UNIVERSITAT DE BARCELONA
Facultat de Matemàtiques i Informàtica

GRAU DE MATEMÀTIQUES

Treball final de grau

FROM THE CLASSICAL TO THE STOCHASTIC HEAT EQUATION

Autor: Jordi Petchamé Guerrero

Directora: Dra. Marta Sanz-Solé
Realitzat a: Departament de Matemàtiques i Informàtica
Barcelona, January 24, 2021

Abstract

This bachelor's thesis revolves around the connection between stochastic processes and the heat equation. The main goal is to carry out a thorough study of the transition from the classical to the stochastic one-dimensional heat equation. In order to develop the mathematical framework for linear stochastic partial differential equations, we use tools of basic probability theory, calculus and functional analysis. We start with a concise study of the classical deterministic heat equation, from its physical derivation to the search for explicit solutions under specific conditions. Then, we describe the mathematical foundations of the stochastic version of this partial differential equation, focusing on Gaussian stochastic processes. On that basis, we define the stochastic heat equation on $\mathbb{R}$. Finally, we conclude this project with a comprehensive analysis of its solutions' continuity properties.

2010 Mathematics Subject Classification. 35K05, 35R60, 60G15, 60H40

Contents

Chapter 1

Introduction

The interplay between physics and mathematics has been known and studied for centuries. Many physics problems have drawn the attention of mathematicians, who seek to solve them by applying their mathematical background. This has given rise to the field of mathematical physics and, in particular, to physical evolution problems governed by partial differential equations. Within this area, one of the most widely studied topics is the (classical) heat equation, also known as diffusion equation, originally developed by Fourier in the early 19th century.

In the first half of the 20th century, the field of probability theory saw remarkable progress when it became rigorously formulated with the language of the recently developed field of measure theory, thanks to Kolmogorov, among others. With this established framework, the area of probability expanded and new fields were born, including stochastic processes. Therefore, new evolution problems involving randomness were proposed, leading the way to stochastic ordinary and partial differential equations such as the stochastic heat equation.

The overall goal of this bachelor's thesis is to describe the transition from the classical to the stochastic heat equation and study the latter's continuity properties, choosing the simplest case, which corresponds to one dimension. While the classical equation is a deterministic evolution problem, the stochastic has a random component and requires a different framework that we aim to develop. We will carry this out with the following structure:

- First, compute explicit solutions to the classical deterministic heat equation, focusing on a particular solution of great importance.
- Next, give a description of the mathematical concepts of stochastic analysis, with an emphasis on Gaussian processes, that allow us to formulate a linear stochastic PDE.
- Finally, define the stochastic heat equation on $\mathbb{R}$ and analyze its continuity properties.

To this end, we draw on basic probability theory and calculus. In addition, we consider a few physical principles to derive the heat equation, and we make use of elementary functional analysis.

This thesis is structured in four chapters. Following the introduction, in Chapter 2 we begin with a brief historical account of the classical heat equation, based on [17]. Then, we give a physical derivation and we compute explicit solutions under specific conditions, both for the homogeneous and inhomogeneous cases. This chapter provides an introduction to the deterministic heat equation that will be necessary to study the stochastic case in the last chapter. For this reason, we focus on the features that will be used later in the thesis, particularly on the fundamental solution to the heat equation and how this notion enables us to obtain solutions to the corresponding initial-value problem. In this context, we prefer

not to dig deeper into qualitative properties of solutions, such as existence or unicity. We refer the interested reader to [22, 7]. Due to the limited extent of this project, we do not illustrate our computations with examples either.

The next chapter introduces the probabilistic framework in which we will develop the subsequent chapters. We start by introducing the notion of stochastic processes in continuous time, and we give a characterization. Then, we briefly define one of the central concepts of this thesis, namely Gaussian random vectors. Next, we describe characteristic functions together with their properties, and we use them to characterize Gaussian random vectors. The last section is an introduction to another key notion: Gaussian stochastic processes. The main references for stochastic processes are [23, 25], and for characteristic functions and Gaussian random vectors, [12, 1, 18].

Chapter 4 outlines the types of Gaussian processes that we will deal with in the study of the stochastic heat equation. In the first section we give an extended description of Brownian motion. After a brief overview of the historical roots of this process, we show one of its possible mathematical constructions. The historical account is based for the most part in [10], which we strongly recommend to anyone interested in this topic. Paul Lévy's construction was the most appealing for me, because of my interest in functional analysis, an area which was practically unknown to me prior to this semester. The main references for this section are [23, 13, 15, 14, 16, 3]. For further reading, the book [20] contains a general description of processes related to Brownian motion along with some interesting examples. In the remaining sections, we expand the notion of Gaussian processes to a wider range of index sets, including Hilbert spaces and Borel sets. This will allow us to define the stochastic heat equation on $\mathbb{R}$. In these sections we follow [5] and, to a lesser extent, [2].

Finally, we apply the probabilistic and analytical tools developed in the previous chapters to give a meaning to the stochastic heat equation on $\mathbb{R}$ with additive noise, obtained by replacing the deterministic external forcing in the classical heat equation by a stochastic space-time white noise. We then focus on the continuity properties of the sample paths of its solutions, proving that they are locally Hölder continuous under certain conditions. In doing so, we introduce an example of SPDEs, a relatively modern and still developing field of mathematics. The basic reference for this chapter is [5].

Chapter 2

Classical Heat Equation

Throughout this thesis, we will consider the heat equation to be one-dimensional.

2.1 Brief History

The heat equation (or diffusion equation) traces its roots back to the early 19th century, when the French mathematician and physicist Jean Baptiste Joseph Fourier (1768-1830) presented in his 1807 manuscript *Théorie de la Propagation de la Chaleur dans les Solides* a formula describing the conduction of heat in solids. After an early unsuccessful attempt at formulating heat conduction in terms of action at a distance, he assumed that the temperature in an infinitesimal lamina was only dependent on the conditions at the lamina and its endpoints, thus formulating heat diffusion in a continuum. The property of linearity enabled Fourier to apply the concept of superposition, which together with separation of variables, generated general solutions in terms of infinite trigonometric series and, eventually, integral transforms that would be named after him. However, skepticism on trigonometric series by some of the most renowned mathematicians of the time resulted in his manuscript not being well received. It would take 15 years for his theories to be accepted by the scientific community, when in 1822 he published *Théorie Analytique de la Chaleur* [9] which brought him international recognition and expanded his concepts to a wider audience. Fourier's method started to be applied as a mathematical framework in other fields such as economics or biology.

2.2 Homogeneous Heat Equation

2.2.1 Physical Derivation

Consider a homogeneous and isotropic cylindrical metal rod lying on the x-axis from the origin to $x = L > 0$ with non-uniform temperature. We assume that the mass density ρ, specific heat at constant volume c_v, thermal conductivity κ and cross-sectional area σ are all constant. In addition, we dictate that the sides of the rod are perfectly insulated and the length of the rod is much larger than σ. We may suppose that heat moves only down the length of the rod and that its transfer intensity is uniformly distributed in each section of the rod. Since heat is thermal energy, we can use the law of conservation of energy as follows:

> Let $V = (x, x + \Delta x)$ be an arbitrary control length inside the rod. *The time rate of change of thermal energy in V equals the net flux of heat through the boundary ∂V of V, due to the conduction.*

Denote by $e = e(x,t)$ the thermal energy per unit mass. The total amount of thermal energy contained in V is given by the function

$$H(t) = \int_V e(x,t)\rho\, dx.$$

Therefore, its time rate of change is

$$\begin{aligned}\frac{dH}{dt} &= \frac{d}{dt}\int_V e(x,t)\rho\, dx \\ &= \int_V e_t(x,t)\rho\, dx.\end{aligned}$$

Let $\mathbf{F}$ be the heat flux vector. If $d\sigma$ is an area element in ∂V with outward pointing unit normal vector $\boldsymbol{\nu}$, then the net heat flux through ∂V is given by

$$-\int_{\partial V} \mathbf{F}\cdot\boldsymbol{\nu}\, d\sigma = -\int_V \mathrm{div}\mathbf{F}\, dx$$

according to the Divergence Theorem[1].

Thus, conservation of energy requires

$$\int_V e_t(x,t)\rho\, dx = -\int_V \mathrm{div}\mathbf{F}\, dx. \tag{2.2.1}$$

Since V is arbitrary, the identity (2.2.1) leads to a basic law of heat conduction, mathematically expressed as the pointwise relation

$$e_t\rho = -\mathrm{div}\mathbf{F}. \tag{2.2.2}$$

In order to find e and $\mathbf{F}$, we apply the following two physical principles:

- **Fourier's law of heat conduction.** *The heat flux is a linear function of the negative temperature gradient*:

  $$\mathbf{F} = -\kappa u_x$$

  where $u = u(x,t)$ is the absolute temperature and κ is strictly positive[2].

  Since κ is constant, we obtain

  $$\mathrm{div}\mathbf{F} = -\kappa u_{xx}. \tag{2.2.3}$$

  This law states that heat flows from hotter to cooler regions at a rate proportional to the temperature gradient.

- *The thermal energy is a linear function of the absolute temperature*:

  $$e = c_\upsilon u. \tag{2.2.4}$$

Inserting (2.2.3) and (2.2.4) in (2.2.2) yields the **homogeneous heat equation**

$$u_t = ku_{xx} \tag{2.2.5}$$

where the coefficient $k = \frac{\kappa}{c_\upsilon\rho}$ is called *thermal diffusivity*.

In the remainder of this section, we will consider the heat equation to be homogeneous.

[1] In one dimension, it is equivalent to integration by parts.

[2] In this document, "positive" means "non-negative", hence the word "strictly" to make the distinction.

2.2.2 Initial and Boundary Conditions

In order to predict the existence of a unique solution to equation (2.2.5), we need to prescribe some conditions:

(i) **Initial Condition (IC).** We set

$$u(x,0) = g(x) \tag{2.2.6}$$

where g is a real-valued function that models the initial temperature distribution.

(ii) **Boundary Conditions (BCs).** To determine a unique evolution, it is necessary to specify the interaction with its surroundings. The most common boundary conditions are the following:

- **Dirichlet boundary conditions:** The temperature is prescribed at both ends of the rod:

$$u(0,t) = h_1(t) \quad , \quad u(L,t) = h_2(t)$$

for any $t > 0$.

- **Neumann boundary conditions:** From Fourier's law we have

$$\begin{aligned} &\text{inward heat flow at } x = 0 : -\kappa u_x(0,t), \\ &\text{inward heat flow at } x = L : \kappa u_x(L,t). \end{aligned}$$

We assign the heat flux at the endpoints through the conditions

$$u_x(0,t) = h_1(t) \quad , \quad u_x(L,t) = h_2(t)$$

at any time $t > 0$.

- **Robin boundary condition:** Assume that the inward heat flux from one extremity of the rod, say $x = L$, depends linearly on $U - u$, where U is the temperature of the surroundings. Thus,

$$\kappa u_x = \gamma(U - u) \qquad , \gamma > 0.$$

Setting $\alpha = \frac{\gamma}{\kappa}$ and $h = \frac{\gamma U}{\kappa}$, the condition at $x = L$ is

$$u_x + \alpha u = h.$$

This condition is also called **radiation boundary condition**.

- **Mixed boundary conditions:** They are given by equations involving the temperature and heat flux at the boundaries. For example, one endpoint with the Dirichlet condition and the other with the Neumann condition.

Thus we obtain the initial/boundary-value problem

$$\begin{cases} u_t = k u_{xx} & , 0 < x < L, \ t > 0 \\ u(x,0) = g(x) & , 0 \le x \le L \\ + \text{ BCs} & , t > 0. \end{cases} \tag{2.2.7}$$

In the sequel we will refer to (2.2.7) as the *homogeneous local Cauchy problem.*

If the rod is of infinite length, there are no longer any boundaries and the only necessary requirement is that the initial condition has a controlled growth at infinity. This case will be considered later in this section.

2.2.3 Solution by Separation of Variables

The technique of *separation of variables* involves seeking a non-trivial explicit solution of the form

$$u(x,t) = X(x)T(t)$$

where X and T are functions to be determined.

Taking the relevant partial derivatives

$$u_{xx} = X''(x)T(t) \quad \text{and} \quad u_t = X(x)T'(t),$$

the heat equation in (2.2.7) becomes

$$XT' - kX''T = 0$$

from which, separating the variables,

$$\frac{T'}{kT} = \frac{X''}{X}. \tag{2.2.8}$$

Note that the left hand side depends only on time, whereas the right hand side only does on space. The equality must hold for any $t > 0$ and $x \in (0, L)$, therefore both sides are equal to a common constant, say λ. Hence, (2.2.8) becomes

$$\frac{T'}{kT} = \frac{X''}{X} = \lambda.$$

In order for u to satisfy the designated boundary conditions, we must find the solutions of the following boundary-value problem for X:

$$\begin{cases} X''(x) = \lambda X(x) & , 0 < x < L \\ X \text{ satisfies the BCs.} \end{cases}$$

In this context, this problem is called an *eigenvalue problem*. The values of λ are the *eigenvalues* and the solutions X are the associated *eigenfunctions*. Moreover, it is an example of a *Sturm-Liouville problem*[3].

We now turn our attention to the time evolution. For a given value of $u(x, 0)$, we obtain the related *Cauchy problem* for ordinary differential equations

$$\begin{cases} T'(t) = k\lambda T(t) & , t > 0 \\ T \text{ satisfies the IC.} \end{cases}$$

The general solution is clearly

$$T(t) = T(0)e^{k\lambda t}.$$

Thus, for each eigenfunction X_n with corresponding eigenvalue λ_n we have a solution T_n such that the function

$$u_n(x,t) = X_n(x)T_n(t) \tag{2.2.9}$$

[3] A second order linear differential equation on a finite interval $[a, b]$ of the form

$$-\frac{d}{dx}\left[p(x)\frac{d}{dx}\right]y + q(x)y = \lambda\omega(x)y$$

with p, q and ω specified such that $p(x), \omega(x) > 0$ for $x \in (a, b)$, subject to boundary conditions at a and b, is called a *Sturm-Liouville problem*.

is a solution to the heat equation on the interval where the boundary conditions are satisfied.

The linearity of the problem allows us to construct a solution by applying the *principle of superposition* to the family of solutions $\{u_n\}_{n\geq 1}$ given by (2.2.9). To solve the initial condition, we will need the whole set of solutions u_n, giving rise to a full solution of the form

$$u(x,t) = \sum_{n=1}^{\infty} u_n(x,t).$$

In particular,

$$u(x,t) = \sum_{n=1}^{\infty} X_n(x)T_n(t) = \sum_{n=1}^{\infty} X_n(x)T_n(0)e^{k\lambda_n t}.$$

In order to find the coefficients $T_n(0)$, we often use the *orthogonality* property of eigenfunctions. Examples with Dirichlet boundary conditions can be found in [22, pp. 24–31],[11].

2.2.4 Steady State

This subsection briefly treats the asymptotic behavior of the solution.

When the temperature is independent of time, the solution is known as the *steady-state* (or *equilibrium*) *solution* u_{eq}. Since by definition $u_{eq} = u_{eq}(x)$, it must satisfy the equation $u''_{eq} = 0$ with our BCs. An elementary computation gives

$$u_{eq}(x) = u_{eq}(0) + \big[u_{eq}(L) - u_{eq}(0)\big]\frac{x}{L}$$

which corresponds to a uniform heat flux along the bar.

Let us introduce the function

$$U(x,t) = u_{eq}(x) - u(x,t).$$

This function represents a *transient regime* that must converge to 0 as $t \to \infty$ with a rate that shows how fast the temperature distribution reaches equilibrium.

2.2.5 Fundamental Solution

The heat equation has a privileged solution that enables us to construct general solutions. In order to find an explicit solution, it is convenient to explore properties of symmetry and invariance. We present hereunder the main transformations that preserve the heat equation.

Let $u = u(x,t)$ be a solution to the *homogeneous heat equation on* $\mathbb{R}$

$$u_t = ku_{xx} \qquad , x \in \mathbb{R}\ ,\ t > 0. \tag{2.2.10}$$

- *Space and time translation*: For $y \in \mathbb{R}$ and $s > 0$ fixed, the function $\tilde{u}$ defined on $\mathbb{R} \times (s, \infty)$ as

  $$\tilde{u}(x,t) = u(x-y, t-s)$$

  is also a solution to (2.2.10).

- *Parabolic dilations*: The heat equation is invariant under the combined stretching

  $$x \mapsto \lambda x \quad , \quad t \mapsto \lambda^2 t \tag{2.2.11}$$

  for $\lambda > 0$.

- *Derivation and integration*: Any derivative (u_x, u_t, u_{xx}, etc.) and integral (assuming proper convergence) of a solution is still a solution.

The transformation given by (2.2.11) leaves the expression $\frac{x}{\sqrt{kt}}$ unchanged and suggests that we look for solutions of the form

$$u(x,t) = v\left(\frac{x}{\sqrt{kt}}\right) \tag{2.2.12}$$

for some function $v\colon \mathbb{R} \mapsto \mathbb{R}$ to be determined. These solutions belong to the class of *self-similar solutions*[4].

The **fundamental solution** $\Phi_k(x,t)$ to (2.2.10) is defined as the positive, spatially even, self-similar solution, with a singularity at the origin, such that

$$\int_{\mathbb{R}} \Phi_k(x,t)\, dx = 1 \qquad \forall t > 0. \tag{2.2.13}$$

We will follow two different approaches to compute Φ_k. In the first one we will develop the idea behind (2.2.12), while the latter is based on the *Fourier transform*.

Let $\xi = \frac{x}{\sqrt{kt}}$. Differentiating the partial derivatives that appear in the heat equation yields

$$u_t(x,t) = -\frac{1}{2}\left(\frac{x}{t\sqrt{kt}}\right) v'\left(\frac{x}{\sqrt{kt}}\right),$$
$$u_{xx}(x,t) = \left(\frac{1}{kt}\right) v''\left(\frac{x}{\sqrt{kt}}\right).$$

Hence

$$u_t - k u_{xx} = -\frac{1}{t}\left[v''(\xi) + \frac{\xi}{2} v'(\xi)\right]$$

and we see that for u to be a solution to (2.2.10) the function v must be a real solution to the ordinary differential equation

$$v''(\xi) + \frac{\xi}{2} v'(\xi) = 0. \tag{2.2.14}$$

Let $\phi = v'$. We can rewrite (2.2.14) as

$$\phi'(\xi) + \frac{\xi}{2}\phi(\xi) = 0$$

and it is clear that the general solution is

$$\phi(\xi) = c e^{-\frac{\xi^2}{4}} \qquad , c \in \mathbb{R}. \tag{2.2.15}$$

Notice that ϕ happens to be a *Gaussian function*. Selecting the constant c such that (2.2.15) is normalized gives us[5]

$$\phi(\xi) = \frac{1}{\sqrt{4\pi}} e^{-\frac{\xi^2}{4}}. \tag{2.2.16}$$

[4] A solution to an evolution problem is called *self-similar* if its graph at a fixed time remains similar to itself at all times during the evolution. One-dimensional self-similar solutions have the general form

$$u(x,t) = a(t) F(x/b(t)).$$

[5] Recall that $\int_{\mathbb{R}} e^{-x^2}\, dx = \sqrt{\pi}$.

We can think of the expression above as the *probability density function* of a random variable with $N(0,2)$ law. Since (2.2.16) is positive, continuous and vanishes at $\pm\infty$, it is integrable and thus the function

$$v(\xi) = \int_{-\infty}^{\xi} \phi(z)\, dz = \frac{1}{\sqrt{4\pi}} \int_{-\infty}^{\xi} e^{-\frac{z^2}{4}}\, dz$$

is the corresponding *cumulative distribution function* of ϕ.

For every fixed $x \in \mathbb{R}_{>0}$, $t \to 0^+$ implies that $\xi \to +\infty$, thus

$$\lim_{t\to 0^+} v(\xi) = 1.$$

On the other hand, if we fix $x \in \mathbb{R}_{<0}$, we have

$$\lim_{t\to 0^+} v(\xi) = 0.$$

By definition, a distribution function is right continuous[6]. Therefore we can formulate the initial condition for u in terms of the *Heaviside function* $\mathcal{H}(x)$ as follows:

$$\lim_{t\to 0^+} u(x,t) = \mathcal{H}(x) = \begin{cases} 1 & \text{if } x \geq 0 \\ 0 & \text{if } x < 0. \end{cases} \tag{2.2.17}$$

Differentiating $u(x,t) = v\left(\frac{x}{\sqrt{kt}}\right)$ with respect to x gives

$$u_x(x,t) = \left(\frac{1}{\sqrt{kt}}\right) v'\left(\frac{x}{\sqrt{kt}}\right) = \left(\frac{1}{\sqrt{kt}}\right) \phi\left(\frac{x}{\sqrt{kt}}\right) = \left(\frac{1}{\sqrt{4\pi kt}}\right) e^{-\frac{x^2}{4kt}}, \tag{2.2.18}$$

which solves the heat equation for $t > 0$.

