, but not the obligation, to purchase a given asset (such as a stock, a currency, or a commodity), called the underlying asset, at a fixed time T (the maturity) and at a predetermined price K (the strike price). If, at maturity, the price of the underlying asset is greater than K, the holder will exercise the option and purchase the asset at price K. Otherwise, the holder will buy the asset directly on the market at a lower price and the option will not be exercised. In any case, this type of contract guarantees that the holder can acquire the underlying asset at a price not exceeding K, thus providing protection against unfavorable market movements.

Similarly, a European put option grants its holder the right, but not the obligation, to sell the underlying asset at time T for the fixed price K.

Other types of derivatives exist, characterized by different exercise rules, such as American options, which allow the holder to exercise the option at any time prior to maturity. In the following, we shall restrict our attention to European options.

The main problem we address is the determination of the fair price of an option. Indeed, the seller of the option is exposed to risk: if at time T the price of the underlying asset exceeds the strike price K, the seller of a call option must purchase the asset on the market and deliver it to the buyer at price K. If S_T denotes the price of the underlying asset at maturity, the loss incurred by the seller in this case is $(S_T - K)^+$.

This suggests that the price of a European call option at the initial time should be related to the expected value of its payoff,

$$\mathbb{E}\bigl[(S_T^{s,x} - K)^+\bigr]. \tag{5.2}$$

In the case of constant coefficients b and σ, the expectation above is easy to compute, since the law of $S_T^{s,x}$ is explicitly known. Indeed, from

$$S_t^{s,x} = x \exp\left(\left(b - \frac{\sigma^2}{2} \right) (t-s) + \sigma (B_t - B_s) \right),$$

it follows that

$$\log S_T^{s,x} \sim \mathcal{N}\left(\log x + \left(b - \frac{\sigma^2}{2} \right) (T-s),\ \sigma^2 (T-s) \right),$$

and therefore $S_T^{s,x}$ is lognormally distributed. However, as we shall see later, this quantity does not provide the correct price of the option.

5.1.2 Financial Markets and Portfolios

We consider a financial market on a filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, \mathbb{P})$ supporting a Brownian motion.

The market consists of one risk–free asset S^0 and m risky assets $S^1, \dots, S^m$.

The price of the risk–free asset satisfies

$$dS_t^0 = r_t S_t^0 \, dt, \tag{5.3}$$

where $(r_t)_{t\geq 0}$ is an adapted process representing the instantaneous interest rate. If $S_s^0 = x$, then

$$S_t^0 = x \exp\left(\int_s^t r_u \, du \right) \tag{5.4}$$

Typically, S_t^0 corresponds to a bank account or a money market account. For this reason, it is also called the *risk–free asset*, although its evolution may still be random through the process (r_t).

For $i = 1, \dots, m$, the prices of the risky assets follow stochastic differential equations of the form (5.1).

Definition 5.1. *The value of the portfolio at time t is*

$$V_t = \sum_{i=0}^m H_t^i S_t^i.$$

where $H_t^0, \dots, H_t^m$ denote respectively the number of units of the risk–free and risky assets held at time t. More precisely $H_t^0, \dots, H_t^m$ are real numbers, in particular they are allowed to take negative values. A negative position corresponds to a short sale of the corresponding asset. In particular, holding $H_t^j < 0$ means that the investor has sold short shares of the risky asset, while $H_t^0 < 0$ represents borrowing at the risk–free rate.

Definition 5.2. *The portfolio is said to be* self–financing *if its value satisfies*

$$dV_t = \sum_{i=0}^m H_t^i \, dS_t^i. \tag{5.5}$$

Equation (5.5) means that changes in portfolio value are solely due to price variations of the assets and not to external cash flows. In the remainder of this chapter we restrict our attention to a market with one risky asset and one risk–free asset.

