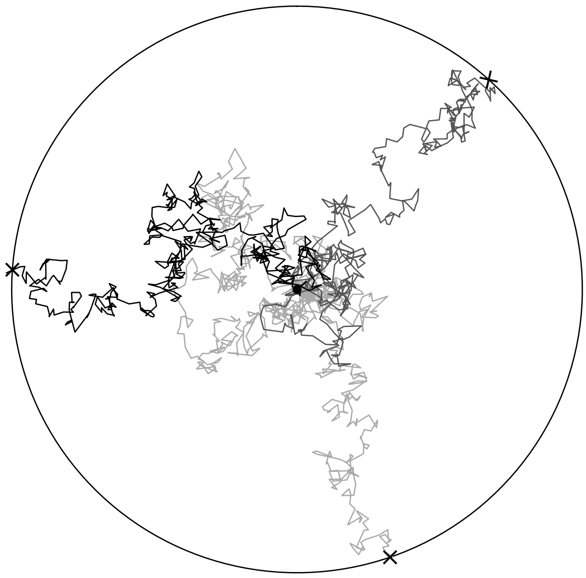
Departamento de Matemáticas, Facultad de Ciencias
Universidad Autónoma de Madrid

Itô diffusions and harmonic functions: deterministic and stochastic differential equations

Degree in Mathematics



Author: Pablo Cuesta Sierra
Tutor: José Manuel Conde Alonso

Academic year 2023-2024

Resumen

Los procesos estocásticos de tiempo continuo son una poderosa herramienta para modelizar fenómenos aleatorios, pero también se pueden utilizar para resolver problemas deterministas.

Una difusión de Itô, el proceso estocástico que resuelve una ecuación diferencial estocástica homogénea en el tiempo, es un tipo especial de proceso estocástico que tiene la propiedad de Markov fuerte: es "sin memoria". La fórmula de Itô vincula cada difusión de Itô a un operador diferencial parcial elíptico asociado. Estas propiedades nos permiten generalizar el concepto tradicional de función armónica, formulando una propiedad de valor medio que depende de la difusión de Itô. Además, la solución de varios problemas deterministas de EDPs puede expresarse en términos de una difusión de Itô utilizando esta relación.

Este documento introduce la teoría de procesos estocásticos de tiempo continuo, incluyendo el movimiento Browniano y las martingalas; que son esenciales para la construcción de la integral de Itô y ecuaciones diferenciales estocásticas. Finalmente, se presentan las propiedades anteriormente comentadas de las difusiones de Itô con el fin de aplicarlas a la resolución de un problema global de ecuación parabólica en derivadas parciales y al problema combinado de Dirichlet-Poisson.

Abstract

Time-continuous stochastic processes are a powerful tool to model random phenomena; but they can also be applied to deterministic problems.

An Itô diffusion, which is the stochastic process that solves a time-homogeneous stochastic differential equation, is a special kind of stochastic process which has the strong Markov property: it is "memoryless". The Itô formula ties each Itô diffusion to an associated elliptic partial differential operator. These properties allow us to generalize the concept of traditional harmonic functions, by formulating a mean value property that depends on the Itô diffusion. Moreover, the solution to several deterministic problems in PDEs can be expressed in terms of an Itô diffusion by using this relationship.

This document introduces the theory of time-continuous stochastic processes, including Brownian motion and martingales, which are essential to the construction of the Itô integral and stochastic differential equations. Finally, the previously discussed properties of the Itô diffusions are presented, along with its application to two PDE problems: a global parabolic PDE problem, and the combined Dirichlet-Poisson problem.

Contents

CHAPTER 1
Introduction

Stochastic processes and, in particular, stochastic differential equations have numerous applications in various fields of applied mathematics. To name a few, they have direct applications in population growth modelling, game theory, filtering problems, queuing theory, optimal control problems, pricing of financial derivatives, and generative modelling [Øks13; Son+20]. This thesis focuses on the connection between Itô diffusions, a certain type of stochastic processes, and elliptic partial differential operators. This relationship allows us to express the solutions of certain PDEs in terms of a stochastic process. In this chapter, we present an overview of the main concepts that will be developed in the sequel, with an emphasis on this connection, which is best illustrated by a generalization of the mean value property of harmonic functions formulated in the language of stochastic processes.

A stochastic process X is a family of $\mathbb{R}^n$-valued random variables $(X_t)_{t\in T}$ indexed by a set T, which is usually interpreted as time. A common choice for T is a countable set; in this case, we say that the process is *time-discrete*. An example of a time-discrete stochastic process is the (1-dimensional) random walk:

$$X_{hn} = \sum_{j=1}^{n} sI_j, \qquad \text{for each } n \in \mathbb{N}, \tag{1.1}$$

where the (I_j) are i.i.d.r.v. with uniform distribution on $\{-1, 1\}$, and the constants $s, h > 0$ denote the space and time-steps, respectively. The index set for this random walk is $T = h\mathbb{N}$.

In this work, however, we focus on *time-continuous* stochastic processes: indexed by $T = \mathbb{R}_+$. The most important example for us will be Brownian motion, which is a stochastic process $(B_t)_{t\geq 0}$ that results from making s and h in the random walk (1.1) infinitesimally small[1]; see Figure 1.1. When putting n independent Brownian motions at each coordinate of an $\mathbb{R}^n$-valued process, we obtain an n-dimensional Brownian motion, which we can think of as a time-continuous random walk in $\mathbb{R}^n$. New considerations regarding the properties of a stochastic process arise in the time-continuous case. For example, given a process $(X_t)_{t\geq 0}$, its *paths*, $t \mapsto X_t(\omega)$, which are functions from $\mathbb{R}_+$ to $\mathbb{R}^n$ for each $\omega \in \Omega$, are now indexed by a linear continuum. Processes with a.s. continuous paths are called *path-continuous*. The family $(\mathscr{X}_t)_{t\geq 0}$ is the increasing sequence of σ-algebras generated by the process X: $\mathscr{X}_t := \sigma(X_s : 0 \leq s \leq t)$. In a way, $\mathscr{X}_t$ contains the information carried by the process up to time t.

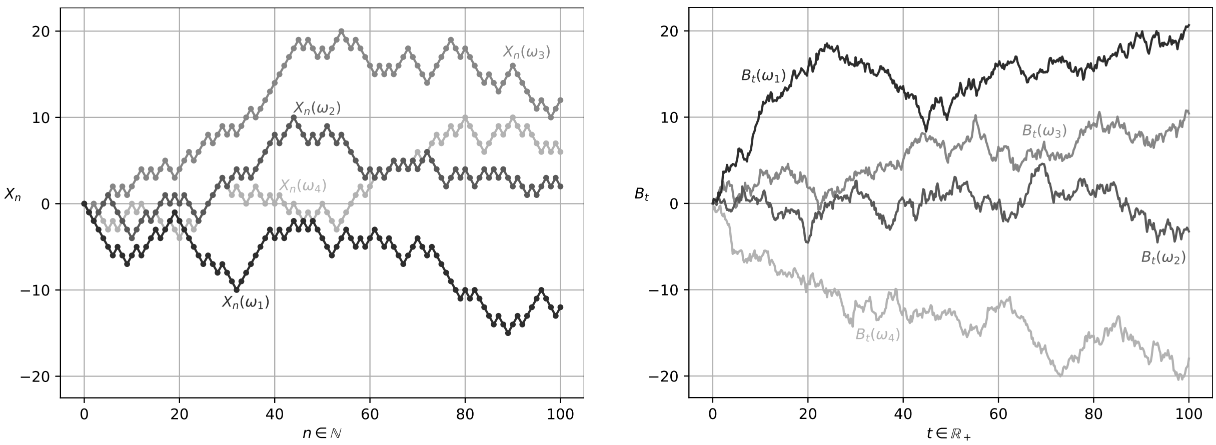


Figure 1.1: Each figure contains four realizations of a 1-dimensional stochastic process. On the left, the process is a random walk (time-discrete); on the right, a Brownian motion (time-continuous).

A special class of stochastic processes is that of martingales, which model fair games: the expected value of a martingale at a future time, conditioned to the present information, is equal to its present value. The symmetry

[1] To obtain Brownian motion, (s, h) has to approach $(0, 0)$ in a specific way. See [SV08, Section 2.4] for details on this construction.

and independence of the increments of a random walk makes it a martingale. Brownian motions inherit this property. We will use this fact to properly define stochastic integration; more precisely, the Itô integral.

The Itô integral is a generalization of the Lebesgue-Stieltjes integral. Given a non-decreasing, right-continuous function $g : [a,b] \to \mathbb{R}$, its Lebesgue-Stieltjes measure dg is given by extending

$$dg((c,d]) = g(d) - g(c), \qquad \forall (c,d] \subseteq [a,b]$$

to the Borel subsets of $[a,b]$. If $g \in C^1$, then the integral of a Borel-measurable function f with respect to dg is

$$\int_a^b f\,dg = \int_a^b f(x)g'(x)\,dx.$$

The Itô integral is defined similarly, exchanging for each ω the function g by the path $t \mapsto B_t(\omega)$:

$$\int_0^t v(s,\omega)\,dB_s(\omega).$$

An *Itô process* X is the sum of a deterministic and a stochastic integral of the above kind:

$$X_t(\omega) = X_0(\omega) + \int_0^t u(s,\omega)\,ds + \int_0^t v(s,\omega)\,dB_s(\omega). \tag{1.2}$$

X can also be denoted in differential form: $dX_t = u\,dt + v\,dB_t$. Similar to the first-order Taylor expansion, given an $f \in C^2(\mathbb{R}_+ \times \mathbb{R}^n)$, the *Itô formula* allows us to express the Itô process $f(t,X_t)$ as in (1.2), in terms of f, its derivatives, and (X_t):

$$d\big(f(t,X_t)\big) = \left(\frac{\partial f}{\partial t} + \sum_{i=1}^n \frac{\partial f}{\partial x_i} u + \frac{1}{2}\sum_{i,j=1}^n \frac{\partial^2 f}{\partial x_i \partial x_j}\big(vv^T\big)_{ij}\right) dt + \sum_{i=1}^n \frac{\partial f}{\partial x_i} v_i\,dB_t, \tag{1.3}$$

where u, v, f and its derivatives are evaluated at (t,X_t). The main difference between the Itô formula and the Taylor expansion is that (1.3) includes a term depending on the second-order derivatives of f. This is due to the formal identity $dB_t \cdot dB_t = dt$, which is a consequence of the Brownian motion having independent, Gaussian centered increments $B_t - B_s$ with variance $t - s$. The Itô formula is the cornerstone of the theory of stochastic differential equations.

Analog to ordinary differential equations (ODEs), a stochastic differential equation (SDE) has the form

$$dX_t(\omega) = b(t,X_t(\omega))\,dt + \sigma(t,X_t(\omega))\,dB_t(\omega). \tag{1.4}$$

We can intuitively understand such an equation as a differential equation with a random perturbation:

$$\frac{dX}{dt} = b(t,X_t) + \sigma(t,X_t)\cdot\text{“noise}(t)\text{”},$$

where the noise is stationary (or memoryless), with mean 0 and independent at different times. A solution of the SDE is a stochastic process (X_t) that satisfies (1.4). Similar theorems to those of ODEs of existence and uniqueness of solutions can be proven for SDEs. We will focus on the time-homogenous case; that is, when b and σ do not depend on t; a processes (X_t) solving such an SDE,

$$dX_t = b(X_t)\,dt + \sigma(X_t)\,dB_t,$$

is called an *Itô diffusion*. The time-homogeneity of Itô diffusions is characterized by the strong Markov property. Before formulating it, we need to introduce stopping times.

Let $(X_t)_{t\geq 0}$ be a path-continuous stochastic process taking values in $\mathbb{R}^n$, and let $H \ni X_0$ be an open subset of $\mathbb{R}^n$. The *first exit time of X from H* is defined as the first time that the process leaves the set:

$$\tau_H(\omega) := \inf\{t \geq 0 : X_t(\omega) \notin H\}.$$

As the process X is path-continuous, $X_{\tau_H} \in \partial H$. This random variable τ_H is a special case of a stopping time. A *stopping time* τ *w.r.t. a process* $(X_t)_{t\geq 0}$ is random variable satisfying the property that the event $\{\tau \leq t\}$ is determined by the information up to time t. As before, given a stopping time τ, we define the σ-algebra $\mathscr{X}_\tau$, which contains the information of the process up to the stopping time.

The *strong Markov property* of an Itô diffusion X states that, for every $h \geq 0$, and every $f \in L^\infty(\mathbb{R}^n)$,

$$E^x\left[f(X_{\tau+h}) \,\middle|\, \mathscr{X}_\tau\right] = E^{X_\tau^x}\left[f(X_h)\right], \tag{1.5}$$

where E^x denotes de expectation with respect to the natural probability law of $(X_t)_{t\geq 0}$ starting at $X_0 = x$. The meaning of (1.5) is: the expected value of $f(X_{\tau+h})$ given the information up to time τ is equal to the expected value of $f(X_h)$ if we start the process in position X_τ. The Markov property has a very powerful corollary: the mean value property.

Let G be a subset of H such that $x = X_0 \in G$ (see Figure 1.2). We can define a measure on each $F \subseteq \partial G$ by measuring the probability of the process exiting G through F:

$$\mu_G^x(F) := P\left(X_{\tau_G} \in F\right).$$

This probability measure μ_G^x on $\mathscr{B}(\partial G)$ is called the *harmonic measure*. By the Markov property, the function $\phi(x) := E^x\left[f\left(X_{\tau_H}\right)\right]$ satisfies the *mean value property in H*:

$$\forall G \text{ open such that } \overline{G} \subseteq H,\ \forall x \in G, \qquad \phi(x) = E^x\left[\phi\left(X_{\tau_G}\right)\right] = \int_{\partial G} \phi(y)\, d\mu_G^x(y). \tag{1.6}$$

That is, the expected value of f at X_{τ_H} when starting at $x \in G$ is the average of the expected values of f at X_{τ_H} starting at $y \in \partial G$, weighted by the probability of exiting G through y. Similar to the traditional definition, we say that a locally bounded function ϕ is X*-harmonic* in H if it satisfies the mean value property (1.6).

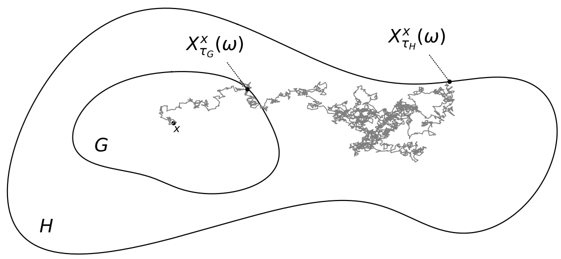


Figure 1.2: Two nested sets $G \subseteq H \subseteq \mathbb{R}^2$ and a realization of a diffusion starting at $x \in G$.

The characteristic operator $\mathfrak{A}$ of an Itô diffusion (X_t) acts on a function $f : \mathbb{R}^n \to \mathbb{R}$ by "differentiating it along the paths of the diffusion":

$$\mathfrak{A}f(x) := \lim_{U \downarrow x} \frac{E^x\left[f\left(X_{\tau_U}\right)\right] - f(x)}{E^x\left[\tau_U\right]}.$$

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

$$\mathfrak{A}f = \sum_{i=1}^n b_i \frac{\partial f}{\partial x_i} + \frac{1}{2}\sum_{i,j=1}^n \left(\sigma\sigma^T\right)_{ij} \frac{\partial^2 f}{\partial x_i \partial x_j}. \tag{1.7}$$

The right-hand side of (1.7) is an elliptic partial differential operator acting on f. Particularly, the characteristic operator of an n-dimensional Brownian motion is $\frac{1}{2}\Delta$, where Δ denotes the Laplacian operator in $\mathbb{R}^n$. Moreover, a function $f \in C^2(H)$ is X-harmonic in H (i.e. satisfies the mean value property) if and only if $\mathfrak{A}f = 0$ in H. In particular, due to the symmetry of Brownian motion, the harmonic measure induced by a Brownian motion exiting an n-dimensional Euclidean ball D from its center is the normalized (Lebesgue) surface measure on the boundary of the ball (see Figure 1.3). Therefore, (B_t)-harmonic functions in D are exactly the standard harmonic functions in D.

This connection between an Itô diffusion and the operator (1.7) is what allows us to prove the existence and uniqueness of problems in partial differential equations by translating them into stochastic terms. We study the Feynman-Kac formula, which solves the equation

$$\frac{\partial v}{\partial t}(t,x) = \mathfrak{A}v(t,x) - q(x)v(t,x) \qquad \left(t \in \mathbb{R}_+, x \in \mathbb{R}^n\right)$$

for a given $q \in C(\mathbb{R}^n)$ and a given initial value $v(0,\cdot) \in C_c^2(\mathbb{R}^n)$. We also consider the combined Dirichlet-Poisson problem, consisting on finding a $u \in C^2(D)$ that solves: $\mathfrak{A}u = -g$, in a domain $D \subseteq \mathbb{R}^n$, with the boundary condition

$$\forall y \in \partial D, \qquad \lim_{D \ni x \to y} u(x) = \phi(y), \tag{1.8}$$

where the g and ϕ are given, and we will see that with our stochastic approach it is convenient to consider exchanging (1.8) for a condition that involves approaching the boundary of D from the interior through the paths of the diffusion.

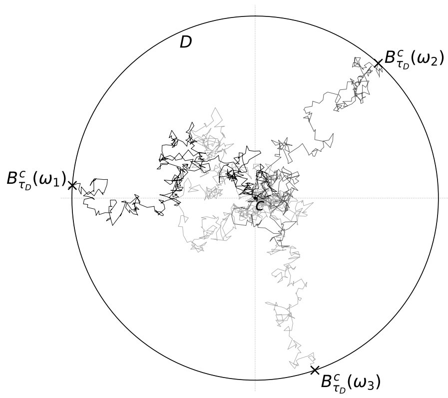


Figure 1.3: Three paths of a 2-dimensional Brownian motion exiting a disk $D \subseteq \mathbb{R}^2$ from the center c.

Goals and scope

The goal of this project is to study the theory of time-continuous stochastic processes, stochastic integration and SDEs, and to present an application of this stochastic theory to deterministic problems in partial differential equations. To this end, this document aims to provide a coherent construction of this theory assuming only the material covered in the probability and analysis courses of the bachelor's degree in mathematics. Due to the broad extension of the material covered in this project and the limited space available for this document, we have chosen to state and prove only results that are necessary for the theorems of the two final chapters of the main text, containing the applications to PDEs.

Structure of the document

Chapter 2 introduces time-continuous stochastic processes, Brownian motion, stopping times, and martingales. Chapter 3 presents the definition of the Itô integral, the Itô formula, and an existence and uniqueness result for the solution of SDEs. Chapter 4 develops the theory of Itô diffusions, their Markov property, the harmonic measure, the mean value property, and the infinitesimal generators of the diffusions. Finally, Chapters 5 and 6 present the applications of Itô diffusions to prove the existence and uniqueness of the solution to global parabolic PDEs (the Feynman-Kac formula) and the combined Dirichlet-Poisson problem, respectively.

The appendix contains the proofs of some technical results that are used in the main text. Appendix A contains basic results from measure theory. Appendix B presents the definition of infinite products of measure spaces and related results. An important theorem in this chapter is the Kolmogorov extension theorem, which guarantees the existence of a probability measure on the infinite product of probability spaces and is used to prove the existence of Brownian motion. Appendix C contains a very technical result that allows us to approximate measurable functions in the product of measure spaces by sums of simple tensors. Appendix D presents a proof of the Kolmogorov continuity criterion, a general criterion for the existence of path-continuous stochastic processes that is used to prove the existence of a path-continuous Brownian motion. Lastly, Appendix E proves the right-continuity of the filtration generated by Brownian motion and sets of measure zero, whose existence is useful for the definition of the Itô integral.

Bibliography

The main references for the elaboration of this document are Revuz and Yor [RY05], for the existence of Brownian motion and general results on stochastic processes and martingales; Ash and Doleans-Dade [AD00], for stopping times, martingales, the Itô integral and the Itô isometry; and Øksendal [Øks13], for the uniqueness and existence of solutions to SDEs, the Itô formula, the theory of Itô diffusions and the applications to PDEs.

Acknowledgments

This work would not have been possible without the help of my advisor, José Manuel Conde Alonso. I want to thank him for giving me complete freedom in choosing the topics for this project and for always answering my endless questions.

I would also like to thank Pablo Soto Martín for allowing me to use the code he developed in his Computer Science bachelor thesis [Sot24], implementing the Euler-Maruyama method for the numerical approximation of SDE solutions. His implementation, integrated within (although not yet merged into) the *scikit-fda* Python package [Ram+24], has been employed to generate the time-continuous paths for the figures in this document.

CHAPTER 2

Stochastic processes and Brownian motion

This chapter introduces the basic concepts of continuous-time stochastic processes (Section 2.1), and the stochastic process that will be central in this thesis: Brownian motion (Section 2.2). The property of Brownian motion that will be of particular interest is the Markov property, which denotes the time-homogeneity of the process (Section 2.3). Finally, we consider a particular kind of stochastic process: martingales (Section 2.4), which will be used in the construction of the stochastic integral.

2.1. Continuous-time stochastic processes

Similar to the discrete case, a stochastic process is a family of random variables indexed by a set. In this case, the set of indices need not be countable.

Definition 2.1 (Stochastic process). *Let T be a set, $(E,\mathscr{E})$ a measurable space (denoted the* state space*), and (Ω,Σ,P) a probability space. A* stochastic process indexed by T, taking values in E *is a family $(X_t)_{t\in T}$ of measurable functions (i.e. random variables) $X_t:(\Omega,\Sigma)\to(E,\mathscr{E})$, $\forall t\in T$.*

The mappings $(t\mapsto X_t(\omega))_{\omega\in\Omega}$ are called the paths *of the process X.*

We may think of the paths of a stochastic process as points chosen randomly in the space of all functions from T to E. A stronger (more restrictive) definition of stochastic process than the one given in Definition 2.1 would be: a random variable taking values in the measurable space of $T\to E$ functions:

$$X:(\Omega,\Sigma,P)\to(E^T,\mathscr{E}^T). \tag{2.1}$$

This definition agrees with Definition 2.1 if we put $X_t:=\pi_t\circ X$, where $\pi_t:E^T\to E$ is the projection of E^T onto the t-th coordinate. The product σ-algebra $\mathscr{E}^T$ is the smallest σ-algebra that makes the projection maps π_t measurable for all $t\in T$, so each $\pi_t\circ X$ is an E-valued random variable. Similarly, if $F\subsetneq T$, we denote by π_F the projection of E^T onto E^F. We denote by $\mathscr{P}_{\text{fin}}(T)$ the collection of all finite subsets of T. See Appendix B for more details on the (possibly infinite) product of measurable spaces.

Let X be a stochastic process indexed by T and taking values in $(E,\mathscr{E})$. For each $F\in\mathscr{P}_{\text{fin}}(T)$, X induces the probability measure $P_{X,F}$ on $(E^F,\mathscr{E}^F)$ given by $P_{X,F}(A)=P(\pi_F(X)\in A)$ for each $A\in\mathscr{E}^F$. More explicitly, if $F=\{t_1,\ldots,t_k\}$ and $A_i\in\mathscr{E}$, for all $1\le i\le k$, then $P_{X,F}(A_1\times\cdots\times A_k)=P(X_{t_1}\in A_1,\ldots,X_{t_k}\in A_k)$; which is extended from the rectangles of $\mathscr{E}^F$ to the whole σ-algebra by Carathéodory's extension theorem. The family of probability measures $(P_{X,F})_{F\in\mathscr{P}_{\text{fin}}(T)}$ is called the *family of finite-dimensional distributions of X*.

Let $X:(\Omega,\Sigma,P)\to E^T$ and $X':(\Omega',\Sigma',P')\to E^T$ be two T-indexed E-valued stochastic processes. We say that X and X' are *equivalent* (or that X *is a version of* X') if they have the same finite-dimensional distributions: $\forall F\in\mathscr{P}_{\text{fin}}(T)$, $P_{X,F}=P'_{X',F}$. If in addition $(\Omega,\Sigma,P)=(\Omega',\Sigma',P')$, we say that:

(a) X and X' are *modifications of each other* if $\forall t\in T$, $P(X_t=X'_t)=1$.

(b) X and X' are *indistinguishable* if $P(X=X')=P(\forall t\in T\ X_t=X'_t)=1$.

Remark 2.2. *(b) $\implies$ (a). Additionally, if T is a separable topological space, X and X' are modifications of each other, and the paths of X and X' are a.e. continuous, then they are indistinguishable.*

The following result, known as Kolmogorov's continuity criterion gives a criterion for the path-continuity of stochastic processes. See Appendix D for a proof.

Corollary 2.3 (Kolmogorov's continuity criterion). *An $\mathbb{R}$-valued, $[0,\infty)$-indexed stochastic process X for which there exist $\gamma,\varepsilon,c>0$ such that $E\left[|X_{t+h}-X_t|^\gamma\right]\le ch^{1+\varepsilon}$ for every $t,h>0$ has a modification that is almost-surely continuous.*

2.2. Brownian motion

Brownian motion generalizes the notion of a symmetric random walk in continuous time. The discovery of this process is often credited to the botanist Robert Brown, who described the seemingly random, jittery motion of pollen particles suspended in water. In this section, we define Brownian motion by its properties. Then, we prove its existence.

From now on, we will denote $\mathbb{R}_+ = [0,\infty)$. If not stated otherwise, we will consider stochastic processes indexed by $\mathbb{R}_+$ and taking values in $\mathbb{R}$.

Definition 2.4 (Brownian motion). *Let $x \in \mathbb{R}$. A* (1-dimensional) Brownian motion (starting at x) *is a stochastic process*

$$B : (\Omega, \Sigma, P) \xrightarrow{\text{measurable}} \left(\mathbb{R}^{\mathbb{R}_+}, \mathscr{B}(\mathbb{R})^{\mathbb{R}_+}\right),$$

with the properties:

1. *The increments on disjoint intervals are independent; that is, if $0 \le t_1 < t_2 < \cdots < t_k$, then the random variables $\{B_{t_2} - B_{t_1}, B_{t_3} - B_{t_2}, \ldots, B_{t_k} - B_{t_{k-1}}\}$ are independent.*
2. *If $0 \le s < t$, then the increment $B_t - B_s$ of the process on the interval $(s,t]$ is a Gaussian random variable with mean 0 and variance $t-s$.*
3. *$B_0 = x$, almost surely.*
4. *The paths $t \mapsto B_t(\omega)$ are almost surely continuous.*

Remark 2.5. *Once we have a Brownian motion B^0 starting at 0, then, for any $x \in \mathbb{R}$, the process $B^x = x + B^0$ is a Brownian motion starting at x.*

We will proceed to prove the existence of Brownian motions.

Lemma 2.6. *Let $\mathscr{H}$ be a separable real Hilbert space. There exists a probability space (Ω, Σ, P) and a linear mapping $\mathfrak{g} : \mathscr{H} \to L^2(\Omega, \Sigma, P)$ such that for each $h \in \mathscr{H}$, $\mathfrak{g}(h)$ is a Gaussian random variable with mean 0 and variance*

$$\|\mathfrak{g}(h)\|^2_{L^2(\Omega,\Sigma,P)} = \|h\|^2_{\mathscr{H}}.$$

In particular, $\mathfrak{g}$ is a linear isometry between Hilbert spaces.

Proof. Let $(e_n)_{n\in J}$ be a countable or finite orthonormal basis of $\mathscr{H}$. Let $(Z_n)_{n\in J}$ be the sequence of independent standard Gaussian random variables given by Theorem *B.7*. By the central limit theorem, $\mathfrak{g}(h) := \sum_{n\in J} \langle h, e_n\rangle Z_n$ is a Gaussian random variable with mean 0 and variance

$$E[|\mathfrak{g}(h)|^2] = \sum_{n\in J} \langle h, e_n\rangle^2 E[Z_n^2] = \sum_{n\in J} \langle h, e_n\rangle^2 = \|h\|^2_{\mathscr{H}}. \qquad \square$$

A closed linear subspace of $L^2(\Omega, \Sigma, P)$ consisting only of centered Gaussian random variables is called a *Gaussian space.* $\mathfrak{g}(\mathscr{H}) = \{\mathfrak{g}(h) : h \in \mathscr{H}\}$ is a Gaussian space.

Remark 2.7. *If G is a Gaussian space, then for each $d \in \mathbb{N}$, every random vector in G^d is a Gaussian random vector in $\mathbb{R}^d$. This follows from the fact that $V : \Omega \to \mathbb{R}^d$ is a Gaussian random vector if and only if $\langle V, a\rangle$ is a Gaussian random variable for all $a \in \mathbb{R}^d$.*

In Lemma 2.6, we want to choose $\mathscr{H} = L^2(A, \mathscr{A}, \mu)$, where μ is σ-finite and $\mathscr{H}$ is separable. When $F \in \mathscr{A}$ and $\mu(F) < \infty$, we will write $\mathfrak{g}(F)$ instead of $\mathfrak{g}(\mathbb{1}_F)$.

Remark 2.8. *By Lemma 2.6, if $F, G \in \mathscr{A}$ have finite measure, then*

$$E[\mathfrak{g}(F)\mathfrak{g}(G)] = \langle \mathfrak{g}(\mathbb{1}_F), \mathfrak{g}(\mathbb{1}_G)\rangle_{L^2(\Omega)} = \langle \mathbb{1}_F, \mathbb{1}_G\rangle_{L^2(A)} = \mu(F \cap G).$$

Therefore, if $F_1, \ldots, F_k \in \mathscr{A}$ are pairwise disjoint and have finite measure, then $\{\mathfrak{g}(F_1), \ldots, \mathfrak{g}(F_k)\}$ are pairwise uncorrelated, hence independent. This is a consequence of Remark 2.7, and the fact that the coordinates of a Gaussian random vector are pairwise uncorrelated iff they are independent.

Theorem 2.9. *There exists a probability space (Ω, Σ, P) and a stochastic process*

$$B : (\Omega, \Sigma, P) \xrightarrow{\text{measurable}} \left(\mathbb{R}^{\mathbb{R}_+}, \mathscr{B}(\mathbb{R})^{\mathbb{R}_+}\right)$$

satisfying conditions 1, 2 and 3 of Definition 2.4 (of a Brownian motion starting at 0).