This is the fundamental solution Φ_k that we sought. Indeed, it verifies the properties of positiveness, even parity in space, and self-similarity. In a similar way to ϕ, it is the probability density function of a Normal random variable, parametrized with time, with mean zero and variance $2kt$. This implies (2.2.13). It follows from (2.2.18) that $u(x,t)$ is the associated cumulative distribution function, also parametrized with time. Furthermore, since

$$\begin{aligned} &\lim_{t\to 0^+} \Phi_k(x,t) = 0 \qquad \text{for any fixed } x \neq 0, \\ &\lim_{t\to 0^+} \Phi_k(0,t) = \infty, \end{aligned}$$

the fundamental solution behaves as the *Dirac distribution at the origin* $\delta(x)$ when $t \to 0^+$. Eventually, the whole probability density is concentrated at $x = 0$. This is consistent with (2.2.17) because

$$\int_{-\infty}^{x} \lim_{t\to 0^+} \Phi_k(y,t)dy = \lim_{t\to 0^+} \int_{-\infty}^{x} \Phi_k(y,t)dy = \lim_{t\to 0^+} u(x,t) = \mathcal{H}(x),$$

the first equality resulting from the integrability of Φ_k. Thus, we can define our fundamental solution as follows:

$$\Phi_k(x,t) = \begin{cases} \frac{1}{\sqrt{4\pi kt}} e^{-\frac{x^2}{4kt}} & , x \in \mathbb{R}\ ,\ t > 0 \\ \delta(x) & , x \in \mathbb{R}\ ,\ t = 0. \end{cases} \tag{2.2.19}$$

[6]Recall that a function $F\colon \mathbb{R} \to \mathbb{R}$ is *right continuous* if $\lim_{y\to x^+} F(y) = F(x)$ for all $x \in \mathbb{R}$.

Consider now the initial-value problem

$$\begin{cases} u_t = k u_{xx} & , x \in \mathbb{R} \ , \ t > 0 \\ u(x,0) = g(x) & , x \in \mathbb{R} \end{cases} \tag{2.2.20}$$

where $g \in C^0(\mathbb{R}) \cap L^\infty(\mathbb{R})$[7] is the initial condition (2.2.6). We will refer to (2.2.20) as the *homogeneous global Cauchy problem*.

The space translation invariance of the solutions of the heat equation implies that the function $(x,t) \mapsto \Phi_k(x-y,t)$ solves the heat equation for every fixed $y \in \mathbb{R}$. Therefore, the *convolution*

$$u(x,t) = (\Phi_{k,t} * g)(x) \coloneqq \int_{\mathbb{R}} \Phi_k(x-y,t) g(y)\, dy = \frac{1}{\sqrt{4\pi k t}} \int_{\mathbb{R}} e^{-\frac{(x-y)^2}{4kt}} g(y)\, dy \tag{2.2.21}$$

must be another solution on $\mathbb{R} \times \mathbb{R}_{>0}$. Under these hypotheses on the initial condition, this solution is C^∞ on $\mathbb{R} \times \mathbb{R}_{>0}$ - see [7, Section 2.3.1.b]. In fact, if g grows at infinity no more than $ce^{a|x|}$ for two constants $a, c > 0$, it is the unique solution to (2.2.20). This is shown in [22, Section 2.8]. The notation $(\Phi_{k,t} * g)(x)$ is adopted in order to emphasize the convolution operation in space.

Observe that if the initial condition g is bounded, continuous and non-zero somewhere, then at any later time, the solution $\Phi_{k,t} * g$ preserves the sign of the initial condition everywhere in space. We can interpret this as the instantaneous diffusion of heat with *infinite propagation speed*.

We now present a different approach. We will introduce beforehand the definition of the Fourier transform along with some basic properties.

If $u \in L^1(\mathbb{R}^n)$, we define its **Fourier transform** at a point $y \in \mathbb{R}^n$ by

$$\hat{u}(y) = \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}^n} e^{-i\langle x,y\rangle} u(x)\, dx \tag{2.2.22}$$

and its **inverse Fourier transform** at $y \in \mathbb{R}^n$ by

$$\check{u}(y) = \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}^n} e^{i\langle x,y\rangle} u(x)\, dx.$$

We can extend these definitions to the functional space $L^2(\mathbb{R}^n)$ by means of

Plancherel's theorem. Assume $u \in L^1(\mathbb{R}^n) \cap L^2(\mathbb{R}^n)$. Then $\hat{u}, \check{u} \in L^2(\mathbb{R}^n)$ and

$$\|\hat{u}\|_{L^2(\mathbb{R}^n)} = \|\check{u}\|_{L^2(\mathbb{R}^n)} = \|u\|_{L^2(\mathbb{R}^n)}.$$

Let $u \in L^2(\mathbb{R}^n)$. We approximate u by a sequence of functions $\{u_k\}_{k=1}^\infty \subset L^1(\mathbb{R}^n) \cap L^2(\mathbb{R}^n)$ such that

$$\|u_k - u\|_{L^2(\mathbb{R}^n)} \to 0 \quad \text{as } k \to +\infty.$$

According to Plancherel's theorem,

$$\|\hat{u}_k - \hat{u}_j\|_{L^2(\mathbb{R}^n)} = \|\widehat{u_k - u_j}\|_{L^2(\mathbb{R}^n)} = \|u_k - u_j\|_{L^2(\mathbb{R}^n)} \to 0 \quad \text{as } k, j \to +\infty$$

and thus $\{\hat{u}_k\}_{k=1}^\infty$ is a Cauchy sequence in $L^2(\mathbb{R}^n)$. Therefore, this sequence converges to some $\hat{u} \in L^2(\mathbb{R}^n)$, which we define to be the Fourier transform of u. This definition does not depend on the choice of $\{\hat{u}_k\}_{k=1}^\infty$. We similarly define $\check{u}$.

Properties of the Fourier transform. Assume $u, v \in L^2(\mathbb{R}^n)$. Then

[7] See Example (ii) in Section B.1.

(i) $\int_{\mathbb{R}^n} u(x)\bar{v}(x)\,dx = \int_{\mathbb{R}^n} \hat{u}(y)\bar{\hat{v}}(y)\,dy$, where $^-$ denotes the complex conjugate.

(ii) $\widehat{\partial_{x_i}^{\alpha_i} u}(y) = (iy)^{\alpha_i}\hat{u}(y)$.

(iii) $\widehat{(u * v)} = \sqrt{(2\pi)^n}\hat{u}\hat{v}$ if $u, v \in L^1(\mathbb{R}^n) \cap L^2(\mathbb{R}^n)$.

(iv) $u = \check{\hat{u}}$.

For more information on Fourier analysis we refer the reader to [8].

Let us apply the above results to the heat equation. We consider again (2.2.20) and now we apply the Fourier transform to the spatial variable x of the heat equation:

$$\widehat{u_t} = k\widehat{u_{xx}} = k(ix)^2\hat{u} = -kx^2\hat{u} \qquad \text{for } t > 0$$

with initial condition

$$\hat{u} = \hat{g}.$$

Solving this Cauchy problem we obtain

$$\hat{u}(x,t) = \hat{g}(x)e^{-ktx^2}.$$

Consequently

$$u = \left(\hat{g}e^{-ktx^2}\right)^{\check{}} = \frac{g * F}{\sqrt{2\pi}} \tag{2.2.23}$$

where $F(x) = \left(e^{-ktx^2}\right)^{\check{}}$. But then

$$\check{F}(y) = \frac{1}{\sqrt{2\pi}}\int_{\mathbb{R}} e^{ixy-ktx^2}\,dx = \frac{1}{\sqrt{2kt}}e^{-\frac{y^2}{4kt}}.$$

The last equality derives from complex analysis results (for a proof, see [7, pp. 191-192]). Using (2.2.23) we reach the solution to (2.2.20), given by

$$u(x,t) = \frac{1}{\sqrt{4\pi kt}}\int_{\mathbb{R}} e^{-\frac{(x-y)^2}{4kt}} g(y)\,dy \qquad , x \in \mathbb{R}\ , t > 0.$$

We have obtained the same expression as in (2.2.21), from which we can deduce the fundamental solution Φ_k.

2.3 Inhomogeneous Heat Equation

In this section, we assume the existence of an external heat source (for example, an electrical current) or, more generally, any source of what the variable u is modeling.

Let $r = r(x,t)$ be the time rate per unit mass of heat supply by the source. The contribution to the total heat flux due to the source is given by

$$\int_V r(x,t)\rho\,dx.$$

Thus, conservation of energy is now formulated as

$$\int_V e_t(x,t)\rho\,dx = -\int_V \mathrm{div}\mathbf{F}\,dx + \int_V r(x,t)\rho\,dx$$

and the pointwise relation (2.2.2) becomes

$$e_t\rho = -\text{div}\mathbf{F} + r\rho.$$

Applying Fourier's law of heat conduction and heat's linear dependence on absolute temperature in the same fashion as in Subsection 2.2.1 yields the **inhomogeneous** (or **forced**) heat equation

$$u_t = ku_{xx} + f$$

where $f = \frac{r}{c_v}$ is the source term.

2.3.1 The Inhomogeneous Problem. Duhamel's Principle

We now focus on the *inhomogeneous global Cauchy problem*

$$\begin{cases} u_t = ku_{xx} + f(x,t) & ,x \in \mathbb{R}\ ,\ t > 0 \\ u(x,0) = g(x) & ,x \in \mathbb{R}. \end{cases} \tag{2.3.1}$$

As shown earlier, the solution to the homogeneous global Cauchy problem (2.2.20) for any real-valued function $g \in C^0(\mathbb{R}) \cap L^\infty(\mathbb{R})$ is given by

$$u(x,t) = \int_{\mathbb{R}} \Phi_k(x-y,t)g(y)\,dy = \frac{1}{\sqrt{4\pi kt}} \int_{\mathbb{R}} e^{-\frac{(x-y)^2}{4kt}} g(y)\,dy \qquad ,x \in \mathbb{R}\ , t > 0.$$

Duhamel's principle affirms that we are able to construct a solution to an inhomogeneous global Cauchy problem at a point $(x,t) \in \mathbb{R} \times (0,\infty)$ out of its homogeneous solutions. His method consists in the following two steps:

1. Construct a family of solutions to homogeneous global Cauchy problems for different initial times $s \in [0,t]$ and initial data $f(x,s)$.
2. Integrate the above family with respect to s, over $(0,t)$.

Therefore, under weak hypotheses on f (for instance, f belongs to $C^2_1(\mathbb{R} \times [0,\infty))$[8] and is compactly supported) the function

$$u(x,t) = \int_{\mathbb{R}} \Phi_k(x-y,t)g(y)\,dy + \int_{\mathbb{R}} \int_0^t \Phi_k(x-y,t-s)f(y,s)\,dy\,ds \qquad ,x \in \mathbb{R}\ , t > 0 \tag{2.3.2}$$

is a solution to (2.3.1).

The above ideas and formulas can be applied in an analogous way to the *inhomogeneous local Cauchy problem*

$$\begin{cases} u_t = ku_{xx} + f(x,t) & ,0 < x < L,\ t > 0 \\ u(x,0) = g(x) & ,0 \le x \le L \\ + \text{ BCs} & ,t > 0. \end{cases}$$

[8] Following the notation on [7, p. 702],

$$C^2_1(\mathbb{R} \times [0,\infty)) = \{f \colon \mathbb{R} \times [0,\infty) \to \mathbb{R} \mid f, f_x, f_{xx}, f_t \in C^0(\mathbb{R} \times [0,\infty))\}.$$

Chapter 3

Fundamental Elements of Gaussian Stochastic Processes

3.1 Stochastic Processes in Continuous Time

Definition 3.1.1. A *stochastic process* is a family of random variables $\{X_t \mid t \in T\}$ defined on a common probability space $\{\Omega, \mathscr{F}, \mathrm{P}\}$, where t is a parameter running over a suitable index set T, often called the *parameter set* or *domain of definition* of the stochastic process. Stochastic processes are also characterized by their *state space* $\{S, \mathscr{L}\}$. Here, S contains the range of possible values of the process, and $\mathscr{L}$ is a σ-algebra of subsets of S. Furthermore, they are distinguished by the dependence relations between the random variables.

Henceforth we will consider continuous parameter stochastic processes with index set $T \subseteq [0, \infty)$ representing time and state space $S \subseteq \mathbb{R}$ (either countable or uncountable) equipped with $\mathscr{L} = \mathfrak{B}(S)$, the *Borel σ-algebra* of S, where $\mathfrak{B}(\Omega)$ denotes the σ-algebra generated by the open sets of a topological space Ω.

3.1.1 Finite-Dimensional Distributions

Definition 3.1.2. Given a stochastic process $\{X_t \mid t \in T\}$, the *finite-dimensional distributions* are the multidimensional probability laws of any finite family of random variables $\{X_{t_1}, \ldots, X_{t_m}\}$, where $t_1, \ldots, t_m \in T$ and $m \in \mathbb{Z}_{>0}$.

The *law* of a stochastic process is said to exist when all its finite-dimensional distribution functions are given.

Any finite-dimensional distribution function must verify the following *consistency conditions*:

(i) For any permutation $\{k_1, \ldots, k_n\}$ of $\{1, \ldots, n\}$,

$$F_{t_{k_1}, \ldots, t_{k_n}}(x_{k_1}, \ldots, x_{k_n}) = F_{t_1, \ldots, t_n}(x_1, \ldots, x_n).$$

(ii) For any $k \in \{1, \ldots, n\}$ and $x_1, \ldots, x_k \in \mathbb{R}$,

$$F_{t_1, \ldots, t_k}(x_1, \ldots, x_k) = F_{t_1, \ldots, t_n}(x_1, \ldots, x_k, \infty, \ldots, \infty).$$

Definition 3.1.3. A stochastic process can be represented as a random vector $X : \Omega \to \mathbb{R}^T$, where $\mathbb{R}^T$ is the set of real-valued functions defined on T. Moreover, with a suitable σ-algebra

of events in $\mathbb{R}^T$ (for instance, the Borel σ-algebra $\mathfrak{B}(\mathbb{R}^T)$ which generalizes the Borel subsets of $\mathbb{R}^n$), the law of the process can be defined similarly to random vectors as follows:

$$\mathrm{P}_X(B) = \mathrm{P}\big(X^{-1}(B)\big) \quad , B \in \mathfrak{B}\big(\mathbb{R}^T\big).$$

To conclude this subsection, we will state the following theorem by Kolmogorov (1933) that ensures the existence of a stochastic process when given a family of consistent distribution functions. For further details, we refer the reader to [13].

Theorem 3.1.4 (**Kolmogorov Consistency Theorem**). *Consider a family of multidimensional distribution functions*

$$\{F_{t_1,\ldots,t_n}(x_1,\ldots,x_n) \mid t_1 < \ldots < t_n, n \geq 1, t_i \in T\} \tag{3.1.1}$$

satisfying both consistency conditions. There exists a stochastic process such that its finite-dimensional distribution functions are given by (3.1.1).

3.1.2 Sample Paths

If we fix the elements of Ω, we obtain functions defined on T of the form $X_\cdot(\omega)$, where $\omega \in \Omega$. This approach gives rise to the next definition.

Definition 3.1.5. A *sample path* (or *trajectory*) of a stochastic process $\{X_t \mid t \in T\}$ is a function indexed by an element $\omega \in \Omega$ and defined on the parameter set T by

$$\begin{aligned} X_\cdot(\omega)\colon T &\to S \\ t &\mapsto X_t(\omega). \end{aligned}$$

3.1.3 Main Classes of Stochastic Processes

We are interested in the different stochastic structures of a process and the properties of its finite-dimensional distributions. We present hereunder three important classes of random processes, each of them having particular features.

Definition 3.1.6. A stochastic process $\{X_t \mid t \in T\}$ is said to have *independent increments* if for any finite subset $\{t_0,\ldots,t_n\}$ of T such that $t_0 < \ldots < t_n$, the increments $X_{t_1} - X_{t_0},\ldots,X_{t_n} - X_{t_n-1}$ are independent random variables.

It follows that any finite-dimensional distribution is completely determined by the laws of X_t and $X_{t_2} - X_{t_1}$ for all $t_1, t_2 \in T$ such that $t_1 < t_2$.

Definition 3.1.7. If for any $t_1, t_2 \in T$ such that $t_1 < t_2$, the law of the increment $X_{t_2} - X_{t_1}$ is a function depending only on $t_2 - t_1$, then the process has *stationary increments.*

Definition 3.1.8. When the finite dimensional distributions have time-shift invariance, $\{X_t \mid t \in T\}$ is a *strictly stationary* process.

3.2 Multivariate Normal Distribution

Let $n \in \mathbb{Z}_{>0}$ throughout this section.

3.2.1 Preliminaries

This subsection is contained in Appendix A. Here we will only give the definition of the multivariate Normal distribution (Definition A.2.3), a central concept of this thesis. This definition is chosen according to [12].

Definition 3.2.1. A real random vector $X = (X_1, \ldots, X_n)$ has a *multivariate Normal* (or *Gaussian*) distribution if every linear combination of its components has a (univariate) Normal distribution.

We denote by $N(\mu, Q)$ its law, where $\mu = (\mu_1, \ldots, \mu_n) \in \mathbb{R}^n$ is the mean vector and Q is the covariance matrix.

3.2.2 Characteristic Functions

Definition 3.2.2. The *Fourier transform* of a probability measure μ is a complex-valued function defined on $\mathbb{R}^n$ given by

$$\hat{\mu}(t) = \int_{\mathbb{R}^n} e^{i\langle t,x\rangle} \mu(dx).$$

Remark 3.2.3. We have already encountered a definition of a Fourier transform in this document - see (2.2.22). Definition 3.2.2 extends this notion to probability measures.

Remark 3.2.4. Note that

$$e^{i\langle t,x\rangle} = \cos(\langle t, x\rangle) + i \sin(\langle t, x\rangle)$$

and in particular

$$\left| e^{i\langle t,x\rangle} \right| = 1.$$

We can then rewrite the Fourier transform as

$$\hat{\mu}(t) = \int_{\mathbb{R}^n} \cos(\langle t, x\rangle)\mu(dx) + i \int_{\mathbb{R}^n} \sin(\langle t, x\rangle)\mu(dx).$$

The last expression is well-defined since both $\cos(\cdot)$ and $\sin(\cdot)$ are bounded and Borel measurable functions, hence integrable in the sense of Lebesgue.

Definition 3.2.5. The *characteristic function* of a probability measure μ on $\mathbb{R}^n$ is defined as its Fourier transform. We denote it by $\varphi_\mu(t)$.

Likewise, real random vectors also possess characteristic functions defined in an analogous fashion.

Definition 3.2.6. Let X be a real n-dimensional random vector. The characteristic function of X is given by the characteristic function of its probability distribution:

$$\varphi_X(t) = \int_{\mathbb{R}^n} e^{i\langle t,x\rangle} \mathrm{P}_X(dx) = \mathrm{E}\left[e^{i\langle t,X\rangle}\right]. \tag{3.2.1}$$

Proposition 3.2.7 (Fundamental Properties of Characteristic Functions). *Let μ be a probability measure on $\mathbb{R}^n$. Then*

(i) $\varphi_\mu(0) = 1$.

(ii) $|\varphi_\mu(t)| \leq 1 \quad \forall t \in \mathbb{R}^n$.

(iii) $\varphi_\mu(-t) = \overline{\varphi_\mu}(t) \quad \forall t \in \mathbb{R}^n$, *where $\overline{\varphi_\mu}$ denotes the complex conjugate function of φ_μ.*

(iv) φ_μ *is uniformly continuous.*

(v) Let X *be a random vector in* $\mathbb{R}^n$*. Let* A *be a* $m \times n$ *real matrix and* $b \in \mathbb{R}^m$*. Then*

$$\varphi_{AX+b}(t) = e^{i\langle t,b\rangle}\varphi_X\big(A^T t\big) \tag{3.2.2}$$

where A^T *denotes the transpose of* A*.*

Remark 3.2.8. Property (ii) shows that φ_μ is a bounded function, while property (iii) implies that φ_μ is Hermitian.

Characteristic functions of random variables are also related with their independence. The next proposition establishes two properties concerning this linkage.

Proposition 3.2.9 (**Independence Properties**)**.** *Let* $X = (X_1, \ldots, X_n)$ *be a real-valued random vector.*

(i) The random variables $X_1, \ldots, X_n$ *are independent if and only if*

$$\varphi_X(t_1, \ldots, t_n) = \varphi_{X_1}(t_1) \cdots \varphi_{X_n}(t_n).$$

(ii) If $X_1, \ldots, X_n$ *are independent random variables, then*

$$\varphi_{X_1+\ldots+X_n}(t) = \varphi_{X_1}(t) \cdots \varphi_{X_n}(t).$$

Proof. It follows from the linearity of the expectation operator. □

The following theorem provides a notable relation between characteristic functions and finite moments. For a proof, see [12, Theorem 13.2].

Theorem 3.2.10. *Let* X *be a random vector in* $\mathbb{R}^n$*. Suppose* $\mathrm{E}\big[|X|^m\big] < \infty$ *for some strictly positive integer* m*. Then,* φ_X *has continuous partial derivatives up to order* m *and*

$$\frac{\partial^m}{\partial t_{j_1} \cdots \partial t_{j_m}} \varphi_X(t) = i^m \mathrm{E}\big[X_{j_1} \cdots X_{j_m} e^{i\langle t,X\rangle}\big].$$

Equivalently, replacing φ_X *and* $\mathrm{E}\big[X_{j_1} \cdots X_{j_m} e^{i\langle t,X\rangle}\big]$ *by* φ_μ *and* $\int_{\mathbb{R}^n} t_{j_1} \cdots t_{j_m} e^{i\langle t,x\rangle}\mu(dx)$*, respectively, produces the same result for any probability measure* μ*.*

Corollary 3.2.11. *In the one-dimensional case, we obtain the following formulae:*

(i) $\varphi_X^{(k)}(0) = i^k \mathrm{E}\big[X^k\big]$*.*

(ii) $\varphi_\mu^{(k)}(0) = i^k \int_{\mathbb{R}} t^k \mu(dx)$*.*

Remark 3.2.12. These expressions provide a direct method to calculate the moments of random variables. For the first two moments, we have

$$\begin{aligned} \mathrm{E}[X] &= i\varphi_X'(0) \qquad \text{if} \quad \mathrm{E}[|X|] < \infty, \\ \mathrm{E}\big[X^2\big] &= -\varphi_X''(0) \qquad \text{if} \quad \mathrm{E}\big[|X|^2\big] < \infty. \end{aligned}$$

where X is a real-valued random variable.