5.1.3 Discounted Prices

In order to compare cash flows occurring at different times, it is convenient to work with *discounted* quantities. Fix $s \geq 0$ and define the discount factor from s to $t \geq s$ by

$$R_{s,t} := \exp\Big(- \int_s^t r_u \, du \Big) . \tag{5.6}$$

Differentiating we get

$$dR_{s,t} = -r_t R_{s,t} \, dt \tag{5.7}$$

For any asset price process S_t, we define the corresponding *discounted price* by

$$\widetilde{S}_t := R_{0,t} \, S_t . \tag{5.8}$$

It represents the value at time 0 of the asset price at time t, since equation (5.8) can be rewritten as $\widetilde{S}_t = \frac{S_t}{S_t^0}$. In fact (take $s = 0$ in equation (5.4))

$$S_t^0 = S_0^0 \exp\left(\int_0^t r_u \, du \right) .$$

This also leads to the observation that for the risk–free asset S_t^0 we have

$$\widetilde{S}_t^0 = R_{0,t} \, S_t^0 = S_0^0 . \tag{5.9}$$

More generally, using multidimensional Itô formula (see Theorem 2.14), the discounted risky asset price $\widetilde{S}_t = R_{0,t} S_t$ satisfies

$$d\widetilde{S}_t = R_{0,t} \, dS_t + S_t \, dR_{0,t} = R_{0,t} S_t \Big((b_t - r_t) \, dt + \sigma_t \, dB_t \Big) = \widetilde{S}_t \Big((b_t - r_t) \, dt + \sigma_t \, dB_t \Big) , \tag{5.10}$$

where we used (5.1) and (5.7).

Finally, if V_t denotes the value of a self–financing portfolio, we define its discounted value by

$$\widetilde{V}_t := R_{0,t} \, V_t . \tag{5.11}$$

Although we have restricted attention to a market with one risky asset and one risk–free asset, this definition of discounted value is completely general and applies to any portfolio value process. Discounted asset and portfolio values will play a central role in the pricing problem, since under a suitable change of probability measure they become martingales.

5.1.4 Market Price of Risk and Change of Measure

Recall that the discounted risky asset price satisfies

$$d\widetilde{S}_t = \widetilde{S}_t\big((b_t - r_t)\,dt + \sigma_t\,dB_t\big).$$

We introduce the *market price of risk* process

$$\theta_t := \frac{b_t - r_t}{\sigma_t}.$$

Define the stochastic exponential

$$Z_t := \exp\Big(-\int_0^t \theta_u\,dB_u - \frac{1}{2}\int_0^t \theta_u^2\,du\Big).$$

We assume that $(Z_t)_{t\geq 0}$ is a martingale. A sufficient condition for this is given, for instance, by Novikov's condition (see Section 3.1).

We define a new probability measure $\mathbb{Q}$ on $(\Omega, \mathcal{F})$ by

$$\frac{d\mathbb{Q}}{d\mathbb{P}} = Z_T. \tag{5.12}$$

By Girsanov's theorem, the process

$$W_t := B_t + \int_0^t \theta_u\,du$$

is a Brownian motion under $\mathbb{Q}$.

Differentiating we obtain

$$dB_t = dW_t - \theta_t\,dt \tag{5.13}$$

and substituting into (5.10), we obtain

$$d\widetilde{S}_t = \sigma_t \widetilde{S}_t\,dW_t.$$

Therefore, under the assumptions on the coefficients, and by Theorem 8.10 in [3], $\widetilde{S} = (\widetilde{S}_t)_{t\geq 0}$ is a process of class H (see Definition 2.10), and since σ is bounded, $(\widetilde{S}_t)_{t\geq 0}$ is the integral of a process of class H, thus a martingale under the probability measure $\mathbb{Q}$ by Theorem 2.9. In particular, since the mean of the prices of $\widetilde{S}_t$ is constant in time with respect to $\mathbb{Q}$, the latter is called a *risk-neutral probability.*

Consider a strategy $H_t = (H_t^0, H_t^1)$ and let

$$V_t = H_t^0 S_t^0 + H_t^1 S_t$$

be the corresponding self financing portfolio value.

