TÉCNICO LISBOA

On the Derivation of the Navier–Stokes Equations from a Nondeterministic Variational Principle

Marcos Jorge Martins Carvalho Latas

Thesis to obtain the Master of Science Degree in

Mathematics and Applications

Supervisor: Professor Ana Bela Ferreira Cruzeiro Zambrini

Examination Committee

Chairperson: Professor António Manuel Pacheco Pires
Supervisor: Professor Ana Bela Ferreira Cruzeiro Zambrini
Members of the Committee: Professor Ana Patrícia Carvalho Gonçalves
Professor José Manuel Vergueiro Monteiro Cidade Mourão

October 2019

Acknowledgement

I thank Professor Ana Bela Cruzeiro, on whose research the third chapter of this work is based, for taking the time to supervise this dissertation.

Abstract

The mathematical description of fluid dynamics, which models the flow of matter in all the possible states with the exception of the solid state, is of fundamental importance to many branches of physics, due to the wide range of phenomena it predicts and explains. The Navier–Stokes equations, among the most important equations of mathematical physics, provide a very accurate description of the motion of viscous, incompressible fluids with uniform density. Discovering a general theory of these equations remains one of the great unsolved problems of mathematics. In the first chapter of this dissertation, we essentially review the derivation of the Navier–Stokes equations from the principles of conservation of matter and conservation of momentum. In the second chapter we derive the diffusion equation and its fundamental solution, using the latter to define the Wiener measure. Next we define the Wiener process and present a concise description of some of its possible constructions. We provide an overview of functional integration and of the construction of integrals with respect to martingales of various types, and subsequently we define stochastic differential equations, with and without reflecting boundaries. We describe the concept of local time and present a few results related to it. In the third chapter we explain how the Navier–Stokes equations with periodic boundary conditions can be derived from a nondeterministic variational principle, and adapt the results obtained to the case when we have Neumann boundary conditions in a different domain.

Keywords: Navier–Stokes equations, Wiener process, stochastic integration, reflecting stochastic differential equations, local time, nondeterministic variational principle.

Resumo

A descrição matemática da dinâmica de fluidos, que modela o fluxo da matéria em todos os estados possíveis com a excepção do estado sólido, é de uma importância fundamental para muitos ramos da física, devido à vasta gama de fenómenos que prevê e explica. As equações de Navier–Stokes, que estão entre as mais importantes equações da física-matemática, fornecem uma descrição muito precisa do movimento de fluidos viscosos e incompressíveis com densidade uniforme. Descobrir uma teoria geral destas equações permanece um dos grandes problemas ainda não resolvidos da matemática. No primeiro capítulo desta dissertação revemos essencialmente a derivação das equações de Navier–Stokes a partir dos princípios da conservação da matéria e da conservação do momento. No segundo capítulo derivamos a equação da difusão e a sua solução fundamental, usando a segunda para definir a medida de Wiener. De seguida definimos o processo de Wiener e fornecemos uma descrição concisa de algumas das suas possíveis construções. Providenciamos um panorama da integração funcional e da construção de integrais em ordem a martingalas de vários tipos, e subsequentemente definimos equações diferenciais estocásticas, com e sem fronteiras que reflectem. Descrevemos o conceito de tempo local e apresentamos alguns resultados relacionados com ele. No terceiro capítulo explicamos como as equações de Navier–Stokes com condições de fronteira periódicas podem ser derivadas a partir de um princípio variacional não-determinístico e adaptamos os resultados obtidos ao caso de termos condições de fronteira de Neumann num domínio diferente.

Palavras-chave: equações de Navier–Stokes, processo de Wiener, integração estocástica, equações diferenciais estocásticas com reflexão, tempo local, princípio variacional não-determinístico.

Contents

Frequently Used Notation

General Notation

$:=$	equal to, by definition
f^+	$\max\{f, 0\}$, where f can be various mathematical objects, such as functions or numbers
f^-	$-\min\{f, 0\}$, where the same proviso as above applies
a.s.	(convergent) almost surely, or almost everywhere (p. 74)
∂D	the boundary of the set D
$\overline{D}$	the closure of the set D

Special Constants

μ	the shear viscosity coefficient (p. 16)
ν	the kinematic viscosity coefficient (p. 18)
c	the specific heat (p. 23)
κ	the thermal conductivity coefficient (p. 23)

Sets and Spaces

$\mathbb{R}_+$	the interval $[0, \infty[$
$\overline{\mathbb{R}}$	the interval $[-\infty, +\infty]$
$(X, \mathcal{M}, \mu)$	a measure space (p. 69)
$(\Omega, \mathcal{F}, \mathcal{F}_t, P)$	a filtered probability space (p. 74)
$C_0(\mathbb{R}_+)$	the Wiener space (p. 27)
$C(\mathbb{R}_+, \mathbb{R}^d)$	the space of continuous functions $f\colon \mathbb{R}_+ \to \mathbb{R}^d$
$L^2([a, b])$	the Hilbert space of all square-integrable functions on $[a, b]$, $a < b$ (p. 30)
$\mathcal{M}_2$	the space of square-integrable martingales (p. 33)
$\mathcal{M}_2^{\mathsf{c}}$	the space of a.s. continuous, square-integrable martingales (p. 33)
$\mathcal{L}_2$	the space of real-valued, $\mathcal{F}_t$-adapted, measurable stochastic processes that verify condition (2.40) (p. 33)
$\mathcal{L}_2^{\mathsf{loc}}$	the space of real-valued, $\mathcal{F}_t$-adapted, measurable stochastic processes that verify condition (2.47) (p. 35)
$\mathcal{L}_0$	the space of all simple processes (p. 34)
$\mathcal{M}^{\mathsf{loc}}$	the space of local martingales (p. 35)
$\mathcal{M}^{\mathsf{c,loc}}$	the space of continuous, local martingales (p. 35)
$\mathcal{M}_2^{\mathsf{loc}}$	the space of local, square-integrable martingales (p. 37)

$\mathcal{M}_2^{\mathrm{c,loc}}$ the space of continuous, local, square-integrable martingales (p. 35)

$\mathcal{L}_2(M)$ the space of real-valued, predictable stochastic processes that verify condition (2.52) (p. 36)

$\mathbb{T}$ a flat torus of dimension d

$\mathbb{H}$ a linear subspace dense in $L^2([0,T]\times\mathbb{T})$ (p. 54)

Functions, Operators and Functionals

u the velocity

p the pressure

ρ the density, i.e. the mass per unit volume

∇u the gradient of u (p. 8)

$\nabla\cdot u$ or $\operatorname{div} u$ the divergence of u (p. 9)

$\mathrm{D}f/\mathrm{D}t$ the material, or Lagrangian, derivative of f (p. 8)

Δu or $\nabla\cdot\nabla u$ the Laplacian of u (p. 17)

E the expectation (p. 76)

$\mathbb{1}_A(f)$ the indicator function of f in the set A, which returns 1 if $f\in A$, and returns 0 otherwise (p. 73)

$\operatorname{sgn}(x)$ the signal function of x, which returns 1 if $x>0$, and returns -1 otherwise

$\delta(x)$ the Dirac delta function evaluated at x (p. 48)

$\Pr$ the probability (p. 69)

S the action functional (p. 51)

δS the variation of the action (p. 51)

$\mathrm{D}_t M$ the generalized derivative, or mean forward derivative, of the martingale M (p. 53)

σ-Algebras, Stochastic Processes and Martingales

W or W_t the Wiener process (p. 28)

$\mathrm{d}W$ the stochastic differential of W (p. 38)

$\mathcal{B}(\mathbb{R}^N)$ the Borel σ-algebra (p. 70)