The next statement is a sort of reciprocal of the previous theorem for $n = 1$.

Theorem 3.2.13. *Let* μ *be a probability measure in* $(\mathbb{R}, \mathfrak{B}(\mathbb{R}))$*. Suppose* φ_μ *is* k *times differentiable in a neighborhood of* 0*. Then* μ *has moments up to order* $2m$*, where* $m \in \mathbb{Z}_{>0}$ *and* $2m \leq k$*.*

We will now apply the above results to the Normal distribution. To this end, we will previously work out the characteristic function of the standard Normal distribution and then make a change of variables for the general case.

Example 3.2.14. It is well known that the density function of a standard Normal random variable is given by

$$f(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}.$$

Thus, for $\mu = N(0,1)$,

$$\varphi_\mu(t) = \int_{\mathbb{R}} e^{itx} \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}\, dx.$$

Since the function $x \mapsto \sin(tx)e^{-\frac{x^2}{2}}$ is odd and integrable over the real line for any $t \in \mathbb{R}$, we have that

$$\varphi_\mu(t) = \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} \cos(tx) e^{-\frac{x^2}{2}}\, dx.$$

By Theorem 3.2.10, we can differentiate both sides with respect to t and commute the derivative with respect to t with the integral over $\mathbb{R}$, yielding

$$\varphi'_\mu(t) = -\frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} x \sin(tx) e^{-\frac{x^2}{2}}\, dx.$$

Integrating by parts gives us an ordinary differential equation of the form

$$\varphi'_\mu = -t\varphi_\mu.$$

Setting Fundamental Property (i) as the initial condition of the Cauchy Problem, the unique solution is therefore

$$\varphi_\mu(t) = e^{-\frac{t^2}{2}}.$$

Note that this is a real-valued function. In fact, it is both cause and consequence of the symmetry of the standard Normal law. A nice proof can be found in [26, pp. 77-78].

Moreover, it is clear that φ_μ is C^∞. By Theorem 3.2.13, μ has finite moments of every order. More precisely, the series expansion

$$e^{-\frac{t^2}{2}} = \sum_{n=0}^{\infty} \frac{(-1)^n t^{2n}}{n! 2^n}$$

yields

$$\varphi_\mu^{(k)}(0) = \begin{cases} 0 & \text{if} \quad k = 2n+1 \\ \frac{(-1)^n (2n)!}{n! 2^n} & \text{if} \quad k = 2n \end{cases}$$

for any choice of $n \in \mathbb{Z}_{\geq 0}$. Hence, by Corollary 3.2.11,

$$\mathrm{E}\big[X^k\big] = \begin{cases} 0 & \text{if} \quad k = 2n+1 \\ \frac{(2n)!}{n! 2^n} & \text{if} \quad k = 2n \end{cases} \tag{3.2.3}$$

for $X \sim N(0,1)$[1].

[1] As stated in Section A.1, $X \sim N(0,1)$ means that the distribution of X is $N(0,1)$. This notation will be used throughout this document for any random variable and distribution.

Let X be a random variable with the standard Normal law. Let Y be the random variable given by the transformation

$$Y = \tilde{\mu} + \sigma X \qquad (\tilde{\mu} \in \mathbb{R}, \sigma > 0).$$

Then, by Proposition A.1.3, Y has a Normal distribution with expectation $\tilde{\mu}$ and variance σ^2. To compute its characteristic function, we use (3.2.2). We obtain

$$\varphi_Y(t) = e^{it\tilde{\mu} - \frac{\sigma^2 t^2}{2}}. \tag{3.2.4}$$

3.2.3 Characterization of the Multivariate Normal Distribution

Our aim is to determine a probability distribution with the information given by the associated characteristic function. In this subsection we will show that the characteristic function serves as a label for the distribution function of a random variable. In addition, we will apply these results to Gaussian random vectors.

The following theorem provides a formula to obtain a distribution function from a characteristic function.

Theorem 3.2.15 (**Inversion Formula**). *Let μ be a probability measure on $(\mathbb{R}, \mathfrak{B}(\mathbb{R}))$ with characteristic function φ and distribution function F. Let a, b be two continuity points of F (i.e. they belong to the dense subset $D = \{x \in \mathbb{R} \mid F(x^-) = F(x)\} \subset \mathbb{R}$) such that $a < b$. Then*

$$F(b) - F(a) = \lim_{\sigma \downarrow 0} \frac{1}{2\pi} \int_{\mathbb{R}} \varphi(t) e^{-\frac{\sigma^2 t^2}{2}} \frac{e^{-ita} - e^{-itb}}{it}\, dt. \tag{3.2.5}$$

Proof. Let X, Y_σ be two independent random variables with laws μ and $N(0, \sigma^2)$, respectively. Applying property (ii) in Proposition 3.2.9 we deduce that the characteristic function of the sum $Z_\sigma = X + Y_\sigma$ is given by

$$\varphi_{Z_\sigma}(t) = \varphi_X(t)\varphi_{Y_\sigma}(t) = \varphi(t) e^{\frac{-\sigma^2 t^2}{2}}.$$

We denote by F_σ the distribution function of Z_σ, and we assume the following equality holds:

$$F_\sigma(b) - F_\sigma(a) = \frac{1}{2\pi} \int_{\mathbb{R}} \varphi(t) e^{-\frac{\sigma^2 t^2}{2}} \frac{e^{-ita} - e^{-itb}}{it}\, dt. \tag{3.2.6}$$

Letting $\sigma \downarrow 0$, we notice that $Z_\sigma \to X$ in L^2 from the fact that $E[(Z_\sigma - X)^2] = E(Y_\sigma^2) = \sigma^2 \to 0$. This implies the convergence in distribution of Z_σ to X. In particular, $F_\sigma(b) - F_\sigma(a)$ converges pointwise to $F(b) - F(a)$.

It remains to show the identity (3.2.6). On one hand,

$$\begin{aligned} F_\sigma(b) &= \mu(X + Y_\sigma \le b) \\ &= \int_{x+y \le b} \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{y^2}{2\sigma^2}}\, \mu(dx)\, dy \\ &= \int_{\mathbb{R}} \left(\int_{-\infty}^{b-x} \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{y^2}{2\sigma^2}}\, dy \right) \mu(dx)\ , \end{aligned}$$

thus,

$$F_\sigma(b) - F_\sigma(a) = \frac{1}{\sqrt{2\pi\sigma^2}} \int_{\mathbb{R}} \int_{a-x}^{b-x} e^{-\frac{y^2}{2\sigma^2}}\, dy\, \mu(dx).$$

On the other hand,

$$\frac{1}{2\pi}\int_{\mathbb{R}} \varphi(t)e^{-\frac{\sigma^2t^2}{2}}\frac{e^{-ita}-e^{-itb}}{it}\,dt = \frac{1}{2\pi}\int_{\mathbb{R}}\int_a^b \varphi(t)e^{-\frac{\sigma^2t^2}{2}}e^{-itx}\,dx\,dt.$$

Moreover,

$$\frac{1}{2\pi}\int_a^b\int_{\mathbb{R}} \varphi(t)e^{-\frac{\sigma^2t^2}{2}}e^{-itx}\,dt\,dx = \frac{1}{2\pi}\int_a^b\int_{\mathbb{R}}\left(\int_{\mathbb{R}} e^{ity}\,\mu(dy)\right)e^{-\frac{\sigma^2t^2}{2}}e^{-itx}\,dt\,dx.$$

It is clear that the function $(t,y)\mapsto e^{it(y-x)-\frac{\sigma^2t^2}{2}}$ is Lebesgue integrable with respect to the product measure $\mu(dy)\,dt$. Hence, we can apply Fubini's theorem to obtain

$$\begin{aligned}\frac{1}{2\pi}\int_a^b\int_{\mathbb{R}} \varphi(t)e^{-\frac{\sigma^2t^2}{2}}e^{-itx}\,dt\,dx &= \frac{1}{2\pi}\int_a^b\int_{\mathbb{R}}\left(\int_{\mathbb{R}} e^{it(y-x)-\frac{\sigma^2t^2}{2}}\,dt\right)\mu(dy)\,dx\\ &= \frac{1}{\sqrt{2\pi\sigma^2}}\int_a^b\int_{\mathbb{R}} e^{-\frac{(y-x)^2}{2\sigma^2}}\,\mu(dy)\,dx\\ &= \frac{1}{\sqrt{2\pi\sigma^2}}\int_{\mathbb{R}}\int_{a-y}^{b-y} e^{-\frac{u^2}{2\sigma^2}}\,du\,\mu(dy).\end{aligned}$$

□

Remark 3.2.16. The factor $e^{-\frac{\sigma^2t^2}{2}}$ ensures convergence of the improper integral in (3.2.5).

Theorem 3.2.17 (**Multidimensional Inversion Formula**). *Let μ be a probability measure on $(\mathbb{R}^n,\mathfrak{B}(\mathbb{R}^n))$ with characteristic function φ. Let $a,b\in\mathbb{R}^n, a<b$.*
If $\{x\in\mathbb{R}^n \mid a<x<b\}$ is a continuity set of μ, then

$$\mu(\{a<x<b\}) = \frac{1}{(2\pi)^n}\lim_{M_1\to\infty}\cdots\lim_{M_n\to\infty}\int_{-M}^{M}\prod_{i=1}^{n}\left(\frac{e^{-it_ia_i}-e^{-it_ib_i}}{it_i}\right)\varphi(t)\lambda(dt) \qquad (3.2.7)$$

where $M=(M_1,\dots,M_n), t=(t_1,\dots,t_n)\in\mathbb{R}^n$ and λ denotes the Lebesgue measure on $\mathbb{R}^n$..

Proof. We refer the reader to [4, Theorem 2.3.1]. □

This leads to the following essential property.

Corollary 3.2.18 (**Injectivity Property**). *If μ and ν are two probability measures on $(\mathbb{R}^n,\mathfrak{B}(\mathbb{R}^n))$ that admit the same characteristic function, then they are identical. In other words, a probability measure on $(\mathbb{R}^n,\mathfrak{B}(\mathbb{R}^n))$ is completely determined by its characteristic function.*

Proof. We give the proof only for the case $n=1$.

Theorem 3.2.15 implies that for any pair of continuity points $a,b\in\mathbb{R}$ such that $a<b$, the following equality holds:

$$F_\mu(b)-F_\mu(a) = F_\nu(b)-F_\nu(a),$$

where F_μ and F_ν are the distribution functions of μ and ν, respectively.

If $a\to\infty$ then $F_\mu(b)=F_\nu(b)$ in a dense subset of $\mathbb{R}$, hence $F_\mu\equiv F_\nu$. □

In the case where the characteristic function φ is integrable according to Lebesgue, the Inversion Formula can be written as a simpler connection between the characteristic function and the density function of a random variable.

Proposition 3.2.19 (**Inversion Formula for Fourier transforms**). *Suppose φ is the characteristic function of a probability measure μ on $(\mathbb{R}, \mathfrak{B}(\mathbb{R}))$. If φ is integrable according to Lebesgue, then it is also absolutely continuous and the associated density function is given by the formula*

$$f(x) = \frac{1}{2\pi} \int_{\mathbb{R}} e^{-itx} \varphi(t)\, dt.$$

Remark 3.2.20. Absolute continuity of μ with respect to Lebesgue measure *does not* imply integrability of φ.

We now have all the necessary ingredients to fully characterize an $\mathbb{R}^n$-valued Normal random variable, as well as the independence of its components.

Theorem 3.2.21. *X is an $\mathbb{R}^n$-valued Normal random variable if and only if its characteristic function is given by*

$$\varphi_X(t) = e^{i\langle t,\mu\rangle - \frac{1}{2}\langle t, Qt\rangle}, \tag{3.2.8}$$

where μ is the mean vector and Q is the covariance matrix.

Proof. Let $Y = \sum_{j=1}^n a_j X_j = \langle a, X\rangle$ be a linear combination of all of the components of X.

Sufficient condition. Assume (3.2.8) holds. Formula (3.2.2) implies that for any $v \in \mathbb{R}$,

$$\varphi_Y(v) = \varphi_X(va) = e^{iv\langle a,\mu\rangle - \frac{v^2}{2}\langle a, Qa\rangle}.$$

Notice that this expression represents the characteristic function of a Normal random variable with mean $\langle a, \mu\rangle$ and variance $\langle a, Qa\rangle$. Thus, by Corollary 3.2.18, Y is Normal and, by definition, X is also Normal.

Necessary condition. Suppose now X is Gaussian. If $Q = \mathrm{Cov}(X)$, then $\mathrm{E}[Y] = \langle a, \mu\rangle$ and $\mathrm{Var}(Y) = \langle a, Qa\rangle$. Since Y is Normal by hypothesis, using (3.2.4) we have

$$\varphi_Y(v) = e^{iv\langle a,\mu\rangle - \frac{v^2}{2}\langle a, Qa\rangle}.$$

Lastly, choosing $v = 1$ yields

$$\varphi_Y(v) = \varphi_Y(1) = \varphi_{\langle a, X\rangle}(1) = \mathrm{E}\left[e^{i\langle a, X\rangle}\right] = \varphi_X(a). \qquad \square$$

Corollary 3.2.22. *Let $X = (X_1, \ldots, X_n)$ be a real-valued Gaussian random vector. The components of X are Normal and independent random variables if and only if the covariance matrix Q is diagonal.*

In particular, two components X_j and X_k are independent if and only if they are uncorrelated.

3.2.4 Fundamental Properties of the Multivariate Normal Distribution

1. Let X be a random vector in $\mathbb{R}^n$ with mean vector μ. Then there exists
   (a) a family of independent Normal random variables $Y_1, \ldots, Y_n$ such that
   $$Y_j \sim N(0, \lambda_j) \quad \forall j \in \{1, \ldots, n\},$$
   where $\lambda_j \geq 0 \quad \forall j \in \{1, \ldots, n\}$;
   (b) an orthogonal matrix A such that
   $$X = \mu + AY.$$

2. A Gaussian random vector X in $\mathbb{R}^n$ has a probability density if and only if the associated covariance matrix Q is *non-degenerate*, that is, $\det(Q) \neq 0$.

   If such a density exists, it is given by

$$f_X(x) = \frac{1}{\sqrt{(2\pi)^n \det(Q)}} e^{-\frac{1}{2}\langle x-\mu, Q^{-1}(x-\mu)\rangle}. \tag{3.2.9}$$

   Remark 3.2.23. This property shows that for $n > 2$ there exist non-constant random variables without associated densities.

   Example 3.2.24. Let us examine the case $n = 2$. Suppose $X \sim N(\mu_X, \sigma_X^2), Y \sim N(\mu_Y, \sigma_Y^2)$ and let ρ be the correlation coefficient. Then

$$Q = \begin{pmatrix} \sigma_X^2 & \rho\sigma_X\sigma_Y \\ \rho\sigma_X\sigma_Y & \sigma_Y^2 \end{pmatrix}$$

   and

$$\det(Q) > 0 \iff |\rho| < 1.$$

   By (3.2.9), the density of the random vector (X, Y) is

$$f_{(X,Y)}(x,y) = \frac{1}{2\pi\sigma_X\sigma_Y\sqrt{1-\rho^2}} e^{-\frac{1}{2(1-\rho^2)}\left[\frac{(x-\mu_X)^2}{\sigma_X^2} - 2\rho\frac{(x-\mu_X)(y-\mu_Y)}{\sigma_X\sigma_Y} + \frac{(y-\mu_Y)^2}{\sigma_Y^2}\right]}.$$

3. Let X be a random vector in $\mathbb{R}^n$ with law $N(\mu, Q)$. Let A be a $r \times n$ real matrix. Then

$$AX \sim N(A\mu, AQA^T).$$

4. Let X and Y be Gaussian random vectors in $\mathbb{R}^n$ and $\mathbb{R}^m$, respectively. If X and Y are independent, then $Z = (X, Y)$ is a Gaussian random vector in $\mathbb{R}^{n+m}$.

Remark 3.2.25. A random vector in $\mathbb{R}^n$ might not be Gaussian despite having marginal Normal distributions.

We will illustrate this with an example [12, Example 2, p. 134].

Example 3.2.26. Let $Y \sim N(0,1)$. Let $Z = Y\mathbb{1}_{|Y|\leq a} - Y\mathbb{1}_{|Y|\geq a}$. Then Z is also $N(0,1)$ but $Y + Z = 2Y\mathbb{1}_{|Y|\leq a}$ is not Normal, since $\mathrm{P}(Y + Z > 2|a|) = 0$ and $Y + Z$ is not a.s. equal to a constant. Thus, the random vector (Y, Z) is not Gaussian, even though its two components are Normal.

3.3 Gaussian Stochastic Processes

Definition 3.3.1. A stochastic process $\{X_t \mid t \in T\}$ is said to be *Gaussian* if any of its finite-dimensional distributions is Gaussian.

The existence of this kind of processes is a consequence of Theorem 3.1.4. One can find a nice proof in [23, pp. 6–7]

3.3.1 Characterization

Let $m \in \mathbb{Z}_{>0}$. Let $\{t_1, \ldots, t_m\}$ be a subset of T such that $t_1 < \ldots < t_m$. The law of the random vector $X_{t_1,\ldots,t_m} = (X_{t_1}, \ldots, X_{t_m})$ is characterized by the following two parameters:

(i) *Mean vector*:

$$\mu_{t_1,\ldots,t_m} = \mathrm{E}\big[X_{t_1,\ldots,t_m}\big] = \big(\mathrm{E}\big[X_{t_1}\big], \ldots, \mathrm{E}\big[X_{t_m}\big]\big)\,.$$

(ii) *Covariance matrix*:

$$Q_{t_1,\ldots,t_m} = \Big(\mathrm{Cov}\big(X_{t_i}, X_{t_j}\big)\Big)_{1\le i,j\le m}.$$

As previously stated, if $\det(Q_{t_1,\ldots,t_m}) > 0$, then the random vector $X_{t_1,\ldots,t_m}$ has a density function given by

$$f_{t_1,\ldots,t_m}(x) = \frac{1}{\sqrt{(2\pi)^m \det(Q_{t_1,\ldots,t_m})}} e^{-\frac{1}{2}\langle x-\mu_{t_1,\ldots,t_m}, Q^{-1}_{t_1,\ldots,t_m}(x-\mu_{t_1,\ldots,t_m})\rangle}.$$

These parameters can be extended to the whole family of random variables that constitute the stochastic process by means of the functions

$$\mu(t) = \mathrm{E}[X_t]$$

and

$$Q(s,t) = \mathrm{Cov}(X_s, X_t) = \mathrm{E}\big[(X_s - \mu(s))(X_t - \mu(t))\big]$$

called, respectively, the *mean function* and the *covariance function* of the stochastic process $\{X_t \mid t \in T\}$.

3.3.2 Centered Gaussian Processes

Definition 3.3.2. We say that a Gaussian process $\{X_t \mid t \in T\}$ is *centered* if for all $t_1, \ldots, t_k \in T$, $k \in \mathbb{Z}_{>0}$, the random vector $(X_{t_1}, \ldots, X_{t_k})$ is centered Normal, that is,

$$\mathrm{E}\big[(X_{t_1}, \ldots, X_{t_k})\big] = 0.$$

This is equivalent to the simpler condition

$$\mathrm{E}\big[X_t\big] = 0 \qquad \forall t \in T.$$

The finite-dimensional distributions of such a stochastic process are completely determined by its covariance function

$$Q(s,t) = \mathrm{E}\big[X_s X_t\big].$$

Chapter 4

Examples of Gaussian Processes

4.1 Brownian Motion

4.1.1 A Little History

What is nowadays known as Brownian motion dates back to the Summer of 1827, when a Scottish botanist called Robert Brown observed a "rapid oscillatory motion" of microscopic pollen grains suspended in water. Although he was not the first person to observe it directly, he stood out for repeating the experiment of observing the irregular, erratic, ceaseless motion of suspended particles, as well as demonstrating its presence in inorganic and organic matter. A similar phenomenon was apparent whenever very small particles were suspended in a fluid medium. He described an experiment in which a minute drop of water containing one single particle, immersed in oil, exhibits the motion unabated, thus refuting some physical explanations such as fluid currents or interactions between particles. He believed matter is composed of small particles moving at an irregular rapid pace due to the particles themselves and not the surrounding fluid.

A successful explanation had to wait until 1905, when Albert Einstein published an article explaining with probabilistic tools the displacements of suspended particles by the collisions with the molecules composing the surrounding liquid. In the years between, not many discoveries were made, mainly because the conducted experiments were incomplete and too qualitative. There was a general lack of interest among physicists, and the statistical methods of kinetic theory of gases were not yet sufficiently developed. Moreover, the interpretations on the origin of Brownian motion differed, among which were the kinetic theory of gases (also called atomic hypothesis) which Einstein would later develop. Nevertheless, some interesting results were published, such as Siegmund Exner's observation of the correlation between the liquid's temperature and the motion's intensity, or Louis Georges Gouy's indication that the atomic hypothesis violates the second principle of thermodynamics.