Proof. Take $(A, \mathscr{A}, \mu) = (\mathbb{R}_+, \mathscr{B}(\mathbb{R}_+), m)$, where m is the Lebesgue measure. Applying Lemma 2.6, there exists a probability space (Ω', Σ', P') and a linear isometry $\mathfrak{g} : L^2(\mathbb{R}_+) \to L^2(\Omega)$ such that for each $f \in L^2(\mathbb{R}_+)$, $\mathfrak{g}(f)$ is a Gaussian random variable with mean 0 and variance $\|f\|^2_{L^2(\mathbb{R}_+)}$.

We define $X_t := \mathfrak{g}([0,t]) = \mathfrak{g}(\mathbb{1}_{[0,t]})$. Note that $X_t - X_s = \mathfrak{g}((s,t])$. By the linearity of $\mathfrak{g}$ and Remark 2.8, conditions 1 and 2 of Definition 2.4 are satisfied. The family $(P_{X,F})_{F \in \mathscr{P}_{\text{fin}}(\mathbb{R}_+)}$ of finite-dimensional distributions of X ($P_{X,F}$ is a probability measure on $\mathscr{B}(\mathbb{R})^F$) is Kolmogorov consistent by definition (see Definition *B*.1). By the Kolmogorov extension theorem (Theorem *B*.6), there exists a probability measure P on $\mathbb{R}^{\mathbb{R}_+}$ such that $P_{X,F} = P \circ \pi_F^{-1}$ for all $F \in \mathscr{P}_{\text{fin}}(\mathbb{R}_+)$. Finally, take $(\Omega, \Sigma, P) = \left(\mathbb{R}^{\mathbb{R}_+}, \mathscr{B}(\mathbb{R})^{\mathbb{R}_+}, P\right)$ and $B = \text{id}_{\mathbb{R}^{\mathbb{R}_+}}$. □

The process B resulting from Theorem 2.9 satisfies $E\left[|B_{t+h} - B_t|^4\right] = 3h^2 \quad \forall t, h > 0$, because the increment is a Gaussian random variable. Kolmogorov's continuity criterion (Corollary 2.3) applies and we get:

Theorem 2.10. *For each $x \in \mathbb{R}$, there exists a Brownian motion starting at x.*

Definition 2.11 (n-dimensional Brownian motion). *Let $n \in \mathbb{N}$. An* n-dimensional Brownian motion starting at *$x = (x_1, \dots, x_n) \in \mathbb{R}^n$ is an $\mathbb{R}^n$-valued stochastic process $(B_t)_{t\geq 0} = B = (B_1, \dots, B_n)$ such that for each $1 \leq i \leq n$, B_i is a Brownian motion starting at x_i and $B_1, \dots, B_n$ are independent.*

We can equivalently define an n-dimensional Brownian motion by its properties, just like we did in the one-dimensional case. An $\mathbb{R}_+$-indexed, $\mathbb{R}^n$-valued stochastic process B is an n-dimensional Brownian motion starting at x iff (1) the increments on disjoint intervals are independent, (2) $0 \leq s < t \implies B_t - B_s \sim N_n(\mathbf{0}, (t-s)\mathbf{I}_n)$, (3) $B_0 = x$ a.e., and (4) the paths $t \mapsto B_t(\omega)$ are a.e. continuous.

When working with an n-dimensional Brownian motion, we will denote the components of the process by $B_1, \dots, B_n$, and the process itself by $B = (B_1, \dots, B_n)$.

2.3. Filtrations, stopping times and the strong Markov property

The time-homogeneity of Brownian motion is modeled by the Markov property, which roughly states that the future of the process is independent of its past given its present; the process has no memory. The strong Markov property is a stronger version of this that will be described after we define the notion of stopping times. To formalize these ideas, we first introduce the concept of a filtration.

Definition 2.12. *Let $I \subseteq \mathbb{R}$. A* filtration $(\mathscr{F}_t)_{t \in I}$ (indexed by I) *on the measurable space (Ω, Σ) is a family of σ-algebras such that $\mathscr{F}_s \subseteq \mathscr{F}_t \subseteq \Sigma$ for all $s \leq t$. In that case, $\left(\Omega, \Sigma, (\mathscr{F}_t)_{t\in I}\right)$ is said to be a* filtered space, *and if (Ω, Σ, P) is a probability space, then $\left(\Omega, \Sigma, (\mathscr{F}_t)_{t\in I}, P\right)$ is said to be a* filtered probability space.

We will usually have $I = \mathbb{R}_+$ or $I = \mathbb{N}$. For practical purposes, we often think of an element $\mathscr{F}_t$ of the filtration $(\mathscr{F}_t)_{t\geq 0}$ as representing the information available at each time t.

Definition 2.13. *A stochastic process $(X_t)_{t\geq 0}$ on (Ω, Σ, P) is* $(\mathscr{F}_t)_{t\geq 0}$-adapted *if for all $t \geq 0$, X_t is $\mathscr{F}_t$-measurable.*

Every stochastic process $(X_t)_{t\geq 0}$ on (Ω, Σ, P) is adapted to its natural filtration $(\mathscr{X}_t)_{t\geq 0}$, defined by $\mathscr{X}_t := \sigma(X_s : 0 \leq s \leq t)$. In fact, this is the minimal filtration to which $(X_t)_{t\geq 0}$ is adapted. Saying that $(X_t)_{t\geq 0}$ is adapted to $(\mathscr{F}_t)_{t\geq 0}$ is equivalent to saying that $(\mathscr{X}_t)_{t\geq 0}$ is a subfiltration of $(\mathscr{F}_t)_{t\geq 0}$, meaning that $\mathscr{X}_t \subseteq \mathscr{F}_t$ for all $t \geq 0$. In a sense, $\mathscr{X}_t$ contains the information carried by the process up to time t.

Definition 2.14. *Given a filtered space $\left(\Omega, \Sigma, (\mathscr{F}_t)_{t\geq 0}\right)$, we define the filtration $(\mathscr{F}_{t+})_{t\geq 0}$ by:*

$$\mathscr{F}_{t+} = \bigcap_{s>t} \mathscr{F}_s.$$

We say that $(\mathscr{F}_t)_{t\geq 0}$ is right-continuous *if $\mathscr{F}_t = \mathscr{F}_{t+}$ for all $t \geq 0$; and define $\mathscr{F}_\infty := \sigma\left((\mathscr{F}_t)_{t\geq 0}\right) = \sigma\left((\mathscr{F}_{t+})_{t\geq 0}\right)$.*

For any filtration $(\mathscr{F}_t)_{t\geq 0}$, the filtration $(\mathscr{F}_{t+})_{t\geq 0}$ is right-continuous. Throughout this section, let $(B_t)_{t\geq 0}$ be a Brownian motion and $(\mathscr{B}_t)_{t\geq 0}$ its natural filtration. The Markov property for Brownian motion is a corollary of its definition:

Theorem 2.15 (Markov property for Brownian motion). *For all $s \geq 0$, the process $Y_t = B_{t+s} - B_s$ is a Brownian motion independent of $\mathscr{B}_s$.*

Definition 2.16 (Stopping time). *Let $\left(\Omega, \Sigma, (\mathscr{F}_t)_{t\geq 0}, P\right)$ be a filtered probability space. A random variable $T : \Omega \to [0, \infty]$ is a $(\mathscr{F}_t)_{t\geq 0}$-*stopping time *if for all $t \geq 0$, $\{T \leq t\} \in \mathscr{F}_t$. We say that a stopping time T is* finite *if $T < \infty$ a.e.*

*The σ-*algebra $\mathscr{F}_T$ of events prior to T *is: $\mathscr{F}_T = \{A \in \Sigma : A \cap \{T \leq t\} \in \mathscr{F}_t \ \forall t \geq 0\}$.*

It is sometimes useful to consider: $\mathscr{F}_{T+} = \{A \in \Sigma : A \cap \{T < t\} \in \mathscr{F}_t \ \forall t \geq 0\}$.

As is usual in this context, from now on we denote by $a \vee b$ the maximum of two real numbers a and b, and by $a \wedge b$ their minimum.

Theorem 2.17. *Let $\left(\Omega, \Sigma, (\mathscr{F}_t)_{t\geq 0}, P\right)$ be a filtered probability space and T, S two $(\mathscr{F}_t)_{t\geq 0}$-stopping times. Then:*

(a) $\mathscr{F}_T \subseteq \mathscr{F}_{T+}$.
(b) T *is* $\mathscr{F}_T$*-measurable.*
(c) $S \leq T \implies (\mathscr{F}_S \subseteq \mathscr{F}_T)$ *and* $(\mathscr{F}_{S+} \subseteq \mathscr{F}_{T+})$.
(d) $S < T \implies \mathscr{F}_{S+} \subseteq \mathscr{F}_T$.
(e) If U is an $\mathscr{F}_T$-measurable random variable and $T \leq U$, then U is an $(\mathscr{F}_t)_{t\geq 0}$-stopping time.
(f) $S+T$, $S \wedge T$ and $S \vee T$ are $(\mathscr{F}_t)_{t\geq 0}$-stopping times, and $\mathscr{F}_{S\wedge T} = \mathscr{F}_S \cap \mathscr{F}_T$.

Proof. (a) Let $A \in \mathscr{F}_T$. Then $A \cap \{T < t\} = \cup_{n\in\mathbb{N}} A \cap \{T \leq t - 1/n\} \in \mathscr{F}_t\ \forall t \geq 0$.

(b) Let $s \geq 0$, then $\forall t \geq 0\ \{T \leq t\} \cap \{T \leq s\} = \{T \leq t \wedge s\} \in \mathscr{F}_t$, so $\{T \leq s\} \in \mathscr{F}_T$.

(c) Let $A \in \mathscr{F}_S$. Then for every $t \geq 0$, $A \cap \{T \leq t\} = A \cap \{T \leq t\} \cap \{S \leq t\} \in \mathscr{F}_t$, which implies that $A \in \mathscr{F}_T$. An almost exact argument works for $\mathscr{F}_{S+} \subseteq \mathscr{F}_{T+}$.

(d) $A \in \mathscr{F}_{S+} \implies \forall t \geq 0$, $A \cap \{T \leq t\} = A \cap \{T \leq t\} \cap \{S < t\} = \{T \leq t\} \cap (\bigcup_n A \cap \{S < t - 1/n\}) \in \mathscr{F}_t$.

(e) $\forall t \geq 0$, $\{U \leq t\} = \{U \leq t\} \cap \{T \leq t\}$ because $T \leq U$, and $\{U \leq t\} \cap \{T \leq t\} \in \mathscr{F}_t$ because $\{U \leq t\} \in \mathscr{F}_T$.

(f) $\forall t \geq 0$, $\{S \vee T \leq t\} = \{S \leq t\} \cap \{T \leq t\} \in \mathscr{F}_t$ and $\{S \wedge T \leq t\} = \{S \leq t\} \cup \{T \leq t\} \in \mathscr{F}_t$; so both $S \wedge T$ and $S \vee T$ are stopping times.

$$\begin{aligned}\forall t \geq 0,\ \{S+T \leq t\}^c &= \{S+T > t\} = \{t - T < S\} = \{S > t\} \cup \{T > t\} \cup \bigcup_{q\in\mathbb{Q}\cap(0,t)} \{t - T < q < S\} \\ &= \{S > t\} \cup \{T > t\} \cup \bigcup_{q\in\mathbb{Q}\cap(0,t)} \left(\{T > t-q\} \cap \{S > q\}\right) \in \mathscr{F}_t;\ \text{so } S+T \text{ is a stopping time.}\end{aligned}$$

Let $A \in \mathscr{F}_{S\wedge T}$; then $\forall t \geq 0$, $A \cap \{T \leq t\} = (A \cap \{T \wedge S \leq t\}) \cap \{T \leq t\} \in \mathscr{F}_t$. This implies that $\mathscr{F}_{S\wedge T} \subseteq \mathscr{F}_T$; by a symmetrical argument, $\mathscr{F}_{S\wedge T} \subseteq \mathscr{F}_S$. For the other inclusion, let $A \in \mathscr{F}_S \cap \mathscr{F}_T$; then $\forall t \geq 0$, $A \cap \{S \wedge T \leq t\} = (A \cap \{S \leq t\}) \cup (A \cap \{T \leq t\}) \in \mathscr{F}_t$; so $A \in \mathscr{F}_{S\wedge T}$. □

Proposition 2.18. *If T is a random variable and $(\mathscr{F}_t)_{t\geq 0}$ a filtration, then:*

$$\forall t \geq 0\ \{T \leq t\} \in \mathscr{F}_t \implies \forall t \geq 0\ \{T < t\} \in \mathscr{F}_t \implies \forall t \geq 0\ \{T \leq t\} \in \mathscr{F}_{t+}.$$

Proof. For the first implication: if T is a stopping time, then $\{T < t\} = \bigcup_{n\in\mathbb{N}} \{T \leq t - 1/n\} \in \mathscr{F}_t$.

For the second implication, suppose that $\{T < s\} \in \mathscr{F}_s$ for all $s \geq 0$. Let $0 \leq t < s$, then $\exists n_{ts} \in \mathbb{N} : \forall m \geq n_{ts}$, $t + 1/m < s$; therefore, $\forall m \geq n_{ts}$, $\{T < t + 1/m\} \in \mathscr{F}_{t+1/m} \subseteq \mathscr{F}_s$ and

$$\{T \leq t\} = \bigcap_{m\geq n_{ts}} \{T < t + 1/m\} \in \mathscr{F}_s.$$

We have proven that $\{T \leq t\} \in \bigcap_{s>t} \mathscr{F}_s = \mathscr{F}_{t+}$. □

Corollary 2.19. *(a) A random variable T is a stopping time with respect to a right-continuous filtration $(\mathscr{F}_t)_{t\geq 0}$ if and only if $\{T < t\} \in \mathscr{F}_t$ for all $t \geq 0$.*
(b) If $(T_n)_{n\in\mathbb{N}}$ is a sequence of stopping times with respect to a right-continuous filtration $(\mathscr{F}_t)_{t\geq 0}$, then $\inf_n T_n$ *is a stopping time.*

Proof. (a) is direct. (b) $\forall t \leq 0 \quad \{\inf_n T_n < t\} = \bigcup_{n\in\mathbb{N}} \{T_n < t\} \in \mathscr{F}_t$, and apply (a). □

Definition 2.20 (Progressively measurable processes). *Given a filtered space $\left(\Omega, \Sigma, (\mathscr{F}_t)_{t\geq 0}\right)$, a stochastic process*

$$X : \mathbb{R}_+ \times \Omega \to \mathbb{R}^n$$

is $(\mathscr{F}_t)_{t\geq 0}$-progressively measurable *(or simply* progressive*) if for all $t \geq 0$, the map $(s, \omega) \mapsto X_s(\omega)$ from $[0, t] \times \Omega$ to $\mathbb{R}^n$ is $(\mathscr{B}([0, t]) \otimes \mathscr{F}_t)$-measurable.*

Proposition 2.21. *An adapted process with right or left continuous paths is progressively measurable.*

Proof. Let X be an adapted process with right-continuous paths (the left-continuous case is almost identical). Fix $t \geq 0$. Let X be an adapted process with right continuous paths. Define for each $n \in \mathbb{N}$:

$$X_s^{(n)}(\omega) := \mathbb{1}_{\{0\}}(s) X_0(\omega) + \sum_{k=0}^{2^n-1} \mathbb{1}_{\left(\frac{k}{2^n}t, \frac{k+1}{2^n}t\right]}(s) X_{\frac{k+1}{2^n}t}(\omega),$$

which are $\mathscr{B}([0, t]) \otimes \mathscr{F}_t$ measurable because the $X^{(n)}$ are simple on the s variable and each of the $X_{\frac{k}{2^n}t}$ is $\mathscr{F}_t$-measurable. Since X is right-continuous, $X_s^{(n)}(\omega) \to X_s(\omega)$ pointwise for all $s \in [0, t]$ and $\omega \in \Omega$. The pointwise limit of measurable functions is measurable, so X is progressively measurable. □

Proposition 2.22. *Let T be an $(\mathscr{F}_t)_{t\geq 0}$-stopping time and X an $(\mathscr{F}_t)_{t\geq 0}$-progressive process. Then, the random variable $\mathbb{1}_{\{T<\infty\}}X_T : \omega \mapsto \mathbb{1}_{\{T<\infty\}}(\omega)X_{T(\omega)}(\omega)$ is $\mathscr{F}_T$-measurable.*

Proof. We have to prove that $\mathbb{1}_{\{T<\infty\}}X_T$ is $\mathscr{F}_T$-measurable. By definition of $\mathscr{F}_T$, this is equivalent to:

$$\begin{aligned}
&\forall t\geq 0\ \forall B\in\mathscr{B}(\mathbb{R}^n) \quad \left\{\mathbb{1}_{\{T<\infty\}}X_T\in B\right\}\cap\{T\leq t\} = \{X_T\in B\}\cap\{T\leq t\}\in\mathscr{F}_t\\
&\iff \forall t\geq 0\ \forall B\in\mathscr{B}(\mathbb{R}^n) \quad \{X_T\in B\}\cap\{T\leq t\}\in\mathscr{F}_t\cap\{T\leq t\}\\
&\iff \forall t\geq 0 \quad X_T|_{\{T\leq t\}} \text{ is } \mathscr{F}_t\cap\{T\leq t\}\text{-measurable.}
\end{aligned}$$

$T|_{\{T\leq t\}} : (\{T\leq t\},\mathscr{F}_t\cap\{T\leq t\}) \to ([0,t],\mathscr{B}([0,t]))$ is measurable, because T is a stopping time:

$$\forall 0\leq a<b\leq t \quad \{T\leq b\}\in\mathscr{F}_b\subseteq\mathscr{F}_t,\ \{T\leq a\}\in\mathscr{F}_a\subseteq\mathscr{F}_t \implies T^{-1}((a,b]) = \{T\leq b\}\setminus\{T\leq a\}\in\mathscr{F}_t\cap\{T\leq t\}.$$

Therefore, the map

$$\begin{aligned}
\eta : (\{T\leq t\},\mathscr{F}_t\cap\{T\leq t\}) &\to ([0,t]\times\Omega,\mathscr{B}([0,t])\otimes\mathscr{F}_t)\\
\omega &\mapsto (T(\omega),\omega)
\end{aligned}$$

is measurable. Since X is progressive,

$$\begin{aligned}
X|_{[0,t]\times\Omega} : ([0,t]\times\Omega,\mathscr{B}([0,t])\otimes\mathscr{F}_t) &\to \left(\mathbb{R}^n,\mathscr{B}(\mathbb{R}^n)\right)\\
(s,\omega) &\mapsto X_s(\omega)
\end{aligned}$$

is measurable. Finally, the composition $X_T|_{\{T\leq t\}} = (X|_{[0,t]\times\Omega})\circ\eta$ is $\mathscr{F}_t\cap\{T\leq t\}$-measurable, as desired. □

Combining Proposition 2.21 and Proposition 2.22, we get:

Corollary 2.23. *If T is an $(\mathscr{F}_t)_{t\geq 0}$-stopping time and X an $(\mathscr{F}_t)_{t\geq 0}$-adapted process of right(or left)-continuous paths. Then, the random variable $\mathbb{1}_{\{T<\infty\}}X_T : \omega\mapsto X_{T(\omega)}(\omega)$ is $\mathscr{F}_T$-measurable.*

This applies to Brownian motion: if T is a finite $(\mathscr{B}_t)_{t\geq 0}$-stopping time, then B_T is $\mathscr{B}_T$-measurable.

Theorem 2.24 (Strong Markov property for Brownian motion). *Let T be a finite $(\mathscr{B}_t)_{t\geq 0}$-stopping time. Then $X_t = B_{T+t}-B_T$ is a Brownian motion independent of $\mathscr{B}_T$.*

Proof. It is clear that X has continuous paths and $X_0 = 0$ a.e. We have to check the other two properties of Brownian motion (independent Gaussian increments), as well as independence of $\mathscr{B}_T$. For each $n\in\mathbb{N}$, define

$$T_n := \sum_{k=1}^{\infty}\frac{k}{n}\mathbb{1}_{\left\{\frac{k-1}{n}\leq T<\frac{k}{n}\right\}}.$$

Using Theorem 2.17(e), the T_n are stopping times. Fix n and let $C\in\mathscr{B}_T$. Then, using Proposition 2.18,

$$C\cap\left\{T_n=\frac{k}{n}\right\} = \left(C\cap\left\{T<\frac{k}{n}\right\}\right)\cap\left\{\frac{k-1}{n}\leq T\right\} = \bigcup_m\left(C\cap\left\{T\leq\frac{k}{n}-\frac{1}{m}\right\}\right)\cap\left\{T<\frac{k-1}{n}\right\}^c\in\mathscr{B}_{k/n}.$$

Now, we apply Theorem 2.15: let $A\in\mathscr{B}(\mathbb{R})$, $0\leq s\leq t$, then

$$\begin{aligned}
&P\left(C\cap\left\{(B_{t+T_n}-B_{s+T_n})\in A\right\}\right) = \sum_{k=1}^{\infty}P\left(C\cap\{T_n=k/n\}\cap\left\{(B_{t+k/n}-B_{s+k/n})\in A\right\}\right)\\
&= \sum_{k=1}^{\infty}P(C\cap\{T_n=k/n\})\,P((B_{t+k/n}-B_{s+k/n})\in A) = \sum_{k=1}^{\infty}P(C\cap\{T_n=k/n\})\,P(B_t-B_s\in A)\\
&= P(C)\,P(B_t-B_s\in A).
\end{aligned} \tag{2.2}$$

Suppose that $A=(a,b)\subseteq\mathbb{R}$. Note that $\left(B_{t+T_n}-B_{s+T_n}\right)\xrightarrow{n\to\infty}(B_{t+T}-B_{s+T})$ a.e., by path-continuity of B and $T_n\to T$ a.e. This implies that

$$\begin{aligned}
\varepsilon>0 \implies P(B_{t+T}-B_{s+T}=a) &\leq P\left(\liminf_n\left\{a-\varepsilon<B_{t+T_n}-B_{s+T_n}<a+\varepsilon\right\}\right)\\
&\leq \liminf_n P\left(a-\varepsilon<B_{t+T_n}-B_{s+T_n}<a+\varepsilon\right) = P(a-\varepsilon<B_t-B_s<a+\varepsilon)\xrightarrow{\varepsilon\to 0^+}0.
\end{aligned}$$

This implies that $\forall a \in \mathbb{R},\ P(B_{t+T} - B_{s+T} = a) = 0$. Moreover,

$$P(C \cap \{a < B_{t+T} - B_{s+T} < b\}) \le P\left(\liminf_n C \cap \left\{a < B_{t+T_n} - B_{s+T_n} < b\right\}\right) \tag{2.3}$$
$$\le \liminf_n P\left(C \cap \left\{a < B_{t+T_n} - B_{s+T_n} < b\right\}\right) \le \limsup_n P\left(C \cap \left\{a < B_{t+T_n} - B_{s+T_n} < b\right\}\right)$$
$$\le P\left(\limsup_n C \cap \left\{a < B_{t+T_n} - B_{s+T_n} < b\right\}\right) \le P(C \cap \{a \le B_{t+T} - B_{s+T} \le b\}).$$

As $P(B_{t+T} - B_{s+T} = a) = P(B_{t+T} - B_{s+T} = b) = 0$, the previous inequalities are all equalities and (by (2.2))

$$P(\{a < X_t - X_s < b\} \cap C) = P(\{a < B_{t+T} - B_{s+T} < b\} \cap C) = P(a < B_t - B_s < b)\, P(C).$$

This proves that $\forall 0 \le s \le t$, $X_t - X_s$ has the same distribution as $B_t - B_s$, and that X is independent of $\mathscr{B}_T$[1]. The proof will be complete once we check that X has independent increments.

Replicating the idea of (2.2), we have, for $0 \le t_0 \le \cdots \le t_j$ and $A_1, \ldots, A_j \in \mathscr{B}(\mathbb{R})$:

$$\begin{aligned}
P\left(\bigcap_{i=1}^{j} \{B_{t_i+T_n} - B_{t_{i-1}+T_n} \in A_i\}\right) &= \sum_{k=1}^{\infty} P\left(\{T_n = k/n\} \cap \bigcap_{i=1}^{j} \{B_{t_i+k/n} - B_{t_{i-1}+k/n} \in A_i\}\right) \\
&= \sum_{k=1}^{\infty} P(T_n = k/n) \prod_{i=1}^{j} P\left(B_{t_i+k/n} - B_{t_{i-1}+k/n} \in A_i\right) \\
&= \prod_{i=1}^{j} P\left(B_{t_i} - B_{t_{i-1}} \in A_i\right) = \prod_{i=1}^{j} P\left(B_{t_i+T_n} - B_{t_{i-1}+T_n} \in A_i\right),
\end{aligned}$$

which proves the independence of the increments of every $\left(B_{t+T_n} - B_{T_n}\right)_{t \ge 0}$. By a very similar argument as that of (2.3), we can infer the independence of the increments of $X = (B_{t+T} - B_T)_{t \ge 0}$. □

2.4. Martingales

In this section, we introduce the concept of martingales, and present a few results that will assist the later definition of the stochastic integral. Martingales are a type of stochastic process that model a fair game: the expected value of the process at a future time is equal to its current value. The name "martingale" was used to denote a certain class of betting strategies, popular in 18th-century France, that claimed to guarantee a profit in a fair game. Part of the motivation for the study of martingales comes from disproving the superiority of such strategies over any other in fair games.

Definition 2.25 (Martingale). *Let $\left(\Omega, \Sigma, (\mathscr{F}_t)_{t \ge 0}, P\right)$ be a filtered probability space. A stochastic process $(X_t)_{t \ge 0}$ is a* martingale *if*

(a) $\forall t \ge 0$, $X_t \in L^1(\mathscr{F}_t)$, and

(b) $\forall 0 \le s \le t$, $X_s = E[X_t \mid \mathscr{F}_s]$.

If we change the equality in (b) to "$\le$", we say that $(X_t)_{t \ge 0}$ is a submartingale*, and if we change it to "$\ge$", we say that $(X_t)_{t \ge 0}$ is a* supermartingale.

Whenever we say that $(X_t)_{t \ge 0}$ is a (sub/super)martingale without specifying the filtration, it is implied that $(X_t)_{t \ge 0}$ is a (sub/super)martingale with respect to its natural filtration $(\mathscr{X}_t)_{t \ge 0}$.

Example 2.26. *Directly from its definition, a Brownian motion $(B_t)_{t \ge 0}$ is a martingale.*

Theorem 2.27 (Doob's martingale inequality). *Let $(X_t)_{t \ge 0}$ be a right-continuous submartingale w.r.t. $(\mathscr{F}_t)_{t \ge 0}$. Then*

$$\forall \lambda > 0 \quad \lambda P\left(\left\{\sup_{0 \le s \le t} X_s > \lambda\right\}\right) \le E\left[\mathbb{1}_{\{\sup_{0 \le s \le t} X_s \ge \lambda\}} X_t\right].$$

Proof. The finite index case is proven in [AD00, Lemma 9.6.4]: fix λ and let $k \in \mathbb{N}$, $0 \le t_1 < t_2 < \cdots < t_k \le t$. Then

$$\lambda P\left(\max_{1 \le i \le k} X_{t_i} > \lambda\right) \le E\left[\mathbb{1}_{\left\{\max_{1 \le i \le k} X_{t_i} > \lambda\right\}} X_{t_k}\right].$$

[1] This is a bit subtle, as we have only proven that the pre-image by $X_t - X_s$ of $\{(a,b) \subseteq \mathbb{R} : a < b\}$ is independent of $\mathscr{B}_T$. However, $\mathscr{B}_T$ and $\left\{(X_t - X_s)^{-1}((a,b)) \subseteq \mathbb{R} : a < b\right\}$ are both π-systems. Applying Theorem A.4, $X_t - X_s$ is independent of $\mathscr{B}_T$.

For the continuous case, let $Q_t := \{t\} \cup (\mathbb{Q} \cap [0,t])$, which is countable. Let $(I_k)_{k=1}^{\infty} \subseteq Q_t$ be an increasing sequence of finite sets such that $Q_t = \cup_{k=1}^{\infty} I_k$. Since $X_t \in L^1(P)$ and the sequence of sets $\left(\{\max_{s\in I_k} X_s > \lambda\}\right)_{k=1}^{\infty}$ increases to $\{\sup_{s\in Q_t} X_s > \lambda\}$, then

$$\forall k \in \mathbb{N} \quad \lambda P\left(\max_{s\in I_k} X_s > \lambda\right) \leq E\left[\mathbb{1}_{\left\{\max_{s\in I_k} X_s > \lambda\right\}} X_t\right]$$

implies (taking limits as $k \to \infty$ and applying the dominated convergence theorem)

$$\lambda P\left(\sup_{s\in Q_t} X_s > \lambda\right) \leq E\left[\mathbb{1}_{\left\{\sup_{s\in Q_t} X_s > \lambda\right\}} X_t\right].$$

By right-continuity of X and the fact that Q_t is dense in $[0,t]$, we have $\sup_{s\in Q_t} X_s = \sup_{0\leq s\leq t} X_s$. □

Lemma 2.28 (Corollary of the layer cake representation theorem). *Let (A,Γ,ν) be a measure space and m be the Lebesgue measure on $\mathscr{B}([0,\infty))$. If $0 < r < \infty$ and f, g are non-negative Γ-measurable functions, then*

$$\int_A g \cdot f^r \, d\nu = r \int_0^{\infty} t^{r-1} \left(\int_{\{f>t\}} g \, d\nu\right) dt.$$

Proof. Apply the layer cake representation theorem (Proposition A.5) to the function f and the (non-negative) measure ν_g given by $\Gamma \ni E \mapsto \int_E g \, d\nu$, which has the property $\int_A h \, d\nu_g = \int_A hg \, d\nu$ for every Γ-measurable non-negative function h. □

Theorem 2.29 (Doob's maximal inequality). *Let $1 < p < \infty$ $(X_t)_{t\geq 0} \subseteq L^p(\Omega)$ be a right-continuous $(\mathscr{F}_t)_{t\geq 0}$-martingale. Then*

$$\left\| \sup_{0\leq s\leq t} |X_s| \right\|_p \leq \frac{p}{p-1} \|X_t\|_p.$$

Proof. As X is a martingale, then $|X|$ is a submartingale by convexity of $|\cdot|$ and Corollary A.7.