$\mathcal{M}\otimes\mathcal{N}$ the product of σ-algebras $\mathcal{M}$ and $\mathcal{N}$ (p. 70)

$\mathcal{F}_t^X$ the filtration generated by the process X (p. 74)

$\{\mathcal{F}_t : t\in\mathbb{R}_+\}$ a filtration (p. 74)

M or $M_t(\omega)$ a martingale with respect to a filtration (p. 32)

$\langle M\rangle$ the quadratic variation process of M (p. 35)

$\langle M,N\rangle$ the cross-variation process of M and N (p. 36)

$L_t(x)$ the local time of a process (p. 46)

Introduction

The concept of fluid encompasses all the possible states of matter that are not the solid state. The mathematical description of its dynamics is of fundamental importance to many branches of physics, because of the wide range of phenomena it predicts and explains. The Navier–Stokes equations, among the most important equations of mathematical physics, are second-order, non-linear partial differential equations that provide a very accurate description of the motion of viscous, incompressible fluids with uniform density. A general theory of these equations is (famously) non-existent yet, and analytic results concerning the existence and uniqueness of smooth solutions are scarce. In particular, the theory of these equations in $\mathbb{R}^3$, the most important case, is incipient, and is likely to remain that way for many decades (if not centuries). Due to the incompleteness of the theory of the Navier–Stokes equations, their applications rely mostly on the solutions obtained via numerical methods.

We present in the first chapter of this work the derivation of the Navier–Stokes equations from the general physical principles of conservation of matter and conservation of momentum. For this derivation we need little more mathematical tools than the divergence theorem and the Reynolds transport theorem (which we will refer to simply as the transport theorem). We describe also the range of applicability of the equations, the concept of weak solution, and we review some of the main analytic results that have been obtained, which are not many. We end the chapter by listing a few of the areas and problems to which the equations are applied, and present a very concise sketch of the history of the deduction of the equations, obtained by 18th and 19th century mathematicians.

In the second chapter we start by deriving the diffusion equation, representative of one of the three main types of linear partial differential equations, namely the parabolic equations, from the principle of conservation of energy and the Fourier law, and obtain its fundamental solution. This equation models, among other phenomena, the distribution of temperature in a bounded region, and is for this reason also known as the heat equation. The fundamental solution we obtained is used in the definition of the Wiener measure, a probability measure that was introduced with the goal of assigning a probability to the random path of a Brownian particle. This measure can in turn be used in the definition of integrals with respect to it, namely the Wiener integral and its extensions. We define next the Wiener process and present a very short description of some of its possible constructions. The construction of the Wiener process was one of the major mathematical results of the 20th century and laid the foundations for a mathematically rigorous theory of stochastic processes. We proceed by giving a brief overview of the two main approaches to the definition of functional integration, one of which, stochastic integration, is

of central importance to our purposes. We offer a short exposition of the construction of integrals of processes with respect to martingales of various types. One of the integrals we describe is the Itô integral, which is defined with respect to the Wiener process, and which we will use extensively in the next chapter. Many other essential concepts and results are also expounded, such as the notion of diffusion process, semimartingale, and the change-of-variable formula, also known as the Itô formula. Since it is of capital importance to us, we define next what a stochastic differential equation is, with and without reflecting boundaries. In both cases we explain the conditions for the existence and uniqueness of solutions. The concept of stochastic differential equation with reflection is connected to the Skorokhod problem, which we also describe. The notion of local time of a continuous semimartingale is defined in the last section of this chapter and we provide several deductions and comments on it, ending the chapter with an adaptation of the arguments that lead us to the conclusion that one of the functions of the Skorokhod problem is a local time at point zero of a certain reflecting stochastic differential equation defined in $\mathbb{R}_+$.

In the third chapter we start by explaining the results obtained in [5, pp. 2–5], namely how the Navier–Stokes equations can be derived from a nondeterministic, or stochastic, variational principle where the pressure is introduced as a Lagrange multiplier. We introduce a certain semimartingale and a certain stochastic differential equation, the latter representing diffusions on the domain being considered, and after various computations we show the central result of this chapter, which states that the above diffusions and a function representing the pressure are critical to a certain action functional (which we also define) if and only if the drift of the previously defined stochastic differential equation satisfies the Navier–Stokes equations with periodic boundary conditions. To obtain this result we assume that the Navier–Stokes equations are defined on the domain $[0, T] \times \mathbb{T}$, where $\mathbb{T}$ is the d-dimensional flat torus, which we identify with $[0, 2\pi]^d$. We conclude the chapter, and this dissertation, by adapting the results previously obtained, assuming that the diffusions are represented by a stochastic differential equation with reflecting boundaries, and by considering the Navier–Stokes equations with Neumann boundary conditions in the domain $[0, T] \times \mathbb{R}^d_+$.

In the appendix we present a bird's eye view of a few fundamental notions of measure theory, probability theory and the theory of stochastic processes, either mentioned throughout the dissertation, or essential to their comprehension. One of the sections of the appendix is a brief note on the problematic of the development of classical integration theory, more particularly, the progressive generalization of the definition of the integral. We succinctly present the origin of the Lebesgue–Stieltjes integral, which features, for instance, in the definition of the general version of the multidimensional Itô formula. This thesis assumes a certain familiarity with some foundational notions of other fields of mathematics, such as topology or functional analysis, and it would be impractical to be completely systematic and provide a definition of every concept encountered. We have refrained from defining concepts from those fields in the appendix, and instead refer the reader to the many textbooks which cover those areas. In reviewing the concepts and results mentioned throughout this dissertation, we used the following textbooks and monographs: [34] on measure theory and on general results on Lebesgue integration; [26], [13] and [38] on topology, on results related to metric spaces and on the basic notions of functional analysis; and

[20] and [17] on stochastic processes and on stochastic calculus. Other general, invaluable references extensively used were [16] and [19].

Chapter 1

Derivation of the Navier–Stokes Equations from Physical Principles

1.1 The Laws of Continuum Mechanics and Classical Hydrodynamics

The science of *mechanics*, the branch of physics that describes the behaviour of bodies when subject to energy and forces, can be divided into two main branches: *classical mechanics*, which describes the motion of bodies at the macroscopic scale (that is, at a scale humans can perceive without the aid of magnifying instruments), and *quantum mechanics*, which describes the behaviour of matter at the scale of atoms and subatomic particles (atoms have a radius of the order of one ångstöm, i.e. of 10^{-10} m). Any physical theory rests on a few essential, universal principles, namely the conservation of mass, the conservation of momentum and the conservation of energy (and related to this last principle, also the first and second laws of thermodynamics). The principle of conservation of mass states simply that mass can be neither created nor destroyed. The principle of conservation of momentum states that the momentum of a closed (or isolated) system does not change in the course of time, that is, it indicates that the interaction of bodies composing an isolated system leads only to an exchange in momentum between them, which does not affect the motion of the system as a whole. Finally, the principle of conservation of energy states that when one of the various forms of energy (kinetic, potential, internal, et cetera) decreases, it reappears as an increase in another form of energy. Also fundamental, in relation to this principle, is the first law of thermodynamics, which states that the heat added to a system and the work done on the system both contribute to an increase in the energy of the system, and the second law of thermodynamics, which states that no process is possible that has as its sole result the transfer of energy in the form of heat from a cold body to a hotter one.