One of Einstein's *Annus Mirabilis* papers was an article on Brownian motion in which he obtained two major results: the relation between the diffusion coefficient and the properties of the medium, and the correspondence between Brownian motion and diffusion. His purpose was to present a test for the validity of the kinetic theory of gases. To this end, he defined a measurable quantity, namely the *mean of the squares of displacements*, formally written as

$$\lambda_x = \sqrt{\frac{RT}{N_A}\frac{1}{3\pi\mu a}}\sqrt{t}$$

where R is the gas constant, T the temperature, N_A Avogadro constant, μ the fluid's viscosity, and a the radius of a particle.
His reasoning followed three steps: first, he related the diffusion coefficient to the properties

of the medium with two physical ingredients: the viscous force from Stokes' law and the van 't Hoff law which relates the pressure increase with the concentration of particles. He obtained the prototype of the *fluctuation-dissipation theorem* by considering the two aforementioned forces at equilibrium. Secondly, he studied the irregular movement of particles suspended in a liquid and, under a series of hypotheses on the particle's motion, he derived the diffusion equation by using probability distributions in the statistical mechanics' way. Lastly, combining the results from the previous steps, he obtained a formula for the mean of the squares of displacements, which is probably the most famous result on Brownian motion. Further to this, he computed its numerical value. Einstein's argument did not provide a dynamical theory of Brownian motion. Instead, he determined the motion's nature and the value of the diffusion coefficient under a series of assumptions. His key idea was the introduction of Brownian motion as a stochastic process. A noteworthy application was to estimate Avogadro's constant.

The description of Brownian motion as a physical concept also included two notable works subsequent to Einstein's. In 1906, Marian von Smoluchowski published a stylistically different article in which he independently obtained a diffusion coefficient comparable to Einstein's. He took into account the experimental work done on the subject and he built his theory to account for these observations. Unlike Einstein, his calculations were directly based on collisions between particles, and he introduced stochasticity by the mean of average quantities. His model consisted on a random walk in which the suspended particle traveled on a straight line at constant velocity between two collisions. Two years later, the French physicist Paul Langevin published his only article on the subject, in full knowledge of Einstein's and Smoluchowski's works. Instead of constructing a deterministic equation with probability distributions as unknowns, he used a probabilistic equation with a fluctuation force described by a random variable, now known as *Langevin's equation*. Mathematically, it can be written as

$$m\frac{d^2x}{dt^2} = -6\pi a\frac{dx}{dt} + X$$

where X is a stochastic fluctuation force. His derivation was simpler and innovative for the time, since it contains a stochastic equation.

A rigorous mathematical treatment of Brownian motion would not appear until 1921, when Norbert Wiener, a young American mathematician, published his first article on the subject. Previously, he became interested in integration theory and decided to apply his ideas on this field to Brownian motion after an unsuccessful attempt at turbulence theory. A pioneer in the field, he remained interested for a decade, being joined eventually by remarkable mathematicians such as Kolmogorov or Lévy. Having studied Einstein's and Smoluchowski's theories, he considered them lacking of the mathematical properties of the curves followed by single particles. His work was guided by the study of trajectories and its functions. We can divide his work into three periods, for each of which he developed a different model of Brownian motion.

The first period occurs from 1920 to 1922, when he developed his ideas on functional averages, along with an axiomatic theory of integration without measure theory, which was not yet fully developed. Percy John Daniell, in 1918, proposed an axiomatic theory of integration independent from measure theory. Wiener adopted his ideas to explicitly compute functional averages over function spaces. His 1920 article exposed his progress on Daniell's integration and established the foundations of his axiomatic theory which he would later apply to Brownian motion. The following year he published two articles in which he developed Einstein's work on Brownian motion. In the first one he explicitly addressed Brownian motion for the first time. He referenced Einstein's work and constructed an idealization of the Brownian motion considering infinitesimally sized molecules and continuously described collisions, thereby obtaining a simplified probability distribution for the position of the par-

ticles. This mathematical model extended the range of validity of the Gaussian distribution. Furthermore, he defined general functionals of functions describing Brownian trajectories, as well as functional means computed with means of functions. The next article was the only one that did not offer mathematical innovations. He took a different position with respect to his other articles by assuming that the velocities of Brownian particles were well-defined. His target was to show that Einstein's hypothesis on the independence of increments over time was not necessary. In 1922, Wiener's ideas on axiomatic integration culminated. He deduced that all bounded, continuous functionals defined on Brownian trajectories were summable.

It was in 1923 when Wiener set forth the modern mathematical foundation of Brownian motion, which would eventually be named as *Wiener process*. In his seminal article *Differential Space*, he proved the existence of standard Brownian motion and provided the first argument of non-differentiability of Brownian trajectories. After a series of exchanges with Paul Lévy, he took an approach based on the French mathematician, in which the averages of functionals were defined as the limit of averages of sphere volumes. Assuming that the increments over regular and disjoint time intervals are independent, Wiener measured the inner region of the spheres from Lévy's formulation, thus giving a prototype of what today is known as *Wiener measure*. Since it had the same form as his probability distribution from 1921, it illustrated the usefulness of Lévy's spheres as a tool to study Brownian trajectories.

The following decade saw more advancements in Brownian motion. Wiener published a thesis in 1930 on harmonic analysis describing Brownian motion in terms of Lebesgue measure. Three years later, along with fellow mathematicians Paley and Zygmund, they published an article aiming to unify their ideas on randomness in analysis. Interestingly, the definition they gave of the displacement carried by moving particles was similar to the way stochastic processes are currently defined. The most important result of the article was the proof of non-differentiability of Brownian trajectories.

4.1.2 Definition and Characterization

The terms Brownian motion and Wiener process will be used interchangeably.

Definition 4.1.1. Let $x_0 \in \mathbb{R}$. A real-valued stochastic process $\{B_t \mid t \geq 0\}$ defined on a probability space $(\Omega, \mathcal{F}, \mathrm{P})$ is a *one-dimensional Brownian motion starting at* x_0 if the process is centered Gaussian with covariance function given by

$$Q(s,t) = \mathrm{E}\big[B_s B_t\big] = s \wedge t \coloneqq \min\{s,t\}. \tag{4.1.1}$$

We say that $\{B_t \mid t \geq 0\}$ is a *one-dimensional standard Brownian motion* if $x_0 = 0$.

Remark 4.1.2. One can define a one-dimensional Brownian motion $\{W_t \mid t \geq 0\}$ starting at an arbitrary point $x \in \mathbb{R}$ in terms of the one-dimensional standard Brownian motion $\{B_t \mid t \geq 0\}$ as follows:

$$W_t \coloneqq x + B_t.$$

Unless otherwise stated, we will consider Brownian motion to be one-dimensional standard.

Proposition 4.1.3 (**Characterization of Brownian Motion**). *A real-valued stochastic process $B = \{B_t \mid t \geq 0\}$ is a Brownian motion if and only if the following holds:*

(i) $B_0 = 0$ and for any given $t > 0$, the distribution of B_t is $N(0,t)$.

(ii) The process has independent increments: for any $0 \leq t_1 \leq t_2 \leq \ldots \leq t_n$, the increments $B_{t_n} - B_{t_{n-1}}, B_{t_{n-1}} - B_{t_{n-2}}, \ldots, B_{t_2} - B_{t_1}$ are independent random variables.

(iii) For any $0 \leq s \leq t$, the increment $B_t - B_s$ has the same distribution as that of the random variable B_{t-s}. Therefore, the process has stationary increments.

Proof. Necessary condition. Let B be a Brownian motion.

(i) Using (4.1.1) we deduce that $\mathrm{E}\big[B_0^2\big] = 0$ and for any $t > 0$, $\mathrm{E}\big[B_t^2\big] = t$. Thus, $B_0 = 0$ and $B_t \sim N(0, t)$.

(ii) Let $0 \leq t_i \leq t_j \leq t_k \leq t_l \leq t_n$. Again, using (4.1.1) we have

$$\mathrm{E}\big[\,\big(B_{t_j} - B_{t_i}\big)\,\big(B_{t_l} - B_{t_k}\big)\,\big] = t_j \wedge t_l - t_j \wedge t_k - t_i \wedge t_l + t_i \wedge t_k = t_j - t_j - t_i + t_i = 0.$$

Since B is centered, this implies

$$\mathrm{E}\big[\,\big(B_{t_j} - B_{t_i}\big)\,\big(B_{t_l} - B_{t_k}\big)\,\big] = \mathrm{E}\big[B_{t_j} - B_{t_i}\big]\mathrm{E}\big[B_{t_l} - B_{t_k}\big].$$

The independence follows from the Gaussian nature of the finite-dimensional distributions of B.

(iii) Using (i), we only need to show that $B_t - B_s \sim N(0, t-s)$. We know that finite linear combinations of Normal random variables are also Normal. Since $\mathrm{E}\big[B_t - B_s\big] = 0$ by linearity, we have

$$\mathrm{E}\Big[\,\big(B_t - B_s\big)^2\,\Big] = t - 2s + s = t - s$$

and thus $\mathrm{Var}\big(B_t - B_s\big) = t - s$.

Sufficient condition. Assume that (i), (ii) and (iii) hold. Let $0 \leq t_1 \leq t_2 \leq \ldots \leq t_n$. By Fundamental Property 4 of the Gaussian distribution, the random vector $Z = \big(B_{t_1}, B_{t_2} - B_{t_1}, \ldots, B_{t_n} - B_{t_{n-1}}\big)$ is Gaussian. Moreover, Fundamental Property 3 of this same distribution implies that any linear transformation of Z preserves the law. Therefore, the finite-dimensional distributions of any linear transformation of Z are Gaussian. Clearly, $\mathrm{E}\big[B_t\big] = 0$ for any $t \geq 0$, hence B is a centered Gaussian process.

It only remains to verify (4.1.1). Let $0 \leq s \leq t$. By independence of the increments, we obtain

$$\begin{aligned}\mathrm{E}\big[B_t B_s\big] &= \mathrm{E}\big[\,\big(B_t - B_s + B_s\big)\,B_s\big] = \mathrm{E}\big[\,\big(B_t - B_s\big)\,B_s\big] + \mathrm{E}\big[B_s^2\big] \\ &= \mathrm{E}\big[B_t - B_s\big]\mathrm{E}\big[B_s\big] + \mathrm{E}\big[B_s^2\big] = \mathrm{E}\big[B_s^2\big] = s = t \wedge s.\end{aligned}$$

This completes the proof. □

Note. The above proposition is a modern version of *Einstein's postulates.*

4.1.3 Construction of Brownian Motion

There are different ways of constructing a Brownian motion. For instance, it can be constructed as the continuous limit of a random walk. Convergence in law is guaranteed by *Donsker's theorem*, which can be described as the infinite dimensional version of the central limit theorem. The interested reader can consult [21] for more details on this construction. Instead, we will follow Paul Lévy's formulation from 1948, based on functional analysis. For the reader's convenience, we present in Appendix B elementary concepts of Banach and Hilbert spaces that we will use in our construction.

Preliminaries

Definition 4.1.4. We denote by $L^2([0,1])$ the linear space of all equivalence classes for a.e. equality of Lebesgue measurable functions $f\colon [0,1] \to \mathbb{R}$ such that

$$\left(\int_0^1 |f(s)|^2 \, ds\right)^{1/2} < \infty.$$

Definition 4.1.5. Let $f, g \in L^2([0,1])$. We define an *inner product* in $L^2([0,1])$ by the formula

$$\langle f, g \rangle_{L^2([0,1])} = \int_0^1 f(s)g(s) \, ds.$$

This leads us to define an induced *norm* for $L^2([0,1])$ as follows:

$$\|f\|^2_{L^2([0,1])} = \langle f, f \rangle_{L^2([0,1])}.$$

Henceforward we will omit the subscript $L^2([0,1])$ in the norm and inner product.

Example 4.1.6. Given two indicator functions $\mathbb{1}_{[0,t]}, \mathbb{1}_{[0,s]}$ in $L^2([0,1])$, their inner product is

$$\langle \mathbb{1}_{[0,t]}, \mathbb{1}_{[0,s]} \rangle = t \wedge s.$$

We now introduce Hilbert space theory. The following results are based on Section B.2.

Theorem 4.1.7. *$L^2([0,1])$ is a Hilbert space.*

Definition 4.1.8. An orthonormal sequence in $L^2([0,1])$ is called a *complete orthonormal system* (CONS) of $L^2([0,1])$ when the set of all finite linear combinations of its elements is dense in $L^2([0,1])$.

Remark 4.1.9. A CONS is also called an *orthonormal basis*. Proposition B.2.9 shows the equivalence.

By Definition B.2.8, if $\{u_i\}_{i \in \mathbb{Z}_{>0}}$ is a CONS of $L^2([0,1])$, then for any $f \in L^2([0,1])$ we have the series expansion

$$f = \sum_{i=1}^{\infty} \langle f, u_i \rangle u_i$$

which converges in $L^2([0,1])$.

As a result, Parseval's identity holds:

$$\|f\|^2 = \sum_{i=1}^{\infty} \langle f, u_i \rangle^2.$$

Wiener's reasoning

Wiener observed that if Brownian motion could be built on a probability space $(\Omega, \mathcal{F}, \mathrm{P})$ then there would be a linear isometry of $L^2([0,1])$ into $L^2(\Omega, \mathcal{F}, \mathrm{P})$, the Hilbert space of equivalence classes for a.s. equality of random variables X such that $\mathrm{E}\left[X^2\right] < \infty$. The underlying idea of his construction is that the isometry described above maps orthonormal bases and series expansions between $L^2([0,1])$ and $L^2(\Omega, \mathcal{F}, \mathrm{P})$. It is now known as *Wiener's isometry*.

Lévy's construction

The first step of the building process is to find a suitable orthonormal basis of $L^2([0,1])$. Paul Lévy made use of the *Haar functions*, defined as follows:

Definition 4.1.10. Let $I(n)$ be the set of odd integers between 0 and 2^n, where n is a strictly positive integer. We define the Haar functions by

$$H_0^{(1)}(t) = 1 \quad \forall t \in [0,1],$$

and for $n \in \mathbb{Z}_{>0}, k \in I(n)$,

$$H_n^{(k)}(t) = \begin{cases} 2^{\frac{n-1}{2}} & , \frac{k-1}{2^n} \leq t < \frac{k}{2^n} \\ -2^{\frac{n-1}{2}} & , \frac{k}{2^n} \leq t < \frac{k+1}{2^n} \\ 0 & , \text{otherwise.} \end{cases}$$

Proposition 4.1.11. *The Haar functions form a complete orthonormal system of $L^2([0,1])$.*

Proof. Orthonormality with respect to the inner product in $L^2(0,1)$ can be verified directly. It suffices to show that this system spans all step functions that are constant on each dyadic interval $[j2^{-n}, (j+1)2^{-n}), j = 0, 1, \ldots, 2^n - 1$. This can be verified by induction on n. □

Corollary 4.1.12. *Every $f \in L^2([0,1])$ can be expressed as an L^2-series expansion of the form*

$$f = \langle f, H_0^{(1)} \rangle H_0^{(1)} + \sum_{n=1}^{\infty} \sum_{k \in I(k)} \langle f, H_n^{(k)} \rangle H_n^{(k)}.$$

Therefore, by Parseval's identity,

$$\|f\|^2 = \langle f, H_0^{(1)} \rangle^2 + \sum_{n=1}^{\infty} \sum_{k \in I(k)} \langle f, H_n^{(k)} \rangle^2.$$

Furthermore, Lévy employed another family of real-valued functions defined on the same space, called *Schauder functions*. These are closely related to the Haar functions.

Definition 4.1.13. For any $n \in \mathbb{Z}_{>0}, k \in I(n), t \in [0,1]$, the Schauder functions are given by

$$G_0^{(1)}(t) = t,$$
$$G_n^{(k)}(t) = \langle \mathbb{1}_{[0,t]}, H_n^{(k)} \rangle = \int_0^t H_n^{(k)}(s)\, ds.$$

For each fixed $n \geq 1$, these functions are positive and non-overlapping for different values of k. Moreover, their graphs are small "tents" of height $2^{-\frac{n+1}{2}}$ and base $\left[\frac{k-1}{2^n}, \frac{k+1}{2^n}\right]$, centered at $\frac{k}{2^n}$.

Proposition 4.1.14. *Let $\{Z_0^{(1)}, Z_n^{(k)} \mid n \in \mathbb{Z}_{>0}\ , k \in I(n)\}$ be i.i.d. standard Normal random variables on a common probability space $(\Omega, \mathcal{F}, \mathrm{P})$. Then*

1. *The mapping*

$$I_W \colon L^2([0,1]) \to L^2(\Omega, \mathcal{F}, \mathrm{P})$$
$$f \mapsto \langle f, H_0^{(1)} \rangle Z_0^{(1)} + \sum_{n=1}^{\infty} \sum_{k \in I(k)} \langle f, H_n^{(k)} \rangle Z_n^{(k)}$$

is a linear isometry.

2. $\forall f \in L^2([0,1]), \quad I_W(f) \sim N\big(0, \|f\|^2\big)$.

Proof. 1. The linearity can be easily deduced from inner product properties. Let us check the following two equalities:

$$\|I_W(f)\|^2_{L^2(\Omega,\mathcal{F},\mathrm{P})} = \mathrm{E}\big[I_W(f)^2\big] = \|f\|^2.$$

The l.h.s. equality stems from the inner product in $L^2(\Omega,\mathcal{F},\mathrm{P})$ defined by $\langle X, Y\rangle_{L^2(\Omega,\mathcal{F},\mathrm{P})} = \mathrm{E}[XY]$ which induces the norm $\|X\|_{L^2(\Omega,\mathcal{F},\mathrm{P})} = \big(\mathrm{E}\big[X^2\big]\big)^{\frac{1}{2}}$. For the r.h.s. equality we will compute the expectation:

$$\begin{aligned}\mathrm{E}\big[I_W(f)^2\big] &= \langle f, H_0^{(1)}\rangle^2\, \mathrm{E}\big[(Z_0^{(1)})^2\big] \\ &\quad + 2\sum_{n=1}^{\infty}\sum_{k\in I(n)} \langle f, H_0^{(1)}\rangle\langle f, H_n^{(k)}\rangle \mathrm{E}\big[Z_0^{(1)} Z_n^{(k)}\big] \\ &\quad + \sum_{n=1}^{\infty}\sum_{k\in I(n)}\sum_{m=1}^{\infty}\sum_{l\in I(m)} \langle f, H_n^{(k)}\rangle\langle f, H_m^{(l)}\rangle \mathrm{E}\big[Z_n^{(k)} Z_m^{(l)}\big].\end{aligned}$$

Since $\mathrm{E}[X_iX_j] = \mathrm{E}[X_i]\mathrm{E}[X_j]$ for any pair of independent Normal random variables defined on the same probability space, we have

$$\mathrm{E}\big[I_W(f)^2\big] = \langle f, H_0^{(1)}\rangle^2 \mathrm{E}\big[(Z_0^{(1)})^2\big] + \sum_{n=1}^{\infty}\sum_{k\in I(n)} \langle f, H_n^{(k)}\rangle^2 \mathrm{E}\big[(Z_n^{(k)})^2\big].$$

Using formula (3.2.3) together with Parseval's identity in Corollary 4.1.12 we obtain

$$\mathrm{E}\big[I_W(f)^2\big] = \langle f, H_0^{(1)}\rangle^2 + \sum_{n=1}^{\infty}\sum_{k\in I(n)} \langle f, H_n^{(k)}\rangle^2 = \|f\|^2.$$

2. We have shown that the series defining $I_W(f)$ converges in $L^2(\Omega,\mathcal{F},\mathrm{P})$. Thus, we have convergence for the expectation and variance of $I_W(f)$. A simple computation using (3.2.1) shows that the limit follows a Normal distribution. Furthermore, $\mathrm{E}\big[I_W(f)\big] = 0$ since $\mathrm{E}\big[Z_n^{(k)}\big] = \mathrm{E}\big[Z_0^{(1)}\big] = 0$. We have already shown that $\mathrm{E}\big[I_W(f)^2\big] = \|f\|^2$, hence the demonstration is complete. □

In order to prove Lévy's construction theorem, we will introduce two lemmas from probability theory.

Lemma 4.1.15 (Tail estimate for the standard Normal distribution). *For any $Z \sim N(0,1)$ and $a > 0$,*

$$\mathrm{P}(|Z| \geq a) \leq \frac{2}{a\sqrt{2\pi}} e^{\frac{-a^2}{2}}.$$

Proof.

$$\mathrm{P}(|Z| \geq a) = \frac{2}{\sqrt{2\pi}}\int_a^{\infty} e^{-\frac{x^2}{2}}\, dx \leq \frac{2}{\sqrt{2\pi}}\int_a^{\infty} \frac{x}{a} e^{-\frac{x^2}{2}}\, dx = \frac{2}{a\sqrt{2\pi}} e^{-\frac{a^2}{2}}.$$ □

Lemma 4.1.16 (First Borel-Cantelli Lemma). *Let $\{A_n\}_{n\geq 1}$ be a sequence of events in $(\Omega,\mathcal{F},\mathrm{P})$. Then*

$$\sum_{n=1}^{\infty} \mathrm{P}(A_n) < \infty \Longrightarrow \mathrm{P}\big(\limsup_{n\to\infty} A_n\big) = 0.$$

Theorem 4.1.17 (**Lévy's Construction Theorem**). *Let* $B = \{B_t = I_W(\mathbb{1}_{[0,t]}) \mid t \in [0,1]\}$.