Define $Y := \sup_{0\leq s\leq t} |X_s|$. Applying the layer cake representation (Proposition A.5) theorem to Y, Doob's martingale inequality (Theorem 2.27) and Lemma 2.28 to $f = Y$, $g = X_t$ and $r = p-1$,

$$\|Y\|_p^p = E\left[Y^p\right] = p\int_0^{\infty} t^{p-1} P(Y > t) \; dt \leq p\int_0^{\infty} t^{p-2} E\left[\mathbb{1}_{\{Y>t\}} X_t\right] \; dt = \frac{p}{p-1} E\left[X_t Y^{p-1}\right].$$

Applying Hölder's inequality with $1 = \frac{1}{p} + \frac{1}{q}$, $E\left[X_t Y^{p-1}\right] \leq \|X_t\|_p \left\|Y^{p-1}\right\|_q = \|X_t\|_p \|Y\|_p^{p-1}$. □

Proposition 2.30. *If $X = (X_n)_{n=0}^{\infty}$ is an $(\mathscr{F}_n)_{n=0}^{\infty}$-submartingale and $S \leq T$ are two bounded $(\mathscr{F}_n)_{n=0}^{\infty}$-stopping times, then $X_S \leq E[X_T \mid \mathscr{F}_S]$ with equality in the case of X being a martingale (and "$\geq$" if X were a supermartingale).*

Proof. Define $d_n := X_n - X_{n-1}$ and suppose that $S \leq T \leq M \in \mathbb{R}$. Then,

$$\forall n \in \mathbb{N} \quad X_n = X_0 + \sum_{j=1}^{n} d_j, \quad \text{and} \quad X_T = X_0 + \sum_{j=1}^{M} d_j \mathbb{1}_{\{j\leq T\}}.$$

Let $A \in \mathscr{F}_S$. We have:

$$\begin{aligned}
E[\mathbb{1}_A (X_T - X_S)] &= E\left[\mathbb{1}_A \sum_{j=0}^{M} d_j \mathbb{1}_{\{S<j\leq T\}}\right] = \sum_{j=0}^{M} E\left[d_j \mathbb{1}_{A\cap\{S<j\}\cap\{j\leq T\}}\right] \\
&= \sum_{j=0}^{M} E\left[E\left[d_j \mathbb{1}_{A\cap\{S<j\}\cap\{j\leq T\}} \;\middle|\; \mathscr{F}_{j-1}\right]\right] = \sum_{j=0}^{M} E\left[\mathbb{1}_{A\cap\{S<j\}\cap\{j\leq T\}} E\left[d_j \mid \mathscr{F}_{j-1}\right]\right] \geq 0.
\end{aligned}$$

The last inequality follows from the fact that X is a submartingale—it would be an equality (or the opposite inequality) if X were a (super)martingale. We have also used that $A \cap \{S < j\}, \{j \leq T\} \in \mathscr{F}_{j-1}$. □

Definition 2.31 (Upcrossings). *Let $f : T \subseteq \mathbb{R} \to \mathbb{R}$. Fix two real numbers $a < b$. For each finite subset $F = \{t_1 < \cdots < t_n\}$ of T, we define an increasing sequence $(s_k)_{k=1}^{\infty} = \left(s_k^{f,F}\right)_{k=1}^{\infty}$ inductively as follows:*

$$s_1 := \inf\{t_i : f(t_i) \leq a\}, \quad s_2 := \inf\{t_i > s_1 : f(t_i) \geq b\},$$

$$\forall k \geq 1 \qquad s_{2k+1} := \inf\{t_i > s_{2k} : f(t_i) \leq a\}, \quad s_{2k+2} := \inf\{t_i > s_{2k+1} : f(t_i) \geq b\},$$

where we put $\inf \emptyset = t_n$. *We define the number of* upcrossings of $[a, b]$ by f on F *as*

$$U_{ab}^f(F) := \sup\{k \geq 1 : s_{2k} < t_n\},$$

That is, the number of times within F that f rises above b after coming from below a. Then, we define the number of upcrossings of $[a, b]$ by f on T *as*

$$U_{ab}^f = U_{ab}^f(T) := \sup\left\{U_{ab}^f(F) : F \in \mathscr{P}_{fin}(T)\right\}.$$

Theorem 2.32 (Upcrossing theorem). *Let* $(X_t)_{t\geq 0}$ *be an* $(\mathscr{F}_t)_{t\geq 0}$*-submartingale,* $F = \{t_1 < \cdots < t_n\}$ *a finite subset of* $\mathbb{R}_+$ *and* $a < b$ *real numbers. Then* $U_{ab} := U_{ab}^X(F)$ *understood as* $\omega \mapsto U_{ab}^{t\mapsto X_t(\omega)}(F)$ *is a random variable. Moreover,*

$$(b-a)E[U_{ab}] \leq E\left[\left(X_{t_n} - a\right)^+\right].$$

Proof. Fix $(s_k)_{k=1}^\infty$ as in Definition 2.31: they are now a non-decreasing sequence of random variables bounded by t_n, and each s_k is an $(\mathscr{F}_t)_{t\geq 0}$-stopping time (because X is adapted). Therefore, we can apply Proposition 2.30 to these stopping times and X.

For each $k \in \mathbb{N}$, define $A_k := \{s_k < t_n\} \in \mathscr{F}_{s_k}$; as $(s_k)_k$ is non-decreasing, then $(A_k)_k$ is non-decreasing. By definition of U_{ab}, $P(A_{2k}) = P(U_{ab} \geq k)$ and, as U_{ab} takes values in $\mathbb{N}$, applying Proposition *A.5* we get $E[U_{ab}] = \sum_1^\infty P(U_{ab} \geq k) = \sum_1^\infty P(A_k)$.

$\mathbb{1}_{A_{2k}} X_{s_{2k}} \geq b$, $\mathbb{1}_{A_{2k+1}} X_{s_{2k+1}} \leq a$ and $\mathbb{1}_{A_{2k+1}^c} X_{s_{2k+1}} = t_n$. Applying Proposition 2.30 to $s_{2k} \leq s_{2k+1} \leq t_n$, we have: $\int_{A_{2k}} X_{s_{2k}}\, dP \leq \int_{A_{2k}} X_{s_{2k+1}}\, dP$. Then, we can write (for every $k \in \mathbb{N}$):

$$\begin{aligned}(b-a)P(A_{2k}) &= \int_{A_{2k}} (b-a)\, dP \leq \int_{A_{2k}} \left(X_{s_{2k+1}} - a\right)\, dP = \left(\int_{A_{2k+1}} + \int_{A_{2k}\setminus A_{2k+1}}\right)\left(X_{s_{2k+1}} - a\right)\, dP \\ &\leq 0 + \int_{A_{2k}\setminus A_{2k+1}} \left(X_{s_{2k+1}} - a\right)\, dP = \int_{A_{2k}\setminus A_{t_n}} \left(X_{s_{t_n}} - a\right)\, dP \leq \int_{A_{2k}\setminus A_{2k+1}} \left(X_{s_{t_n}} - a\right)^+\, dP.\end{aligned}$$

The sets $(A_{2k} \setminus A_{2k+1})_k$ are pairwise disjoint, and both sides of the previous equation are non-negative, so we can sum over $k \in \mathbb{N}$ to obtain:

$$(b-a)E[U_{ab}] = \sum_{k=1}^\infty (b-a)P(A_{2k}) \leq \sum_{k=1}^\infty \int_{A_{2k}\setminus A_{2k+1}} \left(X_{s_{t_n}} - a\right)^+\, dP \leq E\left[\left(X_{t_n} - a\right)^+\right]. \qquad \square$$

Theorem 2.33 (Submartingale convergence theorem). *Let* $(\mathscr{F}_n)_{n=0}^\infty$ *be a (discrete) filtration and* $(X_n)_{n=0}^\infty$ *an* $(\mathscr{F}_n)_{n=0}^\infty$*-submartingale such that* $\sup_n E\left[X_n^+\right] < \infty$*. Then, there exists an* $X_\infty \in L^1(P)$ *such that* $X_n \to X_\infty$ *a.e. (in particular,* $\lim_n X_n$ *exists and is finite a.e.).*

Proof. For each $n \in \mathbb{N}$, denote $F_n := \mathbb{N} \cap [0, n]$. Fix $a < b$ (real numbers) and let

$$A_{ab} := \left\{\liminf_n X_n < a < b < \limsup_n X_n\right\}.$$

The sequence of variables $(U_{ab}(F_n))_n$ is non-negative and increasing to $U_{ab}(\mathbb{N})$. By the monotone convergence theorem, $E[U_{ab}(\mathbb{N})] = \lim_{n\to\infty} E[U_{ab}(F_n)]$, applying Theorem 2.32, we get

$$E[U_{ab}(\mathbb{N})] \leq (b-a)^{-1} \sup_n E\left[(X_n - a)^+\right] \leq (b-a)^{-1}\left(\sup_n E\left[X_n^+\right] + (-a)^+\right) < \infty^{2}.$$

Which implies that $U_{ab}(\mathbb{N}) < \infty$ a.e. $A_{ab} \subseteq \{U_{ab}(\mathbb{N}) = \infty\}$, so $P(A_{ab}) = 0$.

$\{\nexists \lim_n X_n\}$[3] $= \bigcup_{a,b\in\mathbb{Q};\, a<b} A_{ab}$, which implies that the limit $\lim_n X_n$ exists a.e., we will call it X_∞. To prove that $X_\infty \in L^1(P)$, use that $|X_n| = X_n^+ + X_n^- = 2X_n^+ - X_n$ and $E[X_n] \geq E[X_0]$ (because X is a submartingale). By Fatou's lemma, $E[|X_\infty|] = E[\liminf_n |X_n|] \leq \liminf_n E[|X_n|] \leq 2\sup_n E\left[X_n^+\right] - E[X_0] < \infty$. $\square$

Theorem 2.34. *Let* $(\mathscr{F}_n)_{n=-\infty}^0$ *be a (discrete, negative-time) filtration and* $(X_n)_{n=0}^{-\infty}$ *an* $(\mathscr{F}_n)_{n=-\infty}^0$*-submartinale such that* $\inf_n E[X_n] > -\infty$*, then there is an* $X_{-\infty} \in L^1(P)$ *such that* $X_n \xrightarrow{n\to-\infty} X_{-\infty}$ *a.e.*

[2] We have used the inequality: $\forall a, b \in \mathbb{R}\ (a+b)^+ = (a+b) \vee 0 \leq (a \vee 0) + (b \vee 0) = a^+ + b^+$.

[3] By $\nexists \lim_n$ we mean that the limit does not exist in $[-\infty, +\infty]$.

The proof of this theorem is almost identical to the proof of Theorem 2.33; it can be found in [AD00, Theorem (6.4.4)].

A very important example of martingale is the following: let $X \in L^1(P)$ and $(\mathscr{F}_t)_{t\geq 0}$ a filtration, then the process $(E[X \mid \mathscr{F}_t])_{t\geq 0}$ is an $(\mathscr{F}_t)_{t\geq 0}$-martingale.

The proof of the following theorem can be found in [AD00, Theorems (6.6.2) and (6.6.3)]. To prove it, the concept of *uniform integrability* is used. We will not cover uniform integrability in this document, see [AD00, Sections 6.5 and 6.6] for details.

Theorem 2.35. *Let $(\mathscr{F}_n)_{n=0}^{\infty}$ be a filtration and $\mathscr{F}_\infty = \sigma\left((\mathscr{F}_n)_{n=0}^{\infty}\right)$, as usual. If $X \in L^1(P)$, then $E[X \mid \mathscr{F}_n] \to E[X \mid \mathscr{F}_\infty]$ a.e. and in L^1.*

Let $(\mathscr{F}_n)_{n=-\infty}^{0}$ be a filtration and let $\mathscr{F}_{-\infty} = \cap_{n\in -\mathbb{N}}\mathscr{F}_n$. If $X \in L^1(P)$, then $E[X \mid \mathscr{F}_n] \to E[X \mid \mathscr{F}_{-\infty}]$ a.e. and in L^1 as $n \to -\infty$.

Theorem 2.36. *Let (Ω, Σ, P) be a complete probability space and $(\mathscr{F}_t)_{t\geq 0}$ a right-continous filtration such that the σ-algebras $\{\mathscr{F}_t\}_{t\geq 0}$ contain $\mathscr{N} := \{E \in \Sigma : P(E) = 0\}$ (the Σ-nullsets). Then any $(\mathscr{F}_t)_{t\geq 0}$-martingale admits a right-continous modification.*

Proof. (a) Let $0 \leq a < b$ and $n \in \mathbb{N}$. Define $Q_n = \mathbb{Q} \cap [0, n]$. Let $(F_k)_{k=1}^{\infty}$ be a sequence of finite sets increasing to Q_n. Then $\left(U_{ab}^X(F_k)\right)_k$ is pointwise increasing to $U_{ab}^X(Q_n)$. By Theorem 2.32, and monotone convergence theorem,

$$E\left[U_{ab}^X(Q_n)\right] = \lim_{k\to\infty} E\left[U_{ab}^X(F_k)\right] \leq \frac{E\left[(X_n - a)^+\right]}{b-a} < \infty.$$

Then, the (measurable) set $\left\{U_{ab}^X(Q_n) = \infty\right\}$ has probability 0. Moreover,

$$A_{n,ab} := \left\{\omega \in \Omega : \exists t \in [0, n) : \liminf_{\mathbb{Q}\ni s\downarrow t} X_s(\omega) < a < b < \limsup_{\mathbb{Q}\ni s\downarrow t} X_s(\omega)\right\} \subseteq \left\{\lim_k U_{ab}^X(F_k) = \infty\right\},$$

by completeness of (Ω, Σ, P), $A_{n,ab} \in \mathscr{N}$.

(b) Let

$$A := \left\{\omega \in \Omega : \exists t \geq \infty : \liminf_{\mathbb{Q}\ni s\downarrow t} X_s(\omega) < \limsup_{\mathbb{Q}\ni s\downarrow t} X_s(\omega)\right\} \subseteq \left(\bigcup_{n\in\mathbb{N};\ a,b\in\mathbb{Q};\ a<b} A_{n,ab}\right).$$

By (a) and completeness of the probability space, $A \in \mathscr{N}$; so we can define the process $(Y_t)_{t\geq 0}$ as

$$Y_t := \mathbb{1}_{A^c} \lim_{(t,\infty)\cap\mathbb{Q}\ni s\to t} X_s.$$

Fix $t \geq 0$. Y_t is $\mathscr{F}_t$-measurable (because $\mathscr{F}_t = \mathscr{F}_{t+}$ and $A \in \mathscr{N} \subseteq \mathscr{F}_t$). Take an $\mathbb{N} \ni n > t$. Then, for every $t < s < n$, such that $s \in \mathbb{Q}$, $X_s = E[X_n \mid \mathscr{F}_s]$ because X is a martingale. Making $s \downarrow t$, by definition of Y and Theorem 2.35, $X_s \to Y_t$ a.e. and $E[X_n \mid \mathscr{F}_s] \to E[X_n \mid \mathscr{F}_{t+}] = E[X_n \mid \mathscr{F}_t] = X_t$ a.e. Thus, Y (right-continuous by definition) is a modification of X. Y is also a martingale because it is the modification of a martingale. □

Assume now that the hypotheses of Theorem 2.36 are satisfied, and fix $a > 0$.

Definition 2.37. *We denote by $\mathscr{M}_a$ the equivalence classes of right-continuous $(\mathscr{F}_t)_{0\leq t\leq a}$-martingales $(X_t)_{0\leq t\leq a}$ with $X_a \in L^2(\Omega)$, where we declare two martingales equivalent if they are indistinguishable.*

As explained in Remark 2.2, if two right-continuous martingales are a modification of each other, then they are indistinguishable. So, we can also think of $\mathscr{M}_a$ as the space of right-continuous martingales up to modification: $X = Y$ in $\mathscr{M}_a \iff P(X_t = Y_t) = 1 \quad \forall t \in [0, a]$.

We can naturally define an inner product in $\mathscr{M}_a$: $\langle X, Y\rangle_{\mathscr{M}_a} := \langle X_a, Y_a\rangle_{L^2(\Omega)}$ $\forall X, Y \in \mathscr{M}_a$.

Lemma 2.38. *$\langle\cdot,\cdot\rangle_{\mathscr{M}_a}$ is an inner product. $X \mapsto X_a$ is a linear isometry from $\mathscr{M}_a$ to $L^2(\Omega)$.*

Proof. The property $[\langle v, v\rangle = 0 \iff v = 0]$ is given by Theorem 2.29. The rest follow from the fact that $L^2(\Omega)$ is a pre-Hilbert space. □

From Theorem 2.29, we get that if $X^n \to X$ in $\mathscr{M}_a$, then $\sup_{t\leq a}\left|X_t^n - X_t\right| \to 0$ in $L^2(\Omega)$. In particular, $\forall t \leq a$, $X_t^n \to X_t$ in $L^2(\Omega)$.

Theorem 2.39. *$\mathscr{M}_a$ is a Hilbert space.*

Proof. Let $(X^n)_{n\in\mathbb{N}} \subseteq \mathcal{M}_a$ be a Cauchy sequence. Then $(X^n_a)_{\mathbb{N}} \subseteq L^2(\Omega, \mathcal{F}_a)$ is a Cauchy sequence. Since $L^2(\Omega, \mathcal{F}_a)$ is complete, there exists $f \in L^2(\Omega, \mathcal{F}_a)$ such that $X^n_a \to f$ in $L^2(\Omega, \mathcal{F}_a)$. By Theorem 2.36, there exists a $\tilde{X} \in \mathcal{M}_a$ such that $\tilde{X}_a = f$: it is enough to take $\tilde{X}$ to be a right-continuous modification of the martingale $(E[f \mid \mathcal{F}_t])_{0\le t\le a}$. It is trivial that $X^n \to \tilde{X}$ in $\mathcal{M}_a$. □

Theorem 2.40. *The subspace $\mathcal{M}^C_a$ of continuous $(\mathcal{F}_t)_{0\le t\le a}$-martingales is closed in $\mathcal{M}_a$.*

Proof. Let $(X^n)_{n\in\mathbb{N}} \subseteq \mathcal{M}^C_a$ converge to $X \in \mathcal{M}_a$ in the norm of $\mathcal{M}_a$. WLOG, we can assume that $\sum_n \left\|X^n_a - X_a\right\|_2^2 < \infty$ (otherwise, take an adequate subsequence of $(X^n)_{n\in\mathbb{N}}$). Applying Chebyshev's inequality and Theorem 2.29, for every $\varepsilon > 0$

$$\sum_n P\left(\sup_{0\le t\le a}\left|X^n_t - X_t\right| \ge \varepsilon\right) \le \varepsilon^{-2}\sum_n \left\|\sup_{0\le t\le a}\left|X^n_t - X_t\right|\right\|_2^2 \le \varepsilon^{-2}\sum_n \left\|X^n_a - X_a\right\|_2^2 < \infty.$$

By the first Borel-Cantelli lemma, $P\left(\limsup_n \left\{\sup_{0\le t\le a}\left|X^n_t - X_t\right| \ge \varepsilon\right\}\right) = 0$. Define

$$A := \left\{\sup_{0\le t\le a}\left|X^n_t - X_t\right| \xrightarrow{n\to\infty} 0\right\}^c = \left\{\limsup_{n\to\infty}\sup_{0\le t\le a}\left|X^n_t - X_t\right| > 0\right\} = \bigcup_{k\in\mathbb{N}}\left\{\limsup_{n\to\infty}\sup_{0\le t\le a}\left|X^n_t - X_t\right| \ge 1/k\right\},$$

which has probability 0, implying that, for almost every ω (specifically, for every $\omega \in A^c$), $X^n_t \to X_t$ uniformly in $t \in [0,a]$. Therefore, X is a.e. path-continuous. We can obtain a modification of X that is path-continuous by redefining $X = 0$ in A. □

Before ending this chapter, we will prove a result that will become useful when working with a path-continuous process X and the set of $\mathcal{X}_t$-measurable functions, as it will allow us to approximate them by sums of functions of the form $\psi_1(X_{t_1})\cdots\psi_k(X_{t_k})$:

Theorem 2.41. *Let $(X_t)_{t\ge 0}$ be a stochastic process with continuous integrable paths, $t^* \ge 0$ and $g \in L^1(\mathcal{F}_{t^*})$. Then,*

$$g = \lim_{k\to\infty}\sum_{j=1}^{M_k}\left(\psi^k_{j1}(X_{t_1})\cdots\psi^k_{j1}(X_{t_k})\right),$$

a.e. and in L^1; where the ψ^k_{ji} are Borel-measurable and $\{t_i\}_{i=1}^\infty \subseteq [0,t^*]$.

Proof. Let $\{t_i\}_{i=1}^\infty$ be dense in $[0,t^*]$. For each $n \in \mathbb{N}$, define $\mathcal{H}_n := \sigma(X_i : i \le n)$; the $\mathcal{H}_n$ form a discrete filtration. We will now prove that $\mathcal{H}_\infty = \mathcal{X}_{t^*}$. The inclusion $\mathcal{H}_\infty \subseteq \mathcal{X}_{t^*}$ is trivial. For the other inclusion, we only need to prove that, for each $0 \le t \le t^*$, X_t is $\mathcal{H}_\infty$-measurable. By density of $(t_i)_i$ and continuity of the paths of X, we can express X_t as the a.e. limit of a sequence contained in $\{X_{t_i}\}_i$, all of which are $\mathcal{H}_\infty$-measurable; and we are done.

By Theorem 2.35,

$$g = E\left[g \mid \mathcal{H}_\infty\right] = \lim_n E\left[g \mid \mathcal{H}_n\right]$$

a.e. and in L^1. We can assume that $\left\|g - E\left[g \mid \mathcal{H}_n\right]\right\|_{L^1(\Omega)} \le 1/n$; otherwise, take an adequate subsequence. By Proposition C.9, $E\left[g \mid \mathcal{H}_n\right] = g_n\left(X_{t_1},\dots,X_{t_n}\right)$, for some measurable $g_n : \mathbb{R}^n \to \mathbb{R}$. Consider the product of σ-finite measure spaces $\prod_{i=1}^n\left(\mathbb{R},\mathcal{B}(\mathbb{R}),P_{X_i}\right) =: (\Gamma^n,\Sigma^n,P^n)$, which is equipped with the probability measure induced in $(\mathbb{R}^n,\mathcal{B}(\mathbb{R})^n)$ by the random vector $\underline{X}_n := \left(X_{t_1},\dots,X_{t_n}\right)^T$. $g_n \in L^1(\Gamma^n)$ and it therefore satisfies the hypotheses of Remark C.8. As a consequence, there is a sequence $\left(s^n_\ell\right)_\ell$ of functions of the form $s^n_\ell(x_1,\dots,x_n) = \sum_{j=1}^{M_{ln}}\psi^{ln}_{j1}(x_1)\cdots\psi^{ln}_{jn}(x_n)$, where each ψ^{ln}_{ji} is Borel-measurable (in fact, we proved that they can be characteristic functions), such that: $s^n_\ell \to g_n$ in $L^1(\Gamma^n)$. Select one of these $(s^n_\ell)_\ell$: $\tilde{s}_n$, such that $\left\|g_n - \tilde{s}_n\right\|_{L^1(\Gamma^n)} \le 1/n$. Then

$$\begin{aligned}\left\|g - \tilde{s}_n(\underline{X}_n)\right\|_{L^1(\Omega)} &\le \left\|g - g_n(\underline{X}_n)\right\|_{L^1(\Omega)} + \left\|g_n(\underline{X}_n) - \tilde{s}_n(\underline{X}_n)\right\|_{L^1(\Omega)}\\ &= \left\|g - g_n(\underline{X}_n)\right\|_{L^1(\Omega)} + \left\|g_n - \tilde{s}_n\right\|_{L^1(\Gamma^n)} \le 2/n \to 0.\end{aligned}$$

Which proves that $\tilde{s}_n(\underline{X}_n) \to g$ in $L^1(\Omega)$, and $\tilde{s}_n(\underline{X}_n)$ has the desired form. Convergence in L^1 implies that there is a subsequence of $(\tilde{s}_n)_n$ that converges a.e. to g, and we are done. □

Remark 2.42. *By taking a dense countable subset of $\mathbb{R}_+$ (instead of $[0,t^*]$), we can extend the previous theorem to approximate a.e. and in L^1 a function $g \in L^1(\mathcal{F}_\infty)$ by sums of functions of the form $\psi_1(X_{t_1})\cdots\psi_k(X_{t_k})$.*

CHAPTER 3
Stochastic integrals

3.1. The Itô isometry

To apply the results obtained in section 2.4, we need to work with filtrations satisfying the hypotheses of Theorem 2.36. That is, we need a complete σ-algebra and a right-continuous filtration of sub-σ-algebras containing all *nullsets* (sets of measure 0). Throughout this section, let $(B_t)_{t\geq 0}$ be a (1-dimensional) Brownian motion.

Proposition 3.1. *Let $\mathscr{B} = \overline{\sigma(B_t)_{t\geq 0}}$ be the smallest complete σ-algebra generated by $(B_t)_{t\geq 0}$. Define $\mathscr{N} = \{A \in \mathscr{B} : P(A) = 0\}$ (the set of $\mathscr{B}$-*nullsets*) and $\mathscr{B}'_t = \sigma(\mathscr{N}, \mathscr{B}_t)$. Then $\left(\mathscr{B}'_t\right)_{t\geq 0}$ is a right-continuous filtration of sub-σ-algebras of $\mathscr{B}$ containing all $\mathscr{B}$-nullsets.*

Proposition 3.1 is in [AD00, Section 9.5, Exercise 2], and a proof following this exercise is given in Appendix E.

The construction of the Itô integral presented by [AD00, Section 9.7] is done using the filtration $\left(\mathscr{B}'_t\right)_{t\geq 0}$. However, with the purpose of being able to rigorously extend the Itô integral to more dimensions, which will be explained in subsection 3.1.2, we will work with an abstract filtered space $\left(\Omega, \mathscr{F}, (\mathscr{F}_t)_{t\geq 0}\right)$; such that $\mathscr{F}$ is a complete σ-algebra containing $\mathscr{B}$, and $(\mathscr{F}_t)_{t\geq 0}$ a right-continuous filtration of sub-σ-algebras of $\mathscr{F}$ containing all $\mathscr{F}$-nullsets and such that $(B_t)_{t\geq 0}$ is an $(\mathscr{F}_t)_{t\geq 0}$-martingale (this implies that $\mathscr{B}'_t \subseteq \mathscr{F}_t \ \forall t \geq 0$).

3.1.1. 1-dimensional Itô integral

Definition 3.2 (Simple and integrable processes).

1. *We denote by $\mathscr{S}_a$ the set of* (simple processes) *$f(t,\omega) : [0,a] \times \Omega \to \mathbb{R}$ such that:*

$$f(t,\omega) = \mathbb{1}_{\{0\}}(t) f_0(\omega) + \sum_{i=1}^{n-1} \mathbb{1}_{(t_i,t_{i+1}]}(t) f_i(\omega),$$

   where $n \in \mathbb{N}$, $0 = t_1 < t_2 < \cdots < t_n = a$, f_0 is $\mathscr{F}_0$-measurable and $f_i \in L^2\left(\mathscr{F}_{t_i}\right)$ for $1 \leq i \leq n-1$.
   For such a process f, we define its Itô integral $I[f]$ as follows:

$$\forall 1 \leq t \leq a, \qquad I[f]_t = \int_0^t f(s,\omega)\, dB_s(\omega) := \sum_{i=1}^{n-1} f_i(\omega) \left(B_{t_{i+1}\wedge t}(\omega) - B_{t_i \wedge t}(\omega)\right).$$

2. *We denote by $\mathscr{A}_a$ the set of $(\mathscr{F}_t)_{0\leq t\leq a}$-adapted processes $g(t,\omega) : [0,a] \times \Omega \to \mathbb{R}$ such that $g \in L^2([0,a] \times \Omega)$ (i.e. $E\left[\int_0^a g^2(t,\omega)\, dt\right] < \infty$).*

To ease the notation, we will usually write $h(s)$ to denote the random variable $\omega \mapsto h(s,\omega)$. For example,

$$\int_0^t f(s)\, dB_s \quad \text{instead of} \quad \omega \mapsto \int_0^t f(s,\omega)\, dB_s(\omega),$$

or

$$\int_0^t f(s) g(s)\, ds \quad \text{instead of} \quad \omega \mapsto \int_0^t f(s,\omega) g(s,\omega)\, ds.$$

Lemma 3.3. *Let $f, g \in \mathscr{S}_a$. Then:*

(a) $I[f], I[g] \in \mathscr{M}_a^C$.

(b) $I : \mathscr{S}_a \to \mathscr{M}_a^C$ *is a linear mapping.*

(c) $0 \leq t \leq a \implies I[f]_t = I\left[f \mathbb{1}_{[0,t]}\right]_a$.

(d) *The Itô isometry for simple processes:*

$$0 \le t \le a \Longrightarrow \langle I[f]_t, I[g]_t \rangle_{L^2(\Omega)} = \langle f, g \rangle_{L^2([0,t]\times\Omega)},$$
$$\left(i.e.\ E\left[\int_0^t f(s)dB_s \int_0^t g(s)dB_s \right] = E\left[\int_0^t f(s)g(s)ds \right] \right).$$

In particular, $\langle I[f], I[g] \rangle_{\mathcal{M}_a} = \langle f, g \rangle_{L^2([0,a]\times\Omega)}$.

Proof. Only the last point is not immediate. Let $f, (f_i)_{i=1}^n$ be as in Definition 3.2: $f_i \in L^2(\mathcal{F}_{t_i})$ is the value of f for $t \in (t_i, t_{i+1}]$. WLOG, we can assume that g is defined in the same t-intervals as f with value $g_i \in L^2(\mathcal{F}_{t_i})$ for $t \in (t_i, t_{t+1}]$. Then

$$\begin{aligned}
E\left[f_i g_i (B_{t_{i+1}} - B_{t_i})^2 \right] &= E\left[E\left[f_i g_i (B_{t_{i+1}} - B_{t_i})^2 \,\middle|\, \mathcal{F}_{t_i} \right] \right] \\
&= E\left[f_i g_i E\left[(B_{t_{i+1}} - B_{t_i})^2 \,\middle|\, \mathcal{F}_{t_i} \right] \right] = E\left[f_i g_i (t_{i+1} - t_i) \right], \\
\text{for } j < i, \quad E\left[f_i g_j \left(B_{t_{j+1}} - B_{t_j}\right)(B_{t_{i+1}} - B_{t_i}) \right] &= E\left[E\left[f_i g_j \left(B_{t_{j+1}} - B_{t_j}\right)(B_{t_{i+1}} - B_{t_i}) \,\middle|\, \mathcal{F}_{t_i} \right] \right] \\
&= E\left[f_i g_j \left(B_{t_{j+1}} - B_{t_j}\right) \underbrace{E\left[(B_{t_{i+1}} - B_{t_i}) \mid \mathcal{F}_{t_i} \right]}_{0} \right] = 0.
\end{aligned}$$

And the Itô isometry follows. □

Theorem 3.4 (The Itô isometry—definition of the Itô integral for processes in $\mathcal{A}_a$).