All matter that follows the above conservation laws of physics, and there seem to exist phenomena that do not (see the tantalizing footnote on [14, p. 41]), can be classified either as solid, liquid, gas or plasma. In a very informal way, we can define a *fluid* as any material that is not in a solid state

(the fourth state, plasma, which is also a fluid, will not be described here). Since solids and fluids constitute two separate categories, it behoves us to briefly describe their fundamental differences. We do so by indicating how they react to a change in their volume and shape, giving afterwards a slightly more technical description. In the case of solids, we have to apply forces of great intensity to cause a variation of the volume or shape of a body, whereas in the case of gases the variation of the volume can be obtained easily and they offer no resistance to a variation of their shape. In the case of liquids, their volume is almost unchanged when the pressure is increased (in the case of gases, their volume depends largely on the pressure that is exerted on them) and their behaviour when their shape is altered is similar to the behaviour of gases. Since the behaviour of liquids and gases can be similarly described and explained, in opposition to solids, one puts liquids and gases in the same category. A slightly more technical description of the difference between solids and fluids goes as follows: while a solid undergoes a finite angular deformation proportional to the shear force, which is a force that is applied parallel or tangential to the face of the material, a fluid undergoes a continuous angular deformation as long as the shear force is applied, offering no resistance to that force (the source of this description is [14, p. 42]). We shall offer a more mathematical description of the shear force after we define the Cauchy stress tensor in the next section (see page 15 et seq.).

We will concern ourselves in this work with one of the branches of classical mechanics, namely *continuum mechanics*, the theory of dynamics of bodies with the assumption that their substance is infinitely subdividable (i.e. that a body completely fills the space it occupies, as opposed to the fact that it is actually composed of atoms, and therefore not continuous at all), and in this branch we will devote our attention solely to one of its subdivisions: *fluid mechanics*. Fluid mechanics can be further divided into *hydrostatics*, which studies fluids in equilibrium, and *hydrodynamics*, which is concerned with the motion of fluids, without considering their molecular structure. We will concern ourselves only with hydrodynamics in this work.

A fluid is regarded as a continuous medium, because the phenomena considered in hydrodynamics are macroscopic. At the microscopic level the flow of a fluid is governed by models in the field of statistical mechanics, more particularly by the Boltzmann equation. Since it is completely outside the scope of this work we will not elaborate further on the microscopic description of these physical phenomena.

The flow of fluids is only studied in two or three-dimensional cases, since it is nonsensical to study such phenomena in higher dimensions. In this chapter we will only deal with the case of deterministic motion, but in chapter 3 we will consider random motions, namely diffusions, and derive again the equations of classical hydrodynamics from a stochastic variational principle.

There are two classical representations (or methods) for describing the motion of a fluid. One is the *Lagrange method*, or the *Lagrange representation*, where the state of each particle of the fluid at any given time is described with reference to its initial position (that is, a history of the dynamics of a particle, its trajectory, is described). The other is the *Euler method*, or the *Euler representation*, which gives the values of attributes, i.e. the state, of the fluid, such as the velocity, the pressure, or the density, at arbitrary times and positions. If we are given five quantities, namely the components u_i of the velocity u, for $i = 1, 2, 3$, the pressure p and the density ρ, then the state of the fluid is completely determined.

Note that these five quantities are functions of the three spatial coordinates x_i, for $i = 1, 2, 3$, and of the temporal coordinate t.

It is important to realize that when we speak of a particle (or a point) of the fluid, that any such constituent is always supposed to be so large that it contains a very great number of molecules, i.e. it is very small compared with the volume of the body under consideration, but at the same time very large compared with the distances between the molecules. So when we speak of the motion of a particle of the fluid, we are not referring to the displacement of an individual molecule, but instead to the displacement of a small volume element containing many molecules. To make these observations more concrete, consider the fact that a fluid such as water contains around 3×10^{22} water molecules per cubic centimetre (the source of this value is [19, s.v. 'Statistical Mechanics', p. 1512]). Note also that the concept of particle is itself ambiguous and is used simply to ease the descriptions of phenomena in our natural language. In the 19th century particles were thought to be very small units of matter, which were assumed to behave according to the rules of classical mechanics. Since the 20th century we know that they behave in a quantized way and are actually not particles at all but waves (described approximately by the Schrödinger equation).

In the Lagrange representation a particle with coordinates $X = (X_1, X_2, X_3)$ at instant $t = 0$ has coordinates $x = x(t, X)$ at an arbitrary time $t \neq \infty$, i.e. the motion of the fluid is a function of time and is perfectly determined by the function x, whereas in the Euler representation each quantity of the fluid is regarded as a function of a space-time point (t, x). Notice that when X is fixed and t varies, this specifies the path of a particle, whereas for a fixed t and a varying X, it determines a transformation of a region initially occupied by the fluid into its position at time t. When deriving the equations of motion for fluids in the next section, we will be operating in the Euler representation, the representation used in all the monographs we consulted. Despite that fact, concepts from the Lagrange representation will also be used, such as the notion of material derivative (see definition 1.1).

Since fluids obey the general laws of continuum mechanics, namely the conservation of mass, the conservation of linear momentum and the conservation of energy, basing our arguments on these three laws successively, we will be able to derive the system of conservation laws of classical hydrodynamics (note that the second law of thermodynamics is only used in fluid mechanics in the case of a gas, to deduce equations relating unknowns such as pressure, temperature, et cetera). Since our goal is to derive the equations of motion for fluids, also known as the Navier–Stokes equations, it suffices to derive an incomplete system of conservation laws, a system in which the principle of conservation of energy is not accounted for. The reason for this is that after obtaining the full system we would assume that the body forces f do not depend on the temperature, rendering the equation of conservation of energy irrelevant. When the forces depend on the temperature the model based on the Navier–Stokes equations ceases to be applicable. We provide a brief description of the range of applicability of these equations at the beginning of section 1.3.

1.2 Derivation of the Equations

Assuming that the transformation $x = x(t, X)$ is continuous and invertible, and that x is smooth, we denote by $X = X(t, x)$ its inverse, where X is also assumed to be smooth. Naturally, these mathematical restrictions have a physical interpretation. Continuity corresponds to the fact that an element of the fluid in the neighbourhood of a given particle continues in its neighbourhood during motion. Invertibility means simply that a particle cannot occupy two places simultaneously nor can two different particles occupy the same space. From the assumption of invertibility we can conclude that the Jacobian determinant

$$J(t, X) := \det\left(\frac{\partial x_i}{\partial X_j}\right) \tag{1.1}$$

is finite and never zero, since this assumption is equivalent to requiring that J does not vanish. Therefore the element of volume $\mathrm{d}x$ changes according to the formula $\mathrm{d}x = J\,\mathrm{d}X$.

We believe it is important to always provide, wherever possible, the physical interpretation of mathematical concepts. We will do so often throughout this work. The gradient, which maps scalar fields to vector fields, measures the rate and direction of change in a scalar field. More specifically, the gradient of a function f at a point p, usually denoted by $\nabla f(p)$, has the direction of the greatest increase of the function at $f(p)$, and its magnitude is the rate of increase in that direction. For instance, if $f(p)$ is the temperature at point p, then the vector $\nabla f(p)$ is perpendicular to the surface of constant temperature and points in the direction of increasing temperature. The Jacobian is simply the generalization of the gradient for vector-valued functions of several variables.

The time rate at which any physical quantity f varies, such as the density $\rho = \rho(t, x)$, the velocity $u = u(t, x)$, or the temperature $\theta = \theta(t, x)$, while moving with the fluid particle, is given by the material derivative. This derivative will be used in the deduction of the conservation of mass and the conservation of momentum equations.