1. *B is a centered Gaussian stochastic process with covariance function $Q(s,t) = \mathrm{E}[B_s B_t] = s \wedge t$. Hence, B is a Brownian motion on $[0,1]$.*
2. *The series expansion*

$$B_t = I_W(\mathbb{1}_{[0,t]}) = \langle \mathbb{1}_{[0,t]}, H_0^{(1)} \rangle Z_0^{(1)} + \sum_{n=1}^{\infty} \sum_{k \in I(k)} \langle \mathbb{1}_{[0,t]}, H_n^{(k)} \rangle Z_n^{(k)}$$

converges a.s. uniformly for every $t \in [0,1]$. Thus, the sample paths $t \mapsto B_t$ are a.s. continuous.

Proof. 1. It is easy to verify that $B_0 = 0$ a.s.

By Proposition 4.1.14, for every $t \in [0,1], B_t \sim N(0, \|\mathbb{1}_{[0,t]}\|^2)$. Therefore for every $t, s \in [0,1]$,

$$\mathrm{E}[B_t B_s] = \langle B_t, B_s \rangle_{L^2(\Omega,\mathcal{F},\mathrm{P})} = \langle \mathbb{1}_{[0,t]}, \mathbb{1}_{[0,s]} \rangle = t \wedge s.$$

Lastly, we need to prove that the finite-dimensional distributions are Gaussian. We apply the linearity of the isometry I_W to show that any linear combination of the form

$$\sum_{i=1}^{m} a_i B_{t_i} = \sum_{i=1}^{m} a_i I_W(\mathbb{1}_{[0,t_i]}) = I_W\left(\sum_{i=1}^{m} a_i \mathbb{1}_{[0,t_i]}\right) \qquad (a_1, \ldots, a_m \in \mathbb{R})$$

is a Normal random variable. Hence, by definition, the random vector

$$(B_{t_1}, \ldots, B_{t_m})$$

is Gaussian.

2. Introducing the Schauder functions $G_n^{(k)}$ from Definition 4.1.13, we have

$$B_t = G_0^{(1)}(t) Z_0^{(1)} + \sum_{n=1}^{\infty} \sum_{k \in I(k)} \langle G_n^{(k)}(t), Z_n^{(k)} \rangle Z_n^{(k)}.$$

Our aim is to prove that the double series $\sum_{n=1}^{\infty} \sum_{k \in I(k)} \left| G_n^{(k)}(t) Z_n^{(k)} \right|$ converges uniformly a.s. in $[0,1]$.

We know that $\max_{t \in [0,1]} \left| G_n^{(k)}(t) \right| = 2^{-\frac{n+1}{2}} \leq 2^{-\frac{n}{2}}$. Let $m_n = \max_{k \in I(n)} \left| Z_n^{(k)} \right|$. By Lemma 4.1.15, $\forall n \geq 1$, $k \in I(n)$,

$$\begin{aligned} \mathrm{P}\left(\left| Z_n^{(k)} \right| \geq 2^{\frac{n}{4}} \right) &\leq \frac{2}{2^{\frac{n}{4}} \sqrt{2\pi}} e^{-\frac{(2^{n/4})^2}{2}} \\ &= \frac{\sqrt{2}}{2^{\frac{n}{4}} \sqrt{\pi}} e^{-\frac{2^{n/2}}{2}} \\ &= 2^{-\frac{n}{4}} \sqrt{\frac{2}{\pi}} e^{-2^{\frac{n}{2}-1}}. \end{aligned}$$

Therefore,

$$\mathrm{P}\left(m_n \geq 2^{\frac{n}{4}} \right) = \mathrm{P}\left(\bigcup_{k \in I(n)} \left| Z_n^{(k)} \right| \geq 2^{\frac{n}{4}} \right) \leq 2^n \mathrm{P}\left(\left| Z_n^{(1)} \right| \geq 2^{\frac{n}{4}} \right)$$

$$\leq 2^{n-\frac{n}{4}}\sqrt{\frac{2}{\pi}}e^{-2^{n/2-1}}$$
$$\leq 2^{\frac{3}{4}n+\frac{1}{2}}e^{-2^{n/2-1}}.$$

Clearly, $\sum_{n=1}^{\infty} 2^{\frac{3}{4}n+\frac{1}{2}}e^{-2^{n/2-1}} < \infty$ and thus applying Lemma 4.1.16 we have

$$\mathrm{P}\left(\liminf_{n\to\infty}\left\{m_n \geq 2^{\frac{n}{4}}\right\}\right) = 1.$$

This amounts to saying there exists a set $\widetilde{\Omega}$ with $\mathrm{P}(\widetilde{\Omega}) = 1$ such that for every $\omega \in \widetilde{\Omega}$, there exists $n_0 = n_0(\omega) \in \mathbb{Z}_{>0}$ such that $m_n(\omega) \leq 2^{\frac{n}{4}}$ for every $n \geq n_0(\omega)$. Then, for n sufficiently large,

$$\begin{aligned}\sum_{k\in I(n)} \left|G_n^{(k)}(t)Z_n^{(k)}\right| &\leq \sum_{k\in I(n)} 2^{-\frac{n}{2}}2^{\frac{n}{4}} \\ &\leq \sum_{k\in I(n)} 2^{-\frac{n}{4}} \\ &< \infty.\end{aligned}$$

Weierstrass criterion on uniforme convergence of series of functions is verified, so for every $\omega \in \widetilde{\Omega}$, $B_t(\omega)$ converges uniformly in t. Since almost surely the mapping $t \mapsto \sum_{k\in I(n)} G_n^{(k)}(t)Z_n^{(k)}$ is continuous, the sample paths $t \mapsto B_t$ are a.s. continuous. □

Corollary 4.1.18. *There exists a stochastic process $B = \{B_t \mid t \geq 0\}$ on a probability space $(\Omega, \mathcal{F}, \mathrm{P})$ such that B is a Brownian motion.*

Proof. Theorem 4.1.17 assures the existence of a sequence $(B^{(k)})_{k\geq 1}$ of independent Brownian motions indexed by $[0,1]$ and defined on a probability space $(\Omega, \mathcal{F}, \mathrm{P})$. We define B recursively by

$$B_t = \begin{cases} B_t^{(1)} & , t \in [0,1] \\ \sum_{i=1}^{k} B_1^{(i)} + B_{t-k}^{(k+1)} & , t \in [k, k+1]. \end{cases}$$

It is clearly a centered process and one can easily check that $\mathrm{E}\big[B_s B_t\big] = s \wedge t$. The independence of the Brownian motions $B^{(k)}$ implies that B is Gaussian. Hence, it is a Brownian motion according to Definition 4.1.1. □

We end this subsection with the notion of Brownian motion in multiple dimensions.

Definition 4.1.19. Let $d \in \mathbb{Z}_{>0} \setminus \{1\}$. Let $B^{(1)}, B^{(2)}, \ldots, B^{(d)}$ be independent Brownian motions indexed by $\mathbb{R}_{\geq 0}$. The process

$$B \coloneqq \{B_t = (B_t^{(1)}, B_t^{(2)}, \ldots, B_t^{(d)}) \mid t \geq 0\}$$

is a *d-dimensional Brownian motion* on $\mathbb{R}_{\geq 0}$.

4.1.4 Sample Path Properties

For the sake of completeness, in this subsection we present a collection of properties of Brownian motion that hold almost surely. Proofs will be omitted, since they are beyond the scope and objectives of this thesis. They can be found in [13, 23, 15, 16, 6], to name a few references.

Fundamental properties

Proposition 4.1.20. *Let $B = \{B_t \mid t \geq 0\}$ be a Wiener process.*

(i) *Since $\mathrm{E}\big[B_0^2\big] = 0$, B_0 has a Dirac distribution at the origin.*

(ii) $\mathrm{E}\big[|B_t - B_s|^2\big] = |t - s|$.

Proposition 4.1.21 (Invariance Properties). *Let $B = \{B_t \mid t \geq 0\}$ be a Brownian motion. The processes generated by the following transformations are also Brownian motions.*

(i) ***Scaling**: For any $\lambda > 0$, $B^\lambda \coloneqq \{\frac{1}{\lambda} B_{\lambda^2 t} \mid t \geq 0\}$.*

(ii) ***Time inversion**: The process $X = \{X_t \mid t \geq 0\}$ defined by*

$$X_t \coloneqq \begin{cases} 0 & \text{for } t = 0 \\ tB\big(\frac{1}{t}\big) & \text{for } t > 0. \end{cases}$$

(iii) ***Time shift**: For any fixed $a > 0$, $B_{+a} \coloneqq \{B_{a+t} - B_a \mid t \geq 0\}$.*

(iv) ***Symmetry**: $-B \coloneqq \{-B_t \mid t \geq 0\}$.*

Property (i) represents the *fractal* nature of Brownian motion: it identifies a transformation that changes the individual Brownian motions but not their distributions. Property (ii) has the following interesting corollary:

Corollary 4.1.22 (Strong Law of Large Numbers). *Almost surely,*

$$\lim_{t \to \infty} \frac{B_t}{t} = 0.$$

Roughly speaking, this means that for any $\varepsilon > 0$, $|B_t|$ oscillates with an amplitude less than $t\varepsilon$.

Quadratic Variation

Fix $t > 0$. Let $\{\Pi_n\}_{n=1}^\infty$ be a sequence of partitions of the interval $[0, t]$ defined by

$$0 = t_0^{(n)} \leq t_1^{(n)} \leq \ldots \leq t_{r(n)}^{(n)} = t.$$

Suppose

$$\lim_{n \to \infty} \|\Pi\| = 0$$

where $\|\cdot\| = \max_{j=1,2,\ldots,r(n)} |t_j - t_{j-1}|$ is the norm of the partition.

Then, the *quadratic variations*

$$V_{t^{(n)}}^{(2)} \coloneqq \sum_{j=1}^{r(n)} \left| B_{t_j^{(n)}} - B_{t_{j-1}^{(n)}} \right|^2$$

of a Brownian motion over these partitions converge to t in L^2 as $n \to \infty$. That is,

$$\lim_{n \to \infty} V_{t^{(n)}}^{(2)} = t \qquad \text{in } L^2.$$

The previous convergence also holds a.s. if $\sum_{n=1}^\infty \|\Pi_n\| < \infty$.

We now define the *variation* of a Brownian sample path B on $[0,t]$ as follows:

$$V_t \coloneqq \limsup_{\|\Pi\|\to 0} \sum_{j=1}^{n} \left| B_{t_j} - B_{t_{j-1}} \right|$$

where $\Pi = (t_0, t_1, \ldots, t_n)$ is a partition of $[0,t]$ and $\|\cdot\|$ its associated norm.

Almost surely, when $\|\Pi\| \to 0$, the quadratic variations converge to t as $n \to \infty$. Therefore, since the sample paths are a.s. continuous, Brownian motion has unbounded variation almost surely.

Law of the Iterated Logarithm

Our next theorem describes the oscillations of Brownian motion when the time parameter is close to the origin and when it tends to infinity.

Theorem 4.1.23 (Law of the Iterated Logarithm). *For almost every $\omega \in \Omega$,*

$$\limsup_{t\downarrow 0} \frac{B_t(\omega)}{\sqrt{2t\log\log\left(\frac{1}{t}\right)}} = 1 \ , \qquad \limsup_{t\to\infty} \frac{B_t(\omega)}{\sqrt{2t\log\log(t)}} = 1 \ ,$$

$$\liminf_{t\downarrow 0} \frac{B_t(\omega)}{\sqrt{2t\log\log\left(\frac{1}{t}\right)}} = -1 \ , \qquad \liminf_{t\to\infty} \frac{B_t(\omega)}{\sqrt{2t\log\log(t)}} = -1 \ .$$

The function $t \mapsto \sqrt{2t\log\log(t)}$ determines almost surely the asymptotic growth of the process. It is known as the *asymptotic smallest upper envelope* of the Brownian motion.

Continuity properties

Definition 4.1.24. A function $g(\cdot)$ is called a *modulus of continuity* for a function $f\colon [0,T] \to \mathbb{R}$ if for any $0 \le s \le t \le T$,

$$|t-s| \le \delta \Longrightarrow |f(t) - f(s)| \le g(\delta)$$

for every $\delta > 0$ sufficiently small.

The modulus of continuity for a function describes its local smoothness.

Theorem 4.1.25 (Lévy Modulus of Continuity, 1937). *Let $B = \{B_t \mid t \in [0,1]\}$ be a Brownian motion. Let $g\colon (0,1] \to (0,\infty)$ be given by $g(\delta) = \sqrt{2\delta\log(1/\delta)}$. Then*

$$\mathrm{P}\left(\limsup_{\delta\downarrow 0} \frac{1}{g(\delta)} \max_{\substack{0\le s<t\le 1 \\ t-s\le\delta}} |B_t - B_s| = 1 \right) = 1.$$

Definition 4.1.26. Let $\gamma \in (0,1]$. A function $f\colon \mathbb{R} \to \mathbb{R}$ is said to be *locally γ-Hölder continuous* if for any bounded interval $I \subset \mathbb{R}$, there exists a constant $c(\gamma, f, I) > 0$ such that

$$\frac{|f(x) - f(y)|}{|x-y|^\gamma} \le c(\gamma, f, I)$$

for every $x, y \in I$ satisfying $x \neq y$. We refer to γ as *Hölder's exponent* and to $c(\gamma, f, I)$ as *Hölder's constant.*

Remark 4.1.27. Note that γ-Hölder continuity becomes stronger as γ increases.

Theorem 4.1.28. *If $\gamma < \frac{1}{2}$, then, almost surely, the sample paths of a Brownian motion are everywhere locally γ-Hölder continuous.*

Nowhere Differentiability

Definition 4.1.29. For a continuous function $f\colon [0,\infty) \to \mathbb{R}$ we define the *upper and lower (right and left) Dini derivatives* at t by

$$D^{\pm} f(t) = \limsup_{h \to 0^{\pm}} \frac{f(t+h) - f(t)}{h}$$

and

$$D_{\pm} f(t) = \liminf_{h \to 0^{\pm}} \frac{f(t+h) - f(t)}{h} \ ,$$

respectively.

Theorem 4.1.30 (Paley, Wiener and Zygmund, 1933). *Let $B = \{B_t \mid t \geq 0\}$ be a Wiener process defined on a probability space $(\Omega, \mathcal{F}, \mathrm{P})$. For almost every $\omega \in \Omega$, the sample paths $t \mapsto B_t(w)$ are nowhere differentiable.*

Furthermore, almost surely, for every $t \geq 0$,

$$\textit{either} \quad D^{+} B_t = +\infty \quad \textit{or} \quad D_{+} B_t = -\infty \quad \textit{or both.}$$

The following stronger result indicates the value of Hölder's exponent for which the transition between local γ-Hölder sample path continuity and non-continuity happens.

Theorem 4.1.31. *For any $\gamma \in \left[\frac{1}{2}, 1\right]$, the sample paths of a Wiener process are a.s. nowhere locally γ-Hölder continuous.*

4.2 Isonormal Gaussian Processes

In this section we introduce a new class of Gaussian stochastic processes with a wider range of index sets.

4.2.1 Introduction

Definition 4.2.1. A *Gaussian random field* is a Gaussian stochastic process indexed by an arbitrary set of dimensionality at least one.

Remark 4.2.2. The finite-dimensional distributions, as well as the mean and covariance functions, are given analogously to Chapter 3. The only difference lies in the index set. The existence of Gaussian random fields is ensured by [5, Lemma 1.2.2].

Definition 4.2.3. A Gaussian process $W = \{W(h) \mid h \in H\}$ indexed by a real separable Hilbert space[1] H is said to be *H-isonormal* if it is centered with covariance function $\mathrm{E}[W(h)W(g)] = \langle h, g \rangle_H$ for all $h, g \in H$.

Note. When no particular Hilbert space is mentioned, we will speak of *isonormal Gaussian processes.*

Remark 4.2.4. The definition above extends the definition of Brownian motion to Hilbert spaces. In fact, Brownian motion is a particular case of isonormal Gaussian processes. The next theorem gives a more detailed description of this idea.

[1] See Definition B.2.4.

Theorem 4.2.5. *Fix $T > 0$. If W is an $L^2([0,T])$-isonormal Gaussian process, then*

$$W_{0,T} \coloneqq \{W(t) \coloneqq W\big(\mathbb{1}_{[0,t]}\big) \mid t \in [0,T]\}$$

defines a Brownian motion on $[0,T]$.

Proof. $W_{0,T}$ is a centered Gaussian process satisfying $\mathrm{E}[W(s)W(t)] = \langle \mathbb{1}_{[0,s]}, \mathbb{1}_{[0,t]} \rangle_{L^2([0,T])} = s \wedge t$ for all $s, t \in [0,T]$. Therefore, it is a Brownian motion by Definition 4.1.1. □

Lemma 4.2.6. *Let W be an H-isonormal Gaussian process. The mapping $W\colon H \to L^2(\Omega, \mathcal{F}, \mathrm{P})$ is a linear Hilbert space isometry. Therefore, it maps orthogonal elements of H to orthogonal elements of $L^2(\Omega, \mathcal{F}, \mathrm{P})$.*

Proof. W is an isometry since for all $h \in H, \|h\|_H^2 = \mathrm{E}\big[W(h)^2\big]$. Furthermore, for every $a, b \in \mathbb{R}$ and $h, g \in H$,

$$\begin{aligned}\mathrm{E}\Big[\big(W(ah+bg) - aW(h) - bW(g)\big)^2\Big] &= \|ah+bg\|_H^2 + a^2\|h\|_H^2 + b^2\|g\|_H^2 \\ &\quad - 2a\langle ah+bg, h\rangle_H - 2b\langle ah+bg, g\rangle_H + 2ab\langle h, g\rangle_H \\ &= 0.\end{aligned}$$

This proves the linearity of W. □

4.2.2 Construction

The following proposition guarantees the existence of an isonormal Gaussian process.

Proposition 4.2.7. *Let $\{e_n\}_{n\geq 1}$ be a complete orthonormal system of H. Let $W = \{W(h) \mid h \in H\}$ be an isonormal Gaussian process. We define $Z_n = W(e_n) \quad \forall n \geq 1$.*
Then $(Z_n)_{n\geq 1}$ is a sequence of i.i.d. standard Normal random variables and each random variable $W(h)$ is defined by

$$W(h) = \sum_{n=1}^{\infty} \langle h, e_n\rangle_H Z_n \ ,$$

where the series converges in $L^2(\Omega, \mathcal{F}, \mathrm{P})$.

Proof. By Definition 4.2.3, the random variables Z_n are uncorrelated standard Normal. Thus, W being a Gaussian process, they are also independent. According to Definition B.2.8, for any $h \in H$, the series

$$\sum_{n=1}^{\infty} \langle h, e_n\rangle_H e_n$$

converges in H to h. Lemma 4.2.6 ensures that

$$W(h) = \sum_{n=1}^{\infty} \langle h, e_n\rangle_H W(e_n) = \sum_{n=1}^{\infty} \langle h, e_n\rangle_H Z_n$$

and the series converges in $L^2(\Omega, \mathcal{F}, \mathrm{P})$, since

$$\forall n, m \geq 1, n \neq m, \quad \begin{cases} \mathrm{E}\big[|Z_n|^2\big] = 1 \\ \mathrm{E}[|Z_n|]\mathrm{E}[|Z_m|] = 0 \end{cases}$$

and

$$\sum_{n=1}^{\infty} |\langle h, e_n\rangle_H|^2 = \|h\|_H^2$$

by Parseval's identity. □

Remark 4.2.8. The above proposition yields a bijection between the isonormal Gaussian process W and the i.i.d. sequence $(Z_n)_{n\geq 1}$. We can interpret the random variables Z_n as the Hilbert space coordinates of the process W.

4.3 Gaussian White Noise

Firstly, we will introduce the notion of white noise in space and how it can be obtained from an isonormal Gaussian process. Next, we will add another dimension and we will give it a special meaning.

4.3.1 White Noise on $\mathbb{R}^d$

Let D be a non-empy subset of $\mathbb{R}^d$, with $d \in \mathbb{Z}_{>0}$. Let μ be a σ-finite measure on $(D, \mathfrak{B}(D))$. We define the set $\mathfrak{B}_f(D) \coloneqq \{B \in \mathfrak{B}(D) \mid \mu(B) < \infty\}$.

In this subsection we will consider the case $D = \mathbb{R}^d$.

Definition 4.3.1. A (*Gaussian*) *white noise with intensity* μ (μ-GWN) is a centered Gaussian random field of the form

$$W = \{W(A), A \in \mathfrak{B}_f(\mathbb{R}^d)\}$$

defined on a probability space $(\Omega, \mathcal{F}, \mathrm{P})$ with covariance function

$$\mathrm{E}[W(A)W(B)] = \mu(A \cap B) \quad \forall A, B \in \mathfrak{B}_f(\mathbb{R}^d).$$

Note. When μ is the Lebesgue measure on $\mathbb{R}^d$, we will simply name it *white noise*.