Under the probability measure $\mathbb{P}$, the risky asset satisfies

$$dS_t = b_t S_t\,dt + \sigma_t S_t\,dB_t, \tag{5.14}$$

while the risk–free asset satisfies

$$dS_t^0 = r_t S_t^0 \, dt. \tag{5.15}$$

Using the self–financing condition

$$dV_t = H_t^0 \, dS_t^0 + H_t^1 \, dS_t.$$

and substituting with equations (5.14) and (5.15) we obtain

$$dV_t = H_t^0 r_t S_t^0 \, dt + H_t^1 \big(b_t S_t \, dt + \sigma_t S_t \, dB_t \big).$$

Since $H_t^0 S_t^0 = V_t - H_t^1 S_t$, we rewrite it in terms of V_t:

$$dV_t = r_t V_t \, dt + H_t^1 S_t (b_t - r_t) \, dt + H_t^1 \sigma_t S_t \, dB_t.$$

Now note that $b_t - r_t = \sigma_t \, \theta_t$, and we arrive to the equation

$$dV_t = r_t V_t \, dt + H_t^1 \sigma_t S_t \big(\theta_t \, dt + dB_t \big). \tag{5.16}$$

Now consider the discounted wealth process

$$\widetilde{V}_t := R_{0,t} V_t.$$

By Itô's formula and the self–financing property, we obtain

$$d\widetilde{V}_t = H_t^1 \sigma_t \widetilde{S}_t \big(\theta_t \, dt + dB_t \big).$$

Under the risk–neutral probability measure $\mathbb{Q}$ constructed via Girsanov's theorem, the process

$$W_t := B_t + \int_0^t \theta_u \, du$$

is a Wiener process and

$$d\widetilde{V}_t = H_t^1 \sigma_t \widetilde{S}_t \, dW_t.$$

Hence, $\widetilde{V}_t$ is a stochastic integral with respect to the $\mathbb{Q}$–Brownian motion W_t. Under suitable integrability conditions on the strategy, it follows that $(\widetilde{V}_t)_{t \in [0,T]}$ is a $\mathbb{Q}$–martingale. This property is crucial for the valuation problem, which we treat in the following section.

5.1.5 Valuation of European options

A European option with maturity T is a nonnegative random variable Ψ, which is also $\mathcal{F}_T$–measurable.

In the case of a European call option with strike price $K > 0$, obviously

$$\Psi = (S_T - K)^+.$$

A hedging portfolio for an option Ψ is a self–financing portfolio with $V_t \geq 0$ for all $t \geq 0$ whose value process V_t satisfies

$$V_T = \Psi.$$

If such a portfolio exists, the option is said to be *replicable.*

Suppose we have a replicable European option and let V_t be the hedging portfolio. Then $V_T = \Psi$, and since $\widetilde{V}_t$ is a $\mathbb{Q}$–martingale, we obtain

$$\widetilde{V}_t = \mathbb{E}^{\mathbb{Q}}\left[\widetilde{V}_T \mid \mathcal{F}_t\right].$$

In particular, at time $t = 0$,

$$V_0 = \mathbb{E}^{\mathbb{Q}}[R_{0,T}\,\Psi \mid \mathcal{F}_0].$$

Since the initial sigma–algebra $\mathcal{F}_0$ is assumed to be trivial, the conditional expectation reduces to an ordinary expectation, and therefore

$$V_0 = \mathbb{E}^{\mathbb{Q}}\left[R_{0,T}(S_T - K)^+\right]$$

That is the fair price of the option, in fact it is the minimal initial capital needed to construct a self-financing portfolio that replicates the option. We see here what was wrong in formula (5.2): it is not only that we forgot the discount factor, but also the expectation must be taken with respect to the risk neutral probability $\mathbb{Q}$.

5.2 The Black–Scholes Formula

We now specialize the general model to the Black–Scholes model [7].

We assume that the interest rate $r \in \mathbb{R}$ and the volatility $\sigma > 0$ are constant. Under the risk–neutral probability measure the drift of the risky asset coincides with the interest rate. Therefore, the risky asset satisfies

$$\frac{dS_t}{S_t} = r\,dt + \sigma\,dW_t, \tag{5.17}$$

where W is a Brownian motion under $\mathbb{Q}$. To show this, it is sufficient to substitute equation 5.13 in equation 5.1 and use the definition of θ_t (in the case of constant coefficients)

Let $s < T$ and suppose that $S_s = x$. We already treated the case of constant coefficients, showing that the solution is lognormally distributed, the only difference with the previous case is that the drift is now given by the interest rate r. In particular, we recall that the solution of (5.17) is given by

$$S_t^{s,x} = x\exp\left(\left(r - \frac{1}{2}\sigma^2\right)(t-s) + \sigma(W_t - W_s)\right). \tag{5.18}$$

In particular,

$$\log S_t^{s,x} = \log x + \left(r - \frac{1}{2}\sigma^2\right)(t-s) + \sigma(W_t - W_s).$$

Since $W_t - W_s \sim \mathcal{N}(0, t-s)$, it follows that

$$\log S_t^{s,x} \sim \mathcal{N}\left(\log x + \left(r - \frac{1}{2}\sigma^2\right)(t-s),\ \sigma^2(t-s)\right).$$

Hence S_t has a lognormal distribution under $\mathbb{Q}$.