Definition 1.1. The *material*, or *Lagrangian*, *derivative* $\mathrm{D}/\mathrm{D}t$ of f is defined as follows:

$$\frac{\mathrm{D}f}{\mathrm{D}t} := \frac{\partial f}{\partial t} + u \cdot \nabla f, \tag{1.2}$$

which is written in terms of its components thus

$$\frac{\mathrm{D}f_i}{\mathrm{D}t} := \frac{\partial f_i}{\partial t} + \sum_{j=1}^{3} u_j \frac{\partial f_i}{\partial x_j}, \qquad i = 1, 2, 3. \tag{1.3}$$

The expression $u \cdot \nabla f$ is sometimes called the *convective rate of change*. The term $\partial f/\partial t$ corresponds to the local rate of change due to the temporal changes at position x, and the convective rate of change corresponds to the rate of change of f due to the transport of a material element to a different position. We now reveal the origin of the definition of the material derivative. Using the previous definitions we can write

$$f(t, x) = f(t, x(t, X)) := F(t, X). \tag{1.4}$$

It is also clear that the velocity u at time t of a particle initially located at X is

$$u(t,x) := \frac{\mathrm{d}}{\mathrm{d}t}x(t,X). \tag{1.5}$$

Thus

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}F(t,X) &= \frac{\mathrm{d}}{\mathrm{d}t}f(t,x(t,X)) \\ &= \frac{\partial}{\partial t}f(t,x(t,X)) + \frac{\mathrm{d}}{\mathrm{d}x_i}f(t,x(t,X))\frac{\mathrm{d}x_i}{\mathrm{d}t} \\ &= \frac{\partial}{\partial t}f(t,x) + u(t,x)\cdot\nabla f(t,x) \\ &= \frac{\mathrm{D}f}{\mathrm{D}t}(t,x).\end{aligned} \tag{1.6}$$

Later we will encounter $\mathrm{D}u/\mathrm{D}t$, which corresponds to the acceleration of an element of the fluid at (t,x), but only because we are operating in the Euler representation. If we were operating in the Lagrange representation, the acceleration of an element of the fluid would be given by the ordinary derivative $\partial u/\partial t$.

The following theorem concerns the rate of change not of a physical quantity, as in the case of the material derivative, but that of any volume integral. Let $f(t,x)$ be any function, and let $\Omega(t)$ be a closed volume moving with the fluid. The integral

$$\int\limits_{\Omega(t)} f(t,x)\,\mathrm{d}x \tag{1.7}$$

is over the varying volume $\Omega(t)$ so if we want to calculate the rate of change of this integral, that is,

$$\frac{\mathrm{d}}{\mathrm{d}t}\int\limits_{\Omega(t)} f(t,x)\,\mathrm{d}x, \tag{1.8}$$

we cannot simply commute the differentiation and integration operators. Using the previously mentioned transformation formula $\mathrm{d}x = J\,\mathrm{d}X$ we can transform this integral into one over a fixed volume $\Omega(0)$. In this case we can commute the operators. This is the essential idea used on the proof of the theorem.

Theorem 1.1 (Transport)**.** *Let $\Omega(t)$ denote an arbitrary volume that is moving with the fluid. Then*

$$\frac{\mathrm{d}}{\mathrm{d}t}\int\limits_{\Omega(t)} f(t,x)\,\mathrm{d}x = \int\limits_{\Omega(t)}\left(\frac{\mathrm{D}f}{\mathrm{D}t}(t,x) + f(t,x)\operatorname{div}u(t,x)\right)\mathrm{d}x, \tag{1.9}$$

where

$$\operatorname{div}u = \nabla\cdot u := \sum_{i=1}^{3}\frac{\partial u_i}{\partial x_i} \tag{1.10}$$

is the divergence of u.

Proof. Notice that $x = x(t,X)$ is a transformation $x\colon \Omega(0)\to\Omega(t)$. Thus (the source of this sequence of

equalities is [25, pp. 13–14])

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}\int\limits_{\Omega(t)} f(t,x)\,\mathrm{d}x &= \frac{\mathrm{d}}{\mathrm{d}t}\int\limits_{\Omega(0)} f(t,x(t,X))J(t,X)\,\mathrm{d}X\\ &= \int\limits_{\Omega(0)} \frac{\mathrm{d}}{\mathrm{d}t}(F(t,X)J(t,X))\,\mathrm{d}X\\ &= \int\limits_{\Omega(0)} \left(\left(\frac{\mathrm{d}}{\mathrm{d}t}F(t,X)\right)J(t,X) + F(t,X)\frac{\mathrm{d}}{\mathrm{d}t}J(t,X)\right)\mathrm{d}X\\ &= \int\limits_{\Omega(0)} \left(\frac{\mathrm{D}f}{\mathrm{D}t}(t,x)J(t,X) + F(t,X)\,\mathrm{div}\,u(t,x)J(t,X)\right)\mathrm{d}X\\ &= \int\limits_{\Omega(t)} \left(\frac{\mathrm{D}f}{\mathrm{D}t}(t,x) + f(t,x)\,\mathrm{div}\,u(t,x)\right)\mathrm{d}x, \end{aligned} \tag{1.11}$$

where we used successively the usual change of variables in an integral, the definition of material derivative and the fact that

$$\frac{\mathrm{d}}{\mathrm{d}t}J(t,X) = \mathrm{div}\,u(t,x)J(t,X), \tag{1.12}$$

which we prove now.

Since

$$\begin{aligned}J(t,X) &:= \det\left(\frac{\partial x_i}{\partial X_j}\right)\\ &= \frac{\partial x_1}{\partial X_1}\left(\frac{\partial x_2}{\partial X_2}\frac{\partial x_3}{\partial X_3} - \frac{\partial x_2}{\partial X_3}\frac{\partial x_3}{\partial X_2}\right) - \frac{\partial x_1}{\partial X_2}\left(\frac{\partial x_2}{\partial X_1}\frac{\partial x_3}{\partial X_3} - \frac{\partial x_2}{\partial X_3}\frac{\partial x_3}{\partial X_1}\right)\\ &\quad + \frac{\partial x_1}{\partial X_3}\left(\frac{\partial x_2}{\partial X_1}\frac{\partial x_3}{\partial X_2} - \frac{\partial x_2}{\partial X_2}\frac{\partial x_3}{\partial X_1}\right)\\ &= \sum_{i,j,k}\varepsilon_{ijk}\frac{\partial x_i}{\partial X_1}\frac{\partial x_j}{\partial X_2}\frac{\partial x_k}{\partial X_3}, \end{aligned} \tag{1.13}$$

where ε_{ijk} is the Levi-Civita symbol, which is defined by

$$\varepsilon_{ijk} := \begin{cases} +1 & \text{if (i,j,k) is (1,2,3), (2,3,1), or (3,1,2),}\\ -1 & \text{if (i,j,k) is (3,2,1), (1,3,2), or (2,1,3),}\\ 0 & \text{if i=j, or j=k, or k=i,}\end{cases} \tag{1.14}$$

and noting that

$$\frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial x_i}{\partial X_j} = \frac{\partial}{\partial X_j}\frac{\mathrm{d}x_i}{\mathrm{d}t} = \frac{\partial u_i}{\partial X_j} = \frac{\partial u_i}{\partial x_m}\frac{\partial x_m}{\partial X_j}, \tag{1.15}$$

we are finally able to compute

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}J(t,X) &= \sum_{i,j,k}\varepsilon_{ijk}\left(\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial x_i}{\partial X_1}\right)\frac{\partial x_j}{\partial X_2}\frac{\partial x_k}{\partial X_3} + \frac{\partial x_i}{\partial X_1}\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial x_j}{\partial X_2}\right)\frac{\partial x_k}{\partial X_3} + \frac{\partial x_i}{\partial X_1}\frac{\partial x_j}{\partial X_2}\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial x_k}{\partial X_3}\right)\right)\\ &= \sum_{i,j,k}\varepsilon_{ijk}\left(\frac{\partial u_i}{\partial x_m}\frac{\partial x_m}{\partial X_1}\frac{\partial x_j}{\partial X_2}\frac{\partial x_k}{\partial X_3} + \frac{\partial x_i}{\partial X_1}\frac{\partial u_j}{\partial x_m}\frac{\partial x_m}{\partial X_2}\frac{\partial x_k}{\partial X_3} + \frac{\partial x_i}{\partial X_1}\frac{\partial x_j}{\partial X_2}\frac{\partial u_k}{\partial x_m}\frac{\partial x_m}{\partial X_3}\right)\end{aligned}$$