The existence of white noise is a direct consequence of [5, Lemma 1.2.2]. Consider the Hilbert space $H = L^2(\mathbb{R}^d, \mathfrak{B}(\mathbb{R}^d), \mu)$ and assume the existence of an H-isonormal Gaussian process W. We can readily associate W with a white noise $\overline{W}$ with intensity μ by defining $\overline{W}(A) = W(\mathbb{1}_A)$ for every $A \in \mathfrak{B}_f(\mathbb{R}^d)$.

In this setting, given $h \in H$, we can think of $W(h)$ as an $L^2(\Omega, \mathcal{F}, \mathrm{P})$-valued integral of h. In such a case, we adopt the notation

$$W(h) \coloneqq \int_{\mathbb{R}^d} h(x)\overline{W}(dx) \coloneqq \int_{\mathbb{R}^d} h(x)\, d\overline{W}. \tag{4.3.1}$$

Definition 4.3.2. The random variable $W(h)$ defined in (4.3.1) is called the *Wiener integral* of $h \in H$ with respect to the white noise $\overline{W}$ with intensity μ.

We will close this subsection with two propositions containing some interesting properties of white noise with a given intensity. For the proofs, we refer the reader to [5, 15].

Proposition 4.3.3. *Let W be a μ-GWN. Let $A_1, \ldots, A_n \in \mathfrak{B}_f(\mathbb{R}^d)$ be pairwise disjoint. Then for every $i, j \in \{1, \ldots, n\}$,*

(i) if $i \neq j$, then $W(A_i)$ and $W(A_j)$ are independent and $W(A_i \cup A_j) = W(A_i) + W(A_j)$;

(ii) the random vector $\big(W(A_1), \ldots, W(A_n)\big)$ is centered Normal with covariance matrix given by

$$Q_{ij} = \begin{cases} \mu(A_i) & \text{if } \ i = j \\ 0 & \text{otherwise.} \end{cases}$$

Proposition 4.3.4. *Let W be a μ-GWN.*

(i) Let $(A_n)_{n\geq 1} \subset \mathfrak{B}_f(\mathbb{R}^d)$ be a decreasing sequence with $\mu(A_1) < \infty$. Set $A \coloneqq \cap_{n\geq 1} A_n$. Then $W(A_n) \to W(A)$ in $L^2(\Omega, \mathcal{F}, \mathrm{P})$.

(ii) Let $(A_n)_{n\geq 1} \subset \mathfrak{B}_f(\mathbb{R}^d)$ be an increasing sequence. Set $A \coloneqq \cup_{n\geq 1} A_n$ and assume $\mu(A) < \infty$. Then $W(A_n) \to W(A)$ in $L^2(\Omega, \mathcal{F}, \mathrm{P})$.

4.3.2 Space-time White Noise

Let D be a non-empty open subset of $\mathbb{R}^d$. Letting the measure μ described previously be defined in $\mathbb{R}^{d+1}$ by $\mu(dx, dt) = \mathbb{1}_D(x)\mathbb{1}_{[0,\infty)}(t)\, dx\, dt$ yields a particular case of white noise known as *space-time white noise*.

We define the set $\mathfrak{B}_f(D \times \mathbb{R}_{\geq 0}) \coloneqq \{B \in \mathfrak{B}(D \times \mathbb{R}_{\geq 0}) \mid \lambda(B) < \infty\}$, where λ denotes the Lebesgue measure on $D \times \mathbb{R}_{\geq 0}$. In the sequel, we will consider the Hilbert space $H = L^2(D \times \mathbb{R}_{\geq 0}, \mathfrak{B}(D \times \mathbb{R}_{\geq 0}), \lambda)$.

Definition 4.3.5. A *space-time white noise* on $D \times \mathbb{R}_{\geq 0}$ is a centered Gaussian random field defined on some probability space $(\Omega, \mathcal{F}, \mathrm{P})$ by

$$W = \{W(A), A \in \mathfrak{B}_f(D \times \mathbb{R}_{\geq 0})\}$$

such that $\quad \forall A, B \in \mathfrak{B}_f(D \times \mathbb{R}_{\geq 0}), \quad \mathrm{E}[W(A)W(B)] = \lambda(A \cap B).$

Similarly to white noise, any H-isonormal Gaussian process $\widetilde{W}$ has a related space-time white noise given by

$$W(B \times [0, t]) = \widetilde{W}\big(\mathbb{1}_{B\times[0,t]}\big)$$

for every $B \in \mathfrak{B}_f(D), t \in \mathbb{R}_{\geq 0}$. In this case, for all $h \in H$, the Wiener integral of h with respect to W is the random variable

$$\widetilde{W}(h) = \int_{D\times\mathbb{R}_{\geq 0}} h(x, t)W(dx, dt). \tag{4.3.2}$$

Sometimes (4.3.2) is informally written as

$$W(h) = \int_{D\times\mathbb{R}_{\geq 0}} h(x, t)\dot{W}(dx, dt).$$

The next proposition gathers some important properties of space-time white noise. These are analogous to those included in Proposition 4.3.3 and Proposition 4.3.4.

Proposition 4.3.6. *Let W be a space-time white noise. Let $A_1, \ldots, A_n \in \mathfrak{B}_f(D \times \mathbb{R}_{\geq 0})$ be pairwise disjoint. Let $i, j \in \{1, \ldots, n\}$.*

(i) If $i \neq j$, then $W(A_i)$ and $W(A_j)$ are independent and $W(A_i \cup A_j) = W(A_i) + W(A_j)$.

(ii) The random vector $\big(W(A_1), \ldots, W(A_n)\big)$ is centered Normal with covariance matrix given by

$$Q_{ij} = \begin{cases} \lambda(A_i) & \text{if } \ i = j \\ 0 & \text{otherwise.} \end{cases}$$

(iii) Let $(A_n)_{n\geq 1} \subset \mathfrak{B}_f(\mathbb{R}^d)$ be a decreasing sequence with $\lambda(A_1) < \infty$. Set $A \coloneqq \cap_{n\geq 1} A_n$. Then $W(A_n) \to W(A)$ in $L^2(\Omega, \mathcal{F}, \mathrm{P})$.

(iv) Let $(A_n)_{n\geq 1} \subset \mathfrak{B}_f(\mathbb{R}^d)$ be an increasing sequence. Set $A \coloneqq \cup_{n\geq 1} A_n$ and suppose $\lambda(A) < \infty$. Then $W(A_n) \to W(A)$ in $L^2(\Omega, \mathcal{F}, \mathrm{P})$.

Definition 4.3.7. Let W be an H-isonormal Gaussian process. For $t \in \mathbb{R}_{\geq 0}$, $\phi \in L^2(D, \mathfrak{B}(D), \lambda)$, we define

$$W_t(\phi) \coloneqq W\big(\phi \times \mathbb{1}_{[0,t]}\big).$$

Remark 4.3.8. For a fixed time t, W_t is an $L^2(D, \mathfrak{B}(D), \lambda)$-isonormal Gaussian process.

Proposition 4.3.9. *Let $\{e_n\}_{n\geq 1}$ be a complete orthonormal system of $L^2(D, \mathfrak{B}(D), \lambda)$ and fix $T \in \mathbb{R}_{>0}$. Define $B_n \colon t \in [0,T] \mapsto W\big(e_n \times \mathbb{1}_{[0,t]}\big) \quad \forall n \geq 1$.*
Then

1. *$(B_n)_{n\geq 1}$ is a sequence of independent Brownian motions.*
2. *For any $\phi = \sum_{n=1}^{\infty} \phi_n e_n \in L^2(D, \mathfrak{B}(D), \lambda)$ with real coefficients ϕ_n, each random variable $W_t(\phi)$ is defined almost surely by*

$$W_t(\phi) = \sum_{n=1}^{\infty} B_n(t)\phi_n. \tag{4.3.3}$$

Proof. 1. Let $n \geq 1$. By definition, B_n is a centered Gaussian process. Moreover, for $m \geq 1$ and $s, t \in [0,T]$, we have

$$\begin{aligned}
\mathrm{E}\big[B_n(t)B_m(s)\big] &= \mathrm{E}\big[W\big(e_n \times \mathbb{1}_{[0,t]}\big)W\big(e_m \times \mathbb{1}_{[0,s]}\big)\big] \\
&= \langle e_n \times \mathbb{1}_{[0,t]}, e_m \times \mathbb{1}_{[0,s]} \rangle_{L^2(D\times[0,T], \mathfrak{B}(D\times[0,T]), \lambda)} \\
&= \langle e_n, e_m \rangle^2_{L^2(D, \mathfrak{B}(D), \lambda)}(t \wedge s) \\
&= \delta_{n,m}(t \wedge s)
\end{aligned}$$

and the independence follows. Taking $n = m$ in the above formula proves that B_n is a Brownian motion.

2. For any $\phi \in \mathrm{Span}\{e_1, \ldots, e_N\}$[2], the expression (4.3.3) is obtained taking the limit $N \to \infty$. □

Remark 4.3.10. The above result is the space-time version of Proposition 4.2.7, providing an explicit connection between Brownian motion and time-dependent isonormal Gaussian processes.

[2]Recall that the linear span of a set can be defined as the set of all finite linear combinations of its elements.

Chapter 5

A Linear Stochastic Heat Equation

5.1 Introduction

5.1.1 Motivation

In Chapter 2 we examined the inhomogeneous global Cauchy problem (2.3.1) with a procedure to obtain an explicit solution of the form (2.3.2) based on the given data and the fundamental solution Φ_k defined in (2.2.19). In this chapter we will employ the ideas developed in the two previous chapters to study the case where the external forcing $f(x,t)$ is replaced by a space-time white noise $\dot{W}(x,t)$. The resulting equation is known as the *stochastic heat equation on* $\mathbb{R}$. We will examine the associated solutions and their sample path continuity properties.

5.1.2 Random Field Solutions

Let $\mathcal{L}$ be a linear partial differential operator on $\mathbb{R}\times\mathbb{R}_{>0}$ with constant coefficients. A *linear stochastic partial differential equation* (LSPDE) *on* $\mathbb{R}$ *with additive noise* has the form

$$\begin{cases} \mathcal{L}u = \dot{W}(x,t) & ,x\in\mathbb{R}\ ,\ t>0 \\ u(x,0) = u_0(x) & ,x\in\mathbb{R} \end{cases} \tag{5.1.1}$$

where the initial condition is deterministic. This type of equations are solved with tools from the theory of distributions, pioneered by Laurent Schwartz (see [24]). Within this framework, a concise and general description of LSPDEs with additive noise can be found in [5].

In this chapter we will focus on the stochastic heat equation on $\mathbb{R}$ with $k=1$, corresponding to the *heat operator* $\frac{\partial}{\partial t} - \frac{\partial^2}{\partial x^2}$. In this context, we will denote the heat operator by $\mathcal{L}$ and we will consider the LSPDE

$$\mathcal{L}u = \dot{W}(x,t) \quad ,x\in\mathbb{R}\ ,\ t>0 \tag{5.1.2}$$

with a given deterministic initial condition

$$u(x,0) = u_0(x) \quad ,x\in\mathbb{R}. \tag{5.1.3}$$

Since the value of the thermal diffusivity constant k is 1 throughout this chapter, we will omit the subscript in the associated fundamental solution Φ_1 to the classical heat equation.

For all $(x,t)\in\mathbb{R}\times\mathbb{R}_{\geq 0}$, let $\Phi_{x,t}(y,s) \coloneqq \Phi(x-y,t-s)$ be a function defined on $\mathbb{R}\times[0,t]$. We have already seen that for any pair $(x,t)\in\mathbb{R}\times\mathbb{R}_{>0}$, $(y,s)\mapsto\Phi(x-y,t-s)$ belongs to

$L^1\big(\mathbb{R}\times(0,t)\big)$. Let us show that this is also true for $L^2\big(\mathbb{R}\times(0,t)\big)$:

$$\begin{aligned}\int_{\mathbb{R}}\int_0^t \Phi_{x,t}^2(y,s)\,ds\,dy &= \int_{\mathbb{R}}\int_0^t \Phi^2(y,s)\,ds\,dy\\ &= \int_{\mathbb{R}}\int_0^t \frac{1}{4\pi s}e^{-2\frac{y^2}{4s}}\,ds\,dy\\ &= \int_{\mathbb{R}}\int_0^t \frac{1}{\sqrt{16}\pi s}e^{-\frac{y^2}{2s}}\,ds\,dy\\ &= \int_{\mathbb{R}} \frac{1}{\sqrt{2\pi s}}e^{-\frac{y^2}{2s}}\,dy\int_0^t \frac{1}{\sqrt{8\pi s}}\,ds. \end{aligned}\tag{5.1.4}$$

Notice that for any $s>0$, the function $y\mapsto \frac{1}{\sqrt{2\pi s}}e^{-\frac{y^2}{2s}}$ represents the probability density of a centered Normal random variable with variance s. Therefore, the value of the integral with respect to y in (5.1.4) is 1 and we have

$$\int_{\mathbb{R}}\int_0^t \Phi_{x,t}^2(y,s)\,ds\,dy = \int_{\mathbb{R}}\int_0^t \Phi^2(y,s)\,ds\,dy = \int_0^t \frac{1}{\sqrt{8\pi s}}\,ds = \sqrt{\frac{t}{2\pi}}.\tag{5.1.5}$$

We are now able to give the following definition from [5]:

Definition 5.1.1. Let W be a space-time white noise. Let $I_0(x,t)$ be the solution to the homogeneous linear PDE

$$\mathcal{L}u = 0 \quad , x\in\mathbb{R}\ ,\ t>0\tag{5.1.6}$$

with a specified initial condition. The *random field solution* to the stochastic heat equation $\mathcal{L}u=\dot{W}$ on $\mathbb{R}\times\mathbb{R}_{>0}$ with the same initial condition is

$$u(x,t) = I_0(x,t) + \int_0^t\int_{\mathbb{R}} \Phi_{x,t}(y,s)W(dy,ds).\tag{5.1.7}$$

Remark 5.1.2. Equation (5.1.6) corresponds to the classical homogeneous heat equation (2.2.10) with $k=1$. Thus, for every initial value u_0 such that the function $y\mapsto\Phi(x-y,t)u_0(y)$ belongs to $L^1(\mathbb{R})$, the solution I_0 is given by

$$I_0(x,t) = (\Phi_{1,t}*u_0)(x) = \int_{\mathbb{R}}\Phi(x-y,t)u_0(y)\,dy\tag{5.1.8}$$

according to (2.2.21).

Remark 5.1.3. The integral expression in (5.1.7) is equivalent to the Wiener integral $\widetilde{W}(\Phi_{x,t})$ according to (4.3.2). Therefore, the stochastic process $u=\{u(x,t)\mid(x,t)\in\mathbb{R}\times\mathbb{R}_{\geq 0}\}$ is $L^2(\mathbb{R}\times\mathbb{R}_{\geq 0})$-isonormal Gaussian with expectation $\mathrm{E}\big[u(x,t)\big]=I_0(x,t)$ and variance

$$\begin{aligned}\mathrm{Var}\big(u(x,t)\big) &= \mathrm{Var}\big(I_0(x,t)\big) + \mathrm{Var}\left(\int_0^t\int_{\mathbb{R}}\Phi_{x,t}(y,s)W(dy,ds)\right)\\ &\quad + 2\mathrm{Cov}\left(I_0(x,t), \int_0^t\int_{\mathbb{R}}\Phi_{x,t}(y,s)W(dy,ds)\right)\\ &= \mathrm{E}\left[\left(\int_0^t\int_{\mathbb{R}}\Phi_{x,t}(y,s)W(dy,ds)\right)^2\right]\\ &= \int_0^t\int_{\mathbb{R}}\Phi_{x,t}^2(y,s)\,dy\,ds = \sqrt{\frac{t}{2\pi}},\end{aligned}$$

where the last two equalities are obtained from Lemma 4.2.6 and (5.1.5), respectively.

5.2 Hölder Continuity

The aim of this section is to prove that the sample paths of a random field solution to the equation (5.1.2) with initial condition given by (5.1.3) are Hölder continuous. For the sake of simplicity, we will define

$$v(x,t) \coloneqq \int_0^t \int_{\mathbb{R}} \Phi_{x,t}(y,s) W(dy,ds) \tag{5.2.1}$$

and we will denote the random field solution in Definition 5.1.1 by

$$u(x,t) = I_0(x,t) + v(x,t). \tag{5.2.2}$$

We are already familiar with the notion of local Hölder continuity - see Definition 4.1.26. Due to our interest in the global properties of the stochastic heat equation, we require a notion of Hölder continuity that applies to the whole set of real numbers. As a result, we provide the following definition:

Definition 5.2.1. Let $\gamma \in (0,1]$. A function $f\colon \mathbb{R} \to \mathbb{R}$ such that

$$\frac{|f(x) - f(y)|}{|x-y|^\gamma} \leq c(\gamma, f) < \infty$$

for every $x \in \mathbb{R}$ and $y \in \mathbb{R}\setminus\{x\}$ is called *globally* γ-Hölder continuous, with Hölder's constant $c(\gamma, f) = c(\gamma, f, \mathbb{R}) > 0$.

The space of γ-Hölder continuous functions on an interval $I \subseteq \mathbb{R}$ is denoted as $C^{0,\gamma}(I)$. This space becomes a Banach space[1] if we endow it with the γ-*Hölder norm*

$$\|f\|_{C^{0,\gamma}(I)} = \max_I |f| + c(\gamma, f, I).$$

So far we have only considered Hölder continuity in the space variable. We want to extend this idea to space-time dimension and apply it to the random field solutions defined previously. To this end, we introduce another type of Hölder continuity.

Definition 5.2.2. Let $\gamma_1, \gamma_2 \in (0,1]$. A function $g\colon \mathbb{R} \times \mathbb{R}_{\geq 0} \to \mathbb{R}$ is said to be *jointly* locally (γ_1, γ_2)-Hölder continuous if for any $\varepsilon_1, \varepsilon_2 > 0$, there exists a constant $C = C(\gamma_1, \gamma_2, g, \varepsilon_1, \varepsilon_2) > 0$ such that

$$\frac{|g(x,t) - g(y,s)|}{|x-y|^{\gamma_1} + |t-s|^{\gamma_2}} \leq C$$

for every $(x,t), (y,s) \in \mathbb{R} \times \mathbb{R}_{\geq 0}$ satisfying $0 < |x-y| < \varepsilon_1$ and $0 < |t-s| < \varepsilon_2$.

If this property holds for every $(x,t) \in \mathbb{R} \times \mathbb{R}_{\geq 0}$ and $(y,s) \in \big(\mathbb{R} \setminus \{x\}\big) \times \big(\mathbb{R}_{\geq 0} \setminus \{t\}\big)$, the function g is jointly globally (γ_1, γ_2)-Hölder continuous.

Note. To shorten notation, we will simply say (γ_1, γ_2)-*Hölder continuous* when a function is jointly (γ_1, γ_2)-Hölder continuous.

In order to prove local Hölder continuity of the sample paths $(x,t) \mapsto v(x,t)$, we must use a version of the Kolmogorov continuity criterion for Gaussian processes. To this end, we shall first introduce the following definition.

Definition 5.2.3. Let $X = \{X_t \mid t \in T\}$ and $\tilde{X} = \{\tilde{X}_t \mid t \in T\}$ be two stochastic processes indexed by the same set T with values in the same state space. We say that $\tilde{X}$ is a *modification* (or *version*) of X if

$$\forall t \in T, \quad \mathrm{P}\big(\tilde{X}_t = X_t\big) = 1.$$

[1] See Section B.1.

Remark 5.2.4. This implies in particular that two modifications of the same process have the same finite-dimensional distributions, hence the same law.

We now state a famous continuity criterion on sample paths. In the literature it is attributed to Kolmogorov, either exclusively or along with Chentsov. The proof can be found in e.g. [13, 15].

Theorem 5.2.5 (**Kolmogorov Continuity Criterion**). *Let $X = \{X_t \mid t \in I\}$ be a stochastic process on a probability space $(\Omega, \mathcal{F}, \mathrm{P})$ indexed by a bounded interval I of $\mathbb{R}_{\geq 0}$. Assume there exist three constants $\alpha, \beta, C > 0$ such that for every $t, s \in I$,*

$$\mathrm{E}\big[|X_t - X_s|^\alpha\big] \leq C|t-s|^{1+\beta}.$$

Then there exists a continuous modification $\tilde{X}$ of X whose sample paths are locally γ-Hölder continuous for every $\gamma \in \left(0, \frac{\beta}{\alpha}\right)$.

Remark 5.2.6. If $I = \mathbb{R}_{\geq 0}$, we may still apply Theorem 5.2.5 successively with $I = [0,1], [1,2], [2,3], \ldots$

We have already seen that the set of random field solutions to the stochastic heat equation represents a Gaussian stochastic process. Due to our interest in finding estimates of the moments of these solutions, we introduce the following lemma containing some properties of the moments of Normal random variables that we will use on various occasions.