Consider now a European call option with payoff

$$\Psi = (S_T - K)^+.$$

Denote with $C(s,x)$ the price of the option at time s with the condition $S_s = x$. Therefore

$$C(s,x) = e^{-r(T-s)}\mathbb{E}^{\mathbb{Q}}\big[(S_T^{s,x} - K)^+\big]. \tag{5.19}$$

Using (5.18), we write

$$S_T^{s,x} = x\exp\Big(\Big(r - \frac{1}{2}\sigma^2\Big)(T-s) + \sigma(W_T - W_s)\Big).$$

Let $Z \sim \mathcal{N}(0,1)$ and note that

$$W_T - W_s \sim \sqrt{T-s}\,Z.$$

Thus

$$S_T^{s,x} \sim x\exp\Big(\Big(r - \frac{1}{2}\sigma^2\Big)(T-s) + \sigma\sqrt{T-s}\,Z\Big).$$

Substituting into (5.19), we obtain

$$\begin{aligned} C(s,x) &= e^{-r(T-s)}\mathbb{E}\left[\left(xe^{\left(r-\frac{1}{2}\sigma^2\right)(T-s)+\sigma\sqrt{T-s}Z} - K\right)^+\right] \\ &= e^{-r(T-s)}\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{+\infty}\left(xe^{\left(r-\frac{1}{2}\sigma^2\right)(T-s)+\sigma\sqrt{T-s}z} - K\right)^+ e^{-z^2/2}dz. \qquad (5.20)\end{aligned}$$

It is easy to see that the integrand is zero for $z \le \xi_s$, where

$$\xi_s = \frac{1}{\sigma\sqrt{T-s}}\left(\log\frac{K}{x} - \left(r - \frac{\sigma^2}{2}\right)(T-s)\right). \tag{5.21}$$

Let Φ denote the distribution function of the standard normal law $N(0,1)$,

$$\Phi(y) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{y} e^{-z^2/2}\,dz = \frac{1}{\sqrt{2\pi}}\int_{-y}^{+\infty} e^{-z^2/2}\,dz. \tag{5.22}$$

Therefore,

$$\begin{aligned} C(s,x) &= \frac{e^{-r(T-s)}}{\sqrt{2\pi}}\int_{\xi_s}^{+\infty}\left(xe^{\left(r-\frac{\sigma^2}{2}\right)(T-s)+\sigma\sqrt{T-s}\,z} - K\right)e^{-z^2/2}\,dz \qquad (5.23) \\ &= \frac{x}{\sqrt{2\pi}}\int_{\xi_s}^{+\infty}\exp\Big(-\frac{1}{2}\Big(z - \sigma\sqrt{T-s}\Big)^2\Big)\,dz - Ke^{-r(T-s)}\,\Phi(-\xi_s) \qquad (5.24) \\ &= \frac{x}{\sqrt{2\pi}}\int_{\xi_s-\sigma\sqrt{T-s}}^{+\infty} e^{-z^2/2}\,dz - Ke^{-r(T-s)}\,\Phi(-\xi_s) \qquad (5.25) \\ &= x\,\Phi\Big(-\xi_s + \sigma\sqrt{T-s}\Big) - Ke^{-r(T-s)}\,\Phi(-\xi_s). \qquad (5.26)\end{aligned}$$

In conclusion, we obtain the Black–Scholes formula

$$C(s,x) = x\,\Phi\Big(-\xi_s + \sigma\sqrt{T-s}\Big) - Ke^{-r(T-s)}\,\Phi(-\xi_s), \tag{5.27}$$

where ξ_s is given by (5.21).

Concluding remarks. The derivation of the Black–Scholes formula illustrates concretely the role of Girsanov's theorem in financial modeling. By a change of probability measure, the drift of the risky asset is replaced by the interest rate and the discounted price process becomes a martingale. As a consequence, the valuation of a European option reduces to the computation of a conditional expectation under the risk–neutral measure $\mathbb{Q}$.