(a) *The mapping* $I : f \mapsto I[f] = \int_0^a f(s,\omega)dB_s(\omega)$ *is a linear isometry between the Hilbert spaces* $\mathcal{S}_a \subseteq L^2([0,a]\times\Omega)$ *and* $\mathcal{M}_a^C$. *Therefore, and by the closedness of* $\mathcal{M}_a^C$ *in* $\mathcal{M}_a$, *the mapping can be uniquely extended to the closure of* $\mathcal{S}_a$ *in* $L^2([0,a]\times\Omega)$: $I : \overline{\mathcal{S}_a} \to \mathcal{M}_a^C$.

(b) $\mathcal{A}_a$ *is contained in the closure of* $\mathcal{S}_a$ *in* $L^2([0,a]\times\Omega)$.

(c) $\forall f \in \mathcal{A}_a \quad \forall 0 \le t \le a \quad I[f]_t = I\left[f \mathbb{1}_{[0,t]}\right]_a$.

(d) *The Itô isometry:* $\forall f, g \in \mathcal{A}_a \quad \langle I[f], I[g] \rangle_{\mathcal{M}_a} = \langle f, g \rangle_{L^2([0,a]\times\Omega)}$.

Proof. (a) is a direct consequence of Lemma 3.3 and (c) and (d) extend from $\mathcal{S}_a$ to its closure. Let us prove (b).

Let $f \in L^\infty(\mathbb{R})$ with compact support. In particular, $f \in L^2(\mathbb{R})$. Let τ_h be the operator $\tau_h f = f(\cdot + h)$. Then $\tau_h f \to f$ in $L^2(\mathbb{R})$, as $h \to 0$.

Let $g(t,\omega) \in L^\infty([0,a]\times\Omega)$ be $(\mathcal{F}_t)_t$-adapted. For each $n \in \mathbb{N}$, we take

$$\mathbb{R} = \bigsqcup_{j\in\mathbb{Z}} \underbrace{\left(\frac{j}{2^n}, \frac{j+1}{2^n} \right]}_{I_j^n},$$

and define $\alpha_n(t) = \sum_j \frac{j}{2^n} \mathbb{1}_{I_j^n}(t)$,

$$g_{n,s}(t,\omega) = g(s + \alpha_n(t-s), \omega) = \sum_j g\left(s + j2^{-n}, \omega\right) \mathbb{1}_{s+I_j^n}(t).$$

$g_{n,s} \in \mathcal{S}_a$, because even if the sum looks like it has infinite terms, g has compact support, so only a finite number of terms are non-zero.

$$\forall \omega \in \Omega \quad \lim_{h\to 0} \int_{\mathbb{R}} |g(s+h,\omega) - g(s,\omega)|^2 \, ds = 0$$

$$\Longrightarrow \forall \omega \in \Omega, t \in \mathbb{R} \quad \lim_{h\to 0} \int_{\mathbb{R}} |g(s+t+h,\omega) - g(s+t,\omega)|^2 \, ds = 0$$

(now change h by $\alpha_n(t) - t \xrightarrow{n\to\infty} 0$)

$$\Longrightarrow \forall \omega \in \Omega, t \in \mathbb{R} \quad \lim_{n\to \infty} \int_{\mathbb{R}} |g(s+\alpha_n(t),\omega) - g(s+t,\omega)|^2 \, ds = 0.$$

Let $c < d$ be any pair of real numbers; for every $t \in [c,d]$ and $\omega \in \Omega$

$$\forall n \in \mathbb{N} \quad \int_{\mathbb{R}} |g(s+\alpha_n(t),\omega) - g(s+t,\omega)|^2 \, ds \le 2\|g\|_\infty (a+2) \in L^1([c,d]\times\Omega).$$

Thus, we can apply the dominated convergence theorem (along with Tonelli's theorem) to get

$$\begin{aligned}
&\lim_{n\to\infty}\int_{\mathbb{R}} E\left[\int_c^d \left|g(s+\alpha_n(t),\omega)-g(s+t,\omega)\right|^2\,dt\right]\,ds = \\
&= \lim_{n\to\infty} E\left[\int_c^d \int_{\mathbb{R}} \left|g(s+\alpha_n(t),\omega)-g(s+t,\omega)\right|^2\,ds\,dt\right] = \\
&= E\left[\int_c^d \lim_{n\to\infty}\left(\int_{\mathbb{R}} \left|g(s+\alpha_n(t),\omega)-g(s+t,\omega)\right|^2\,ds\right)\,dt\right] = E\left[\int_c^d 0\,dt\right] 0.
\end{aligned}$$

That is, $E\left[\int_c^d \left|g(s+\alpha_n(t),\omega)-g(s+t,\omega)\right|^2\,dt\right] \to 0$ in $L^1(\mathbb{R})$ (on the variable s). Therefore, there exists a subsequence $(n_k)_k$ such that $E\left[\int_c^d \left|g(s+\alpha_{n_k}(t),\omega)-g(s+t,\omega)\right|^2\,dt\right] \to 0$ a.e. What we have proven is true for every $c<d$. We can make a change of variables $t\mapsto t-s$ and choose suitable c,d to get that $\left\|g_{n_k,s}-g\right\|^2_{L^2([0,a]\times\Omega)} = E\left[\int_0^a \left|g(s+\alpha_{n_k}(t-s),\omega)-g(s,\omega)\right|^2\,dt\right] \to 0$ a.e. (i.e. g is in the closure of $\mathscr{S}_a$.)

Once proven the result for bounded $g\in\mathscr{A}_a$, let $h\in\mathscr{A}_a$ and define $g_n = h\mathbb{1}_{|h|\le n}\in\mathscr{A}_a$, which are bounded (so they are contained in the closure of $\mathscr{S}_a$) and $g_n\to h$ in $L^2([0,a]\times\Omega)$ (by the bounded convergence theorem, for example). Therefore, h is in the closure of $\mathscr{S}_a$. □

We have defined the stochastic integrals as limits of elements in $\mathscr{M}_a^C$. Even though the Itô integral of a process $f\in\mathscr{A}_a$ is an equivalence class in $\mathscr{M}_a^C$, we will always assume we are working with a continuous representative of the class.

Now that we have defined the Itô integral for (some) $[0,a]$-indexed processes, we can extend the definition to processes indexed by $\mathbb{R}_+$.

Definition 3.5. *Let $\mathscr{A}$ be the set of measurable, $(\mathscr{F}_t)_{t\ge 0}$-adapted processes $g(t,\omega):\mathbb{R}_+\times\Omega\to\mathbb{R}$ such that $g|_{[0,a]\times\Omega}\in L^2(\mathbb{R}_+,\Omega)$ for every $a>0$.*

*We will call $\mathscr{A}$ the set of $(\mathscr{F}_t)_{t\ge0}$-*Itô-integrable processes.

Theorem 3.6. *Let $f,g\in\mathscr{A}$. Then:*

(a) There exists a continuous martingale $J[f]$ such that $J[f]_t = \int_0^t f(s)dB_s$ a.e. for all $t\ge0$. That is, $J[f]$ is equal to the stochastic integral up to modification of processes.

(b) J is linear: $\forall\alpha,\beta\in\mathbb{R}\ \forall t\ge0\quad \int_0^t(\alpha f(s)+\beta g(s))\,dB_s = \alpha\int_0^t f(s)dB_s + \beta\int_0^t g(s)dB_s$ a.e.

(c) $\forall 0\le t\le a<\infty\quad \int_0^t f(s)dB_s = \int_0^a f(s)\mathbb{1}_{[0,t]}(s)dB_s$ a.e.

(d) $\forall t\ge0\quad E\left[\left(\int_0^t f(s)dB_s\right)\left(\int_0^t g(s)dB_s\right)\right] = E\left[\int_0^t f(s)g(s)ds\right]$.

Proof. For the first property, let $n\in\mathbb{N}$, $t\in[0,n]$. Define $Y_t^n := \int_0^t f(s)\mathbb{1}_{[0,n]}(s)dB_s\in\mathscr{M}_n^C$. By the uniqueness of the stochastic integral (in $\mathscr{M}_n^C$—meaning up to indistinguishability or equivalently, by right-continuity, up to modification) given by the isometry from Theorem 3.4, the martingales $(Y_t^n)_{0\le t\le n}$ and $(Y_t^{n+1})_{0\le t\le n}$ are indistinguishable. Therefore, we can define $J[f]_t = Y_t^n$ for all $0\le t\le n$, which is a right-continuous $(\mathscr{F}_t)_{t\ge0}$-martingale equal to the stochastic integral up to indistinguishability. The other properties extend trivially from the previous theorems for processes in $\mathscr{S}_a$ $(a<\infty)$. □

3.1.2. Multi-dimensional Itô integral

Let $B=(B_1,\ldots,B_n)$ be an n-dimensional Brownian motion. In a similar way as we did in Proposition 3.1, we will take $\mathscr{F} := \overline{\sigma(B_k(s):1\le k\le n,\ s\ge0)}$ to be the smallest complete σ-algebra generated by the B_k. Define $\mathscr{N}=\{A\in\mathscr{F}:P(A)=0\}$ and let $\mathscr{F}_t=\sigma(\mathscr{N},\mathscr{B}_t)$. Then $(\mathscr{F}_t)_{t\ge0}$ is a right-continuous filtration of sub-σ-algebras of $\mathscr{F}$ containing all $\mathscr{F}$-nullsets. This can be proven in a similar way as Proposition 3.1.

Furtheremore, for each $1\le k\le n$ B_k is an $(\mathscr{F}_t)_{t\ge0}$-martingale, because $B_k(s)-B_k(t)$ is independent of $\mathscr{F}_t$ for every $0\le t\le s$:

$$E[B_s\mid\mathscr{F}_t] = E[B_s-B_t\mid\mathscr{F}_t]+E[B_t\mid\mathscr{F}_t] = E[B_s-B_t]+B_t = B_t.$$

Definition 3.7 (Multi-dimensional Itô integral). *Let $\mathscr{A}$ be the set of $(\mathscr{F}_t)_{t\ge0}$-Itô-integrable processes, and $B=(B_1,\ldots,B_n)$ an n-dimensional Brownian motion. We will call $\mathscr{A}^{m\times n}$ the set of $m\times n$ matrices $v(t,\omega)=[v_{ij}(t,\omega)]:[0,\infty)\times\Omega\to\mathbb{R}^{m\times n}$ where each coordinate v_{ij} is in $\mathscr{A}$.*

For each $v\in\mathscr{A}^{m\times n}$:

$$\int_0^t v\,dB := \int_0^t \begin{bmatrix} v_{11} & \cdots & v_{1n}\\ \vdots & \ddots & \vdots\\ v_{m1} & \cdots & v_{mn}\end{bmatrix}\begin{bmatrix} dB_1\\ \vdots\\ dB_n\end{bmatrix}.$$

That is, for each $t \geq 0$, the k-th coordinate of $\left(\int_0^t v dB\right) : \Omega \to \mathbb{R}^m$ is given by the following sum of 1-dimensional Itô integrals:

$$\pi_k\left(\int_0^t v dB\right) = \sum_{j=1}^{n} \int_0^t v_{kj} dB_j.$$

Remark 3.8. *The sum of $(\mathscr{F}_t)_{t\geq 0}$-martingales is an $(\mathscr{F}_t)_{t\geq 0}$-martingale. Therefore, the Itô integral of a process in $\mathscr{A}^{m\times n}$ is an $(\mathscr{F}_t)_{t\geq 0}$-martingale. By the path-continuity of the 1-dimensional Itô integrals, the $m \times n$ Itô integral is path-continuous (up to modifications).*

3.2. The Itô formula

Similar to the definition of $(\mathscr{F}_t)_{t\geq 0}$-Itô-integrable processes, we define:

Definition 3.9. *Let $\mathscr{A}_{dm}$ be the set of measurable $(\mathscr{F}_t)_{t\geq 0}$-adapted processes $g(t,\omega) : \mathbb{R}_+ \times \Omega \to \mathbb{R}$ that satisfy $g|_{[0,a]\times\Omega} \in L^1(\mathbb{R}_+ \times \Omega)$ for every $a < 0$. We call $\mathscr{A}_{dm}$ the set of $(\mathscr{F}_t)_{t\geq 0}$-$dm$-*integrable processes.[1]

Again, we can define $\mathscr{A}_{dm}^{n\times m}$ as the set of $n \times m$ matrices of $(\mathscr{F}_t)_{t\geq 0}$-$dm$-integrable processes.

Definition 3.10. *Let B be an m-dimensional Brownian motion on $(\Omega, \mathscr{F}, P)$. If $u(t,\omega) \in \mathscr{A}_{dm}^{1\times n}$ and $v(t,\omega) \in \mathscr{A}^{n\times m}$, we call the stochastic process*

$$X(t,\omega) := X_0(\omega) + \int_0^t u(s,\omega) ds + \int_0^t v(s,\omega) dB_s(\omega) \tag{3.1}$$

an Itô process, *where X_0 is a random variable.* (3.1) *meaning that the k-th coordinate of $X(t)$ has the form*

$$X_k(t) = (X_0)_k + \int_0^t u_k(s) ds + \sum_{j=1}^{n} \int_0^t v_{kj}(s) dB_j(s). \tag{3.2}$$

A more concise way of writing (3.1) *is $dX = u dt + v dB$.*

Theorem 3.11 (The multi-dimensional Itô formula). *Let $dX = u dt + v dB$ be an n-dimensional Itô process as in Definition* 3.10. *Let $g(t,x) = \left(g_1(t,x), \ldots, g_p(t,x)\right) \in C^2(\mathbb{R}_+ \times \mathbb{R}^n, \mathbb{R}^p)$*[2]*. Then the process*

$$Y(t,\omega) := g(t, X(t))$$

is an Itô process whose k-th component is given by

$$dY_k(t) = \frac{\partial g_k}{\partial t}(t, X(t))dt + \sum_j \frac{\partial g_k}{\partial x_j}(t, X(t)) dX_j(t) + \frac{1}{2}\sum_{i,j} \frac{\partial^2 g_k}{\partial x_i \partial x_j}(t, X(t)) dX_i(t)\, dX_j(t); \tag{3.3}$$

where the expression $dX_i dX_j$ has the meaning

$$dX_i dX_j = \left(u_i dt + \sum_{\ell=1}^{n} v_{i\ell} dB_\ell\right)\left(u^j dt + \sum_{k=1}^{n} v_{jk} dB_k\right),$$

and the formal multiplication between dt and dB_i objects is given by

$$dB_i \cdot dB_j = \delta_{ij} dt, \quad dB_i \cdot dt = dt \cdot dB_i = dt \cdot dt = 0, \qquad \forall 1 \leq i,j \leq m. \tag{3.4}$$

In order to concisely express (3.3), we will make use of the usual differential operators

$$D_t h = \frac{\partial h}{\partial t}, \qquad D_x h = \begin{bmatrix} \frac{\partial h}{\partial x_1} & \ldots & \frac{\partial h}{\partial x_n} \end{bmatrix}^T, \qquad D_x^2 h = \left[\frac{\partial^2 h}{\partial x_i \partial x_j}\right]_{i,j}.$$

We will also denote:

$$dX = \begin{bmatrix} dX_1 & \cdots & dX_n \end{bmatrix}^T, \quad (dX)\cdot(dX)^T = \begin{bmatrix} dX_1 dX_1 & \ldots & dX_1 dX_n \\ \vdots & \ddots & \vdots \\ dX_n dX_1 & \ldots & dX_n dX_n \end{bmatrix}.$$

[1] The dm stands for the Lebesgue measure in $\mathbb{R}$.

[2] For $D \subseteq \mathbb{R}^q$, we denote by $C^k(D, \mathbb{R}^p)$ the set of k-times continuously differentiable functions from D to $\mathbb{R}^p$. Whenever $\mathbb{R}^p$ is ommited: $C^k(D)$, we assume $p = 1$.

$\langle\cdot,\cdot\rangle$ will be used to denote the formal inner product between vectors and between matrices, which is essentially the same:

$$\langle u,v\rangle = u^T v = \sum_{i=1}^n a_i b_i, \quad u,v \text{ being } n\text{-vectors},$$
$$\langle A,B\rangle = \mathrm{Tr}(A^T B) = \sum_{i,j=1}^n A_{ij}B_{ij}, \quad A,B \text{ being } n\times n \text{ matrices}.$$

With this notation, we can rewrite (3.3) as

$$dY_k = D_t g_k dt + \langle D_x g_k, dX\rangle + \frac{1}{2}\langle D_x^2 g_k, (dX)\cdot(dX)^T\rangle, \tag{3.5}$$

where all derivatives of g_k are evaluated at $(t, X(t))$. We can go a bit further than (3.5) and unfold dX:

$$\begin{aligned} dX &= u dt + v dB \\ \implies dX dX^T &= (u dt + v dB)(u dt + v dB)^T \\ &= v dB (v dB)^T && \text{(using (3.4))} \\ &= v dB (dB)^T v^T \\ &= v v^T dt && (\text{as } dB_i dB_j = \delta_{ij} dt). \end{aligned}$$

So,

$$dY_k = D_t g_k dt + \langle D_x g_k, u dt + v dB\rangle + \frac{1}{2}\langle D_x^2 g_k, v v^T\rangle dt, \tag{3.6}$$

which clearly has the form of an Itô process.

Sketch of proof of Theorem 3.11. It is enough to show the case of $p=1$, the others are identical but repeating the calculations for each coordinate of g. We will assume that g and its first and second-order derivatives are bounded; see [Øks13, Exercise 4.9] on how to extend the proof for a general $g \in C^2(\mathbb{R}_+ \times \mathbb{R}^n)$. We will assume that u,v are bounded and simple in the t variable. The general argument follows by approximating u,v by simple processes; the Itô isometry and the definition of the Lebesgue integral along with the continuity of g and its derivatives will allow us to pass to the limit.

Let $0 = t_1 < t_2 < \cdots < t_J = t$ be the common grid on whose intervals u,v are defined (WLOG, the grid is the same for both u and v). As the grid becomes finer, the simple processes approximate the *still bounded* general u,v.

$$u(s,\omega) = \mathbb{1}_{\{0\}}(s)u_0(\omega) + \sum_{j=1}^{J-1}\mathbb{1}_{(t_j,t_{j+1}]}(s)u^j(\omega); \qquad v(s,\omega) = \mathbb{1}_{\{0\}}(s)v_0(\omega) + \sum_{j=1}^{J-1}\mathbb{1}_{(t_j,t_{j+1}]}(s)v^j(\omega).$$

Remember that for each $j \geq 1$, u^j and v^j are $\mathscr{F}_{t_j}$-measurable. We will use Δ^j to denote the increment of a process between times t_j and t_{j+1}:

$$\Delta^j t = t_{j+1} - t_j, \quad \Delta^j B_k = B_k(t_{j+1}) - B_k(t_j), \quad \Delta^j g = g\left(t_{j+1}, X_{t_{j+1}}\right) - g\left(t_j, X\left(t_j\right)\right), \text{ etc.}$$

Using the Taylor expansion of $g(t, X(t))$ around t_j:

$$\begin{aligned} g(t,X(t)) &= g(0,X(0)) + \sum_{j=1}^{J-1}\Delta^j g \\ &= g(0,X(0)) + \sum_{j=1}^{J-1} g_t\left(t_j, X\left(t_j\right)\right)\Delta^j t + \sum_{i=1}^n\sum_{j=1}^{J-1} g_{x_i}\left(t_j, X\left(t_j\right)\right)\Delta^j X_i + \frac{1}{2}\sum_{j=1}^{J-1} g_{tt}\left(t_j, X\left(t_j\right)\right)\left(\Delta^j t\right)^2 \\ &\quad + \frac{1}{2}\sum_{i=1}^n\sum_{j=1}^{J-1} g_{tx_i}\left(t_j, X\left(t_j\right)\right)\left(\Delta^j X_i\right)\left(\Delta^j t\right) + \underbrace{\frac{1}{2}\sum_{i,k=1}^n\sum_{j=1}^{J-1} g_{x_i x_k}\left(t_j, X\left(t_j\right)\right)\left(\Delta^j X_k\right)\left(\Delta^j X_i\right)}_{\text{(I)}} + \sum_{j=1}^{J-1} R_j, \end{aligned}$$

with $R_j = o\left((\Delta^j t)^2 + \sum_{i=1}^n \left(\Delta^j X_i\right)^2\right)$. From now on, g and its derivatives are evaluated at $\left(t_j, X\left(t_j\right)\right)$ when they are inside a sum over j.

We have that

$$\sum_{j=1}^{J-1} g_t \Delta^j t \xrightarrow{J\to\infty} \int_0^t g_t(s, X(s))\, ds, \qquad \sum_{j=1}^{J-1} g_{tt}\left(\Delta^j t\right)^2 \xrightarrow{J\to\infty} 0,$$

$$\sum_{i=1}^{n}\sum_{j=1}^{J-1} g_{x_i}\Delta^j X_i \xrightarrow{J\to\infty} \sum_{i=1}^{n}\int_0^t g_{x_i}(s, X(s))\, dX_i, \qquad \sum_{j=1}^{J-1} R_j \xrightarrow{J\to\infty} 0.$$

For the rest of the terms, we can expand $\Delta^j X_i = (u^j)_i\left(\Delta^j t\right) + \sum_{\ell=1}^m (v^j)_{i\ell}\left(\Delta^j B_\ell\right)$, so that all can be expressed as a sum of processes of the form

$$\sum_{j=1}^{J-1} f_j\left(\Delta^j t\right)^2, \quad \sum_{j=1}^{J-1} f_j\left(\Delta^j t\right)\left(\Delta^j B_\ell\right), \quad \text{or} \quad \sum_{j=1}^{J-1} f_j\left(\Delta^j B_{l_1}\right)\left(\Delta^j B_{l_2}\right),$$

where the f_j are $\mathscr{F}_{t_j}$-measurable and bounded by a constant $C > 0$ that does not depend on j or J. When $J\to\infty$, only the terms of the third form, with $l_1 = l_2$ will be non-zero:

$$E\left[\left(\sum_{j=1}^{J-1} f_j\left(\Delta^j t\right)^2\right)^2\right] \le C^2\sum_{j_1=1}^{J-1}\sum_{j_2=1}^{J-1}\left(\Delta^{j_1} t\right)^2\left(\Delta^{j_2} t\right)^2 = C^2\left(\sum_{j=1}^{J-1}\left(\Delta^j t\right)^2\right)^2 \xrightarrow{J\to\infty} 0;$$

$$\begin{aligned} E\left[\left(\sum_{j=1}^{J-1} f_j\left(\Delta^j t\right)\left(\Delta^j B_\ell\right)\right)^2\right] &\le C^2\sum_{j_1=1}^{J-1}\sum_{j_2=1}^{J-1}\left(\Delta^{j_1} t\right)\left(\Delta^{j_2} t\right)\underbrace{E\left[\left(\Delta^{j_1} B_\ell\right)\left(\Delta^{j_2} B_\ell\right)\right]}_{\delta_{j_1,j_2}\cdot(\Delta^{j_1} t)} \\ &= C^2\sum_{j=1}^{J-1}\left(\Delta^j t\right)^3 \xrightarrow{J\to\infty} 0; \end{aligned}$$

if $l_1 \ne l_2$, then $E\left[\left(\Delta^{j_1} B_{l_1}\right)\left(\Delta^{j_2} B_{l_2}\right)\right] = 0$ by independence of the coordinates of multidimensional Brownian motion. This justifies the formal operations: $dt\cdot dt = dt\cdot dB_\ell = dB_{l_1}\cdot dB_{l_2} = 0$, (when $l_1 \ne l_2$).

The only thing left to consider is the case when $l_1 = l_2$, and the sum has the same limit as $\sum_{j=1}^{J-1} f_j\left(\Delta^j t\right)$ when $J\to\infty$ in that case; i.e., the limit is $\int_0^t f\, ds$ (f being the function that $\sum_j f_j \mathbb{1}_{(t_j,t_{j+1}]}$ tends to as $J\to\infty$).

$$(\text{II}) := E\left[\left(\sum_{j=1}^{J-1} f_j\left(\Delta^j B_\ell\right)^2 - \sum_{j=1}^{J-1} f_j\left(\Delta^j t\right)\right)^2\right] = \sum_{j_1=1}^{J-1}\sum_{j_2=1}^{J-1} E\Big[\underbrace{f_{j_1} f_{j_2}\left(\left(\Delta^{j_1} B_\ell\right)^2 - \left(\Delta^{j_1} t\right)\right)}_{(\text{II.1})}\underbrace{\left(\left(\Delta^{j_2} B_\ell\right)^2 - \left(\Delta^{j_2} t\right)\right)}_{(\text{II.2})}\Big].$$

If $j_1 < j_2$, then (II.1) and (II.2) are independent, and the expectation of their product is zero, because $E[(\text{II.2})] = 0$. Therefore,

$$\begin{aligned} \text{II} &= \sum_{j=1}^{J-1} E\left[f_j^2\left(\left(\Delta^j B_\ell\right)^2 - \left(\Delta^j t\right)\right)^2\right] \\ &\le C^2\left(\sum_{j=1}^{J-1}\underbrace{E\left[\left(\Delta^j B_\ell\right)^4\right]}_{3(\Delta^j t)^2} + \sum_{j=1}^{J-1}\left(\Delta^j t\right)^2 - 2\sum_{j=1}^{J-1}\underbrace{E\left[\left(\Delta^j B_\ell\right)^2\right]}_{(\Delta^j t)}\left(\Delta^j t\right)\right) \xrightarrow{J\to\infty} 0; \end{aligned}$$

which is why we write $dB_\ell\cdot dB_\ell = dt$. Putting everything together, (I) has the same limit as:

$$\sum_{i,k=1}^{n}\sum_{j=1}^{J-1} g_{x_i x_k}\sum_{\ell=1}^{m}\left(v^j\right)_{i\ell}\left(v^j\right)_{kl}\left(\Delta^j B_\ell\right)^2,$$

and it is

$$\sum_{i,k=1}^{n}\sum_{j=1}^{J-1} g_{x_i x_k}\left(\left(v^j\right)\left(v^j\right)^T\right)_{ik}\left(\Delta^j t\right) = \sum_{j=1}^{J-1}\left\langle D_x^2 g, \left(v^j\right)\left(v^j\right)^T\right\rangle\left(\Delta^j t\right) = \int_0^t \langle D_x^2 g, vv^T\rangle\, ds. \qquad \square$$

3.3. An existence and uniqueness theorem for stochastic differential equations

Now that we have developed a theory for stochastic integration and Itô processes (consisting of the sum of a stochastic and a deterministic integral), it is only natural to try to solve the matching differential equation problems. That is, for some b and σ, is there an Itô process that satisfies the stochastic differential equation (SDE)

$$dX(t) = b(t, X(t))dt + \sigma(t, X(t))dB(t)?$$

To answer this question, we will use the following result from [Øks13, Theorem 5.2.1].

Let $|\cdot|$ be the Euclidean norm in $\mathbb{R}^n$ or $\mathbb{R}^{n\times m}$, given by $|b|^2 = \sum_i |b_i|^2$ or $|\sigma|^2 = \sum_{ij} \left|\sigma_{ij}\right|^2$, respectively.

Theorem 3.12. *Let $T > 0$ and let $b(t,x) : [0,T]\times\mathbb{R}^n \to \mathbb{R}^n$ and $\sigma(t,x) : [0,T]\times\mathbb{R}^n \to \mathbb{R}^{n\times m}$ be two measurable functions such that*

$$\exists C > 0 : \ \forall (t,x) \in [0,T]\times\mathbb{R}^n \qquad |b(t,x)| + |\sigma(t,x)| \le C(1+|x|), \qquad \textit{and} \tag{3.7}$$

$$\exists D > 0 : \ \forall t \in [0,T], x, y \in \mathbb{R}^n \qquad \left|b(t,x)-b(t,y)\right| + \left|\sigma(t,x)-\sigma(t,y)\right| \le D\left|x-y\right|. \tag{3.8}$$

Let $Z \in L^2(P)$ be independent of $\mathcal{F}_\infty$[3]*. Then the stochastic differential equation*

$$\begin{cases} dX_t = b(t,X_t)dt + \sigma(t,X_t)dB_t & \text{(3.9a)} \\ X_0 = Z & \text{(3.9b)} \end{cases}$$

has a unique (up to indistinguishability) path-continuous solution X satisfying that $(X_t)_{t\ge 0}$ is adapted to the filtration $\left(\mathcal{F}_t^Z\right)_t$ given by $\mathcal{F}_t^Z = \sigma(Z, \mathcal{F}_t)$.

Proof. *(Uniqueness)* Let X and $\hat{X}$ be solutions of (3.9a) with initial values Z and $\hat{Z}$, respectively. We will make an estimation of the L^2 distance between X and $\hat{X}$.

Let $a(s,\omega) := b(s,X_s) - b(s,\hat{X}_s)$ and $\gamma(s,\omega) := \sigma(s,X_s) - \sigma(s,\hat{X}_s)$. The Itô isometry can be extended to the multidimensional case:

$$\left\|\int_0^t \gamma\, dB\right\|_{L^2(\Omega)}^2 = E\left[\left|\int_0^t \gamma\, dB\right|^2\right] = E\left[\int_0^t |\gamma|^2\, ds\right] = \|\gamma\|^2_{L^2([0,t]\times\Omega)}; \tag{3.10}$$

this follows from the one dimensional Itô isometry and the fact that if $l_1 \neq l_2$, then $\int_0^t \gamma_{il_1}\, dB_{l_1}$ and $\int_0^t \gamma_{il_2}\, dB_{l_2}$ are independent with mean 0. Then,

$$\begin{aligned}
\|X_t - \hat{X}_t\|_2^2 &= \left\|Z - \hat{Z} + \int_0^t a\, ds + \int_0^t \gamma\, dB\right\|_2^2 \\
&\le \left(\|Z-\hat{Z}\|_2 + \left\|\int_0^t a\, ds\right\|_2 + \left\|\int_0^t \gamma\, dB\right\|_2\right)^2 \\
&\overset{4}{\le} 3\left(\|Z-\hat{Z}\|_2^2 + E\left[\left|\int_0^t a\, ds\right|^2\right] + \left\|\int_0^t \gamma\, dB\right\|_2^2\right) \\
&\le 3\left(\|Z-\hat{Z}\|_2^2 + E\left[\int_0^t \left(t|a|^2 + |\gamma|^2\right) ds\right]\right) \qquad \text{(Jensen's ineq. \& Itô isometry)} \\
&\le \underbrace{3\,\|Z-\hat{Z}\|_2^2}_{F} + \underbrace{3(1+T)D^2}_{A} \int_0^t \|X_s - \hat{X}_s\|_2^2\, ds \qquad \text{(by (3.8) \& Tonelli's thm.)}.
\end{aligned} \tag{3.11}$$

That is, $v(t) := \|X_t - \hat{X}_t\|_2^2 \le F + A\int_0^t v(s)\, ds$. By Grönwall's inequality, $v(t) \le Fe^{At}$. This proves that if $Z = \hat{Z}$, then $X_t = \hat{X}_t$ a.e. for all $t \ge 0$; i.e. they are modifications of each other, which is equivalent to indistinguishable as they are path-continuous. This proves the uniqueness of a solution.