Lemma 5.2.7. *Let $p \in (-1, \infty)$. Let Γ be Euler's Gamma function restricted to the real line, thus having the form*

$$\Gamma(x) = \int_0^\infty e^{-t} t^{x-1}\, dt \qquad (x \in \mathbb{R}_{>0}).$$

*We will call it **Euler's real Gamma function**.*

1. *Let $Z \sim N(0, \sigma^2)$ with $\sigma > 0$. Then*

$$\mathrm{E}\big[|Z|^p\big] = c_p \left(\mathrm{E}\big[Z^2\big]\right)^{\frac{p}{2}}, \qquad \text{where} \quad c_p = \frac{2^{\frac{p}{2}}}{\sqrt{\pi}} \Gamma\left(\frac{p+1}{2}\right). \tag{5.2.3}$$

2. *For every $t > 0$,*

$$\int_{\mathbb{R}} |x|^p \Phi(x,t)\, dx = \frac{2^p}{\sqrt{\pi}} t^{\frac{p}{2}} \Gamma\left(\frac{p+1}{2}\right). \tag{5.2.4}$$

3. *Let $Y \sim N(\mu, \sigma^2)$ with $\mu \in \mathbb{R}, \sigma > 0$. Let c_p be defined as in (5.2.3). Then for every $p > 0$,*

$$\mathrm{E}\big[|Y|^p\big] \leq 2^p (1 + c_p) \left(\mathrm{E}\big[Y^2\big]\right)^{\frac{p}{2}}. \tag{5.2.5}$$

Proof. 1. By definition,

$$\mathrm{E}\big[|Z|^p\big] = \int_{\mathbb{R}} |x|^p f_{\sigma^2}(x)\, dx$$

where $f_{\sigma^2}(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{x^2}{2\sigma^2}}$ is the probability density function of Z.
The even parity of this function implies that

$$\int_{\mathbb{R}} |x|^p \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{x^2}{2\sigma^2}}\, dx = \frac{2}{\sqrt{2\pi\sigma^2}} \int_0^\infty x^p e^{-\frac{x^2}{2\sigma^2}}\, dx$$

$$= \left(\frac{2}{\pi}\right)^{\frac{1}{2}} \sigma^{-1} \int_0^\infty x^p e^{-\frac{x^2}{2\sigma^2}}\, dx.$$

With the change of variables $w = \frac{x^2}{2\sigma^2}$ we obtain

$$\begin{aligned}\int_0^\infty x^p e^{-\frac{x^2}{2\sigma^2}}\, dx &= \int_0^\infty 2^{\frac{p-1}{2}} \sigma^{p+1} w^{\frac{p-1}{2}} e^{-w}\, dw \\ &= 2^{\frac{p-1}{2}} \sigma^{p+1} \Gamma\left(\frac{p+1}{2}\right),\end{aligned}$$

therefore

$$\mathrm{E}\big[|Z|^p\big] = \frac{2^{\frac{p}{2}}}{\sqrt{\pi}} \sigma^p\, \Gamma\left(\frac{p+1}{2}\right).$$

The last expression is equivalent to (5.2.3) since $\sigma^2 = \mathrm{Var}(Z) = \mathrm{E}\big[Z^2\big]$.

2. Let $t > 0$. The function $x \mapsto \Phi(x,t)$ is the density of a centered Normal random variable with variance $2t$. Hence we can apply the previous change of variables replacing σ^2 by $2t$ in order to reach (5.2.4).

3. Let $p > 0$. By definition, $Y = \mu + Z$, where $Z \sim N(0, \sigma^2)$. Using the triangle inequality and (5.2.3) we deduce that

$$\mathrm{E}\big[|Y|^p\big] \le 2^p \left(|\mu|^p + \mathrm{E}\big[|Z|^p\big]\right) = 2^p \left(|\mu|^p + c_p \left(\mathrm{E}\big[Z^2\big]\right)^{\frac{p}{2}}\right). \tag{5.2.6}$$

Notice that $\mathrm{E}\big[Y^2\big] = \mu^2 + \mathrm{E}\big[Z^2\big]$ by linearity of the expectation operator. Therefore both $|\mu|^p$ and $\left(\mathrm{E}\big[Z^2\big]\right)^{\frac{p}{2}}$ are bounded by $\left(\mathrm{E}\big[Y^2\big]\right)^{\frac{p}{2}}$ and thus (5.2.6) is bounded by the r.h.s. of (5.2.5). This completes the proof. □

Theorem 5.2.8. *Let I, J be bounded intervals of $\mathbb{R}$ and $\mathbb{R}_{\ge 0}$, respectively. Fix $\alpha_1, \alpha_2 \in (0, 1]$ and define*

$$\Delta(x,t;y,s) = |x-y|^{\alpha_1} + |t-s|^{\alpha_2} \quad \text{and} \quad Q = \frac{1}{\alpha_1} + \frac{1}{\alpha_2}.$$

Let $u = \{u(x,t) \mid \mathbb{R} \times \mathbb{R}_{\ge 0}\}$ be a Gaussian random field and suppose there exists a constant $K \ge 0$ such that for all $(x,t), (y,s) \in I \times J$,

$$\mathrm{E}\left[\big(u(x,t) - u(y,s)\big)^2\right] \le K\big(\Delta(x,t;y,s)\big)^2.$$

Let Γ be Euler's real Gamma function. Then for every $p > 0$, the following inequality holds for all $(x,t), (y,s) \in I \times J$:

$$\mathrm{E}\big[|u(x,t) - u(y,s)|^p\big] \le K_p\big(\Delta(x,t;y,s)\big)^p,$$

where

$$K_p = 2^p \left(1 + \left(\frac{2^p}{\pi}\right)^{\frac{1}{2}} \Gamma\left(\frac{p+1}{2}\right)\right) K^{\frac{p}{2}}.$$

Therefore, u has a continuous modification $\tilde{u}$, and for every $\alpha \in \left(\frac{Q}{p}, 1\right)$, there exist a constant $c > 0$, a bounded function $a(I, J, \alpha, \alpha_1, \alpha_2, p)$ and a positive random variable Y such that

$$\mathrm{E}\big[Y^p\big] \le K_p a(I, J, \alpha, \alpha_1, \alpha_2, p) < \infty$$

and for all $(x,t), (y,s) \in I \times J$,

$$|\tilde{u}(x,t) - \tilde{u}(y,s)| \le cY\big(\Delta(x,t;y,s)\big)^{\alpha - \frac{Q}{p}}.$$

Furthermore, for any pair $(\beta_1, \beta_2) \in (0, \alpha_1) \times (0, \alpha_2)$, $\tilde{u}$ is (β_1, β_2)-Hölder continuous on $I \times J$.

We refer to [5, Section A.4] for a proof and further details.

In addition to Theorem 5.2.8, we need an upper bound on the moments of order 2 of the increments of the fundamental solution $\Phi_{x,t}$. This is provided by Lemma 5.2.10 below. Let us introduce before an algebraic property of Φ concerning semigroups of linear operators that will prove very useful in this section. We refer to [19] for a thorough account of this topic.

Lemma 5.2.9 (**Semigroup Property**). *The fundamental solution Φ to the classical homogeneous heat equation with $k = 1$ satisfies the (strongly continuous one-parameter) semigroup property*

$$\Phi(x,t) = \int_{\mathbb{R}} \Phi(x-y, t-s)\Phi(y,s)\,dy \qquad \forall\, 0 < s < t. \tag{5.2.7}$$

Proof. It follows from a straightforward integration. This property generalizes the convolution product of Normal probability measures. □

Lemma 5.2.10. *Let Φ be the fundamental solution* (2.2.19) *with $k = 1$.*

(i) For all $h \in \mathbb{R}$,

$$\int_0^\infty \int_{\mathbb{R}} \left[\Phi(z,r) - \Phi(z+h,r)\right]^2 \,dz\,dr = \frac{|h|}{2}. \tag{5.2.8}$$

(ii) For all $\tilde{h} \geq 0$,

$$\int_0^\infty \int_{\mathbb{R}} \Big[\Phi(z, r+\tilde{h}) - \Phi(z,r)\Big]^2 \,dz\,dr = \frac{\sqrt{2}-1}{\sqrt{2\pi}}\sqrt{\tilde{h}}. \tag{5.2.9}$$

(iii) For all $(x,t),(y,s) \in \mathbb{R} \times \mathbb{R}_{\geq 0}$,

$$\int_0^\infty \int_{\mathbb{R}} \left[\Phi(x-z, t-r) - \Phi(y-z, s-r)\right]^2 \,dz\,dr \leq \left(2^{-\frac{1}{2}}|x-y|^{\frac{1}{2}} + \pi^{-\frac{1}{4}}|t-s|^{\frac{1}{4}}\right)^2. \tag{5.2.10}$$

Proof. (i) Developing the square of the integrand and using the semigroup property (5.2.7) together with (5.1.5), we obtain

$$\begin{aligned}
&\int_0^\infty \int_{\mathbb{R}} \left[\Phi(z,r) - \Phi(z+h,r)\right]^2 \,dz\,dr \\
&\qquad = \int_0^\infty \int_{\mathbb{R}} \left[\Phi^2(z,r) + \Phi^2(z+h,r) - 2\Phi(z,r)\Phi(z+h,r)\right] \,dz\,dr \\
&\qquad = \int_0^\infty \left[2\frac{1}{\sqrt{8\pi r}} - 2\Phi(-h,2r)\right] \,dr \\
&\qquad = 2\int_0^\infty \left(\frac{1}{\sqrt{8\pi r}} - \frac{1}{\sqrt{8\pi r}} e^{-\frac{(-h)^2}{8r}}\right) \,dr \\
&\qquad = 2\int_0^\infty \frac{1}{\sqrt{8\pi r}} \left(1 - e^{-\frac{h^2}{8r}}\right) \,dr.
\end{aligned}$$

The change of variables $w = \frac{|h|}{2\sqrt{2r}}$ gives

$$2\int_0^\infty \frac{1}{\sqrt{8\pi r}} \left(1 - e^{-\frac{h^2}{8r}}\right) \,dr = \frac{|h|}{2\sqrt{\pi}} \int_0^\infty \frac{1 - e^{-w^2}}{w^2}\,dw. \tag{5.2.11}$$

Integrating by parts the integral in the r.h.s. of (5.2.11) yields

$$\int_0^\infty \frac{1-e^{-w^2}}{w^2}\,dw = 2\int_0^\infty e^{-w^2}\,dw = \sqrt{\pi}.$$

This implies (5.2.8).

(ii) We use a similar procedure to the one used in (i):

$$\begin{aligned}
&\int_0^\infty \int_{\mathbb{R}} \Big[\Phi(z, r+\tilde{h}) - \Phi(z,r)\Big]^2\,dz\,dr \\
&\qquad = \lim_{t\to\infty} \int_0^t \int_{\mathbb{R}} \Big[\Phi^2(z,r+\tilde{h}) + \Phi^2(z,r) - 2\Phi(z,r+\tilde{h})\Phi(z,r)\Big]\,dz\,dr \\
&\qquad = \lim_{t\to\infty} \int_0^t \left[\frac{1}{\sqrt{8\pi(r+\tilde{h})}} + \frac{1}{\sqrt{8\pi r}} - 2\Phi(0, 2r+\tilde{h})\right] dr \\
&\qquad = \lim_{t\to\infty} \int_0^t \left(\frac{1}{\sqrt{8\pi(r+\tilde{h})}} + \frac{1}{\sqrt{8\pi r}} - \frac{2}{\sqrt{4\pi(2r+\tilde{h})}}\right) dr \\
&\qquad = \lim_{t\to\infty} \frac{1}{\sqrt{8\pi}} \int_0^t \left(\frac{1}{\sqrt{r+\tilde{h}}} + \frac{1}{\sqrt{r}} - \frac{\sqrt{8}}{\sqrt{2r+\tilde{h}}}\right) dr \\
&\qquad = \frac{1}{\sqrt{2\pi}} \lim_{t\to\infty} \left[\sqrt{t+\tilde{h}} - \sqrt{\tilde{h}} + \sqrt{t} - \sqrt{2}\left(\sqrt{2t+\tilde{h}} - \sqrt{\tilde{h}}\right)\right] \\
&\qquad = \frac{1}{\sqrt{2\pi}} \lim_{t\to\infty} \left[\sqrt{t+\tilde{h}} + \sqrt{t} - \sqrt{2}\sqrt{2t+\tilde{h}} + \sqrt{\tilde{h}}\left(\sqrt{2}-1\right)\right].
\end{aligned}$$

Using standard calculus techniques to find limits we easily check that

$$\lim_{t\to\infty}\left(\sqrt{t+\tilde{h}} + \sqrt{t} - \sqrt{2}\sqrt{2t+\tilde{h}}\right) = 0.$$

As a result, (5.2.9) holds.

(iii) Assume $s \le t$. Since the time variable in Φ is defined on $\mathbb{R}_{\ge 0}$, we have

$$\begin{aligned}
\int_0^\infty \int_{\mathbb{R}} [\Phi(\tilde{x}, t-r) - \Phi(\tilde{x}, s-r)]^2\,d\tilde{x}\,dr &= \int_0^s \int_{\mathbb{R}} [\Phi(\tilde{x}, t-r) - \Phi(\tilde{x}, s-r)]^2\,d\tilde{x}\,dr \\
&\quad + \int_s^t \int_{\mathbb{R}} \Phi(\tilde{x}, t-r)^2\,d\tilde{x}\,dr. \qquad (5.2.12)
\end{aligned}$$

Applying the triangle inequality to the integral in (5.2.10) we obtain

$$\left[\int_0^\infty \int_{\mathbb{R}} [\Phi(x-z, t-r) - \Phi(y-z, s-r)]^2\,dz\,dr\right]^{\frac{1}{2}} \le T_1 + T_2$$

with

$$\begin{aligned}
T_1^2 &\coloneqq \int_0^\infty \int_{\mathbb{R}} [\Phi(x-z, t-r) - \Phi(x-z, s-r)]^2\,dz\,dr, \\
T_2^2 &\coloneqq \int_0^\infty \int_{\mathbb{R}} [\Phi(x-z, s-r) - \Phi(y-z, s-r)]^2\,dz\,dr.
\end{aligned}$$

The identities (5.1.5), (5.2.9) and (5.2.8), together with the expression (5.2.12), imply

$$T_1^2 = \frac{\sqrt{2}-1}{\sqrt{2\pi}}\sqrt{t-s} + \frac{\sqrt{t-s}}{\sqrt{2\pi}} = \frac{1}{\sqrt{\pi}}\sqrt{t-s},$$

$$T_2^2 = \frac{|y-x|}{2} = \frac{1}{2}|x-y|.$$

Therefore,

$$\left[\int_0^\infty \int_{\mathbb{R}} \left[\Phi(x-z,t-r) - \Phi(y-z,s-r)\right]^2 \, dz\, dr\right]^{\frac{1}{2}} \leq \pi^{-\frac{1}{4}}|t-s|^{\frac{1}{4}} + 2^{-\frac{1}{2}}|x-y|^{\frac{1}{2}}$$

and we have (5.2.10). □

We are now able to prove local Hölder continuity of the sample paths of v and, moreover, to give separate upper bounds on Hölder's exponents.

Proposition 5.2.11. *For all $(x,t),(y,s) \in \mathbb{R}\times\mathbb{R}_{\geq 0}$,*

$$\mathrm{E}\left[\big(v(x,t)-v(y,s)\big)^2\right] \leq \left(2^{-\frac{1}{2}}|x-y|^{\frac{1}{2}} + \pi^{-\frac{1}{4}}|t-s|^{\frac{1}{4}}\right)^2 \tag{5.2.13}$$

and there exists a continuous modification $\tilde{v}$ of the process $v = \{v(x,t) \mid (x,t) \in \mathbb{R}\times\mathbb{R}_{\geq 0}\}$ whose sample paths are locally (α,β)-Hölder continuous, with $\alpha \in \left(0,\frac{1}{2}\right)$ and $\beta \in \left(0,\frac{1}{4}\right)$.

Proof. By (5.2.1) and (5.2.10),

$$\begin{aligned}
\mathrm{E}\left[\big(v(x,t)-v(y,s)\big)^2\right] &= \mathrm{E}\left[\left(\int_0^t \int_{\mathbb{R}} \Phi_{x,t}(z,r)W(dz,dr) - \int_0^s \int_{\mathbb{R}} \Phi_{y,s}(z,r)W(dz,dr)\right)^2\right] \\
&= \int_0^\infty \int_{\mathbb{R}} \left[\Phi_{x,t}(z,r) - \Phi_{y,s}(z,r)\right]^2 \, dz\, dr \\
&\leq \left(2^{-\frac{1}{2}}|x-y|^{\frac{1}{2}} + \pi^{-\frac{1}{4}}|t-s|^{\frac{1}{4}}\right)^2.
\end{aligned}$$

Since the process v is Gaussian, we can apply Theorem 5.2.8 with $K = \max\left\{2^{-\frac{1}{2}}, \pi^{-\frac{1}{4}}\right\}$, $\alpha_1 = \frac{1}{2}$ and $\alpha_2 = \frac{1}{4}$. □

Our next step is to prove that the sample paths of the homogeneous solution $I_0(x,t)$ are globally Hölder continuous. This is ensured by Lemma 5.2.12 below, which, moreover, provides a ratio between the space-time Hölder's exponents.

Lemma 5.2.12. *Assume the initial condition u_0 given by* (5.1.3) *belongs to $C^{0,\eta}(\mathbb{R})$ for a certain $\eta \in (0,1]$. Then the function*

$$(x,t) \mapsto I_0(x,t) = \int_{\mathbb{R}} \Phi(x-y,t)u_0(y)\, dy$$

defined on $\mathbb{R}\times\mathbb{R}_{\geq 0}$ is $\left(\eta,\frac{\eta}{2}\right)$-Hölder continuous.

Proof. Let $(x,t) \in \mathbb{R}\times\mathbb{R}_{\geq 0}$. According to Remark 5.1.2, $I_0(x,t)$ is given by a convolution operation in space. Let us denote it by $(\Phi_t * u_0)(x)$.

Let $h > 0$ be the increment in the spatial variable. The corresponding increment in $I_0(x,t)$ is thus

$$\begin{aligned}
I_0(x+h,t) - I_0(x,t) &= (\Phi_t * u_0)(x+h) - (\Phi_t * u_0)(x) \\
&= \int_{\mathbb{R}} \left[\Phi(x+h-y,t) - \Phi(x-y,t)\right] u_0(y)\, dy \\
&= \int_{\mathbb{R}} \Phi(y,t)\left[u_0(x+h-y) - u_0(x-y)\right]\, dy
\end{aligned}$$

by the commutativity property of convolutions.

By hypothesis, $u_0 \in C^{0,\eta}(\mathbb{R})$. We denote its Hölder's constant by $c(\eta, u_0)$. It follows that

$$\begin{aligned}\forall\, T > 0, \quad \sup_{t\in[0,T]} |I_0(x+h,t) - I_0(x,t)| &\leq \sup_{t\in[0,T]} \int_{\mathbb{R}} \Phi(y,t)|u_0(x+h-y) - u_0(x-y)|\, dy \\ &\leq \sup_{t\in[0,T]} \int_{\mathbb{R}} \Phi(y,t) c(\eta,u_0) h^\eta\, dy\end{aligned}$$

and, since $\int_{\mathbb{R}} \Phi(y,t)\, dy = 1$ by definition,

$$\sup_{t\in[0,T]} |I_0(x+h,t) - I_0(x,t)| \leq c(\eta,u_0) h^\eta. \tag{5.2.14}$$

Next, we study the time shift. Let $x \in \mathbb{R}, t > 0, s \in [0,t)$. We have

$$\begin{aligned}|I_0(x,t) - I_0(x,s)| &= |(\Phi_t * u_0)(x) - (\Phi_s * u_0)(x)| \\ &= \left| \int_{\mathbb{R}} \Phi(x-y,t) u_0(y)\, dy - \int_{\mathbb{R}} \Phi(x-y,s) u_0(y)\, dy \right|.\end{aligned}$$

Applying the semigroup property (5.2.7) and Fubini gives

$$\begin{aligned}\int_{\mathbb{R}} \Phi(x-y,t) u_0(y)\, dy &= \int_{\mathbb{R}} \left(\int_{\mathbb{R}} \Phi(x-z,t-s)\Phi(z-y,s)\, dz \right) u_0(y)\, dy \\ &= \int_{\mathbb{R}} \Phi(x-z,t-s) \left(\int_{\mathbb{R}} \Phi(z-y,s) u_0(y)\, dy \right) dz \\ &= \big(\Phi_{t-s} * I_0(z,s)\big)(x).\end{aligned}$$

In addition,

$$\begin{aligned}\int_{\mathbb{R}} \Phi(x-y,s) u_0(y)\, dy &= \int_{\mathbb{R}} \Phi(x-z,t-s) \left(\int_{\mathbb{R}} \Phi(x-y,s) u_0(y)\, dy \right) dz \\ &= \big(\Phi_{t-s} * I_0(x,s)\big)(x).\end{aligned}$$

given that $\int_{\mathbb{R}} \Phi(x-z,t-s)\, dz = 1$.

Therefore,

$$\begin{aligned}|I_0(x,t) - I_0(x,s)| &= \big|\big(\Phi_{t-s} * I_0(z,s)\big)(x) - \big(\Phi_{t-s} * I_0(x,s)\big)(x)\big| \\ &= \big|\Phi_{t-s} * \big(I_0(z,s) - I_0(x,s)\big)(x)\big|\end{aligned}$$

by distributivity of convolutions.