$$= \sum_{i,j,k} \varepsilon_{ijk} \left(\frac{\partial u_1}{\partial x_1} \frac{\partial x_i}{\partial X_1} \frac{\partial x_j}{\partial X_2} \frac{\partial x_k}{\partial X_3} + \frac{\partial x_i}{\partial X_1} \frac{\partial u_2}{\partial x_2} \frac{\partial x_j}{\partial X_2} \frac{\partial x_k}{\partial X_3} + \frac{\partial x_i}{\partial X_1} \frac{\partial x_j}{\partial X_2} \frac{\partial u_3}{\partial x_3} \frac{\partial x_k}{\partial X_3} \right)$$
$$= \frac{\partial u_1}{\partial x_1} J(t, X) + \frac{\partial u_2}{\partial x_2} J(t, X) + \frac{\partial u_3}{\partial x_3} J(t, X)$$
$$= \operatorname{div} u(t, x) J(t, X). \qquad \text{QED}$$

The divergence, which maps vector fields to scalar fields, can be interpreted physically as giving the intensity of the outward flux of a vector field around a given point. As an example, consider the case of the velocity of air at each point, which defines a vector field, and the heating and cooling of air. When the air is heated it expands in all directions and when it is cooled it contracts. Therefore, the divergence will have a positive value in the former case and a negative value in the latter case.

Is is clear that the mass of any finite volume $\Omega(t)$ is

$$m := \int_{\Omega(t)} \rho \, \mathrm{d}x, \tag{1.16}$$

where $\rho = \rho(t, x)$ is the density, i.e. the mass per unit volume. By the *principle of conservation of mass*

$$\frac{\mathrm{d}}{\mathrm{d}t} \int_{\Omega(t)} \rho \, \mathrm{d}x = 0, \tag{1.17}$$

that is, the mass of a fluid in a volume $\Omega(t)$ does not change as $\Omega(t)$ moves with the fluid. Applying the transport theorem we obtain

$$\int_{\Omega(t)} \left(\frac{\mathrm{D}\rho}{\mathrm{D}t} + \rho \operatorname{div} u \right) \mathrm{d}x = 0, \tag{1.18}$$

and using the definition of material derivative this equation becomes

$$\int_{\Omega(t)} \left(\frac{\partial \rho}{\partial t} + u \cdot \nabla \rho + \rho \operatorname{div} u \right) \mathrm{d}x = 0, \tag{1.19}$$

if and only if

$$\int_{\Omega(t)} \left(\frac{\partial \rho}{\partial t} + \operatorname{div}(\rho u) \right) \mathrm{d}x = 0. \tag{1.20}$$

Since equations (1.18)–(1.20) are valid for an arbitrary material volume, it follows from equation (1.20) that

$$\frac{\partial \rho}{\partial t} + \operatorname{div}(\rho u) = 0. \tag{1.21}$$

This is the *equation of continuity*. See [3, pp. 135–136] for similar arguments for the deduction of this equation.

A different approach to deducing the equation of continuity, simpler and avoiding the transport theorem can be seen in [3, pp. 73–74] and in [24, pp. 1–2]. Since it is interesting and straightforward, we present it next briefly.

If S is a closed surface which encloses a fixed volume V entirely occupied by a fluid, then the mass

of fluid enclosed by the surface at any instant is

$$\int_V \rho \, \mathrm{d}V. \tag{1.22}$$

Since the mass of fluid flowing per unit time through an element $\mathrm{d}S$ of the surface is $\rho u \cdot \boldsymbol{n} \, \mathrm{d}S$ (this value is positive if the fluid is flowing out of the volume and negative otherwise), the total mass of fluid flowing out of the volume across the surface is

$$\int_S \rho u \cdot \boldsymbol{n} \, \mathrm{d}S, \tag{1.23}$$

where $\boldsymbol{n}$ is the unit outward normal of the area of the surrounding surface. We know also that the decrease per unit time in the mass of fluid in volume V is

$$-\frac{\mathrm{d}}{\mathrm{d}t} \int_V \rho \, \mathrm{d}V, \tag{1.24}$$

therefore the quantities (1.23) and (1.24) must be equal in the absence of other sources of fluid. Using this fact and the divergence theorem, which states that

$$\int_V \nabla \cdot F \, \mathrm{d}V = \int_S F \cdot \boldsymbol{n} \, \mathrm{d}S, \tag{1.25}$$

and which allows us to transform (1.23) into

$$\int_V \nabla \cdot (\rho u) \, \mathrm{d}V, \tag{1.26}$$

we obtain

$$\int_V \frac{\partial \rho}{\partial t} \, \mathrm{d}V = -\int_V \nabla \cdot (\rho u) \, \mathrm{d}V. \tag{1.27}$$

Note that we were able to put the differentiation operator inside the integral since the volume V is fixed in space. This last equation is obviously equivalent to

$$\int_V \left(\frac{\partial \rho}{\partial t} + \nabla \cdot (\rho u) \right) \mathrm{d}V = 0, \tag{1.28}$$

which is valid for all choices of the volume V lying in the fluid (possible only if the integrand is zero everywhere in the fluid). Therefore we again obtain the equation of continuity

$$\frac{\partial \rho}{\partial t} + \nabla \cdot (\rho u) = 0. \tag{1.29}$$

Definition 1.2. We call a fluid *incompressible* if for any domain $\Omega(0)$ and any t, the density is constant, that is, if

$$\mathrm{vol}(\Omega(0)) = \mathrm{vol}(\Omega(t)). \tag{1.30}$$

Taking $f(t,x)=1$ on the left-hand side of equation (1.9), using again the change of variables in the integral that we used in the sequence of equalities (1.11), and applying also equation (1.12), we obtain successively

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}\operatorname{vol}(\Omega(t)) &= \frac{\mathrm{d}}{\mathrm{d}t}\int\limits_{\Omega(t)}\mathrm{d}x\\ &= \frac{\mathrm{d}}{\mathrm{d}t}\int\limits_{\Omega(0)}J(t,X)\,\mathrm{d}X\\ &= \int\limits_{\Omega(0)}\frac{\mathrm{d}}{\mathrm{d}t}J(t,X)\,\mathrm{d}X\\ &= \int\limits_{\Omega(0)}\operatorname{div}u(t,x)J(t,X)\,\mathrm{d}X\\ &= \int\limits_{\Omega(t)}\operatorname{div}u(t,x)\,\mathrm{d}x=0. \end{aligned}\tag{1.31}$$

Since the choice of the domain $\Omega(t)$ is arbitrary, we deduce that a fluid is incompressible if and only if

$$\operatorname{div}u(t,x)=0.\tag{1.32}$$

To deduce this we could have also observed that when a fluid is incompressible the dependence of the density on pressure can be neglected, so equation (1.32) can simply be deduced from the equation of continuity (1.21).

We will now work towards obtaining the conservation of momentum equation by adapting and extending the exposition in [25, pp. 16–17] and in [3, pp. 137–138].

The forces acting on an element of a fluid can be *external*, such as gravitational or electromagnetic forces, or *internal* (also called *contact forces*), which act on an element of volume $\Omega(t)$ through its bounding surface.

Definition 1.3. The *normal stress* t_n is defined as follows:

$$t_n(t,x,n):=\boldsymbol{n}(t,x)\boldsymbol{T}(t,x),\tag{1.33}$$

where $\boldsymbol{n}$ is the unit outward normal at a point of the boundary $\partial\Omega(t)$, and $\boldsymbol{T}=\boldsymbol{T}_{ij}$ is the Cauchy stress tensor. The normal stress corresponds to the force per unit area exerted at a point of the surface $\partial\Omega(t)$.