(Existence) The proof of existence is similar to the ordinary differential equations solution existence. Define $Y_t^{(0)} = X_0$ and $Y_t^{(k)}(\omega)$ inductively for $k \ge 1$:

$$Y_t^{(k+1)}(\omega) = X_0 + \int_0^t b\left(s, Y_s^{(k)}\right) ds + \int_0^t \sigma\left(s, Y_s^{(k)}\right) dB_s. \tag{3.12}$$

[3]This means that Z is independent of B_t for every $t \ge 0$.
[4]$\forall a,b,c \in \mathbb{R}$, $(a+b+c)^2 \le 3(a^2+b^2+c^2)$.

Similar estimations as done in (3.11) yield

$$\left\| Y_t^{(k+1)} - Y_t^{(k)} \right\|_{L^2(\Omega)}^2 \leq 3(1+T)D^3 \int_0^t \left\| Y_t^{(k)} - Y_t^{(k-1)} \right\|_2^2 ds, \text{ and}$$
$$\left\| Y_t^{(1)} - Y_t^{(0)} \right\|_{L^2(\Omega)}^2 \leq 2C^2(1+T)\left(1+E\left[|X_0|^2\right]\right) t = A_1 t,$$

where the constant A_1 depends on T, C and $E\left[|X_0|^2\right]$. By induction on k,

$$\left\| Y_t^{(k+1)} - Y_t^{(k)} \right\|_2^2 \leq \frac{A_2^{k+1} t^{k+1}}{(k+1)!}, \tag{3.13}$$

where A_2 is a constant that depends on T, D and A_1.

If $m > n \geq 1$, then

$$\left\| Y_t^{(m)} - Y_t^{(n)} \right\|_{L^2([0,T]\times\Omega)} \leq \sum_{k=n}^{m-1} \left\| Y_t^{(k+1)} - Y_t^{(k)} \right\|_{L^2([0,T]\times\Omega)} \tag{3.14}$$
$$= \sum_{k=n}^{m-1} \left(\int_0^T \left\| Y_t^{(k+1)} - Y_t^{(k)} \right\|_{L^2(\Omega)} dt \right)^{1/2} \leq \sum_{k=n}^{m-1} \left(\int_0^T \frac{A_2^{k+1} t^{k+1}}{(k+1)!} dt \right)^{1/2} = \sum_{k=n}^{m-1} \left(\frac{A_2^{k+1} T^{k+2}}{(k+2)!} \right)^{1/2} \xrightarrow{m,n\to\infty} 0.$$

Hence, $\left\{ Y_t^{(k)} \right\}_k$ converges in $L^2([0,T]\times\Omega)$ to a process $(X_t)_t$.

Since $Y_t^{(k)}$ is $\mathscr{F}_t^Z$-measurable (for all $t \in [0,T]$ and $k \in \mathbb{N}$), the limit X_t is also $\mathscr{F}_t^Z$-measurable.

By the Hölder inequality,

$$\int_0^t b\left(s, Y_s^{(k)}\right) ds \to \int_0^t b(s, X_s)\, ds \quad \text{in } L^2(\Omega);$$

and by the Itô isometry,

$$\int_0^t \sigma\left(s, Y_s^{(k)}\right) dB_s \to \int_0^t \sigma(s, X_s)\, dB_s \quad \text{in } L^2(\Omega).$$

From these two facts and (3.12), we conclude that X_t satisfies (3.9a) and (3.9b). Moreover, a modification of X can be chosen to be path-continuous by definition of the Itô integral. This proves the existence of a solution. □

The uniqueness given by the previous theorem is in the sense of modifications (or indistinguishable processes, which is equivalent, as we require path-continuity). That is, if we have two $\left(\mathscr{F}_t^Z\right)_t$-adapted solutions X and Y of (3.9a)-(3.9b) then $[X_t(\omega) = Y_t(\omega)\ \forall t \leq T]$ P-a.e. These kind of solutions, where the modification of the Brownian motion B is given in advance, are called *strong solutions*. If we are only given b, σ and are asked to find a pair of processes and a filtration $((X,B),(\mathscr{F}_t)_{t\geq 0})$[5] satisfying the SDE (3.9a)-(3.9b) on a probability space, then we call $((X,B),(\mathscr{F}_t)_{t\geq 0})$ a *weak solution of the SDE*. Two weak solutions may not be modifications of each other. In fact, one may not be adapted to the filtration of the other. However, two weak solutions will be the identical in law: they will have the same finite-dimensional distributions. We call this property of having the same finite-dimensional distributions *weak uniqueness*.

Lemma 3.13. *If b, σ satisfy the conditions of Theorem 3.12, then any solution of the SDE (3.9a)-(3.9b) is weakly unique.*

See [Øks13, Lemma 5.3.1] for a proof of Lemma 3.13. The idea is to show that if $Y^{(0)}$ and $\hat{Y}^{(0)}$ have the same law, then the inductively defined processes from (3.12) have the same probability laws.

[5] Of course, requiring B to be a Brownian motion and being a martingale with respect to the filtration $(\mathscr{F}_t)_{t\geq 0}$ σ-algebras containing all nullsets.

CHAPTER 4
Itô diffusions

In order to model the evolution of processes such as heat diffusion, we need the analog of autonomous differential equations in the stochastic setting. In the context of ordinary differential equations, an autonomous differential equation is one where the right-hand side of the equation does not depend explicitly on the independent variable:

$$\frac{dy(t)}{dt} = F(y(t)).$$

In stochastic calculus, the processes solving this kind of equations are called *Itô diffusions*.

In this chapter, we define Itô diffusions and prove that they satisfy the Markov property. Then, we show how the Markov property yields the mean value property, which is a key ingredient in the applications that will be presented in the following chapters. Finally, we introduce the generator and the characteristic operator of Itô diffusions, which are the infinitesimal operators that link Itô diffusions with partial differential equations.

Definition 4.1. *A* (time-homogenous) Itô diffusion *is a stochastic process*

$$X_t(\omega) = X(t,\omega) : [s,\infty) \times \Omega \to \mathbb{R}^n$$

satisfying an SDE of the form

$$\begin{cases} dX_t = b(X_t)dt + \sigma(X_t)dB_t, \ t \geq s, \\ X_s = x, \end{cases} \tag{4.1}$$

where B is an m-dimensional Brownian motion, and $b : \mathbb{R}^n \to \mathbb{R}^n$, $\sigma : \mathbb{R}^n \to \mathbb{R}^{n\times n}$ satisfy the conditions of Theorem 3.12, which can be simplified (due to the lack of the parameter t) to both b and σ being Lipschitz continuous.

In the language of theorem Theorem 3.12, the solutions of (4.1) have constant $Z = x$. Therefore, $\mathscr{F}_t^Z = \mathscr{F}_t$ and the solutions $(X_t)_{t\geq 0}$ are $(\mathscr{F}_t)_{t\geq 0}$-adapted.

We will denote the (unique) solution of (4.1) by $\left(X_t^{s,x}\right)_{t\geq s}$ or simply $X_t^{s,x}$. Whenever $s = 0$, we will omit the s superscript and write X_t^x instead of $X_t^{0,x}$. These processes have the property of being time-homogenous in the following sense. Let T be a finite $(\mathscr{B}_t)_{t\geq 0}$-stopping time. By the strong Markov property for Brownian Motion (Theorem 2.24), $\tilde{B}_v = B_{T+v} - B_T$ is a Brownian motion independent of $\mathscr{B}_T$ and

$$\begin{aligned} X_{T+t}^{T,x} &= x + \int_T^{T+t} u\left(X_u^{T,x}\right) du + \int_T^{T+t} \sigma\left(X_u^{T,x}\right) dB_u \\ &= x + \int_0^t u\left(X_{T+v}^{T,x}\right) dv + \int_0^t \sigma\left(X_{T+v}^{T,x}\right) dB_{T+v} = x + \int_0^t u\left(X_{T+v}^{T,x}\right) dv + \int_0^t \sigma\left(X_{T+v}^{T,x}\right) d\tilde{B}_v. \end{aligned}$$

$\left(X_{T+t}^{T,x}\right)_{t\geq 0}$ is independent of $\mathscr{F}_T$. Also,

$$X_t^{0,x} = x + \int_0^t u\left(X_v^{0,x}\right) dv + \int_0^t \sigma\left(X_v^{0,x}\right) dB_v.$$

Since B and $\tilde{B}$ have the same finite-dimensional distributions, by weak uniqueness (Lemma 3.13), we have that $\left(X_{T+t}^{T,x}\right)_{t\geq 0}$ and $\left(X_t^{0,x}\right)_{t\geq 0}$ have the same finite-dimensional distributions. This hints at the fact that Itô diffusions satisfy the Markov property, just like Brownian motion. To prove this, we will first introduce some useful notation.

From now on, $(X_t)_{t\geq 0}$ will be an Itô diffusion solving $dX_t = b(X_t)dt + \sigma(X_t)dB_t$, and Q^x will denote its probability law when the initial value of the process is x and E^x the expectation with respect to Q^x. For example, $E^x\left[f_1\left(X_{t_1}\right)\cdots f_k\left(X_{t_k}\right)\right] = E\left[f_1\left(X_{t_1}^x\right)\cdots f_k\left(X_{t_k}^x\right)\right]$, whenever $\left(f_k\right)_k \subseteq L^\infty(\mathbb{R}^n)$ and $(t_k)_k \subseteq \mathbb{R}_+$, where (as usual), E denotes the expectation with respect to the probability law of the Brownian motion.

4.1. Markov properties

Itô diffusions inherit the Markov property from Brownian motion.

Theorem 4.2 (The Markov property for Itô diffusions).

$$f \in L^{\infty}\left(\mathbb{R}^n\right) \implies \forall t, h \geq 0 \quad E^x\left[f\left(X_{t+h}\right) \mid \mathscr{F}_t\right](\omega) = E^{X_t^x(\omega)}\left[f\left(X_h\right)\right]. \tag{4.2}$$

The Markov property is a special case of the strong Markov property:

Theorem 4.3 (The strong Markov property for Itô diffusions). *Let T be a finite $(\mathscr{F}_t)_{t\geq 0}$-stopping time and $f \in L^{\infty}(\mathbb{R}^n)$. Then,*

$$\forall h \geq 0 \quad E^x\left[f\left(X_{T+h}\right) \mid \mathscr{F}_T\right](\omega) = E^{X^x_{T(\omega)}(\omega)}\left[f\left(X_h\right)\right].$$

Proof. Since $X_r^x(\omega) = X_t^x(\omega) + \int_t^r b\left(X_u\right) du + \int_t^r \sigma\left(X_u\right) dB_u$ for every $r \geq t$, by uniqueness of the solution, $X_r^x = X_r^{t,X_t^x}$. Define $F(x,t,r,\omega) := X_r^{t,x}(\omega)$ for $r \geq t$. $F(x,t,r,\omega)$ is measurable with respect to its four arguments jointly and is independent of $\mathscr{F}_t$ (see [Øks13, Exercise 7.6] for more details).

Define $g = f \circ F$, which is the composition of measurable functions and is therefore measurable. By Theorem 2.41, we can write g as the limit:

$$g(x,t,r,\omega) = \lim_{n\to\infty} \sum_{k=1}^{M_n} \phi_k^n(x) \psi_k^n(t,r,\omega),$$

where the limit commutes with the (conditional) expectation and the ϕ_k^n, ψ_k^n are measurable. Then,

$$\begin{aligned}
E\left[f\left(X_{T+h}^x\right) \mid \mathscr{F}_T\right] &= E\left[f\left(X_{T+h}^{T,X_T^x}\right) \,\middle|\, \mathscr{F}_T\right] = E\left[g\left(X_T^x, T, T+h, \omega\right) \mid \mathscr{F}_T\right] \\
&= \lim_{n\to\infty} \sum_{k=1}^{M_n} E\left[\phi_k^n\left(X_T^x\right) \psi_k^n(T, T+h, \omega) \mid \mathscr{F}_T\right] = \lim_{n\to\infty} \sum_{k=1}^{M_n} \phi_k^n\left(X_T^x\right) E\left[\psi_k^n(T, T+h, \omega) \mid \mathscr{F}_T\right] \\
&= \lim_{n\to\infty} \sum_{k=1}^{M_n} E\left[\phi_k^n(y) \psi_k^n(T, T+h, \omega) \mid \mathscr{F}_T\right]_{y = X_T^x} = E\left[g\left(y, T, T+h, \omega\right) \mid \mathscr{F}_T\right]_{y = X_T^x} \\
&= E\left[f\left(X_{T+h}^{T,y}(\omega)\right) \,\middle|\, \mathscr{F}_T\right]_{y = X_T^x} = E\left[f\left(X_{T+h}^{T,y}(\omega)\right)\right]_{y = X_T^x} = E\left[f\left(X_h^{0,y}(\omega)\right)\right]_{y = X_T^x}.
\end{aligned} \tag{4.3}$$

□

There is a stronger, more general version of the strong Markov property for Itô diffusions. Let $L(\mathscr{X}_\infty)$ be the set of all $\mathscr{X}_\infty$-measurable functions. For each $t \geq 0$, we define the *shift operator* $\theta_t : L(\mathscr{X}_\infty) \to L(\mathscr{X}_\infty)$ as follows: If $\eta = \prod_{j=1}^k g_j\left(X_{t_j}\right)$, where $\{t_j\}_{j=1}^k \subseteq \mathbb{R}_+$ and $\left(g_j\right)_j$ are Borel-measurable, then

$$\theta_t \eta := \prod_{j=1}^k g_j\left(X_{t_j+t}\right).$$

We call $\mathscr{P}$ the subset of $L(\mathscr{X}_\infty)$ of all functions like η. Every function in $L(\mathscr{X}_\infty)$ is the pointwise limit of a sequence of functions in $\operatorname{span}(\mathscr{P})$ (Theorem 2.41). We extend the definition of θ_t to $L(\mathscr{X}_\infty)$ by taking limits. See [RY05, Chapter I, p. 36] for a more detailed explanation of the meaning and proper definition of the shift operator.

By induction, it can be proven that, if T is a finite $(\mathscr{B}_t)_{t\geq 0}$-stopping time and $\eta = \prod_{j=1}^k g_j\left(X_{t_j}\right)$ is in $\mathscr{P}$,

$$E^x\left[\prod_{j=1}^k g_j\left(X_{t_j+T}\right) \,\middle|\, \mathscr{F}_T\right] = E^{X_T^x}\left[\prod_{j=1}^k g_j\left(X_{t_j}\right)\right]. \tag{4.4}$$

We do the case for $k = 2$: let $0 \leq h \leq \ell$, then

$$\begin{aligned}
E^x\left[f\left(X_{T+h}\right) g\left(X_{T+\ell}\right) \mid \mathscr{F}_T\right] &= E^x\left[f\left(X_{T+h}\right) E\left[g\left(X_{T+\ell}\right) \mid \mathscr{F}_{T+h}\right] \mid \mathscr{F}_T\right] \\
&= E^x\left[f\left(X_{T+h}\right) E^{X_{T+h}}\left[g\left(X_{\ell-h}\right)\right] \mid \mathscr{F}_T\right] = E^{X_T}\left[f\left(X_h\right) E^{X_h}\left[g\left(X_{\ell-h}\right)\right]\right] \\
&= E^{X_T}\left[f\left(X_h\right) E\left[g\left(X_\ell\right) \mid \mathscr{F}_h\right]\right] = E^{X_T}\left[f\left(X_h\right) g\left(X_\ell\right)\right].
\end{aligned}$$

It follows from (4.4) that the strong Markov property holds for all $\eta \in L^{\infty}(\mathscr{X}_\infty)$:

$$E^x\left[\theta_T \eta \mid \mathscr{F}_T\right] = E^{X_T^x}\left[\eta\right]. \tag{4.5}$$

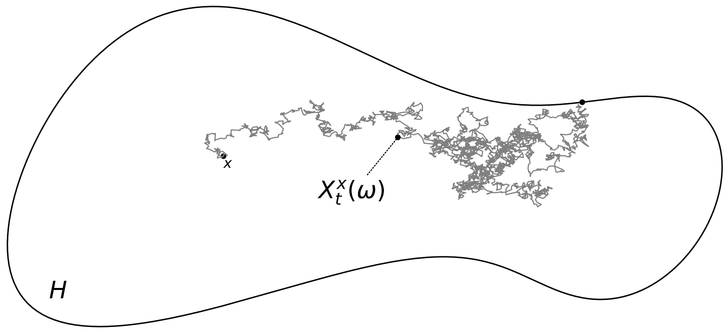


Figure 4.1: A set $H \subseteq \mathbb{R}^2$ and a realization of a diffusion starting at $x \in H$ and exiting H for the first time.

4.2. Harmonic measure and mean value property

This section introduces the concepts of harmonic measure and the mean value property. Although the results presented here may seem to be a simple consequence of the Markov property, they provide an insightful interpretation of the connection between diffusions and elliptic differential operators, which will be further explored in the following sections and chapters. As in the previous section X is a diffusion with values in $\mathbb{R}^n$.

Proposition 4.4. *If $C \subseteq \mathbb{R}^n$ is closed, then*

$$D_C(\omega) := \inf\{t \geq 0 : X_t(\omega) \in C\}^{1}$$

is called the first hitting time of C by X *and it is an $(\mathscr{X}_t)_{t\geq 0}$-stopping time.*

Proof. Let $t \geq 0$. $x \mapsto d(x,C) := \inf_{c\in C} d(x,c)$ is continuous and X_s is $\mathscr{X}_t$-measurable for every $s \leq t$, so $\omega \mapsto d(X_s(\omega),C)$ is $\mathscr{X}_t$-measurable. Thus, $\omega \mapsto \inf_{s\in\mathbb{Q}\cap[0,t]} d(X_s(\omega),C)$ is $\mathscr{X}_t$-measurable and $\{D_C \leq t\}$. By path-continuity of X,

$$\{D_C \leq t\} = \left\{\omega \in \Omega : \inf_{s\in\mathbb{Q}\cap[0,t]} d\left(X_s(\omega),C\right) = 0\right\} \in \mathscr{X}_t.$$

To prove the above set equality:

$$\begin{aligned}
D_C \leq t &\iff \exists r \leq t : \ d(X_r,C) = 0 && \text{(here we use that } C \text{ is closed)}\\
&\iff \min_{s\in[0,t]} d(X_s,C) = 0 && \text{(going up, the inf. is reached by cont. of } s \mapsto d(X_s,C))\\
&\iff \inf_{s\in\mathbb{Q}\cap[0,t]} d(X_s,C) = 0.
\end{aligned}$$

□

Remark 4.5. *As X is $(\mathscr{F}_t)_{t\geq 0}$-adapted, all $(\mathscr{X}_t)_{t\geq 0}$-stopping times are $(\mathscr{F}_t)_{t\geq 0}$-stopping times.*

The opposite of a hitting time is an *exit time* (see Figure 4.1):

Definition 4.6. *Let $H \subseteq \mathbb{R}^n$ be open. The* first exit time of H by X *is defined as*

$$\tau_H := \inf\{t \geq 0 : X_t \notin H\} = D_{H^c}.$$

Lemma 4.7. *Let $H \in \mathscr{B}(\mathbb{R}^n)$, then τ_H is a stopping time with respect to any filtration $\left(\mathscr{J}_t\right)_{t\geq 0}$ of complete σ-algebras satisfying $\mathscr{X}_t \subseteq \mathscr{J}_t$ $\forall t \geq 0$.*

A complete proof for this lemma can be found in [RY05, Chapter III, (2.17)]. We will only need the case for open H anyways. We will apply this result for the case in which $\left(\mathscr{J}_t\right)_{t\geq 0} = (\mathscr{F}_t)_{t\geq 0}$.

Proof. Suppose that H is closed. For each $j \in \mathbb{N}$, define the open set $A_j = \cup_{p\in H} B(p,1/j)$. $H = \cap_{j\in\mathbb{N}} A_j$, because H is closed. Then, the sequence of stopping times $(\tau_{A_j})_{j\in\mathbb{N}}$ is pointwise decreasing. We define $\tau = \inf_j \tau_{A_j}$, which is an $\left(\mathscr{J}_t\right)_{t\geq 0}$-stopping time by Corollary 2.19. We are going to prove that $\tau = \tau_H$.

Fix $\omega \in \Omega$. Until the end of the proof assume that all random variables are evaluated at ω.

$\tau = \inf_j \inf\{t \geq 0 : X_t \notin A_j\}$. Therefore, there exists a decreasing sequence $\left(t_j\right)_j \subseteq \mathbb{R}_+$ such that $t_j \to \tau$ and $X_{t_j} \notin A_j$ for every $j \in \mathbb{N}$. This implies that $X_{t_j} \notin H$ $\forall j \in \mathbb{N}$, so $\tau \geq \tau_H$.

Now suppose that $\tau > \tau_H$. Then there is an s such that $\tau_H < s < \tau \leq \tau_{A_j}$ $\forall j \in \mathbb{N}$. So $\forall v \leq s$ $\forall j \in \mathbb{N}$ $X_v \in A_j$; this implies $\forall v \leq s$ $X_v \in H$, which contradicts $\tau_H < s$. We conclude that $\tau \leq \tau_H$. □

[1] We use the convention: $\inf\emptyset = +\infty$.

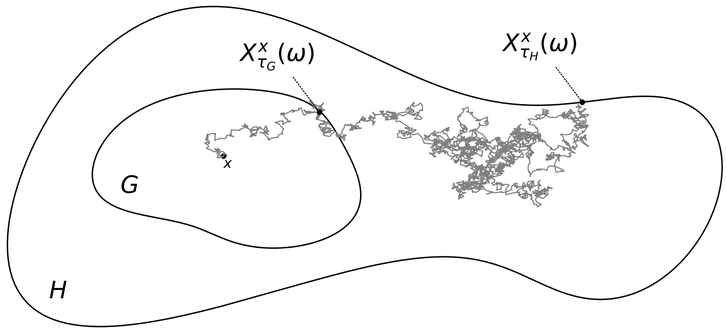


Figure 4.2: Two nested sets $G \subseteq H \subseteq \mathbb{R}^2$ and a realization of a diffusion starting at $x \in G$.

Let α be an $(\mathscr{F}_t)_{t\geq 0}$-stopping time, and define

$$\tau_H^\alpha := \inf\{t > \alpha : X_t \notin H\}, \tag{4.6}$$

which is also an $(\mathscr{F}_t)_{t\geq 0}$-stopping time for every $H \in \mathscr{B}(\mathbb{R}^n)$ [RY05, p. 95].

Let $H \in \mathscr{B}(\mathbb{R}^n)$, $g \in C_b(\mathbb{R}^n)$ and define $\eta := g\left(X_{\tau_H}\right) \mathbb{1}_{\{\tau_H<\infty\}}$. Then,

$$\theta_\alpha \eta \mathbb{1}_{\{\alpha<\infty\}} = g\left(X_{\tau_H^\alpha}\right) \mathbb{1}_{\{\tau_H^\alpha<\infty\}}. \tag{4.7}$$

Proof of (4.7). Approximate η by the sequence $\eta_k := \sum_j g\left(X_{t_j}\right) \mathbb{1}_{[t_j,t_{j+1})}(\tau_H)$, where $t_j = j2^{-k}$ for each $j \in \mathbb{N}\cup\{0\}$.

$$\begin{aligned}\theta_t \mathbb{1}_{[t_j,t_{j+1})}(\tau_H) &= \theta_t \mathbb{1}_{\{\forall r\in(0,t_j)\ X_r\in H\}\cap\{\exists s\in(t_j,t_j+1):\ X_s\notin H\}} = \mathbb{1}_{\{\forall r\in(0,t_j)\ X_{r+t}\in H\}\cap\{\exists s\in(t_j,t_j+1):\ X_{s+t}\notin H\}} \\ &= \mathbb{1}_{\{\forall r\in(t,t+t_j)\ X_t\in H\}\cap\{\exists s\in(t+t_j,t+t_j+1):\ X_s\notin H\}} = \mathbb{1}_{[t+t_j,t+t_{j+1})}\left(\tau_H^t\right).\end{aligned} \tag{4.8}$$

Then,

$$\theta_t\eta = \lim_k \sum_j g\left(X_{t+t_j}\right) \mathbb{1}_{[t+t_j,t+t_{j+1})}(\tau_H^t) = g\left(X_{\tau_H^t}\right) \mathbb{1}_{\{\tau_H^t<\infty\}}. \qquad \square$$

Definition 4.8 (Harmonic measure). *Let G be a Borel set in $\mathbb{R}^n$, and let $x \in G$. The* harmonic measure of X on G starting at x *is defined as*

$$\mu_G^x(F) := Q^x\left(X_{\tau_G} \in F\right) \quad \forall F \in \mathscr{B}(\partial G)^{2}.$$

The particular case that interests us is when G is bounded, $G \subseteq H$ (both G and H are Borel-measurable), and $\alpha = \tau_G$, where τ_H is finite (see Figure 4.2). This implies $\tau_H^\alpha = \tau_H$ and so by (4.7),

$$\theta_{\tau_G} g\left(X_{\tau_H}\right) = g\left(X_{\tau_H}\right). \tag{4.9}$$

Let $f \in L^\infty(\mathbb{R}^n)$ and define $\phi : G \to \mathbb{R}$ by $\phi(x) := E^x\left[f\left(X_{\tau_H}\right)\right]$. Then

$$\begin{aligned} E^x\left[f\left(X_{\tau_H}\right)\right] &= E^x\left[E^x\left[f\left(X_{\tau_H}\right) \mid \mathscr{F}_{\tau_G}\right]\right] \overset{(4.9)}{=} E^x\left[E^x\left[\theta_{\tau_G} f\left(X_{\tau_H}\right) \mid \mathscr{F}_{\tau_G}\right]\right] \\ &\overset{(4.5)}{=} E^x\left[E^{X_{\tau_G}}\left[f\left(X_{\tau_H}\right)\right]\right] = \int_{\partial G} E^y\left[f\left(X_{\tau_H}\right)\right] \cdot Q^x\left(X_{\tau_G} \in dy\right) = \int_{\partial G} E^y\left[f\left(X_{\tau_H}\right)\right] \cdot d\mu_G^x(y).\end{aligned} \tag{4.10}$$

That is, ϕ satisfies the *mean value property*: for every $G \in \mathscr{B}(H)$ and every $x \in G$

$$\phi(x) = \int_{\partial G} \phi(y)\, d\mu_G^x(y).$$

We will come back to this important property later.

[2] For every $A \in \mathscr{B}(\mathbb{R}^n)$, we denote by $\mathscr{B}(A)$ the set $\mathscr{B}(\mathbb{R}^n) \cap A$, which is the same as the σ-algebra generated by the relative open sets of A.

4.3. The infinitesimal generator of an Itô diffusion

In this section, we present one of the main concepts of this work; the relationship between an Itô diffusion and a certain differential operator. The link between the two is the generator of the diffusion, which can be intuitively thought of as differentiating a function along the paths of the diffusion:

Definition 4.9. *The* (infinitesimal) generator A of an Itô diffusion X *is defined by*

$$Af(x) := \lim_{t\to 0} \frac{E^x[f(X_t)] - f(x)}{t}; \quad x \in \mathbb{R}^n. \tag{4.11}$$

We will denote by $\mathcal{D}_A(x)$ *the set of functions* $f:\mathbb{R}^n \to \mathbb{R}$ *such that the limit* (4.11) *exists at* x. *And* $\mathcal{D}_A = \cap_{x\in\mathbb{R}^n}\mathcal{D}_A(x)$.

Lemma 4.10. *Let* $f \in C_c^2(\mathbb{R}^n)$ *(i.e.* $f \in C^2(\mathbb{R}^n)$ *and* f *has compact support). Let* $Y_t = Y_t^x$ *be an Itô process in* $\mathbb{R}^n$ *of the form:*

$$Y_t^x(\omega) = x + \int_0^t u(s,\omega)ds + \int_0^t v(s,\omega)dB_s(\omega),$$

as in Definition 3.10 *where* B *is an* m*-dimensional Brownian motion and* u, v *are bounded on* $\{(t,\omega) : Y(t,\omega) \in \operatorname{supp}(f)\}$. *Let* τ *be an integrable* $(\mathcal{F}_t)_{t\geq 0}$*-stopping time. Then*

$$E^x[f(Y_\tau)] = f(x) + E^x\left[\int_0^\tau \left(\langle u(s,\omega), D_x f(Y_s)\rangle + \frac{1}{2}\langle (vv^T)(s,\omega), D_x^2 f(Y_s)\rangle\right) ds\right], \tag{4.12}$$

where E^x *denotes the expectation with respect to the natural probability law of* $(Y_t)_t$ *starting at* x.