We can now use two previous results to obtain an upper bound on the time increment. Using (5.2.14),

$$\begin{aligned}\big|\Phi_{t-s} * \big(I_0(z,s) - I_0(x,s)\big)(x)\big| &= \left| \int_{\mathbb{R}} \Phi(x-z,t-s)\big[I_0(z,s) - I_0(x,s)\big]\, dz \right| \\ &\leq \left(\int_{\mathbb{R}} \Phi(x-z,t-s)\, |I_0(z,s) - I_0(x,s)|\, dz \right) \\ &\leq c(\eta,u_0) \int_{\mathbb{R}} \Phi(x-z,t-s)|x-z|^\eta\, dz. \end{aligned} \tag{5.2.15}$$

The value of the last integral in (5.2.15) is given by (5.2.4). Hence, we have the upper bound

$$|I_0(x,t) - I_0(x,s)| \leq c(\eta,u_0) \frac{2^\eta}{\sqrt{\pi}} |t-s|^{\frac{\eta}{2}} \Gamma\left(\frac{\eta+1}{2}\right). \tag{5.2.16}$$

From (5.2.14) and (5.2.16) we conclude that there exists a constant $\tilde{c} > 0$ such that

$$|I_0(x+h,t) - I_0(x,s)| \leq c(\eta,u_0)\tilde{c}\left(h^\eta + |t-s|^{\frac{\eta}{2}}\right). \tag{5.2.17}$$

This proves the claim on Hölder continuity of I_0. □

Our final step is to prove local Hölder continuity of the sample paths of random field solutions. We will consider two cases.

Proposition 5.2.13. *Let u_0 be an initial value to* (5.1.6) *defined under the assumptions of Remark 5.1.2. Let $u = \{u(x,t) \mid (x,t) \in \mathbb{R} \times \mathbb{R}_{\geq 0}\}$ be the stochastic process given by* (5.2.2), (5.2.1). *Fix $L, T > 0$.*

1. ***Continuity without the initial time**. Fix $0 < t_0 \leq T$. For every $p \geq 2$, there exists a constant $C = C(p, t_0, T, L, u_0) > 0$ such that for all $(x,t), (y,s) \in [-L, L] \times [t_0, T]$,*

$$\mathrm{E}\big[|u(x,t) - u(y,s)|^p\big] \leq C\left(|x-y|^{\frac{1}{2}} + |t-s|^{\frac{1}{4}}\right)^p. \tag{5.2.18}$$

Hence, there exists a continuous modification $\tilde{u}$ of u with (α, β)-Hölder continuous sample paths on $[-L, L] \times (0, T]$, where $\alpha \in \left(0, \frac{1}{2}\right)$ and $\beta \in \left(0, \frac{1}{4}\right)$.

2. ***Continuity including the initial time**. Suppose u_0 satisfies the assumption in Lemma 5.2.12. Then for every $p \geq 2$, there exists a constant $C = C(p, \eta, T, L, u_0) > 0$ such that for all $(x,t), (y,s) \in [-L, L] \times [0, T]$,*

$$\mathrm{E}\big[|u(x,t) - u(y,s)|^p\big] \leq C\left(|x-y|^{\frac{1}{2}\wedge\eta} + |t-s|^{\frac{1}{4}\wedge\frac{\eta}{2}}\right)^p. \tag{5.2.19}$$

Moreover, there exists a continuous modification $\tilde{u}$ of u with Hölder continuous sample paths on $[-L, L] \times [0, T]$ with exponents (α, β) subject to the following conditions:

$$\begin{cases} \alpha \in \left(0, \frac{1}{2}\right), \beta \in \left(0, \frac{1}{4}\right) & \text{if} \quad \eta \geq \frac{1}{2} \\ \alpha \in (0, \eta), \beta \in \left(0, \frac{\eta}{2}\right) & \text{if} \quad \eta < \frac{1}{2}. \end{cases} \tag{5.2.20}$$

Proof. Let $p \geq 2$.

1. Let $(x,t), (y,s) \in [-L, L] \times [t_0, T]$. Since $u(x,t) = I_0(x,t) + v(x,t)$, the triangle inequality implies

$$\|u(x,t) - u(y,s)\|_{L^p(\Omega,\mathcal{F},\mathrm{P})} \leq |I_0(x,t) - I_0(y,s)| + \|v(x,t) - v(y,s)\|_{L^p(\Omega,\mathcal{F},\mathrm{P})}. \tag{5.2.21}$$

We know that I_0 is C^∞ on $\mathbb{R} \times \mathbb{R}_{>0}$ with uniformly bounded derivatives of all orders on $[-L, L] \times [t_0, T]$. Therefore, I_0 is Lipschitz continuous relative to space and time on the compact $[-L, L] \times [t_0, T]$. Consequently, there exists a constant $C_0 = C_0(t_0, T, L, u_0) > 0$ such that

$$|I_0(x,t) - I_0(y,s)| \leq C_0\left(|x-y| + |t-s|\right). \tag{5.2.22}$$

By definition, $v = \{v(x,t) \mid (x,t) \in \mathbb{R} \times \mathbb{R}_{\geq 0}\}$ is a centered Gaussian random field. Therefore, by (5.2.3), every random variable $v(x,t)$ satisfies the relation

$$\|v(x,t) - v(y,s)\|^p_{L^p(\Omega,\mathcal{F},\mathrm{P})} = c_p\|v(x,t) - v(y,s)\|^p_{L^2(\Omega,\mathcal{F},\mathrm{P})} = c_p\mathrm{E}\Big[\big(v(x,t) - v(y,s)\big)^2\Big]^{\frac{p}{2}}. \tag{5.2.23}$$

Inserting (5.2.22) and (5.2.23) in (5.2.21) yields

$$\begin{aligned} \mathrm{E}\big[|u(x,t) - u(y,s)|^p\big] &\leq \left(C_0\left(|x-y| + |t-s|\right) + c_p^{1/p}\mathrm{E}\Big[\big(v(x,t) - v(y,s)\big)^2\Big]^{\frac{1}{2}}\right)^p \\ &\leq \left[C_0\left(|x-y| + |t-s|\right) + c_p^{1/p}\left(2^{-\frac{1}{2}}|x-y|^{\frac{1}{2}} + \pi^{-\frac{1}{4}}|t-s|^{\frac{1}{4}}\right)\right]^p \end{aligned}$$

where the second inequality stems from the upper bound (5.2.13).

It is easy to see that the last expression is bounded from above by

$$C_1 \left(|x-y| + |t-s| + |x-y|^{\frac{1}{2}} + |t-s|^{\frac{1}{4}} \right)^p$$

for some constant $C_1 = C_1(p, t_0, T, L, u_0) > 0$. Given that $|x-y| \leq 2L$ and $|t-s| < T$, there exists a constant $C = C(p, t_0, T, L, u_0) > 0$ such that (5.2.18) holds.

Lastly, since (5.2.18) is satisfied with $p = 2$, the statement on local Hölder continuity of sample paths arises from Theorem 5.2.8.

2. Let $(x,t), (y,s) \in [-L, L] \times [0, T]$. Applying (5.2.21) and (5.2.23) from the proof of Claim 1, together with the estimates (5.2.17) and (5.2.13), we obtain

$$\mathrm{E}\big[|u(x,t) - u(y,s)|^p\big] \leq \left[c(\eta, u_0)\tilde{c} \left(|x-y|^\eta + |t-s|^{\frac{\eta}{2}} \right) + c_p^{1/p} \left(\frac{|x-y|^{\frac{1}{2}}}{\sqrt{2}} + \frac{|t-s|^{\frac{1}{4}}}{\pi^{\frac{1}{4}}} \right) \right]^p$$

and we can see that there exists a constant $C = C(p, \eta, T, L, u_0) > 0$ that implies (5.2.19).

The claim concerning local Hölder continuity of sample paths with exponents given by (5.2.20) is obtained as in Claim 1. □

Chapter 6

Conclusions

The purpose of this thesis is to show the transition from the classical to the stochastic one-dimensional heat equation and study the latter's continuity properties. To this end, we introduce the necessary ingredients for the definition of the stochastic heat equation on $\mathbb{R}$ and the associated random field solutions.

From simple physical principles, we derive the classical heat equation and subsequently the associated Cauchy problems. Then we focus on the computation and study of the fundamental solution, which we obtain by means of two different approaches. The relevance of this function in the study of the stochastic heat equation on $\mathbb{R}$ is reflected in the last chapter.

In parallel, we introduce the notions of Gaussian random variables, characteristic functions, and Gaussian processes. The latter is the key concept of the second half of this thesis, from the definition of Brownian motion to the study of local Hölder continuity of the sample paths of the random field solutions to the stochastic heat equation on $\mathbb{R}$. Throughout the thesis we focus on the notions that allow us to define a linear stochastic partial differential equation on $\mathbb{R}$. For this reason, we do not deal with other interesting properties of Brownian motion, such as the strong Markov property or the reflection principle, nor we describe more classes of Gaussian processes.

In the last part, we bring together the notions of isonormal Gaussian processes, space-time white noise, and fundamental solution Φ to define the concept of random field solutions, and thereafter we carefully prove the local Hölder continuity of their sample paths.

This thesis highlights the link between the fields of stochastic processes and PDEs by means of the heat equation. This relation, together with the importance of the heat equation in physics, captured my interest and drove me to choose this topic.

Appendix A

Normal and Multivariate Normal Distributions

A.1 Normal Distribution

Definition A.1.1. Let $\mu \in \mathbb{R}$ and $\sigma \in \mathbb{R}_{>0}$. A real-valued random variable X is said to have a *Normal distribution with parameters μ and σ^2* if its probability density function is given by

$$f_X(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}.$$

We denote this law by $N\big(\mu, \sigma^2\big)$. If a random variable X has this distribution, we write $X \sim N\big(\mu, \sigma^2\big)$.

If $\mu = 0$ and $\sigma^2 = 1$, it is known as the *standard* Normal distribution.

Proposition A.1.2 (**Basic Properties**). *Let $X \sim N\big(\mu, \sigma^2\big)$.*

- $\mathrm{E}[X] = \mu$.
- $\mathrm{Var}(X) = \sigma^2$.
- *f_X is symmetric with respect to $x = \mu$.*

Proposition A.1.3 (**Normal Transformation**). *Let $X \sim N(0,1)$. We fix two parameters $\mu \in \mathbb{R}$ and $\sigma \in (0, \infty)$ and we define the random variable $Y = \mu + \sigma X$. Then*

$$f_Y(y) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(y-\mu)^2}{2\sigma^2}} \tag{A.1.1}$$

and therefore $Y \sim N\big(\mu, \sigma^2\big)$.

Proof. The random variable Y is obtained through the function $g\colon \mathbb{R} \to \mathbb{R}, g(x) = \mu + \sigma x$. Since g is C^1, bijective and strictly increasing, we may apply the well-known *density transformation formula*

$$f_Y(y) = f_X\big(g^{-1}(y)\big) \left|\big(g^{-1}\big)'(y)\right| = f_X\big(g^{-1}(y)\big) \frac{1}{\left|g'\big(g^{-1}(y)\big)\right|}$$

to obtain (A.1.1). □

A.2 Multivariate Normal Distribution

Throughout this section, n is a strictly positive integer and $X = (X_1, \ldots, X_n)$ is a real random vector.

Definition A.2.1. The *covariance matrix* Q of X is the $n \times n$ matrix whose general term is

$$Q_{i,j} = \mathrm{Cov}\big(X_i, X_j\big).$$

Proposition A.2.2. *The covariance matrix Q of X is symmetric and positive semidefinite.*

Proof. The symmetry is trivial, since $\mathrm{Cov}\big(X_i, X_j\big) = \mathrm{Cov}\big(X_j, X_i\big) \quad \forall i, j \in \{1, \ldots, n\}$. Moreover, for all $a_1, \ldots, a_n \in \mathbb{R}$,

$$\sum_{i=1}^{n} \sum_{j=1}^{n} a_i a_j Q_{i,j} = \mathrm{Var}\left(\sum_{i=1}^{n} a_i X_i \right),$$

and since variances are always positive, the proof is complete. □

Definition A.2.3. X has a *multivariate Normal* (or *Gaussian*) distribution if every linear combination of its components has a (univariate) Normal distribution.

We denote by $N(\mu, Q)$ its law, where $\mu = (\mu_1, \ldots, \mu_n) \in \mathbb{R}^n$ is the mean vector and Q is the covariance matrix.

Appendix B

Functional Analysis

B.1 Banach Spaces

Definition B.1.1. Let X be a real linear space. A mapping $\|\cdot\|: X \to \mathbb{R}$ is called a *norm* if for each scalar λ and every $x, y \in X$, the following properties hold:

(i) *positivity*: $\|x\| \geq 0$ and $\|x\| = 0$ if and only if $x = 0$;

(ii) *homogeneity*: $\|\lambda x\| = \lambda \|x\|$;

(iii) *triangular inequality*: $\|x + y\| \leq \|x\| + \|y\|$.

Definition B.1.2. A *normed space* is a linear space endowed with a norm.

Hereafter we assume X is a normed linear space.

Definition B.1.3. If every Cauchy sequence in X converges, X is called *complete*.

Definition B.1.4. A *Banach space* is a complete, normed linear space.

Examples. Let $U \subset \mathbb{R}^n$ be a bounded open set.

(i) The space $C^0\big(\overline{U}\big)$ of continuous functions on $\overline{U}$, endowed with the *maximum norm*

$$\|f\|_{C^0\left(\overline{U}\right)} = \max_{\overline{U}} |f|.$$

(ii) *L^p spaces.* Let $p \in [1, \infty]$. For every Lebesgue measurable function $f: U \to \mathbb{R}$, we define the *L^p norm*

$$\|f\|_{L^p(U)} := \begin{cases} \left(\int_U |f|^p\right)^{1/p} & \text{if } p < \infty \\ \operatorname{ess\,sup}_U |f| & \text{if } p = \infty, \end{cases}$$

where $\operatorname{ess\,sup}_U |f|$ denotes the *essential supremum* of $|f|$, i.e. the infimum of all numbers $M \geq 0$ such that $|f(x)| \leq M$ a.e. in U[1].

We denote by $L^p(U)$ the linear space of all Lebesgue measurable functions $f: U \to \mathbb{R}$ such that $\|f\|_{L^p(U)} < \infty$. This space becomes a Banach space when we identify any pair of functions equal a.e. in U and we equip it with the L^p norm. If $p \in [1, \infty)$, $L^p(U)$ is the set of real-valued functions defined on U such that $|f|^p$ is Lebesgue integrable in U. In this case, the L^p norm is called *integral norm of order p*.

[1] A property holds *almost everywhere* (a.e.) in a set if it is true for all points in the set except for a subset of Lebesgue measure zero.

B.2 Hilbert Spaces

Let H be a real linear space.

Definition B.2.1. An *inner product* in H is a mapping $\langle \cdot, \cdot \rangle \colon H \times H \to \mathbb{R}$ with the following properties: for all $x, y, z \in H$ and $\lambda, \mu \in \mathbb{R}$,

(i) *positivity*: $\langle x, x \rangle \geq 0$ and $\langle x, x \rangle = 0$ if and only if $x = 0$;

(ii) *symmetry*: $\langle x, y \rangle = \langle y, x \rangle$;

(iii) *bilinearity*: $\langle \mu x + \lambda y, z \rangle = \mu \langle x, z \rangle + \lambda \langle y, z \rangle$.

A linear space endowed with an inner product is called an *inner product space*.

Definition B.2.2. The *norm* induced by an inner product $\langle \cdot, \cdot \rangle$ is given by

$$\|x\| = \langle x, x \rangle^{1/2} \quad \forall x \in H.$$

Definition B.2.3. A *Hilbert space* H is a Banach space endowed with an inner product which induces the norm.

Definition B.2.4. A Hilbert space is said to be *separable* if it contains a countable dense subset.

Example B.2.5. The space $L^2(U)$ is a separable Hilbert space with respect to the inner product

$$\langle f, g \rangle_{L^2(U)} = \int_U f g \, dx.$$

Definition B.2.6. Two Hilbert spaces H_1 and H_2 are isometric if there exists a bijective linear mapping $L \colon H_1 \to H_2$, called *isometry*, that preserves the norm, that is,

$$\|x\|_{H_1} = \|Lx\|_{H_2} \quad \forall x \in H_1.$$

Definition B.2.7. Two elements x, y belonging to a Hilbert space H endowed with an inner product $\langle \cdot, \cdot \rangle$ are *orthonormal* if

$$\begin{cases} \langle x, y \rangle = 0 \\ \|x\| = \|y\| = 1. \end{cases}$$

A set of mutually orthonormal elements is an *orthonormal set*.

Definition B.2.8. An *orthonormal basis* in a separable Hilbert space H is a countable sequence $\{u_i\}_{i \geq 1} \subset H$ such that its elements are pairwise orthonormal and every $x \in H$ may be expanded in the form

$$x = \sum_{i=1}^{\infty} \langle x, u_i \rangle u_i.$$

The series above is called the *generalized Fourier series*. Moreover, the following expression, known as *Parseval's identity*, holds:

$$\|x\|^2 = \sum_{i=1}^{\infty} \langle x, u_i \rangle^2.$$

We end this appendix with two useful propositions from [22, Section 6.4.2].

Proposition B.2.9 (**Characterization of Orthonormal Bases**). *Let H be a separable Hilbert space. An orthonormal sequence $\{u_i\}_{i\geq 1} \subset H$ is a basis for H if and only if one of the following conditions is satisfied:*

(i) if $x \in H$ is orthogonal to u_i for every $i \geq 1$, then $x = 0$;

(ii) the set of all finite linear combinations of the elements of $\{u_i\}_{i\geq 1}$ is dense in H.

Proposition B.2.10. *Every separable Hilbert space admits a countable orthonormal basis.*

Bibliography

[1] A. Alabert. *Mesura i Probabilitat*. Servei de Publicacions de la UAB, 1997. [https://gent.uab.cat/aureli_alabert/sites/gent.uab.cat.aureli_alabert/files/mesuraprobabilitat.pdf].

[2] C.-E. Bréhier. *A short introduction to Stochastic PDEs*, 2014. hal-00973887v2. [https://hal.archives-ouvertes.fr/hal-00973887v2/document].

[3] Z. Brzeźniak and T. Zastawniak. *Basic Stochastic Processes*. Springer-Verlag London, 1999.

[4] R. Cuppens. *Decomposition of Multivariate Probabilities*. Elsevier, 1975.

[5] R. C. Dalang and M. Sanz-Solé. *An Introduction to Stochastic Partial Differential Equations*. Preliminary version 09.12.2020.

[6] A. Eberle. *Stochastic Analysis. An Introduction*, June 2015. [https://wt.iam.uni-bonn.de/fileadmin/WT/Inhalt/people/Andreas_Eberle/StoAn1011/StoAnSkriptneu.pdf].

[7] L. C. Evans. *Partial Differential Equations*. American Mathematical Society, Providence, Rhode Island, second edition, 2010.

[8] G. B. Folland. *Fourier Analysis and Its Applications*. American Mathematical Society, 1992.

[9] J. B. J. Fourier. *Théorie Analytique de la Chaleur*. F. Didot, Paris, 1822.

[10] A. Genthon. *The concept of velocity in the history of Brownian motion*. [https://arxiv.org/pdf/2006.05399v1.pdf].

[11] M. J. Hancock. *The 1-D Heat Equation*. 18.303 Linear Partial Differential Equations. Fall 2006. Massachusetts Institute of Technology: MIT OpenCourseWare, [https://ocw.mit.edu]. License: Creative Commons BY-NC-SA. [https://ocw.mit.edu/courses/mathematics/18-303-linear-partial-differential-equations-fall-2006/lecture-notes/heateqni.pdf].

[12] J. Jacod and P. Protter. *Probability Essentials*. Springer-Verlag Berlin Heidelberg, second edition, 2004.

[13] I. Karatzas and S. E. Shreve. *Brownian Motion and Stochastic Calculus*. Springer-Verlag, second edition, 1991.

[14] D. Khoshnevisan. *Probability*. American Mathematical Society, Providence, Rhode Island, 2007.

[15] J.-F. Le Gall. *Brownian Motion, Martingales, and Stochastic Calculus*. Springer International Publishing Switzerland, 2016.

[16] P. Mörters and Y. Peres. *Brownian Motion*, volume 30 of *Cambridge Series in Statistical and Probabilistic Mathematics*. Cambridge University Press, 2010. [https://people.bath.ac.uk/maspm/book.pdf].

[17] T. Narasimhan. Fourier's heat conduction equation: History, influence and connections. *Reviews of Geophysics*, 37(1):151–172, February 1999. [https://agupubs.onlinelibrary.wiley.com/doi/epdf/10.1029/1998RG900006].

[18] D. Nualart i M. Sanz-Solé. *Curs de Probabilitats*. PPU, Barcelona, 1990.

[19] A. Pazy. *Semigroups of Linear Operators and Applications to Partial Differential Equations*. Springer-Verlag New York, 1983.

[20] M. A. Pinsky and S. Karlin. *An Introduction to Stochastic Modeling*. Elsevier, fourth edition, 2011.

[21] D. Revuz and M. Yor. *Continuous Martingales and Brownian Motion*. Springer-Verlag Berlin Heidelberg, third edition, 1999.

[22] S. Salsa. *Partial Differential Equations in Action. From Modelling to Theory*. Springer International Publishing Switzerland, third edition, 2016.

[23] M. Sanz-Solé. *An Introduction to Stochastic Calculus*. Universitat de Barcelona, October 2020.

[24] L. Schwartz. *Théorie des Distributions*. Hermann, Paris, 1966.

[25] P. Todorovic. *An Introduction to Stochastic Processes and Their Applications*. Springer-Verlag New York, 1992.

[26] J. Vives. Notes de curs de l'assignatura *Probabilitats Avançades*. Universitat de Barcelona, 2018.