We will define the Cauchy stress tensor below, after obtaining the conservation of momentum equation, so as not to break the flow of the argument.

If we denote by f an external force per unit mass, then f acts on $\Omega(t)$ as

$$\int\limits_{\Omega(t)}\rho f\,\mathrm{d}x,\tag{1.34}$$

and the surface force exerted on the volume Ω can be represented by

$$\int_{\partial\Omega(t)} t_n \,\mathrm{d}S, \tag{1.35}$$

where dS is a surface element.

By the *Cauchy principle* we know that, at any given time, t_n depends only on the position and orientation of dS, i.e. $t_n = t_n(t, x, n)$, and we know also by the *principle of conservation of linear momentum* that the rate of change of the linear momentum of a volume equals the resultant force on the volume, so

$$\frac{\mathrm{d}}{\mathrm{d}t}(mu) = \int_{\Omega(t)} \rho f \,\mathrm{d}x + \int_{\partial\Omega(t)} t_n \,\mathrm{d}S. \tag{1.36}$$

From equation (1.16) it is clear that

$$\frac{\mathrm{d}}{\mathrm{d}t}(mu) = \frac{\mathrm{d}}{\mathrm{d}t}\int_{\Omega(t)} \rho u \,\mathrm{d}x. \tag{1.37}$$

Before we can proceed, we need to prove a useful equality.

Lemma 1.1.

$$\frac{\mathrm{d}}{\mathrm{d}t}\int_{\Omega(t)} \rho f \,\mathrm{d}x = \int_{\Omega(t)} \rho \frac{\mathrm{D}f}{\mathrm{D}t} \,\mathrm{d}x. \tag{1.38}$$

Proof. Expanding the term $\nabla \cdot (\rho u)$ on the equation of continuity, that equation becomes

$$\frac{\partial \rho}{\partial t} + \rho \nabla \cdot u + u \cdot \nabla \rho = 0. \tag{1.39}$$

Noticing that the first and third terms on the left-hand side of this equation correspond to the material derivative of ρ, we obtain

$$\frac{\mathrm{D}\rho}{\mathrm{D}t} + \rho \nabla \cdot u = 0. \tag{1.40}$$

By the transport theorem, we obtain successively

$$\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}t}\int_{\Omega(t)} \rho f \,\mathrm{d}x &= \int_{\Omega(t)} \left(\frac{\mathrm{D}}{\mathrm{D}t}(\rho f) + \rho f \nabla \cdot u\right) \mathrm{d}x \\
&= \int_{\Omega(t)} \left(\rho \frac{\mathrm{D}f}{\mathrm{D}t} + f\frac{\mathrm{D}\rho}{\mathrm{D}t} + \rho f \nabla \cdot u\right) \mathrm{d}x \\
&= \int_{\Omega(t)} \rho \frac{\mathrm{D}f}{\mathrm{D}t} \,\mathrm{d}x,
\end{aligned} \tag{1.41}$$

where we used equation (1.40) to obtain the last equality. QED

By lemma 1.1 we conclude that

$$\frac{\mathrm{d}}{\mathrm{d}t}\int_{\Omega(t)} \rho u \,\mathrm{d}x = \int_{\Omega(t)} \rho \frac{\mathrm{D}u}{\mathrm{D}t} \,\mathrm{d}x, \tag{1.42}$$

where $\mathrm{D}u/\mathrm{D}t$ is the acceleration. Thus using equations (1.37) and (1.42), equation (1.36) becomes

$$\int\limits_{\Omega(t)} \rho \frac{\mathrm{D}u}{\mathrm{D}t}\,\mathrm{d}x = \int\limits_{\Omega(t)} \rho f\,\mathrm{d}x + \int\limits_{\partial\Omega(t)} t_n\,\mathrm{d}S. \tag{1.43}$$

By the divergence theorem we conclude that the second term on the right-hand side of this last equation is

$$\int\limits_{\partial\Omega(t)} t_n\,\mathrm{d}S = \int\limits_{\partial\Omega(t)} \boldsymbol{n}(t,x)\boldsymbol{T}(t,x)\,\mathrm{d}S = \int\limits_{\Omega(t)} \operatorname{div}\boldsymbol{T}\,\mathrm{d}x. \tag{1.44}$$

Therefore equation (1.43) becomes

$$\int\limits_{\Omega(t)} \rho \frac{\mathrm{D}u}{\mathrm{D}t}\,\mathrm{d}x = \int\limits_{\Omega(t)} \rho f\,\mathrm{d}x + \int\limits_{\Omega(t)} \operatorname{div}\boldsymbol{T}\,\mathrm{d}x. \tag{1.45}$$

Since the domain of integration is arbitrary we obtain finally the *Cauchy equation of motion*, or the *conservation of momentum equation*, in differential form:

$$\rho \frac{\mathrm{D}u}{\mathrm{D}t} = \rho f + \operatorname{div}\boldsymbol{T}, \tag{1.46}$$

which can be written in terms of its components thus

$$\rho\left(\frac{\partial u_i}{\partial t} + \sum_{j=1}^{3} u_j \frac{\partial u_i}{\partial x_j}\right) = \rho f_i + \sum_{j=1}^{3} \frac{\partial \boldsymbol{T}_{ij}}{\partial x_j}, \qquad i = 1,2,3. \tag{1.47}$$

Further transformations of the conservation of momentum equation require additional assumptions. Before we proceed with further modifications of this fundamental equation, it behoves us to define the Cauchy stress tensor (a more detailed exposition of the stress tensor can be seen in [3, pp. 141–147]).

Definition 1.4. The *Cauchy stress tensor* $\boldsymbol{T}_{ij}$, a tensor of rank-3, symmetric by assumption, is the sum of the viscous stress tensor $\boldsymbol{\varepsilon}_{ij}$, the elastic stress tensor $\boldsymbol{\tau}_{ij}$ and the hydrostatic pressure p, i.e.

$$\boldsymbol{T}_{ij} := \boldsymbol{\varepsilon}_{ij} + \boldsymbol{\tau}_{ij} - p\delta_{ij}. \tag{1.48}$$

The *viscous stress tensor* models the part of the stress related to the rate at which it is deforming around a point of the fluid, and the *elastic stress tensor* describes the internal forces in an elastic material due to its deformation. Recall that an *elastic material* is a material that when deformed due to an external force experiences internal resistance to the deformation and restores itself to its original state if the external force is no longer applied. The *hydrostatic pressure* is a pressure that acts perpendicularly to the local surface (independently of the orientation). The minus sign on the last expression on the right-hand side of (1.48) is chosen so that when $p > 0$, which represents the pressure exerted from the outside, the contact forces compress the fluid inside.

Definition 1.5. An *ideal*, or *perfect*, *fluid* is an incompressible fluid in which there is no internal friction.

Definition 1.6. A *viscous fluid* is a fluid in which the phenomena of internal friction and thermal conduction cannot be neglected, i.e. it has the property of resisting flow caused by intermolecular forces.

Perfect fluids, or non-viscous fluids, provide a close approximation of the behaviour of actual fluids. *Strain* is also a fundamental concept that can defined as the deformation of the continuum at a point. Strain can occur in two ways: via elongation, i.e. via a change in length, or via shear, which produces an angle change in the body.

In a perfect fluid, $\boldsymbol{\tau}_{ij}$ is zero so

$$\boldsymbol{T}_{ij} = \boldsymbol{\varepsilon}_{ij} - p\delta_{ij}, \tag{1.49}$$

and in the case of an hydrostatic fluid in equilibrium conditions, $\boldsymbol{\varepsilon}_{ij}$ is zero.