Proof. By the Itô formula (3.6) (with $D_t f = 0$),

$$E^x[f(Y_\tau)] = f(Y_0) + E^x\left[\int_0^\tau \langle D_x f, u\rangle dt + \int_0^\tau \langle D_x f, v dB\rangle + \frac{1}{2}\int_0^\tau \langle D_x^2 f, vv^T\rangle dt\right].$$

We only need to show that $E\left[\int_0^\tau \langle D_x f, vdB\rangle\right] = 0$. It is enough to prove that if $\hat{B}$ is one coordinate of the Brownian motion B and g is a bounded and Itô-integrable process, then $E\left[\int_0^\tau g(s,\omega)\, d\hat{B}_s\right] = 0$. To show this, let $k \in \mathbb{N}$, then $E^x\left[\int_0^{\tau\wedge k} g d\hat{B}_s\right] = E^x\left[\int_0^k \mathbb{1}_{\{s<\tau\}} g d\hat{B}_s\right] = 0$, because $\mathbb{1}_{\{s<\tau\}}g(s,\omega)$ is Itô-integrable and its Itô integral is a martingale. Also,

$$E^x\left[\left(\int_0^\tau g d\hat{B}_s - \int_0^{\tau\wedge k} g d\hat{B}_s\right)^2\right] = E^x\left[\left(\int_{\tau\wedge k}^\tau g d\hat{B}_s\right)^2\right] = E^x\left[\int_{\tau\wedge k}^\tau g^2 ds\right] \leq \|g\|_\infty^2 E^x[\tau - \tau\wedge k] \xrightarrow{k\to\infty} 0.$$

Therefore, $0 = \lim_k E^x\left[\int_0^{\tau\wedge k} g d\hat{B}_s\right] = E^x\left[\int_0^\tau g d\hat{B}_s\right]$. □

Theorem 4.11. *Let* X *be the Itô diffusion* $dX_t = b(X_t)dt + \sigma(X_t)dB_t$. *Then* $C_c^2(\mathbb{R}^n) \subseteq \mathcal{D}_A$ *and* $\forall f \in C_c^2(\mathbb{R}^n)$

$$Af(x) = \langle b(x), D_x f(x)\rangle + \frac{1}{2}\langle \sigma(x)\sigma(x)^T, D_x^2 f(x)\rangle \qquad \forall x \in \mathbb{R}^n. \tag{4.13}$$

Proof. We apply Lemma 4.10 to $u(t,\omega) = b(X_t(\omega))$, $v(t,\omega) = \sigma(X_t(\omega))$ and the constant stopping time $\tau = t$. b and σ are Lipschitz, so $b(X_s)$ and $\sigma(X_s)$ are bounded whenever X_s is in the compact support of f. By the definition of Af (4.11), the dominated convergence theorem and the Lebesgue differentiation theorem (X is path-cont.):

$$\begin{aligned}
Af(x) &= E^x\left[\lim_{t\to 0}\frac{1}{t}\int_0^t \left(\langle b(X_s), D_x f(X_s)\rangle + \frac{1}{2}\langle \sigma(X_s)\sigma(X_s)^T, D_x^2 f(X_s)\rangle\right) ds\right] \\
&= E^x\left[\langle b(X_0), D_x f(X_0)\rangle + \frac{1}{2}\langle \sigma(X_0)\sigma(X_s)^T, D_x^2 f(X_0)\rangle\right] \\
&= \langle b(x), D_x f(x)\rangle + \frac{1}{2}\langle \sigma(x)\sigma(x)^T, D_x^2 f(x)\rangle.
\end{aligned}$$

□

Combining (4.12) and (4.13):

Theorem 4.12 (Dynkin's formula). *Let* $f \in C_c^2(\mathbb{R}^n)$ *and* τ *a finite* $(\mathcal{F}_t)_{t\geq 0}$*-stopping time. Then*

$$E^x[f(X_\tau)] = f(x) + E^x\left[\int_0^\tau Af(X_s)ds\right]. \tag{4.14}$$

Example 4.13. *The generator of* n*-dimensional Brownian motion is* $\frac{1}{2}\Delta$, *where* Δ *denotes is the Laplace operator in* $\mathbb{R}^n$.

Example 4.14. *The generator of the graph of* n*-dimensional Brownian motion:* $X_t = [t, B_t]^T \in \mathbb{R}^{n+1}$, *is the* backward heat operator $\frac{\partial}{\partial t} + \frac{1}{2}\Delta_x$, *where* Δ_x *is the Laplace operator in* $\mathbb{R}^n$.

4.4. The characteristic operator of an Itô diffusion

The notion of the generator of an Itô diffusion can be extended to a more general concept, the characteristic operator. This characteristic operator is defined in terms of the exit time of the diffusion from open sets. We will show that the characteristic operator of an Itô diffusion coincides with the differential operator associated with its generator when applied to functions in $C^2(\mathbb{R}^n)$. This is an improvement over Theorem 4.11, which only considers $C^2_c(\mathbb{R}^n)$.

Definition 4.15. *The* characteristic operator $\mathfrak{A}$ of an Itô diffusion X *is defined by*

$$\mathfrak{A}f(x) := \lim_{U \downarrow x} \frac{E^x\left[f\left(X_{\tau_U}\right)\right] - f(x)}{E^x\left[\tau_U\right]}; \quad x \in \mathbb{R}^n, \tag{4.15}$$

Where the U's are open sets decreasing to $\{x\}$, and the limit is taken in the sense of nets[3]*. Once again, we will denote by $\mathscr{D}_{\mathfrak{A}}$ the set of functions $f : \mathbb{R}^n \to \mathbb{R}$ such that the limit (4.15) exists for every $x \in \mathbb{R}^n$. If $E^x\left[\tau_U\right] = \infty$ for every open $U \ni x$, we define $\mathfrak{A}f(x) = 0$.*

In [Dyn65, Theorem 5.5 (p. 142)] it is proven that $\mathscr{D}_A \subseteq \mathscr{D}_{\mathfrak{A}}$ and that $\mathfrak{A}|_{\mathscr{D}_A} = A$, but we will need and prove a waker result.

A point $x \in \mathbb{R}^n$ is called a *trap for X* if $\tau_{\{x\}} = \infty$ Q^x-a.e. We will make use of the following Lemma, which is proven in [Dyn65, Lemma 5.5 (p. 139)].

Lemma 4.16. *If $x \in \mathbb{R}^n$ is not a trap for X, then there exists an open (bounded) set $U \ni x$ such that $E^x[\tau_U] < \infty$.*

Theorem 4.17. *Let $f \in C^2(\mathbb{R}^n)$. Then $f \in \mathscr{D}_{\mathfrak{A}}$ and $\mathfrak{A}f = \left\langle b, D_x f\right\rangle + \frac{1}{2}\left\langle \sigma\sigma^T, D_x^2 f\right\rangle$. In particular, $\mathfrak{A} = A$ in $C^2_c(\mathbb{R}^n)$.*

Proof. Let L denote the operator $Lf = \left\langle b, D_x f\right\rangle + \frac{1}{2}\left\langle \sigma\sigma^T, D_x^2 f\right\rangle$.

If x is a trap for X, then $\mathfrak{A}f(x) = 0$. Choose a bounded open set $V \ni x$, and a function $f_0 \in C^2_c(\mathbb{R}^n)$ such that $f|_V = f_0|_V$. Then $f_0 \in \mathscr{D}_A(x)$ and $0 = Af_0(x) = Lf_0(x) = Lf(x)$. If x is not a trap for X, then there exists an open set $U \ni x$ such that $E^x[\tau_U] < \infty$. Put $\tau = \tau_U$ and apply Dynkin's formula (Theorem 4.12):

$$\left|\frac{E^x\left[f\left(X_\tau\right)\right] - f(x)}{E^x[\tau]} - Af(x)\right| = \frac{1}{E^x[\tau]}\left|E^x\left[\int_0^\tau \left(Af(X_s) - Af(x)\right)\, ds\right]\right| \le \sup_{y \in U}\left|Af(y) - Af(x)\right| \to 0,$$

as $U \downarrow x$, because Af is continuous (4.13). This proves that $\mathfrak{A}f(x) = Af(x)$. □

[3] The limit $\lim_{U \downarrow x} h(U)$ exists and is equal to ℓ iff for every $\varepsilon > 0$, there exists an open set $U \ni x$ such that for all open $V \subseteq U$ with $x \in V$, $|h(V) - \ell| < \varepsilon$.

CHAPTER 5

Application to parabolic partial differential equations

The first application of the theory of Itô diffusions that we are going to present is a result on existence and uniqueness of the solution of global linear differential equations of the form $\frac{\partial v}{\partial t} = Av$ with initial condition $v(0,x) = f(x)$, where A is the generator of an Itô diffusion. As we have seen, Ag coincides with Lg, when $g \in C_c^2(\mathbb{R}^n)$, for a differential operator L of the form

$$Lg = \sum_{i=1}^n b_i(x)\frac{\partial g}{\partial x_i} + \frac{1}{2}\sum_{i,j=1}^n \left(\sigma\sigma^T\right)_{ij}(x)\frac{\partial^2 g}{\partial x_i \partial x_j}.$$

A classic example of this kind of equation is the heat equation $\frac{\partial v}{\partial t} = \frac{1}{2}\Delta v$ for $t > 0$, $x \in \mathbb{R}^n$, where Δ is the Laplace operator in $\mathbb{R}^n$.

As before, we will consider the Itô diffusion X given by $dX_t = b(X_t)dt + \sigma(X_t)dB_t$.

5.1. The Feynman-Kac formula

Theorem 5.1 (The Feynman-Kac formula). *Let $f \in C_c^2(\mathbb{R}^n)$ and $q \in C(\mathbb{R}^n)$. Assume that q is lower bounded*[1]*.*

(a) Define

$$v(t,x) = E^x\left[\exp\left(-\int_0^t q(X_s)\,ds\right) f(X_t)\right]. \tag{5.1}$$

Then $v(t,\cdot) \in \mathscr{D}_A$ $\forall t > 0$ and

$$\frac{\partial v}{\partial t}(t,x) = Av(t,x) - q(x)v(t,x); \qquad \forall t > 0, x \in \mathbb{R}^n \tag{5.2}$$

$$v(0,x) = f(x); \qquad \forall x \in \mathbb{R}^n \tag{5.3}$$

where Av denotes A applied to $x \mapsto v(t,x)$, for every $t > 0$.

(b) If $w \in C^{1,2}(\mathbb{R}_+ \times \mathbb{R}^n)$ is bounded on $K \times \mathbb{R}^n$ for every compact set $K \subseteq \mathbb{R}_+$ and w is a solution of (5.2) & (5.3)*, then $w = v$.*

Proof of (a). Let $Y_t = f(X_t)$, $Z_t = \exp\left(-\int_0^t q(X_s)\ ds\right)$. Then,

$$dZ_t = -q(X_t)Z_t\ dt$$

$$dY_t = \left\langle D_x f(X_t), u(X_t)\right\rangle\ dt + \frac{1}{2}\left\langle D_x^2 f(X_t), (vv^T)(X_t)\right\rangle\ dt + \left\langle D_x f(X_t), v(X_t)dB_t\right\rangle,$$

and so $dY_t \cdot dZ_t = 0$. Therefore, $d(Z_tY_t) = Z_t dY_t + Y_t dZ_t$ (Itô's formula). Since Z_tY_t is an Itô process, by Lemma 4.10 $v(t,x) = E^x[Z_tY_t]$ is differentiable w.r.t. t.

$$\begin{aligned}
\frac{1}{r}\left(E^x[v(t,X_r)] - v(t,x)\right) &\overset{\text{(def)}}{=} \frac{1}{r}E^x\left[E^{X_r}[Z_t f(X_t)] - E^x[Z_t f(X_t)]\right] \\
&\overset{(4.5)}{=} \frac{1}{r}E^x\left[E^x\left[f(X_{r+t})\exp\left(-\int_0^t q(X_{r+s})\ ds\right)\,\middle|\,\mathscr{F}_r\right] - E^x\left[Z_t f(X_t)\mid\mathscr{F}_r\right]\right] \\
&= \frac{1}{r}E^x\left[f(X_{r+t})\exp\left(-\int_0^t q(X_{r+s})\ ds\right) - Z_t f(X_t)\right]
\end{aligned}$$

[1]There exists a constant $C \in \mathbb{R}$ such that $\forall x \in \mathbb{R}^n$ $q(x) \geq C$.

$$
\begin{aligned}
&= \frac{1}{r} E^x \left[f(X_{r+t}) \exp\left(-\int_r^{r+t} q(X_s)\, ds \right) - Z_t f(X_t) \right] \\
&= \frac{1}{r} E^x \left[Z_{r+t} f(X_{r+t}) \exp\left(\int_0^{r} q(X_s)\, ds \right) - Z_t f(X_t) \right] \\
&= \underbrace{\frac{1}{r} E^x \left[Z_{r+t} f(X_{r+t}) - Z_t f(X_t) \right]}_{\text{(I)}} + E^x \Bigg[\underbrace{\frac{1}{r} Z_{r+t} f(X_{r+t}) \left(\exp\left(\int_0^r q(X_s)\, ds \right) - 1 \right)}_{\text{(II)}} \Bigg].
\end{aligned}
$$

(I) $\xrightarrow{r\to 0} v_t(t,x)$, and (II) $\xrightarrow{r\to 0} Z_t f(X_t) q(X_0)$ pointwise boundedly, so

$$
E^x [\text{(II)}] \xrightarrow{r\to 0} E^x \left[Z_t f(X_t) q(X_0) \right] = q(X_0) v(t,x). \qquad \square
$$

Proof of (b). Let w be as in the theorem statement. Define $\hat{A}w(t,x) := -\frac{\partial w}{\partial t} + Aw - qw = 0$, and, for $(s,x,z) \in \mathbb{R} \times \mathbb{R}^n \times \mathbb{R}$, the Itô diffusion $H_t = \left(s-t, X_t^{0,x}, z + \int_0^t q(X_s) ds \right)$, which has a generator A_H given by $A_H \phi(s,x,z) = -\frac{\partial \phi}{\partial s} + A\phi + q(x) \frac{\partial \phi}{\partial z}$ for every $\phi \in C_c^2(\mathbb{R}^{n+2})$. Let $\phi(s,x,z) := e^{-z} w(s,x)$, $R > 0$ and $\alpha_R := \inf\{t > 0 \colon |H_t| \geq R\}$. Dynkin's formula (4.12) gives

$$
E^{s,x,z} \left[\phi\left(H_{t\wedge \alpha_R} \right) \right] = \phi(s,x,z) + E^{s,x,z} \left[\int_0^{t\wedge \alpha_R} A_H \phi(H_s)\, ds \right].
$$

This choice of ϕ implies

$$
A_H \phi = e^{-z} \left(-\frac{\partial w}{\partial t} + Aw - qw \right) = 0,
$$

so

$$
\begin{aligned}
w(s,x) = \phi(s,x,0) &= E^{s,x,0} \left[\phi\left(H_{t\wedge \alpha_R} \right) \right] \\
&= E^{s,x,0} \left[\exp\left(-\int_0^{t\wedge\alpha_R} q(X_r) dr \right) w\left(s - (t\wedge \alpha_R), X_{t\wedge\alpha_R} \right) \right] \\
&\to E^{s,x,0} \left[\exp\left(-\int_0^{t} q(X_r) dr \right) w\left(s - t, X_t \right) \right] \quad \text{as } R \to \infty,
\end{aligned}
$$

because $w(r,x)$ is bounded in $K \times \mathbb{R}^n$, and q is lower bounded. Choosing $t = s$, we are done. $\square$

5.2. Kolmogorov's backward equation

Corollary 5.2 (Kolmogorov's backward equation). *Let $f \in C_c^2(\mathbb{R}^n)$.*

(a) Define

$$
u(t,x) = E^x \left[f(X_t) \right]. \tag{5.4}
$$

Then $u(t,\cdot) \in \mathcal{D}_A$ $\forall t > 0$ and

$$
\frac{\partial u}{\partial t} = Au; \qquad \forall t > 0, x \in \mathbb{R}^n \tag{5.5}
$$

$$
u(0,x) = f(x); \qquad \forall x \in \mathbb{R}^n \tag{5.6}
$$

where Au denotes A applied to $u(t,\cdot)$.

(b) If $w \in C^{1,2}(\mathbb{R}_+ \times \mathbb{R}^n)$ is bounded on $K \times \mathbb{R}^n$ for every compact set $K \subseteq \mathbb{R}_+$ and w is a solution of (5.6) & (5.5), *then $w = u$.*

CHAPTER 6
Application to the combined Dirichlet-Poisson problem

In this chapter, we apply the preceding results to solve a stochastic version of the Dirichlet-Poisson problem.

6.1. Uniqueness in the combined stochastic Dirichlet-Poisson problem

Let $D \subseteq \mathbb{R}^n$ be a domain (open connected set) and let L denote the *semi-elliptic* partial differential operator on $C^2(\mathbb{R}^n)$ with the form

$$Lf = \sum_{i=1}^{n} b_i(x)\frac{\partial f}{\partial x_i}(x) + \sum_{i,j=1}^{n} a_{ij}(x)\frac{\partial^2 f}{\partial x_i \partial x_j}(x),$$

where $b_i, a_{ij} \in C(D)$ and $a_{ij} = a_{ji}$ for all i, j, and the symmetric matrix $a(x) = \left[a_{ij}\right]_{ij}$ has non-negative eigenvalues for all x.

The combined Dirichlet-Poisson problem: Let $\phi \in C(\partial D)$ and $g \in C(D)$ be two given functions. Find $w \in C^2(D)$ satisfying

$$Lw = -g \quad \text{in } D, \tag{6.1}$$

$$\lim_{D \ni x \to y} w(x) = \phi(y) \quad \forall y \in \partial D. \tag{6.2}$$

Due to the nature of our stochastic techniques, the second condition of the problem has to be translated. We will later define a stochastic version of this problem. The idea to solve these kind of problems is the following:

1. Find an Itô diffusion X whose generator A coincides with L on $C_c^2(\mathbb{R}^n)$.
   (i) To do this, we need to find a matrix $\sigma : \mathbb{R}^n \to \mathbb{R}^{n\times m}$ such that $(1/2)\sigma\sigma^T = a$, and such that b and σ satisfy the usual conditions for the existence and uniqueness of solutions of SDEs (Theorem 3.12) (for example, if each $a_{ij} \in C^2(D)$ and its first and second derivatives are bounded, then a square root σ of $2a$ exists [FR75], but this is outside of the reach of this project).
   (ii) Define $dX_t = b(X_t)dt + \sigma(X_t)dB_t$, where B is an m-dimensional Brownian motion.
2. The characteristic operator $\mathfrak{A}$ of X coincides with L on $C^2(D)$, and our candidate for the solution is

$$w(x) = E^x\left[\phi\left(X_{\tau_D}\mathbb{1}_{\{\tau_D<\infty\}}\right)\right] + E^x\left[\int_0^{\tau_D} g(X_t)\,dt\right],$$

   provided some conditions of regularity on g and ϕ are satisfied.

Definition 6.1 (The stochastic Dirichlet-Poisson problem). *Let $\phi \in C(\partial D)$ and $g \in C(D)$ be two given functions. Find $w \in C^2(D)$ satisfying*

$$\mathfrak{A}w = -g \quad \text{in } D, \tag{6.3}$$

$$\lim_{t\uparrow\tau_D} w(X_t) = \phi(X_{\tau_D})\mathbb{1}_{\{\tau_D<\infty\}} \quad Q^x\text{-a.e. } \forall x \in D. \tag{6.4}$$

Condition (6.4) is the stochastic boundary condition. It is a substitute for the boundary condition (6.2) in the stochastic setting. The need for this new formulation comes from the fact that the Itô diffusion X may not approach points in ∂D from arbitrary directions. Consider the 2-dimensional diffusion $X_t = X_0 + [B_t, 0]^T$, where B is 1-dimensional Brownian motion; paths of X will be contained in horizontal lines, so that the solutions of (6.3) & (6.4) will only be guaranteed to be continuous at the boundary in the horizontal direction. For a more extreme case, $X_t = X_0 + [t, 0]^T$ only approaches the boundary in the horizontal direction moving to the right.

Theorem 6.2 (Uniqueness of the solutions of the stochastic Dirichlet-Poisson problem). *Let $\phi \in C_b(\partial D)$ and $g \in C(D)$ satisfying*

$$E^x\left[\int_0^{\tau_D} |g(X_t)|\,dt\right] < \infty \quad \forall x \in D. \tag{6.5}$$

Suppose $w \in C_b^2(D)$ is a solution of the stochastic Dirichlet-Poisson problem (6.3) & (6.4). *Then*

$$w(x) = E^x\left[\phi\left(X_{\tau_D}\mathbb{1}_{\{\tau_D<\infty\}}\right)\right] + E^x\left[\int_0^{\tau_D} g(X_t)\,dt\right] \quad \forall x \in D. \tag{6.6}$$

Proof. Let $\{D_k\}_{k\in\mathbb{N}}$ be a sequence of open bounded sets increasing to D. Let $\alpha_k := k \wedge \tau_{D_k}$. By (6.3) and the Dynkin formula (4.12),

$$w(x) = E^x\left[w\left(X_{\alpha_k}\right)\right] + E^x\left[\int_0^{\alpha_k} g(X_t)\,dt\right].$$

By (6.4), $w\left(X_{\alpha_k}\right) \to \phi\left(X_{\tau_D}\right)\mathbb{1}_{\{\tau_D<\infty\}}$ a.e. boundedly, so

$$E^x\left[w\left(X_{\alpha_k}\right)\right] \xrightarrow{k\to\infty} E^x\left[\phi\left(X_{\tau_D}\right)\mathbb{1}_{\{\tau_D<\infty\}}\right].$$

On the other hand, $\int_0^{\alpha_k} g(X_t)\,dt \to \int_0^{\tau_D} g(X_t)\,dt$ and $\left|\int_0^{\alpha_k} g(X_t)\,dt\right| \le \int_0^{\tau_D} |g(X_t)|\,dt$, which is integrable by (6.5). By the dominated convergence theorem,

$$E^x\left[\int_0^{\alpha_k} g(X_t)\,dt\right] \xrightarrow{k\to\infty} E^x\left[\int_0^{\tau_D} g(X_t)\,dt\right]. \qquad \square$$

To find a solution of the stochastic Dirichlet-Poisson problem, we will solve the Dirichlet ($g = 0$) and the Poisson ($\phi = 0$) problems separately. The solution to the general problem will be the sum of the solutions to these two problems.

6.2. The Dirichlet problem

The classic Dirichlet problem is the particular case of (6.1) where $g = 0$. For example, if the diffusion is Brownian motion, the differential equation becomes the Laplace equation $\Delta w = 0$ en $\mathbb{R}^n$.

For simplicity, in this section we assume that $\tau_D < \infty$ a.e.

6.2.1. Harmonic functions

Definition 6.3. $f \in L^\infty_{loc}(D)$[1] *is called* X-harmonic in D *if*

$$f(x) = E^x\left[f\left(X_{\tau_U}\right)\right] \quad \textit{for all } x \in D \textit{ and all bounded open sets } U \ni x \textit{ such that } \overline{U} \subseteq D.$$

Lemma 6.4. *(a) f is X-harmonic in $D \implies \mathfrak{A}f = 0$ in D.*

(b) $\left[\left(f \in C^2(D)\right), \left(\mathfrak{A}f = 0 \text{ in } D\right)\right] \implies f$ is X-harmonic in D.

Proof. (a) Direct from the definition (4.15) of $\mathfrak{A}$. (b) Let $U \ni x$ be a bounded open set as in Definition 6.3. We want to apply Dynkin's formula, so consider the finite stopping times $\tau_U \wedge k$ for each $k \in \mathbb{N}$. $f(X_{\tau_U\wedge k}) \to f(X_{\tau_U})$ a.e. and $\left|f(X_{\tau_U\wedge k})\right| \le \sup f|_{\overline{U}} < \infty$ by compactness of $\overline{U}$ and continuity of f. We can apply the bounded convergence theorem:

$$E^x\left[f\left(X_{\tau_U}\right)\right] = \lim_{k\to\infty} E^x\left[f\left(X_{\tau_U\wedge k}\right)\right] = f(x) + \lim_{k\to\infty} E^x\left[\int_0^{\tau_U\wedge k} \mathfrak{A}f(X_s)\,ds\right] = f(x) + 0,$$

where we have used Dynkin's formula (Theorem 4.12) in the second equality. $\square$

As showcased in Example 4.13, the Laplace operator in $\mathbb{R}^n$ is the generator of n-dimensional Brownian motion, except for a constant. Basic theory of classic harmonic functions, along with the results of Lemma 6.4, imply that harmonic functions in D are exactly those which are B-harmonic in D. The harmonic measure induced on the boundary of a ball by a Brownian motion starting at its center is the normalized Lebesgue measure on the sphere. See Figure 6.1.

The following Lemma presents the most important example of X-harmonic functions.

[1] $L^\infty_{\text{loc}}(D)$ is the set of locally bounded measurable functions $D \to \mathbb{R}$.

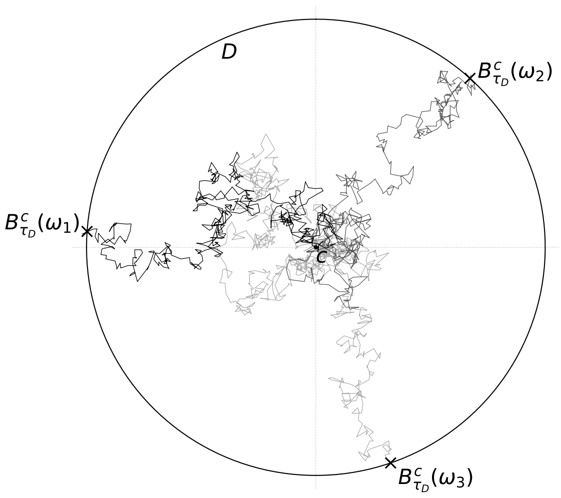


Figure 6.1: Three paths of a 2-dimensional Brownian motion exiting a disk $D \subseteq \mathbb{R}^2$ from the center c.

Lemma 6.5. *Let $\phi \in L^\infty(\partial D)$ and*

$$u(x) := E^x\left[\phi\left(X_{\tau_D}\right)\right], \quad \forall x \in D.$$

Then u is X-harmonic in D.

Proof. If $\overline{V} \subseteq D$ and $x \in V$, by the mean value property (4.10),

$$u(x) = E^x\left[\phi\left(X_{\tau_D}\right)\right] = E^x\left[E^{X_{\tau_V}}\left[\phi\left(X_{\tau_D}\right)\right]\right] = E^x\left[u\left(X_{\tau_V}\right)\right]. \qquad \square$$

6.2.2. The stochastic Dirichlet problem

Definition 6.6 (The stochastic Dirichlet problem). *Let $\phi \in L^\infty(\partial D)$. Find $u : D \to \mathbb{R}$ satisfying*

$$u \text{ is } X\text{-harmonic in } D, \tag{6.7}$$

$$\lim_{t \uparrow \tau_D} w(X_t) = \phi(X_{\tau_D}) \qquad Q^x\text{-a.e. } \forall x \in D. \tag{6.8}$$

By Lemma 6.4, (6.7) is equivalent to $\mathfrak{A}u = 0$ in D if $u \in C^2(D)$.

Theorem 6.7 (Solution of the stochastic Dirichlet problem). *Let $\phi \in L^\infty(\partial D)$. Then the function*

$$u(x) = E^x\left[\phi\left(X_{\tau_D}\right)\right], \quad \forall x \in D$$

is the unique bounded function on D that satisfies (6.7) and (6.8).

Proof. *(Existence)* By Lemma 6.5 u satisfies (6.7). Let $\{D_k\}_{k\in\mathbb{N}}$ be a sequence of open bounded sets increasing to D. Let $\alpha_k := k \wedge \tau_{D_k}$. By the strong Markov property (4.5) and (4.9),

$$u(X_{\alpha_k}) = E^{X_{\alpha_k}}\left[\phi\left(X_{\tau_D}\right)\right] = E^x\left[\theta_{\alpha_k}\phi\left(X_{\tau_D}\right) \,\middle|\, \mathscr{F}_{\alpha_k}\right] = E^x\left[\phi\left(X_{\tau_D}\right) \,\middle|\, \mathscr{F}_{\alpha_k}\right] =: M_k. \tag{6.9}$$

$(M_k)_{k\in\mathbb{N}}$ is a bounded (discrete time) martingale. By the martingale convergence theorem (Theorem 2.35),

$$u\left(X_{\alpha_k}\right) \to u\left(S_{\tau_D}\right) \text{ a.e. and in } L^1. \tag{6.10}$$

By (6.9), $N_t := u\left(X_{\alpha_k \vee (t\wedge\alpha_{k+1})}\right) - u\left(X_{\alpha_k}\right)$ is a martingale with respect to the filtration $\mathscr{G}_t = \mathscr{F}_{\alpha_k \vee (t\wedge\alpha_{k+1})}$. By Doob's martingale inequality (Theorem 2.27),

$$Q^x\left[\sup_{\alpha_k \le r \le \alpha_{k+1}} \left|u(X_r) - u\left(X_{\alpha_k}\right)\right| > \varepsilon\right] \le \frac{1}{\varepsilon^2} E^x\left[\left|u\left(X_{\alpha_{k+1}}\right) - u\left(X_{\alpha_k}\right)\right|^2\right] \to 0, \text{ as } k \to \infty \text{ for all } \varepsilon > 0.$$

Putting this together with (6.10), we conclude that u satisfies (6.8).

(Uniqueness) Let g be another solution. Since it is X-harmonic, $g(x) = E^x\left[g\left(X_{\alpha_k}\right)\right]$. By (6.8) and bounded convergence, $g(x) = \lim_k E^x\left[g\left(X_{\alpha_k}\right)\right] = E^x\left[\phi\left(X_{\tau_D}\right)\right]$. $\square$

6.3. The stochastic Poisson problem

The classic Poisson problem is the particular case of (6.1) where $\phi = 0$.

Definition 6.8 (The stochastic Poisson problem). *Let $g \in C(D)$. Find $v \in C^2(D)$ satisfying*

$$\mathfrak{A} v = -g \qquad \text{in } D, \tag{6.11}$$

$$\lim_{t \uparrow \tau_D} v(X_t) = 0 \qquad Q^x\text{-a.e. } \forall x \in D. \tag{6.12}$$

Theorem 6.9 (Existence theorem for the stochastic Poisson problem). *Let $g \in C(D)$ satisfying (6.5) (this condition is met if, for example, g is bounded and $E^x[\tau_D] < \infty$ for every $x \in D$). Then*

$$v(x) := E^x\left[\int_0^{\tau_D} g(X_s)\, ds\right] \quad \forall x \in D$$

satisfies (6.11) and (6.12).