In Black–Scholes, where the coefficients are constant, the distribution of the asset price at maturity can be determined explicitly, and the pricing formula follows from an elementary computation involving the normal distribution. It is worth observing that the original drift of the asset under the physical probability measure does not appear in the final expression of the price. The price depends only on the volatility σ and on the interest rate r.

The Black–Scholes formula therefore provides a concrete example of the general risk–neutral valuation principle developed in the previous section and shows how the abstract probabilistic framework with Girsanov's theorem leads to explicit pricing results in a specific model.

Bibliography

[1] L. V. Ahlfors, *Complex Analysis*, 3rd ed., McGraw-Hill, New York, 1979.

[2] S. V. Anulova, A. Y. Veretennikov, N. V. Krylov, R. Liptser, and A. N. Shiryaev, *Stochastic Calculus. Probability Theory III*, Encyclopaedia of Mathematical Sciences, Vol. 45, Springer, Berlin, 1998.

[3] P. Baldi, *Stochastic Calculus: An Introduction Through Theory and Exercises*, Springer, Cham, 2017.

[4] L. Bachelier, *Théorie de la spéculation*, *Annales scientifiques de l'École Normale Supérieure*, 3e série, **17** (1900), 21–86.

[5] P. Billingsley, *Probability and Measure*, 3rd ed., Wiley, New York, 1995.

[6] P. Billingsley, *Convergence of Probability Measures*, 2nd ed., Wiley, New York, 1999.

[7] Black, F., Scholes, M. (1973). *The Pricing of Options and Corporate Liabilities.* Journal of Political Economy, **81**(3), 637–654. DOI: 10.1086/260062

[8] H. Brezis, *Functional Analysis, Sobolev Spaces and Partial Differential Equations*, Springer, New York, 2011.

[9] R. H. Cameron and W. T. Martin, Transformations of Wiener integrals under translations, *Annals of Mathematics* **45** (1944), no. 2, 386–396.

[10] A. M. Davie, Uniqueness of solutions of stochastic differential equations, *International Mathematics Research Notices* (2007), no. 24, rnm124, 26 pp. doi:10.1093/imrn/rnm124.

[11] F. Flandoli, Regularizing properties of Brownian paths and a result of Davie, *Stochastics and Dynamics* **11** (2011), no. 2–3, 323–331. doi:10.1142/S0219493711003310.

[12] I. V. Girsanov, On transforming one class of stochastic processes by absolutely continuous substitution of measures, *Theory of Probability and Its Applications* **5** (1960), no. 3, 285–301.

[13] I. Karatzas and S. E. Shreve, *Brownian Motion and Stochastic Calculus*, Springer, New York, 1991.

[14] N. V. Krylov, *Introduction to the Theory of Random Processes*, American Mathematical Society, Providence, RI, 2002.

[15] N. V. Krylov, A simple proof of a result of A. Novikov, *arXiv:math/0207013v2 [math.PR]*, 2009. `https://arxiv.org/abs/math/0207013v2`.

[16] N. V. Krylov and M. Röckner, Strong solutions of stochastic equations with singular time dependent drift, *Probability Theory and Related Fields* **131** (2005), 154–196. doi:10.1007/s00440-004-0361-z.

[17] A. A. Novikov, On an identity for stochastic integrals, *Theory of Probability and Its Applications* **17** (1972), no. 4, 717–720.

[18] W. Rudin, *Real and Complex Analysis*, McGraw-Hill, New York, 1987.

[19] R. Shaposhnikov, Some remarks on Davie's uniqueness theorem, *Proceedings of the Edinburgh Mathematical Society* **59** (2016), no. 4, 1019–1035. doi:10.1017/S001309151600004X. `https://arxiv.org/abs/1401.5455`.

[20] Yu. A. Veretennikov, On strong solution and explicit formulas for solutions of stochastic integral equations, *Mathematics of the USSR-Sbornik* **39** (1981), 387–403.

[21] A. K. Zvonkin, A transformation of the phase space of a diffusion process that will remove the drift, *Matematicheskii Sbornik* **93** (1974), 129–149. (English transl.: *Math. USSR Sbornik* **22** (1974), 129–149.)