The components $\boldsymbol{T}_{ii}$ of the stress tensor are called *normal*, or *compressive*, *stresses* and the components $\boldsymbol{T}_{ij}$, with $i \neq j$, *shear stresses*. To simplify the physical description of the various components of the stress tensor, let us assume we are dealing with a small cube of fluid. The normal components either compress or dilate the cube, while the shear components produce deformation and vary from point to point over the surface of the cube, acting tangentially to the surface.

The index i indicates that the force lies in a plane perpendicular to the i axis and the index j denotes the direction in which the stress, i.e the force, acts. A positive value for one of the diagonal components of the tensor shows that there is a pulling motion by a force, whereas a negative value indicates compression. In general, a component is positive if it acts in the positive direction of the coordinate axes and if the plane where it acts has an outward normal vector pointing in the positive coordinate direction.

The components of the stress tensor can be obtained from those of the rate-of-strain tensor $\boldsymbol{e}_{ij}$, using the relation (for an incompressible fluid)

$$\boldsymbol{T}_{ij} = 2\mu\boldsymbol{e}_{ij} - p\delta_{ij}, \tag{1.50}$$

where μ is a constant called the *shear viscosity coefficient*.

Definition 1.7. The *rate-of-strain tensor*, or *rate of deformation tensor*, $\boldsymbol{e}_{ij}$ is defined as the derivative of the strain tensor $\boldsymbol{E}_{ij}$ with respect to time, that is,

$$\boldsymbol{e}_{ij} := \frac{\partial \boldsymbol{E}_{ij}}{\partial t} = \frac{1}{2}\left(\frac{\partial u_j}{\partial x_i} + \frac{\partial u_i}{\partial x_j}\right), \tag{1.51}$$

where we may also write

$$\frac{\partial u_j}{\partial x_i} = (\nabla u)_{ji}. \tag{1.52}$$

This tensor describes the rate of change of the deformation of a material in the neighbourhood of a certain point, that is, the rate of change of the relative distance between two particles in a fluid at a certain moment in time. This tensor is what distinguishes fluid motion from rigid body motion, i.e. from the motion of bodies in which the relative distance between two points in the body does not change.

We will not go into the details of the deduction of $\boldsymbol{e}_{ij}$, but will nevertheless provide a brief description of the physical notions that underpin it. The displacement vector u, composed of three components, has

contributions from four parts: pure translation (three directions), pure rotation (three planes), elongation (three axes or directions) and shear (three planes). To determine the two ways, mentioned above, in which strain can occur (via elongation or shear), we need to specify the changes in length, to describe the elongation, and the changes in angles, to describe shear, to three sides of the body. Since $\boldsymbol{E}_{ij}$, like the stress tensor, is a symmetric tensor, and representing by $\partial u_i/\partial x_j$ the way component i of the displacement vector changes in the x_j axis, $\boldsymbol{e}_{ij}$ can be deduced (for more details see [24, pp. 47–49, 79–81].

Substituting now (1.51) in (1.50) we obtain the expression for the stress tensor that we will use from now on:

$$\boldsymbol{T}_{ij} = \mu\left(\frac{\partial u_j}{\partial x_i} + \frac{\partial u_i}{\partial x_j}\right) - p\delta_{ij}. \tag{1.53}$$

Definition 1.8. A fluid is said to be *Newtonian* if the viscous stress is a linear function of the strain rate.

Since for most cases fluids can be regarded as Newtonian, and it is not the goal of this thesis to delve into these physical concepts, we will not go into further details. We mention only some examples of non-Newtonian fluids, which include blood, ink, milk and honey.

Equation (1.40), which is equivalent to the equation of continuity, as we have shown in the proof of lemma 1.1, and which expresses the conservation of mass, and the conservation of momentum equation (1.46), constitute the *(partial) system of conservation laws of classical hydrodynamics*

$$\begin{cases} \dfrac{\mathrm{D}\rho}{\mathrm{D}t} + \rho\nabla\cdot u = 0, & (1.54)\\ \rho\dfrac{\mathrm{D}u}{\mathrm{D}t} = \nabla\cdot\boldsymbol{T} + \rho f. & (1.55) \end{cases}$$

Substituting the expression for the stress tensor (1.53) in equation (1.55) of this system we obtain

$$\begin{aligned} \rho\frac{\mathrm{D}u}{\mathrm{D}t} &= \mu\nabla\cdot\nabla u + \mu\nabla\cdot\operatorname{div} u - \nabla p + \rho f\\ &= \mu\Delta u + \mu\nabla\cdot\operatorname{div} u - \nabla p + \rho f, \end{aligned} \tag{1.56}$$

where

$$\Delta u = \nabla\cdot\nabla u = \sum_{i=1}^{3}\frac{\partial^2 u}{\partial x_i^2} \tag{1.57}$$

is the Laplacian of u. If we now assume that the fluid is incompressible, then by equation (1.32), $\operatorname{div} u = 0$, so equation (1.54) becomes

$$\frac{\mathrm{D}\rho}{\mathrm{D}t} = 0, \tag{1.58}$$

and equation (1.56) becomes

$$\rho\frac{\mathrm{D}u}{\mathrm{D}t} = \mu\Delta u - \nabla p + \rho f. \tag{1.59}$$

Using the definition of material derivative on equations (1.58) and (1.59) we transform the system of

equations (1.54) and (1.55) into the new system

$$\begin{cases} \dfrac{\partial \rho}{\partial t} + u \cdot \nabla \rho = 0, & (1.60) \\ \rho \left(\dfrac{\partial u}{\partial t} + u \cdot \nabla u \right) = \mu \Delta u - \nabla p + \rho f, & (1.61) \\ \operatorname{div} u = 0. & (1.62) \end{cases}$$

Assuming that ρ is uniform and defining the *kinematic viscosity coefficient* as $\nu := \mu/\rho$ (μ is the shear viscosity coefficient, already mentioned above), equation (1.60) disappears and therefore we transform this system into the following one:

$$\begin{cases} \dfrac{\partial u}{\partial t} + u \cdot \nabla u = \nu \Delta u - \dfrac{1}{\rho} \nabla p + f, & (1.63) \\ \operatorname{div} u = 0, & (1.64) \end{cases}$$

where u and p, the velocity and the pressure, respectively, are the dependent variables and t and x_i, for $i = 1, 2,$ or 3, are the independent variables. These are the *Navier–Stokes equations* of motion of viscous, incompressible fluids with uniform density. Note that equation (1.63) only makes sense if u is differentiable in t and twice-differentiable in x. The equations are written in terms of their components thus

$$\begin{cases} \dfrac{\partial u_i}{\partial t} + \displaystyle\sum_{j=1}^{3} u_j \frac{\partial u_i}{\partial x_j} = \nu \sum_{j=1}^{3} \frac{\partial^2 u_i}{\partial x_j^2} - \frac{1}{\rho} \frac{\partial p}{\partial x_i} + f_i, \qquad i = 1, 2, 3, & (1.65) \\ \displaystyle\sum_{i=1}^{3} \frac{\partial u_i}{\partial x_i} = 0. & (1.66) \end{cases}$$

The first equation corresponds to the Newton law $f = ma$ for a fluid element subject to the external force f and to the forces arising from pressure and friction. Notice that these equations are second order, non-linear partial differential equations on account of the presence of the Laplace operator and of the product between u_j and the partial derivatives of u_i with respect to x_j, respectively. One very important point to keep in mind is that, unlike the case for linear equations, the non-linearity implies that small changes in a phenomenon can lead to radically different space-asymptotic behaviour. This fact contributes greatly to the difficulty of finding general solutions to these equations.