Proof. Let $U \ni x$ be open, bounded and such that $\overline{U} \subseteq D$. Define $\eta = \int_0^{\tau_D} g(X_s)\, ds$ and $\tau := \tau_U$. By the strong Markov property (4.5),

$$\frac{E^x[v(X_\tau)] - v(x)}{E^x[\tau]} = \frac{E^x[E^{X_\tau}[\eta]] - E^x[\eta]}{E^x[\tau]} = \frac{E^x[E^x[\theta_\tau \eta \mid \mathscr{F}_\tau]] - E^x[\eta]}{E^x[\tau]} = \frac{E^x[\theta_\tau \eta - \eta]}{E^x[\tau]}.$$

Approximate η by sums of the form $\eta_k = \sum_j g\left(X_{t_j}\right) \mathbb{1}_{\{t_j < \tau_D\}} \Delta t_j$, where $\Delta t_j = t_{j+1} - t_j \xrightarrow{k \to \infty} 0$. By a derivation almost identical to (4.8),

$$\theta_t \eta_k = \sum_j g\left(X_{t+t_j}\right) \mathbb{1}_{\{t+t_j < \tau_D^t\}} \Delta t_j$$

for all $k \in \mathbb{N}$; which implies (making $k \to \infty$) that $\theta_\tau \eta = \int_\tau^{\tau_D} g(X_s)\, ds$ (recall that $\tau_D^\tau = \tau_D$ because $U \subseteq D$). Therefore,

$$\frac{E^x[v(X_\tau)] - v(x)}{E^x[\tau]} = \frac{-1}{E^x[\tau]} E^x\left[\int_0^\tau g(X_s)\, ds\right] \to -g(x) \qquad \text{as } U \downarrow x,$$

by continuity of g. This proves (6.11).

Let $\{D_k\}_{k \in \mathbb{N}}$ be a sequence of open bounded sets increasing to D. Let $H(x) := E^x\left[\int_0^{\tau_D} \left|g(X_s)\right| ds\right]$ for $x \in D$ and $\alpha_k := k \wedge \tau_{D_k}$. Repeating the same argument as above,

$$E^x\left[H\left(X_{\alpha_k \wedge t}\right)\right] = E^x\left[\int_{\alpha_k \wedge t}^{\tau_D} \left|g(X_s)\right| ds\right] \to 0 \qquad \text{as } k \to \infty, t \to \tau_D,$$

by dominated convergence; which implies (6.12). □

The solution of the combined stochastic Dirichlet-Poisson problem follows from Theorems 6.7 and 6.9.

Theorem 6.10 (Existence of a solution to the stochastic Dirichlet-Poisson problem). *Let $\phi \in C_b(\partial D)$, $g \in C(D)$ satisfying (6.5) and $\tau_D < \infty$ a.e. Then*

$$w(x) = E^x\left[\phi\left(X_{\tau_D}\right)\right] + E^x\left[\int_0^{\tau_D} g(X_t)\, dt\right] \quad \forall x \in D$$

solves the stochastic Dirichlet-Poisson problem: (6.3) and (6.4).

Bibliography

[AB06] C.D. Aliprantis and K.C. Border. *Infinite Dimensional Analysis: A Hitchhiker's Guide*. Springer Berlin, Heidelberg, 2006. DOI: 10.1007/3-540-29587-9.

[AD00] Robert B. Ash and Catherine A. Doleans-Dade. *Probability and Measure Theory*. Second edition. Harcourt Academic Press, 2000.

[Dur10] Rick Durrett. *Probability: Theory and Examples, 4th Edition*. Cambridge University Press, 2010. DOI: 10.1017/CBO9780511779398.

[Dyn65] E. B. Dynkin. *Markov Processes. Volume I*. Springer Berlin, Heidelberg, 1965. DOI: 10.1007/978-3-662-00031-1.

[Fol13] G.B. Folland. *Real Analysis: Modern Techniques and Their Applications*. Pure and Applied Mathematics: A Wiley Series of Texts, Monographs and Tracts. Wiley, 2013.

[FR75] Wendell Fleming and Raymond Rishel. *Deterministic and Stochastic Optimal Control*. Springer New York, NY, 1975. DOI: 10.1007/978-1-4612-6380-7.

[Øks13] Bernt K. Øksendal. *Stochastic Differential Equations*. Sixth edition. Universitext. Springer Berlin, Heidelberg, 2013. DOI: 10.1007/978-3-642-14394-6.

[Ram+24] Carlos Ramos-Carreño et al. "scikit-fda: A Python Package for Functional Data Analysis". In: *Journal of Statistical Software* 109 (May 2024). DOI: 10.18637/jss.v109.i02. URL: https://github.com/GAA-UAM/scikit-fda (visited on 05/25/2024).

[RY05] Daniel Revuz and Marc Yor. *Continuous martingales and Brownian motion*. Third edition. Grundlehren der mathematischen Wissenschaften. Springer Berlin, Heidelberg, 2005. DOI: 10.1007/978-3-662-06400-9.

[Son+20] Yang Song et al. "Score-Based Generative Modeling through Stochastic Differential Equations". In: *CoRR* abs/2011.13456 (2020). arXiv: 2011.13456. URL: https://arxiv.org/abs/2011.13456.

[Sot24] Pablo Soto-Martín. 2024. URL: https://github.com/psotom/scikit-fda (visited on 05/25/2024).

[SV08] Sandro Salsa and Gianmaria Verzini. *Partial Differential Equations in Action: From Modelling to Theory*. UNITEXT. Springer Cham, 2008. DOI: 10.1007/978-3-031-21853-8.

Appendices

APPENDIX A
Basic results of measure theory

A *monotone class* is a family $\mathscr{M}$ of sets that is closed under countable monotone unions and under countable monotone intersections.

Theorem A.1. *Let $\mathscr{A}$ be an algebra of subsets of Γ and $\mathscr{M}$ a monotone class of subsets of Γ. If $\mathscr{A} \subseteq \mathscr{M}$, then $\sigma(\mathscr{A}) \subseteq \mathscr{M}$.*

Proof in [AD00, Theorem 1.3.9].

We say that $\mathscr{P} \subseteq 2^{\Gamma}$ is a *π-system* if it is closed under intersection; and $\mathscr{L} \subseteq 2^{\Gamma}$ is a *λ-system* if

(a) $\Gamma \in \mathscr{L}$,
(b) $[A, B \in \mathscr{L},\ A \subseteq B] \implies B \setminus A \in \mathscr{L}$, and
(c) $\mathscr{L}$ is closed under countable monotone unions.

Theorem A.2 ($\pi - \lambda$-theorem). *If $\mathscr{P}$ is a π-system, $\mathscr{L}$ is a λ-system and $\mathscr{P} \subseteq \mathscr{L} \subseteq 2^{\Gamma}$, then $\sigma(\mathscr{P}) \subseteq \mathscr{L}$.*

Proof in [Dur10, Theorem A.1.4].

Let (Γ, Σ) be a measurable space. For every $\mathscr{H} \subseteq \Gamma^{\mathbb{R}}$, we will denote by $\mathscr{H}_+$ the subset of all non-negative functions in $\mathscr{H}$, and by $\mathscr{H}_-$ its subset of all non-positive functions. For example, $L^{\infty}_{+}(\Gamma, \Sigma)$ is the set of all non-negative, bounded Σ-measurable functions.

Theorem A.3 (Functional monotone class theorem). *Let $\mathscr{P}$ be a π-system on Γ (i.e. $\mathscr{P} \subseteq 2^{\Gamma}$ is closed under finite intersections) and let $\mathscr{H} \subseteq \Gamma^{\mathbb{R}}$ be a vector space of bounded functions such that*

(i) $\{\mathbb{1}_P : P \in \mathscr{P}\} \subseteq \mathscr{H}$,
(ii) $\mathbb{1}_{\Gamma} \in \mathscr{H}$, and
(iii) $\mathscr{H}_+$ is closed under bounded increasing pointwise limits: if $(f_n)_{n\in\mathbb{N}} \subseteq \mathscr{H}_+$ such that $f_n \uparrow f$ pointwise and f is bounded, then $f \in \mathscr{H}_+$.

Then $L^{\infty}(\sigma(\mathscr{P})) \subseteq \mathscr{H}$.

Proof. Let $\mathscr{L} = \{A \subseteq \Gamma : \mathbb{1}_A \in \mathscr{H}\}$. By assumption, $\mathscr{P} \subseteq \mathscr{L}$. Moreover, $\mathscr{L}$ is a λ-system:

(a) $\Gamma \in \mathscr{L}$, by assumption.
(b) If $A, B \in \mathscr{L}$ and $A \subseteq B$, then $B \setminus A \in \mathscr{L}$, because $\mathscr{H}$ is a vector space and $\mathbb{1}_{B\setminus A} = \mathbb{1}_B - \mathbb{1}_A$.
(c) $\mathscr{L}$ is closed under increasing unions, because $\mathscr{H}_+$ is closed under bounded increasing limits.

By the π-λ theorem (Theorem A.2), $\sigma(\mathscr{P}) \subseteq \mathscr{L}$. As $\mathscr{H}$ is a vector space, this implies that every $\sigma(\mathscr{P})$-simple function is in $\mathscr{H}$. Every function in $L^{\infty}_{+}(\sigma(\mathscr{P}))$ is the increasing pointwise limit of $\sigma(\mathscr{P})$-simple functions (see [Fol13, Theorem (2.10)]), so $L^{\infty}_{+}(\sigma(\mathscr{P})) \subseteq \mathscr{H}$. Finally, every $f \in L^{\infty}(\sigma(\mathscr{P}))$ is the difference of two non-negative bounded measurable functions $f^+ := f \vee 0$ and $f^- := (-f) \vee 0$; therefore, $L^{\infty}(\sigma(\mathscr{P})) \subseteq \mathscr{H}$. □

Theorem A.4. *Let I be an arbitrary index set and $\{P_i\}_{i\in I}$ a collection of π-systems on Γ. If $\{P_i\}_{i\in I}$ are independent, then $\{\sigma(P_i)\}_{i\in I}$ are independent.*

Proof in [Dur10, Theorem 2.1.7].

Proposition A.5 (Layer cake representation). *Let (A, Γ, ν) be a measure space and $dm(t) = dt$ be the Lebesgue measure on $\mathscr{B}([0, \infty))$. If $0 < p < \infty$ and f is a non-negative Γ-measurable function, then*

$$\int_A f^p \, d\nu = p \int_0^{\infty} t^{p-1} \nu\left(f > t\right) \, dt.$$

Proof. For each $a \in A$, $f(a) = m([0, f(a))) = \int_0^{\infty} \mathbb{1}_{\{t<f\}}(a) dt$. Applying the Fubini-Tonelli theorem for non-negative functions, we get

$$\int_A f \, d\nu = \int_A \int_0^{\infty} \mathbb{1}_{\{t<f\}}(a) \, dt \, d\nu(a) = \int_0^{\infty} \int_A \mathbb{1}_{\{t<f\}} \, d\nu \, dt = \int_0^{\infty} \nu\left(f > t\right) \, dt.$$

For the case $p \neq 1$, we have

$$\int_A f^p \, d\nu = \int_0^\infty \nu\left(f^p > t\right) \, dt = \int_0^\infty \nu\left(f^p > s^p\right) \, p s^{p-1} \, ds = p \int_0^\infty s^{p-1} \nu\left(f > s\right) \, ds,$$

where we have applied the change of variable theorem for $t = s^p$. □

Theorem A.6 (Conditional Jensen's inequality). *Let $\varphi : \mathbb{R} \to \mathbb{R}$ be a convex function and let $f \in L^1(\Omega, \Sigma, P)$. If Γ is a sub-σ-algebra of Σ, then $\varphi\left(E\left[f \mid \Gamma\right]\right) \leq E\left[\varphi\left(f\right) \mid \Gamma\right]$ a.e.*

For a proof of this theorem, see [AD00, Theorem (9.3.5)]. A direct consequence is the following:

Corollary A.7. *Let $T \subseteq \mathbb{R}_+$.*

(a) Let $\varphi : \mathbb{R} \to \mathbb{R}$ be a convex, increasing function and $(X_t)_{t\in T}$ an $(\mathscr{F}_t)_{t\in T}$-submartingale such that $\left(\varphi(X_t)\right)_{t\in T} \subseteq L^1(P)$. Then $\left(\varphi(X_t)\right)_{t\in T}$ is an $(\mathscr{F}_t)_{t\in T}$-submartingale.

(b) Let $\varphi : \mathbb{R} \to \mathbb{R}$ be a convex function and $(X_t)_{t\in T}$ an $(\mathscr{F}_t)_{t\in T}$-martingale such that $\left(\varphi(X_t)\right)_{t\in T} \subseteq L^1(P)$. Then $\left(\varphi(X_t)\right)_{t\in T}$ is an $(\mathscr{F}_t)_{t\in T}$-submartingale.

APPENDIX B
Product of measurable spaces

Let I be a non-empty set, let $\mathscr{P}_{\text{fin}}(I)$ be the collection of finite subsets of I and $\{(\Omega_i, \Sigma_i)\}_{i\in I}$ a family of measurable spaces. For each $H \subseteq I$, define

$$\Omega_H := \prod_{i\in H} \Omega_i \quad \text{and} \quad \Omega_{-H} := \prod_{i\in I\setminus H} \Omega_i.$$

For each $G \subsetneq H \subseteq I$, denote by π_{HG} *the natural projection of* Ω_H *onto* Ω_G: $\pi_{HG}((\omega_i)_{i\in H}) := (\omega_i)_{i\in G}$. To ease the notation, we will write $\pi_H := \pi_{IH}$, $\pi_i := \pi_{\{i\}}$ and $\Omega_{-i} := \Omega_{-\{i\}}$.

We define the product σ-algebra $\Sigma_I = \bigotimes_{i\in I} \Sigma_i$ as the smallest σ-algebra on Ω_I such that π_i is measurable for all $i \in I$. The following definitions are equivalent:

$$\begin{aligned}
\Sigma_I &= \sigma(\pi_i : i \in I) \\
&= \sigma\left(\pi_i^{-1}(A) : i \in I, A \in \Sigma_i\right) \\
&= \sigma\left(\pi_F^{-1}\left(\prod_{i\in F} A_i\right) : F \in \mathscr{P}_{\text{fin}}(I), A_i \in \Sigma_i\right) \\
&= \sigma\left(\pi_F^{-1}(A) : F \in \mathscr{P}_{\text{fin}}(I), A \in \Sigma_F\right) \qquad (\star) \\
&= \sigma(\pi_F : F \in \mathscr{P}_{\text{fin}}(I)).
\end{aligned}$$

Whenever all of the (Ω_i, Σ_i) are the same measure space (Ω, Σ), we will write $\left(\Omega^I, \Sigma^I\right)$ instead of (Ω_I, Σ_I). If $F \in \mathscr{P}_{\text{fin}}(I)$ and $A \in \Sigma_F$, we call $\pi_F^{-1}(A) = A \times \Omega_{-F}$ the *F-cylinder of base A*.

Definition B.1 (Kolmogorov consistent family of finite dimensional probability measures). *If for each $F \in \mathscr{P}_{fin}(I)$, P_F is a probability measure on (Ω_F, Σ_F), then we say that the family $(P_F)_{F\in\mathscr{P}_{fin}(I)}$ (of* finite dimensional probability measures on (Ω_I, Σ_I)*) is* Kolmogorov consistent *if*

$$\forall F, G \in \mathscr{P}_{fin}(I) \quad F \subsetneq G \implies P_G \circ \pi_{GF}^{-1} = P_F\,;$$

that is, for each $A \in \Sigma_F$, $P_G\left(\pi_{GF}^{-1}(A)\right) = P_G(A \times \Omega_{G\setminus F}) = P_F(A)$. [1]

Definition B.2 (Kolmogorov extension). *We say that a Kolmogorov consistent family $(P_F)_{F\in\mathscr{P}_{fin}(I)}$ of finite dimensional probabilities has a* Kolmogorov extension P *if there exists a probability measure P on (Ω_I, Σ_I) such that $P_F = P \circ \pi_F^{-1}$ for all $F \in \mathscr{P}_{fin}(I)$.*

Theorem B.3. *Let $\mathscr{A}$ be an algebra, and $\mu : \mathscr{A} \to [0, +\infty]$ be finitely additive. Then:*

(a) If μ is continuous from below at each $A \in \mathscr{A}$ (i.e. if $(A_n)_{n=1}^{\infty} \subseteq \mathscr{A}$ and $A_n \uparrow A \in \mathscr{A}$, then $\mu(A_n) \to \mu(A)$), then μ is σ-additive.

(b) If μ is continuous from above at $\emptyset$ (i.e. if $(A_n)_{n=1}^{\infty} \subseteq \mathscr{A}$ and $A_n \downarrow \emptyset$, then $\mu(A_n) \to 0$), then μ is σ-additive.

Proof. (a) Let $(B_k)_{k\in\mathbb{N}} \in \mathscr{A}$ be pairwise disjoint and such that $\cup_{k\in\mathbb{N}} B_k =: A \in \mathscr{A}$. Define $A_n = \cup_{k=1}^{n} B_k$, then $A_n \uparrow A$, and $\mu(A_n) \to \mu(A)$ by hypothesis. Moreover, $\mu(A_n) = \sum_{k=1}^{n} \mu(B_k)$ for all $n \in \mathbb{N}$, and so $\mu(A) = \sum_{k=1}^{\infty} \mu(B_k)$.

(b) Let $(B_k)_{k\in\mathbb{N}} \in \mathscr{A}$ be pairwise disjoint and such that $\cup_{k\in\mathbb{N}} B_k =: A \in \mathscr{A}$. Define $A_n = A \setminus \cup_{k=1}^{n} B_k$, then $A_n \downarrow \emptyset$ and $\mu(A) - \sum_{k=1}^{n} \mu(B_k) = \mu(A_n) \to 0$ by hypothesis. □

Theorem B.4 (Carathéodory Extension Theorem). *Let μ be a (non-negative) σ-finite measure on an algebra $\mathscr{A}$. Then μ has a unique extension to a measure on $\sigma(\mathscr{A})$.*

A proof of the Carathéodory Extension Theorem can be found in [AD00, Theorem 1.3.10].

Theorem B.5. *If μ is a σ-finite measure on $(\mathbb{R}^n, \mathscr{B}(\mathbb{R}^n))$, then for each $B \in \mathscr{B}(\mathbb{R}^n)$,*

$$\mu(B) = \sup\left\{\mu(K) : K \subseteq B,\ K \text{ compact}\right\}.$$

[1] By the definition ($\star$) of the product σ-algebra Σ_G, π_{GF} is measurable.

Idea of the proof. We assume first that μ is finite. The family of Borel-measurable sets satisfying this property is a monotone class that contains all compact sets; by the monotone class theorem, it contains all Borel-measurable sets. If μ is σ-finite, take a countable partition of $\mathbb{R}^n$ into sets of finite measure, and apply the previous result to each of them to obtain a sequence of compact sets whose measures converge to the measure of the set. Here we use that the finite union of compact sets is compact. See the details in [AD00, Theorem 1.4.11]. $\square$

Theorem B.6 (Kolmogorov extension theorem). *Every Kolmogorov consistent family of finite dimensional probability measures on $\left(\mathbb{R}^I, \mathscr{B}(\mathbb{R})^I\right)$ has a unique Kolmogorov extension.*

See [AB06, Theorem 15.26] for a generalization of this theorem to arbitrary measurable spaces (not necessarily $\mathbb{R}$).

Proof. To ease the notation, we denote $\mathscr{B} = \mathscr{B}(\mathbb{R})$. Let P_F be the probability measure on $\left(\mathbb{R}^F, \mathscr{B}^F\right)$ for each $F \in \mathscr{P}_{\text{fin}}(I)$; such that that the family $(P_F)_{F\in\mathscr{P}_{\text{fin}}(I)}$ is Kolmogorov consistent. We define the hoped-for probability measure P on the measurable cylinders of $\mathbb{R}^I$:

$$F \in \mathscr{P}_{\text{fin}}(I),\ A \in \mathscr{B}^F \implies P\left(A \times \mathbb{R}^{I\setminus F}\right) = P_F(A).$$

P is well defined on measurable cylinders because of the Kolmogorov consistency of the family $(P_F)_{F\in\mathscr{P}_{\text{fin}}(I)}$: if $C \in \mathscr{B}^I$ is an F-cylinder and a G-cylinder, for $F, G \in \mathscr{P}_{\text{fin}}(I)$, then WLOG $F \subseteq G$, and

$$C = A \times \mathbb{R}^{I\setminus F} = \left(A \times \mathbb{R}^{G\setminus F}\right) \times \mathbb{R}^{I\setminus G}, \text{ for some } A \in \mathscr{B}^F \text{ and } P(C) = P_F(A) = P_G\left(A \times \mathbb{R}^{G\setminus F}\right).$$

Moreover, P is finitely additive on the algebra $\mathscr{B}_0^I$ of measurable cylinders of $\mathbb{R}^I$: let $C_1, \ldots, C_n \in \mathscr{B}_0^I$ be pairwise disjoint, then there exists an $F \in \mathscr{P}_{\text{fin}}(I)$ and pairwise disjoint $A_1, \ldots, A_n \in \mathscr{B}^F$ such that $C_i = A_i \times \mathbb{R}^{I\setminus F}$ for all $i \in \{1, \ldots, n\}$. Then

$$P\left(\bigcup_{i=1}^n C_i\right) = P\left(\bigcup_{i=1}^n A_i \times \mathbb{R}^{I\setminus F}\right) = P_F\left(\bigcup_{i=1}^n A_i\right) = \sum_{i=1}^n P_F(A_i) = \sum_{i=1}^n P(C_i).$$

If we show that P is continuous from above at $\emptyset$, then by Theorem *B.3* it is σ-additive, which implies that P is a finite measure on $\mathscr{B}_0^I$ and by the Carathéodory Extension Theorem, it has a unique extension to a measure on $\sigma\left(\mathscr{B}_0^I\right) = \mathscr{B}^I$.

Let $\{C_j\}_{j\in\mathbb{N}} \subseteq \mathscr{B}_0^I$ decrease to $\emptyset$. Suppose that $P\left(C_j\right)$ does not decrease to 0, then there exists $\varepsilon > 0$ such that $P\left(C_j\right) \geq \varepsilon > 0$ for all $j \in \mathbb{N}$. Suppose that $C_j = B_j \times \mathbb{R}^{I\setminus F_j}$ for some $F_j \in \mathscr{P}_{\text{fin}}(I)$ and $B_j \in \mathscr{B}^{F_j}$. WLOG, the sequence $\left(F_j\right)_{j\in\mathbb{N}}$ is increasing.

We now approximate the cylinders by cylinders of compact bases. By Theorem *B.5*, there exists a compact set $K_j \in B_j$ such that $P_{F_k}(B_j \setminus K_j) \leq \varepsilon/2^{j+1}$. Define

$$C_j' := K_j \times \mathbb{R}^{I\setminus F_j} \subseteq C_j, \quad D_j := \left(\bigcap_{i=1}^j C_i'\right) \subseteq \left(\bigcap_{i=1}^j C_i\right) = C_j;$$

then $P\left(C_j \setminus C_j'\right) = P_{F_k}(B_j \setminus K_j) \leq \varepsilon/2^{j+1}$, and

$$P\left(C_j \setminus D_j\right) = P\left(C_j \cap \left(\bigcup_{i=1}^j \left(C_j'\right)^c\right)\right) \leq \sum_{i=1}^j P\left(C_j \cap \setminus C_i'\right) \leq \sum_{i=1}^j P\left(C_i \cap \setminus C_i'\right) \leq \sum_{i=1}^j \frac{\varepsilon}{2^{i+1}} \leq \frac{\varepsilon}{2}.$$

Since $D_j \subseteq C_j$, the previous inequality implies that $P\left(D_j\right) \geq P\left(C_j\right) - \varepsilon/2 > 0$. In particular, $D_j \neq \emptyset$, so we can pick $x^j \in D_j \subseteq C_j' = K_j \times \mathbb{R}^{F_j}$.

The sequence $\left(\pi_{F_1}\left(x^j\right)\right)_{j\in\mathbb{N}}$ is contained in the compact set $K_1 \in \mathbb{R}^{F_1}$. Thus, there is a subsequence $\left(x^{j_{1n}}\right)_n$ such that $\pi_{F_1}\left(x^{j_{1n}}\right) \to x_{F_1} \in K_1$. A tail of $\left(\pi_{F_2}\left(x^{j_{1n}}\right)\right)_n$ is contained in the compact set $K_2 \subseteq \mathbb{R}^{F_2}$, which implies that there is a subsequence $\left(j_{2n}\right)_n$ of $\left(j_{1n}\right)_n$ such that $\pi_{F_2}\left(x^{j_{2n}}\right) \to x_{F_2} \in K_2$; moreover, $\pi_{F_1}\left(x_{F_2}\right) = x_{F_1}$ by uniqueness of the limit in $\mathbb{R}^{F_1}$ and continuity of the projection. We can continue this process to obtain a sequence $\left(j_{in}\right)_n$ such that $\pi_{F_i}\left(x^{j_{in}}\right) \to x_{F_i} \in K_i$ for all $i \in \mathbb{N}$, and

$$i_1 < i_2 \implies \pi_{F_{i_1}}\left(x_{F_{i_2}}\right) = x_{F_1}. \tag{B.1}$$

Then, we select an $\omega \in \mathbb{R}^I$ such that

$$\pi_{F_i}(\omega) = x_{F_i}, \tag{B.2}$$

which is well defined by (B.1). By (B.2), $\pi_{F_i}(\omega) \in K_i$, which implies that $\omega \in C_i' \subseteq C_i$ for every $i \in \mathbb{N}$. Therefore, $\omega \in \bigcap_{i\in\mathbb{N}} C_i = \emptyset$, which is a contradiction.

We conclude that $\left[\left(C_j\right)_{\mathbb{N}} \subseteq \mathscr{B}_0^I\right] \wedge \left[C_j \downarrow \emptyset\right] \wedge \left[P\left(C_j\right) \not\to 0\right]$ is false, and so P is continuous from above at $\emptyset$. We are done, □

Theorem B.7. *Given a probability measure μ on $(\mathbb{R}, \mathscr{B}(\mathbb{R}))$, there exists a probability space (Ω, Σ, P) and a sequence of independent random variables $(X_n : \Omega \to \mathbb{R})_{n\in\mathbb{N}}$, such that $P_{X_n} = \mu$ for all $n \in \mathbb{N}$.*

Proof. Choose $(\Omega, \Sigma) = \left(\mathbb{R}^{\mathbb{N}}, \mathscr{B}(\mathbb{R})^{\mathbb{N}}\right)$ and let the probability measure P on $\mathscr{B}(\mathbb{R})^{\mathbb{N}}$ be the product measure of μ on each coordinate. The existence of such a P is guaranteed by Theorem *B*.6, the Kolmogorov consistent family of finite dimensional distributions being the product of finite $(\mathbb{R}, \mathscr{B}(\mathbb{R}), \mu)$'s. Define $X_n = \pi_n$. Then $P_{X_n} = \mu$ for all $n \in \mathbb{N}$, and $(X_n)_{n\in\mathbb{N}}$ are independent by definition. □

APPENDIX C

Approximate measurable functions by sums of simple tensors

For many of the results that will be proved in this document, we will need to approximate measurable functions in finite product spaces such as $\Omega_1 \times \Omega_2$ by taking limits of functions of the form $\sum_{j=1}^{m_n} \phi_j(\omega_1)\psi_j(\omega_2)$.

Let $(\Gamma, \Sigma, \mu) = \left(\prod_{i=1}^n \Gamma_i, \otimes_{i=1}^n \Sigma_i, \mu\right)$ be a (finite) product of measurable spaces equipped with a measure μ. We denote by $\mathscr{R}$ the set of measurable rectangles of Γ:

$$\mathscr{R} = \left\{ \prod_{i=1}^n A_i : \forall 1 \le i \le n \; A_i \in \Sigma_i \right\}.$$

Then, the smallest algebra containing $\mathscr{R}$, which will be denoted by $a(\mathscr{R})$, consists of all finite unions of measurable rectangles, which we can take to be disjoint:

$$a(\mathscr{R}) = \left\{ \bigsqcup_{j=1}^m R_j : m \in \mathbb{N}, \; \{R_j\}_{j=1}^m \subseteq \mathscr{R} \right\}.$$

Proposition C.1. *Let $(\Gamma, \Sigma, \mu) = \left(\prod_{i=1}^n \Gamma_i, \otimes_{i=1}^n \Sigma_i, \mu\right)$ be a product of measurable spaces equipped with a finite measure. Then, for every $Y \in \Sigma$ and $\varepsilon > 0$, there exists $Z \in a(\mathscr{R})$ such that $\mu(Y \triangle Z) < \varepsilon$.*

Proof. Let $\mathscr{M} := \{Y \in \Sigma : \forall \varepsilon > 0 \; \exists Z \in a(\mathscr{R}) : \; \mu(Y \triangle Z) < \varepsilon\}$. We want to show that $\mathscr{M} \supseteq \Sigma = \sigma(a(\mathscr{R}))$.

By definition $a(\mathscr{R}) \subseteq \mathscr{M}$. Moreover, $\mathscr{M}$ is a monotone class:

1. $\mathscr{M}$ is closed under increasing unions: let $(Y_n)_n \subseteq \mathscr{M}$ increase to $Y \in \Sigma$. Then $(\mu(Y_n))_n$ increases to $\mu(Y)$. Given $\varepsilon > 0, \exists n : \; \mu(Y \setminus Y_n) < \varepsilon/2$. On the other hand, $\exists Z \in a(\mathscr{R}) : \; \mu(Y_n \triangle Z) < \varepsilon/2$. Therefore, $\mu(Y \triangle Z) = \mu(Y \setminus Z) + \mu(Z \setminus Y) \le \mu(Y \setminus Y_n) + \mu(Y_n \setminus Z) + \mu(Z \setminus Y_n) < \varepsilon$.
2. $\mathscr{M}$ is closed under decreasing intersections. The proof is similar to the previous case, but for $\lim_n \mu(Y_n) = \mu(Y)$ to be true when $(Y_n)_n$ decreases to Y, we need to use the fact that μ is finite.

By the monotone class theorem (Theorem A.1), $\mathscr{M} \supseteq \sigma(a(\mathscr{R}))$. □

We now define the set of simple functions:

$$\mathscr{S}(\Gamma, \Sigma, \mu) = \mathscr{S} := \left\{ \sum_{j=1}^n a_j \mathbb{1}_{A_j} : n \in \mathbb{N}, \; (a_j)_1^n \subseteq \mathbb{R}, \; (A_j)_1^n \subseteq \Sigma \text{ pairwise disjoint and of finite measure} \right\},$$

and the set of simple functions with support in $a(\mathscr{R})$:

$$\mathscr{S}^{\mathscr{R}}(\Gamma, \Sigma, \mu) = \mathscr{S}^{\mathscr{R}} := \left\{ \sum_{j=1}^n a_j \mathbb{1}_{R_j} : n \in \mathbb{N}, \; (a_j)_1^n \subseteq \mathbb{R}, \; (R_j)_1^n \subseteq \mathscr{R} \text{ pairwise disjoint and of finite measure} \right\}.$$

Note that if $R = \prod_{i=1}^n A_i$ is a measurable rectangle, then $\mathbb{1}_R(x_1, \ldots, x_n) = \mathbb{1}_{A_1}(x_1) \cdots \mathbb{1}_{A_n}(x_n)$. Our aim is to show that (under certain conditions on the spaces), measurable functions can be approximated pointwise by taking limits of functions in $\mathscr{S}^{\mathscr{R}}$.

Definition C.2. *Let (Γ, Σ, μ) be a general*[1] *measure space. We say that a sequence of measurable functions $(f_n)_{n=1}^\infty$ converges* almost uniformly (a.u.) *to a measurable function f if for every $\varepsilon > 0$, there exists an $N \in \Sigma$ such that $\mu(N) < \varepsilon$ and $\left\| \mathbb{1}_{\Gamma \setminus N}(f_n - f) \right\|_\infty \xrightarrow[n \to \infty]{} 0$.*

We will denote this convergence by writing $f_n \to f$ a.u.

Remark C.3. *In general measure spaces, a.u. convergence implies a.e. convergence: in the a.u. convergence definition we can replace the condition "$\forall \varepsilon > 0$" by "$\forall \varepsilon \in \{2^{-n} : n \in \mathbb{N}\}$", and we can also take the sets $(N_n)_n$ such that $\mu(N_n) \le 2^{-n}$ to be decreasing. Then, there is a.e. convergence in $(\cap_n N_n)^c$, and $\mu(\cap_n N_n) = 0$*

[1] By "general" we only emphasize that no special condition is requested of the measure space, like being (σ-)finite, for example.

Lemma C.4. *Let $(\Gamma,\Sigma,\mu)=\left(\prod_{i=1}^n\Gamma_i,\otimes_{i=1}^n\Sigma_i,\mu\right)$ be a finite measure space and $g\in L^\infty(\Gamma)$. Then, there exists a sequence $(\phi_n)_{n=1}^\infty\subseteq\mathscr{S}^{\mathscr{R}}$ such that $\phi_n\to g$ a.u.*

Proof. By the boundedness of g, there exists a sequence of simple functions $(g_n)\subseteq\mathscr{S}$ that converges to g uniformly $\left(\|g-g_n\|_\infty\to 0\right)$ and $\sup_n|g_n|\le|g|$ (see [Fol13, Theorem (2.10)] for details). For each $n\in\mathbb{N}$, let

$$g_n=\sum_{j=1}^{m_n}a_j^n\mathbb{1}_{E_j^n},$$

where $\{a_j^n\}_{j,n}\subseteq\mathbb{R}$, $\{E_j^n\}_{j,n}\subseteq\Sigma$ are pairwise disjoint. By Proposition C.1,

$$\forall\, j,n\in\mathbb{N}\;\exists R_j^n\in a(\mathscr{R}):\;\mu(E_j^n\triangle R_j^n)<2^{-n-j}.$$

Define, for each $n\in\mathbb{N}$:

$$\mathscr{S}^{\mathscr{R}}\ni\phi_n:=\sum_{j=1}^{m_n}a_j^n\mathbb{1}_{R_j^n},$$

and let $A_n:=\bigcup_{j=1}^{m_n}\left(E_j^n\triangle R_j^n\right)$, which satisfies $\mu(A_n)\le\sum_{j=1}^{m_n}2^{-n-j}\le 2^{-n}$.

$|g_n-\phi_n|\le\sum_{j=1}^{m_n}\left(\left|a_j^n\right|\mathbb{1}_{E_j^n\triangle R_j^n}\right)$, so $|g_n-\phi_n|\,\mathbb{1}_{A_n^c}=0$. Finally, if $\varepsilon>0$, then $\exists n\in\mathbb{N}:\;2^{-n}<\varepsilon$ so that $\mu(A_n)<\varepsilon$ and $\|(g-\phi_n)\,\mathbb{1}_{A_n^c}\|_\infty\le\|(g-g_n)\,\mathbb{1}_{A_n^c}\|_\infty+\|(g_n-\phi_n)\,\mathbb{1}_{A_n^c}\|_\infty\to 0.$ □

Remark C.5. *Let $g,(\Gamma,\Sigma,\mu)$ be as in Lemma C.4. As the measure space is finite, $g\in L^\infty(\Gamma)\subseteq L^1(\Gamma)$, and the sequence $(\phi_n)_{n=1}^\infty\subseteq\mathscr{S}^{\mathscr{R}}$ can be chosen such that $\forall n\in\mathbb{N}$ $\|g-\phi_n\|_1\le\alpha_n$ for any $\alpha\in c_0(\mathbb{N})$. (That is, $\phi_n\to g$ in L^1 as fast as we want). Indeed, in the proof of Lemma C.4, $\|g-g_n\|_1\le\mu(\Gamma)\|g-g_n\|_\infty\to 0$. $\|g_n-\phi_n\|_1\le\sum_{j=1}^{m_n}\left|a_j^n\right|\mu(E_j^n\triangle R_j^n)\le\|g\|_\infty\sum_{j=1}^\infty 2^{-n-j}=\|g\|_\infty 2^{-n}\to 0$. Therefore, $\|g-\phi_n\|_1\le\|g-g_n\|_1+\|g_n-\phi_n\|_1\to 0$. To ensure that $\|g-\phi_n\|_1\le\alpha_n$, we can take an appropriate subsequence of $(\phi_n)_n$.*

Whenever $\mathscr{M}$ is a σ-algebra and $A\in\mathscr{M}$, we will denote by $\mathscr{M}\cap A$ the σ-algebra on A: $\{A\cap M : M\in\mathscr{M}\}$. Note that a function $\mathbb{1}_Af$ is $\mathscr{M}$-measurable iff $f|_A$ is $\mathscr{M}\cap A$-measurable.

Remark C.6. *Let $(\Gamma,\Sigma,\mu)=\left(\prod_{i=1}^n\Gamma_i,\otimes_{i=1}^n\Sigma_i,\times_{i=1}^n\mu_i\right)$ such that each of the $(\Gamma_i,\Sigma_i,\mu_i)$ is σ-finite. Then there is a sequence of finite and pairwise disjoint subspaces $\left((E_m,\Sigma\cap E_m,\mu)\right)_{m=1}^\infty$ such that $\Gamma=\bigsqcup_{m=1}^\infty E_m$ and such that each of the spaces E_m is a product space (i.e. satisfying the hypothesis of Lemma C.4's measure space). It is enough to take the $\{E_m\}_m$ to be all the products of the countable decomposition of each Γ_i in pairwise disjoint subsets of finite measure.*

Theorem C.7. *Let $(\Gamma,\Sigma,\mu)=\left(\prod_{i=1}^n\Gamma_i,\otimes_{i=1}^n\Sigma_i,\times_{i=1}^n\mu_i\right)$ be as in Remark C.6. Then every measurable function $g:\Gamma\to\mathbb{R}$ is the pointwise limit of functions in $\mathscr{S}^{\mathscr{R}}$.*

Proof. Take $(E_m)_{m=1}^\infty$ to be the decomposition of Γ as in Remark C.6, and $g:\Gamma\to\mathbb{R}$ measurable. For each $m,i\in\mathbb{N}$, define $F_{im}:=E_m\cap\left\{i-1\le|g|<i\right\}$ and $g_{im}:=\mathbb{1}_{F_{im}}g$. $(F_{im})_{i,m}$ is a pairwise disjoint decomposition of Γ. Then, $g=\sum_{i,m=1}^\infty g_{im}$ (as a pointwise limit). Observe now that $g_{im}\in L^\infty(E_m)$, for every $i,m\in\mathbb{N}$. By Lemma C.4, $\exists\left(\phi_k^{im}\right)_{k=1}^\infty\subseteq\mathscr{S}^{\mathscr{R}}(E_m)\subseteq\mathscr{S}^{\mathscr{R}}(\Gamma):\;\phi_k^{im}\xrightarrow{k\to\infty}g_{im}$ a.u.

Let $\ell,i,m\in\mathbb{N}$. There is a measurable set $A_\ell^{im}\subseteq E_m$ such that $\mu(A_\ell^{im})<2^{-\ell-i-m}$ and

$$\left\|\left(g_{im}-\phi_k^{im}\right)\mathbb{1}_{E_m\setminus A_\ell^{im}}\right\|_\infty\xrightarrow{k\to\infty}0.$$

Take $k(\ell)\in\mathbb{N}$ such that $\left\|\left(g_{im}-\phi_{k(\ell)}^{im}\right)\mathbb{1}_{E_m\setminus A_\ell^{im}}\right\|_\infty<2^{-\ell}$ and define $h_\ell:=\sum_{i,m=1}^\infty\phi_{k(\ell)}^{im}$, for every $\ell\in\mathbb{N}$. h_ℓ is not necessarily contained in $\mathscr{S}^{\mathscr{R}}$, because it may be an infinite sum. For each $\ell\in\mathbb{N}$, define $A_\ell:=\bigcup_{i,m=1}^\infty A_\ell^{im}$ and $B_L=\bigcup_{\ell=L}^\infty A_\ell$. Then, $\mu(A_\ell)\le\sum_{i,m=1}^\infty 2^{-\ell-i-m}=2^{-\ell}$ and $\mu(B_L)\le\sum_{L=\ell}^\infty 2^{-\ell}=2^{-L+1}$. Finally, $\forall\ell\ge L$,

$$\begin{aligned}\left\|(g-h_\ell)\,\mathbb{1}_{B_L^c}\right\|_\infty&\le\left\|(g-h_\ell)\,\mathbb{1}_{A_\ell^c}\right\|_\infty=\left\|\sum_{i,m=1}^\infty\left(g_{im}-\phi_{k(\ell)}^{im}\right)\mathbb{1}_{F_{im}\setminus A_\ell}\right\|_\infty\\&=\sup_{i,m\in\mathbb{N}}\left\|\left(g_{im}-\phi_{k(\ell)}^{im}\right)\mathbb{1}_{F_{im}\setminus A_\ell}\right\|_\infty\le\sup_{i,m\in\mathbb{N}}\left\|\left(g_{im}-\phi_{k(\ell)}^{im}\right)\mathbb{1}_{E_m\setminus A_\ell^{im}}\right\|_\infty\le 2^{-\ell}\xrightarrow{\ell\to\infty}0.\end{aligned}$$

That is, $h_\ell\to g$ a.u.

To end the proof, let $s_\ell:=\sum_{i,m=1}^\ell\phi_{k(\ell)}^{im}\in\mathscr{S}^{\mathscr{R}}(\Gamma)$. $s_\ell\to g$ a.e. because $\forall x\in\Gamma\;\exists\tilde{\ell}\in\mathbb{N}:\;\forall\ell\ge\tilde{\ell}\;s_\ell(x)=h_\ell(x)$ and $h_\ell\to g$ a.e. □

Remark C.8. *Let g be as in Theorem C.7. If $g \in L^1(\Gamma)$, then the sequence $(s_\ell)_{\ell=1}^\infty \subseteq \mathscr{S}^{\mathscr{R}}$ converging to g a.e. can be chosen such that $\|g - s_\ell\|_1 \to 0$. In particular, $\int s_\ell \to \int g$. Indeed, in the proof of Theorem C.7, the supports of the $\{g_{im}\}_{i,m}$ are pairwise disjoint, so $\|g\|_1 = \sum_{i,m=1}^\infty \|g_{im}\|_1 < \infty$. Which implies $\sum_{i,m\geq j} \|g_{im}\|_1 \xrightarrow[j\to\infty]{} 0$.*

By Remark C.5, we can take (for each $i, m \in \mathbb{N}$) the sequence $\left(\phi_k^{im}\right)_{k=1}^\infty$ to be such that $\left\|g_{im} - \phi_k^{im}\right\|_1 \leq 2^{-k-i-m}$ for every $k \in \mathbb{N}$ and $k(\ell) \geq \ell$ for every $\ell \in \mathbb{N}$. Then,

$$\|g - h_\ell\|_1 \leq \sum_{i,m=1}^\infty \left\|g_{im} - \phi_{k(\ell)}^{im}\right\|_1 \leq \sum_{i,m=1}^\infty 2^{-k(\ell)-i-m} \leq 2^{-\ell} \to 0.$$

Moreover,

$$\left\|\sum_{i,m\geq j+1} \phi_{k(\ell)}^{im}\right\|_1 = \left\|\sum_{i,m\geq j+1} \left(\phi_{k(\ell)}^{im} - g_{im} + g_{im}\right)\right\|_1 \leq \sum_{i,m\geq j+1} \left\|\phi_{k(\ell)}^{im} - g_{im}\right\|_1 + \sum_{i,m\geq j+1} \|g_{im}\|_1 \xrightarrow[j\to\infty]{} 0, \tag{C.1}$$

because $\sum_{i,m\geq j} \left\|\phi_{k(\ell)}^{im} - g_{im}\right\|_1 \leq 2^{-\ell-2j}$. Therefore, we can choose $j(\ell)$ such that $\|h_\ell - s_{j(\ell)}\|_1 = \left\|\sum_{i,m=j(\ell)+1}^\infty \phi_{k(\ell)}^{im}\right\|_1 \leq 2^{-\ell}$. Then, $\|g - s_{j(\ell)}\|_1 \leq \|g - h_\ell\|_1 + \|h_\ell - s_{j(\ell)}\|_1 \leq 2^{-\ell+1} \to 0$, and $\left(s_{j(\ell)}\right)_{\ell\geq l_0} \subseteq \mathscr{S}^{\mathscr{R}}$ is as desired.

Before finishing this section we will present (but not prove) the following elementary measure theory result.

Proposition C.9. *Let Γ_0 be a set and (Γ_1, Σ_1) a measurable space. Then, the following hold:*

(a) Let $X : \Gamma_0 \to (\Gamma_1, \Sigma_1)$ and $Y : \Gamma_0 \to (\mathbb{R}, \mathscr{B}(\mathbb{R}))$. Then Y is $\sigma(X)$-measurable if and only if there exists a measurable function $g : (\Gamma_1, \Sigma_1) \to (\mathbb{R}, \mathscr{B}(\mathbb{R}))$ such that $Y = g \circ X$.

(b) Let $X_1, \ldots, X_n : \Gamma_0 \to (\mathbb{R}, \mathscr{B}(\mathbb{R}))$. Consider the function $\underline{X} = (X_1, \ldots, X_n)^T : \Gamma_0 \to (\mathbb{R}^n, \mathscr{B}(\mathbb{R})^n)$ given by $\underline{X}(\omega) = (X_1(\omega), \ldots, X_n(\omega))^T$. Then, $\sigma\left(\underline{X}\right) = \sigma(X_1, \ldots, X_n)$.

(c) Let $Y : \Gamma_0 \to (\mathbb{R}, \mathscr{B}(\mathbb{R}))$ and $\underline{X} = (X_1, \ldots, X_n)^T$ as in (b). Then Y is $\sigma(X_1, \ldots, X_n)$-measurable if and only if there exists a measurable function $g : (\mathbb{R}^n, \mathscr{B}(\mathbb{R})^n) \to (\mathbb{R}, \mathscr{B}(\mathbb{R}))$ such that $Y = g \circ \underline{X} = g(X_1, \ldots, X_n)$.

In the previous proposition, (a) can be proven by taking first Y to be simple and then extending the result to general measurable functions. (b) follows directly from the definition of measurability of functions and (c) is a direct consequence of (a) and (b).

APPENDIX D
Kolmogorov's continuity criterion

[RY05, Chapter 1, Theorem (2.1)] gives a criterion for the continuity of stochastic processes:

Theorem D.1 (Kolmogorov's continuity criterion). *Let $(E, \|\cdot\|)$ be a Banach space and $R > 0$. If X is an E-valued, $[0,R)^d$-indexed stochastic process X for which there exist $\gamma, \varepsilon, c > 0$ such that*

$$E\left[\|X_t - X_s\|^\gamma\right] \le c\,|t-s|^{d+\varepsilon}, \qquad \forall\, t, s \in [0,R)^d;$$

then, there is a modification $\tilde{X}$ of X such that

$$E\left[\left(\sup_{s \neq t}\left(\|\tilde{X}_t - \tilde{X}_s\| / |t-s|^\alpha\right)\right)^\gamma\right] < \infty, \qquad \forall\, \alpha \in [0, \varepsilon/\gamma).$$

In particular, the paths of $\tilde{X}$ are almost surely α-Hölder continuous. ($|\cdot|$ is any norm on $\mathbb{R}^d$).

Proof. WLOG, assume $R = 1$ and $|\cdot|$ to be the sup-norm on $\mathbb{R}^d$. For $m \in \mathbb{N}$:

$$D_m = \left\{\left(2^{-m} i_1, \ldots, 2^{-m} i_d\right) : i_1, \ldots, i_d \in \mathbb{Z} \cup \left[0, 2^m\right)\right\}, \;\; \Delta_m = \left\{(s,t) \in D_m^2 : \; |s-t| = 2^{-m}\right\}, \;\; D = \cup_m D_m.$$

We define a partial order on D by defining $s \le t$ iff every coordinate of s is less than or equal to the corresponding coordinate of t.

Let $K_i = \sup_{(s,t) \in \Delta_i} \|X_s - X_t\|$; by the hypothesis, $E\left[K_i^\gamma\right] \le c 2^{-i(d+\varepsilon)} = c 2^{-id} 2^{-i\varepsilon} \le c 2^{-i\varepsilon}$, $\forall\, i \in \mathbb{N}$.

For each point $s \in D$, there is an increasing sequence $(s_n)_{n \in \mathbb{N}}$ such that $s_n \in D_n$ and $s_n = s$ from some n on. Let $s, t \in D$ such that $|s-t| \le 2^{-m}$, then either $s_m = t_m$ or $(s_m, t_m) \in \Delta_m$ and so

$$X_s - X_t = \sum_{i=m}^\infty \left(X_{s_{i+1}} - X_{s_i}\right) + X_{s_m} - X_{t_m} + \sum_{i=m}^\infty \left(X_{t_i} - X_{t_{i+1}}\right),$$

where the sums have finite non-zero terms. Therefore, $\|X_s - X_t\| \le 3 \sum_{i=m}^\infty K_i$. Then, for $0 \le \alpha < \varepsilon/\gamma$,

$$\begin{aligned} M_\alpha := \sup_{s,t \in D,\; s \neq t}\left(\|X_s - X_t\| / |s-t|^\alpha\right) &= \sup_{m \in \mathbb{N}}\left(\sup_{s,t \in D,\; 2^{-m-1} < |s-t| \le 2^{-m}}\left(\|X_s - X_t\| / |s-t|^\alpha\right)\right) \\ &\le \sup_{m \in \mathbb{N}}\left(2^{(m+1)\alpha} \sup_{s,t \in D,\; |s-t| \le 2^{-m}} \|X_s - X_t\|\right) \le \sup_{m \in \mathbb{N}}\left(3 \cdot 2^{(m+1)\alpha} \sum_{i=m}^\infty K_i\right) \le 3 \cdot 2^\alpha \sum_{i=0}^\infty 2^{i\alpha} K_i. \end{aligned}$$

For $0 \le \alpha < \varepsilon/\gamma$, $\gamma \ge 1$, and $J = c^{1/\gamma} \cdot 3 \cdot 2^\alpha$, we apply Minkowski's inequality for the $L^\gamma(P)$ norm to get

$$\|M_\alpha\|_\gamma \le 3 \cdot 2^\alpha \sum_{i=0}^\infty 2^{i\alpha} \|K_i\|_\gamma \le J \sum_{i=0}^\infty 2^{i(\alpha - \varepsilon/\gamma)} < \infty.$$

For the case $\gamma \in (0,1)$, we can apply the inequality $\left(\sum_{i=1}^\infty a_i\right)^\gamma \le \sum_{i=1}^\infty a_i^\gamma$ for sequences $\{a_i\}_{\mathbb{N}} \subseteq \mathbb{R}_+$:

$$E\left[(M_\alpha)^\gamma\right] \le E\left[\left(3 \cdot 2^\alpha \sum_{i=0}^\infty 2^{i\alpha} K_i\right)^\gamma\right] \le 3^\gamma \cdot 2^{\alpha\gamma} \sum_{i=0}^\infty 2^{i\alpha\gamma} E\left[K_i^\gamma\right] \le J^\gamma \sum_{i=0}^\infty 2^{i(\alpha\gamma - \varepsilon)} < \infty.$$

This implies that $M_\alpha < \infty$ a.e. so that the paths of X are a.e. α-Hölder continuous and therefore, a.e. uniformly continuous on the dense set $D \subseteq [0,R)^d$. We can define the by-definition a.e. path-continuous process $\tilde{X}_t(\omega) = \lim_{D \ni s \to t} X_s(\omega)$. By the hypothesis and Chebyshev's inequality, for every $t \in [0,R)^d$, X_s converges in probability to X_t as $D \ni s \to t$, and, by definition, X_s converges a.e. (and therefore in probability) to $\tilde{X}_t$; so $X_t = \tilde{X}_t$ a.e., by the uniqueness of the limit in probability. $\square$

The continuity of the modification in Kolmogorov's continuity criterion above can be extended to unbounded index subsets of $\mathbb{R}^d$ by making $R \to \infty$. The paths of the resulting process will be locally α-Hölder continuous, and therefore still continuous. This idea is explained in the proof of Theorem 3.6, so we will not go into further detail here.

Corollary D.2 (Kolmogorov's continuity criterion for $\mathbb{R}$-valued, $\mathbb{R}_+$-indexed processes). *An $\mathbb{R}$-valued, $\mathbb{R}_+$-indexed stochastic process X for which there exist $\gamma, \varepsilon, c > 0$ such that $E\left[|X_{t+h} - X_t|^\gamma\right] \le ch^{1+\varepsilon}$ for every $t, h > 0$ has a modification that is almost-surely continuous.*

APPENDIX E
The augmented filtration of a Brownian motion

Proposition E.1 (Augmented filtration of a Brownian motion). *Let $\mathscr{B} = \overline{\sigma(B_t)_{t\geq 0}}$ be the smallest complete σ-algebra generated by $(B_t)_{t\geq 0}$. Define $\mathscr{N} = \{A \in \mathscr{B} : P(A) = 0\}$ (the set of $\mathscr{B}$-*nullsets*) and $\mathscr{B'}_t = \sigma(\mathscr{N}, \mathscr{B}_t)$. Then $\left(\mathscr{B'}_t\right)_{t\geq 0}$ is a right-continuous filtration of sub-σ-algebras of $\mathscr{B}$ containing all $\mathscr{B}$-nullsets.*

Proof. First of all, we claim that B is a Brownian motion with respect to $(\mathscr{B}'_t)_t$. To prove this, the only non-trivial property to check is the independence of $B_t - B_s$ and $\mathscr{B}'_s$ for every $0 \leq s < t$. $B_t - B_s$ is independent of the π-system $\mathscr{B}_s \cup \mathscr{N}$ because B is a Brownian motion and every set is independent of any nullset. By Theorem *A.4*, $B_t - B_s$ is independent of $\sigma(\mathscr{B}_s \cup \mathscr{N}) = \mathscr{B}'_s$.

Let $g_u(x) : \mathbb{R} \to \mathbb{R}$ denote the density of a centered normal distribution with variance $u : N(0, u)$:

$$g_u(x) = \frac{1}{\sqrt{2\pi u}} \exp\left(-\frac{x^2}{2u}\right).$$

The density of a $N(y, u)$ is given by $x \mapsto g_u(x - y)$. Define, for $u > 0, x \in \mathbb{R}, f \in L^\infty(\mathscr{B})$:

$$p_u(x, f) := \int_{\mathbb{R}} f(y) g_u(x - y) dy = f * g_u(x).$$

1. The function $h_f : (u, x) \mapsto p_u(x, f)$ is continuous in $\mathbb{R}_+ \times \mathbb{R}$, because $\big((u, x) \mapsto g_u(x)\big) \in C^\infty(\mathbb{R}_+ \times \mathbb{R})$ and by the dominated convergence theorem we can commute the integral with the derivatives to get $\frac{\partial h_f}{\partial x}, \frac{\partial h_f}{\partial u} \in C(\mathbb{R}_+ \times \mathbb{R})$, implying differentiability (and therefore continuity) of h_f.

2. Denote by B^y a Brownian motion starting at $y \in \mathbb{R}$. Claim:

$$\forall 0 \leq s < t, A \in \mathscr{B}(\mathbb{R}), \quad E\left[\mathbb{1}_A(B_t) \,\middle|\, \mathscr{B}'_s\right] = E\left[\mathbb{1}_A\left(B^y_{t-s}\right)\right]_{y=B_s}.$$

The proof of this claim is essentially identical to the proof of the Markov property for Itô diffusions: Theorem 4.2, so we will not repeat it here. It is important to note that we are in no way falling into a circular argument here, even if this claim is strictly weaker than Theorem 4.2. We are just stating that the same proof can be easily adapted to work in this case.

3. If $0 \leq s < t$ and $A \in \mathscr{B}(\mathbb{R})$, then $E\left[\mathbb{1}_A(B_t) \,\middle|\, \mathscr{B}'_s\right] = p_{t-s}(B_s, \mathbb{1}_A)$. This can be seen by using the previous point and the fact that $B^y_{t-s} \sim N(y, t-s)$, so

$$E\left[\mathbb{1}_A\left(B^y_{t-s}\right)\right] = P\left(B^y_{t-s} \in A\right) = \int_A g_{t-s}(x - y) dx = p_{t-s}(y, \mathbb{1}_A).$$

4. If $0 \leq s < t$ and $f \in L^\infty(\mathscr{B}(\mathbb{R}))$, then $E\left[f(B_t) \,\middle|\, \mathscr{B}'_s\right] = p_{t-s}(B_s, f)$ a.e. The set of functions that have this property contains the characteristic functions (by the previous point), is a vector space (because $p_u(x, \cdot)$ and $E\left[\cdot \,\middle|\, \mathscr{B}'_s\right]$ are linear), and is closed under increasing, bounded, pointwise limits of non-negative functions (by the monotone convergence theorem). Applying Theorem *A.3*, we conclude that the set of functions that have this property is all of $L^\infty(\mathscr{B}(\mathbb{R}))$.

5. If $0 \leq s < t$ and $f \in L^\infty(\mathscr{B}(\mathbb{R}))$, then $E\left[f(B_t) \,\middle|\, \mathscr{B}'_{s+}\right] = p_{t-s}(B_s, f)$. Making use of Theorem 2.34, continuity of $(u, x) \mapsto p_u(x, f)$ and right-continuity of the paths of B; if $s_n \downarrow s$, then

$$E\left[f(B_t) \,\middle|\, \mathscr{B}'_{s_n}\right] \to E\left[f(B_t) \,\middle|\, \mathscr{B}'_{s+}\right] \text{ a.e.},$$

$$E\left[f(B_t) \,\middle|\, \mathscr{B}'_{s_n}\right] = p_{t-s_n}(B_{s_n}, f) \to p_{t-s}(B_s, f) = E\left[f(B_t) \,\middle|\, \mathscr{B}'_s\right] \text{ a.e.};$$

and we are done, by (a.e) uniqueness of the a.e. limit.

6. If $0 \le s < t_1 < t_2 < \cdots < t_m$ and $\left(f_i\right)_{i=1}^m \subseteq L^\infty(\mathscr{B}(\mathbb{R}))$, then

$$E\left[\prod_{i=1}^m f_i\left(B_{t_i}\right) \middle| \mathscr{B}'_{s+}\right] = E\left[\prod_{i=1}^m f_i\left(B_{t_i}\right) \middle| \mathscr{B}'_s\right] \text{ a.e.}$$

To prove this, we are going to make a series of transformations on $E\left[\prod_{i=1}^m f_i\left(B_{t_i}\right) \mid \mathscr{B}'_s\right]$ to arrive at an expression. The exact same transformations are true for $E\left[\prod_{i=1}^m f_i\left(B_{t_i}\right) \mid \mathscr{B}'_{s+}\right]$ by the previous points, so that will give us the desired result.

$$\begin{aligned} E\left[\prod_{i=1}^m f_i\left(B_{t_i}\right) \middle| \mathscr{B}'_s\right] &= E\left[E\left[\prod_{i=1}^m f_i\left(B_{t_i}\right) \middle| \mathscr{B}'_{t_{m-1}}\right] \middle| \mathscr{B}'_s\right] = E\left[E\left[f_m(B_{t_m}) \mid \mathscr{B}'_{t_{m-1}}\right] \prod_{i=1}^{m-1} f_i\left(B_{t_i}\right) \middle| \mathscr{B}'_s\right] \\ &= E\left[\underbrace{p_{t_m - t_{m-1}}(B_{t_{m-1}}, f_m) f_{m-1}\left(B_{t_{m-1}}\right)}_{\tilde{f}_{m-1}(B_{t_{m-1}})} \prod_{i=1}^{m-2} f_i\left(B_{t_i}\right) \middle| \mathscr{B}'_s\right], \end{aligned}$$

repeating this procedure, we can iteratively define $\tilde{f}_k(x) := f_k(x)\, p_{t_{k+1}-t_k}\left(x, \tilde{f}_{k+1}\right)$ for $m-2 \ge k \ge 1$. This yields

$$E\left[\prod_{i=1}^m f_i\left(B_{t_i}\right) \middle| \mathscr{B}'_s\right] = E\left[\tilde{f}_1\left(B_{t_1}\right) \mid \mathscr{B}'_s\right] = p_{t_1 - s}(B_s, \tilde{f}_1).$$

7. The previous point is true even if some of the t_i are smaller than s (those $f_i(B_{t_i})$'s can be extracted from the conditional expectation on both sides).

8. If $Z \in L^\infty\left(\mathscr{B}'_\infty\right)$, then $E\left[Z \mid \mathscr{B}'_{s+}\right] = E\left[Z \mid \mathscr{B}'_s\right]$ a.e. $\forall s \ge 0$ (apply the previous point and Theorem 2.41).

9. The previous point applied to $Z = \mathbb{1}_A$ for $A \in \mathscr{B}'_{s+}$ imples that $A \in \mathscr{B}'_s$, because $\mathscr{N} \subseteq \mathscr{B}'_s$. □