If the fluid is inviscid, that is, if $\nu = 0$, and incompressible, then equation (1.63) is reduced to

$$\frac{\partial u}{\partial t} + u \cdot \nabla u = -\frac{1}{\rho} \nabla p + f. \tag{1.67}$$

This is the *Euler equation*. Note that the equations of hydrodynamics, whether having the viscosity term or not, have as their domain $\mathbb{R}_+ \times \mathbb{R}^n$, for n equal to 2 or n equal to 3 only, since it does not make sense to study these equations in higher dimensions (as we mentioned in section 1.1, p. 6).

1.3 Range of Applicability, Initial and Boundary Conditions, and the Existence and Uniqueness of Solutions

It is of fundamental importance to quantify the range of applicability of the Navier–Stokes equations. To do so, the *Knudsen number* is frequently employed. This number helps in determining whether the formulation of fluid dynamics based on the principles of continuum mechanics, in which the Navier–Stokes equations are valid, or statistical mechanics should be used to model a phenomenon.

Definition 1.9. The *Knudsen number* Kn is a dimensionless quantity defined as $Kn := \lambda/L$, where λ is the mean free path of the molecules that make up the fluid and L is the representative physical length scale of the system, which may be, for instance, the radius of a body in a fluid. The molecular mean free path λ corresponds to the average distance travelled between intermolecular collisions.

The Navier–Stokes equations are only valid if L is much larger than λ, that is, when Kn is much smaller than one. For $Kn < 0.01$, the equations constitute an adequate model of the physical phenomenon being studied. When the Knudsen number exceeds 0.01, the description provided by the equations collapses and the methods of statistical mechanics must be used (incidentally, it is clear that statistical mechanics can be applied to all ranges of Knudsen numbers). Interestingly, according to [15, p. 2], the physical phenomena which verify $0.1 \lesssim Kn \lesssim 10$ are some of the most challenging to model. Explicit formulas for the estimation of the mean free path exist and an example can be seen in [15, p. 2].

All physical problems described by partial differential equations have a domain of application, which includes boundaries of some kind. To adequately describe a problem it is essential to define both initial conditions and boundary conditions. The *initial condition* for the equations of hydrodynamics describes the initial distribution of velocities and is of the form

$$u(0, x) = u_0(x), \qquad x \in \Omega, \tag{1.68}$$

where u_0 is a given function and Ω is the bounded domain. For the given u_0, the maximum time T allowed is called the *blow-up time*, which can be infinite. Note that the initial condition must also satisfy the boundary conditions defined below. We call a problem *well-posed* if it has a unique solution and if small changes in the initial conditions lead to small changes in the solutions. *Boundary conditions* can be of the first, second, or third kind, also called Dirichlet, Neumann, or Robin boundary conditions, respectively (other boundary conditions can be defined, but these three kinds are the most important ones). As far as Robin boundary conditions are concerned, they can be described simply as a mixture of the other two kinds of conditions. When dealing with the Navier–Stokes equations, it is common to call the *Dirichlet boundary condition* the *no-slip* condition, which corresponds to the case where the fluid fills a smooth, bounded domain Ω with a rigid boundary $\partial\Omega$. Since for this condition we specify a solution on the boundary of a region, we write it as

$$u = \varphi, \qquad u \in \partial\Omega, \tag{1.69}$$

where φ is the velocity at the boundary. If the boundary is at rest, this condition is written

$$u = 0, \qquad u \in \partial\Omega. \tag{1.70}$$

In the case of *Neumann boundary conditions* (the only boundary conditions we will consider in chapter 3 of this work), we have a partial differential equation that holds in some region of space and we specify on the boundary the directional derivative of u in the outward normal direction. This derivative is represented as $\partial u/\partial \boldsymbol{n} = \nabla u \cdot \boldsymbol{n}$, and is proportional to the inward flux. These types of conditions are common in problems where the flux, for example of heat, is given across the boundary (if the outward normal derivative is zero, then there is no flux across the boundary). Since we will deal with the diffusion, or heat, equation in section 2.1, it is instructive to present here the physical interpretation of the outward normal derivative, which is simply the change in temperature in the direction $\boldsymbol{n}$. Consequently, in a certain region of the boundary where $\partial u/\partial \boldsymbol{n} > 0$, we know that heat is flowing in, i.e. the temperature is higher outside than inside the particular region of the domain close to this area of the boundary (since heat flows from hot to cold regions). We have exactly the opposite situation when $\partial u/\partial \boldsymbol{n} < 0$. If $\partial u/\partial \boldsymbol{n} = 0$, then we know that no heat flows across the boundary. In the context of the Navier–Stokes equations, the Neumann boundary condition is usually written as

$$\nabla u \cdot \boldsymbol{n} = 0 \qquad \text{on } [0,T] \times \partial\Omega, \tag{1.71}$$

which means that the substance is enclosed in a domain so that none can enter or escape.

Finding general solutions of the Navier–Stokes equations, given initial and boundary conditions, would allow us to predict the behaviour of a fluid for a limited time interval, or ideally, for all times. To find those solutions we would need to solve the following system (assuming here for purposes of illustration that we have Neumann boundary conditions), composed of equations (1.63), (1.64), (1.68) and (1.71):

$$\begin{cases} \frac{\partial u}{\partial t} + u \cdot \nabla u = \nu \Delta u - \frac{1}{\rho}\nabla p + f & \text{in } [0,T] \times \Omega, \\ \operatorname{div} u = 0 & \text{in } [0,T] \times \Omega, \\ u(0,x) = u_0(x) & \text{for } x \in \Omega, \\ \nabla u \cdot \boldsymbol{n} = 0 & \text{on } [0,T] \times \partial\Omega, \end{cases} \tag{1.72}$$

for a certain domain Ω.

Several types of solutions for the Navier–Stokes equations have been defined, but only a few have been found and even then only for very specific and limited cases. The two main kinds of solutions are the so-called *weak solutions* and the *strong solutions*. We define the former concept now. If we multiply both sides of equation (1.63) by a smooth vector field $\theta(t,x)$ and integrate with respect to space and then to time, the equation becomes

$$\int_0^T \int \frac{\partial u}{\partial t}\theta \,\mathrm{d}x\mathrm{d}t + \int_0^T \int (u \cdot \nabla u)\theta \,\mathrm{d}x\mathrm{d}t = \nu \int_0^T \int \Delta u\theta \,\mathrm{d}x\mathrm{d}t - \frac{1}{\rho}\int_0^T \int \nabla p\theta \,\mathrm{d}x\mathrm{d}t + \int_0^T \int f\theta \,\mathrm{d}x\mathrm{d}t. \tag{1.73}$$

If we next formally integrate by parts all terms of equation (1.73) except the last term on the right-hand side, so as to make the derivatives fall on the vector field θ, we obtain

$$-\int_0^T \int u \frac{\partial \theta}{\partial t} \, \mathrm{d}x\mathrm{d}t - \int_0^T \int u(u \cdot \nabla \theta) \, \mathrm{d}x\mathrm{d}t = \nu \int_0^T \int u \Delta \theta \, \mathrm{d}x\mathrm{d}t + \frac{1}{\rho} \int_0^T \int p \cdot \operatorname{div} \theta \, \mathrm{d}x\mathrm{d}t + \int_0^T \int f\theta \, \mathrm{d}x\mathrm{d}t. \quad (1.74)$$

We omitted in the above equation the first term of each integration by parts, since it is always zero, assuming θ is zero on the upper and lower limits of integration. Note also that T can be infinite. Applying the same procedure to equation (1.64), but this time multiplying it by a smooth function $\varphi(t, x)$, we obtain

$$\int_0^T \int (\nabla \cdot u)\varphi \, \mathrm{d}x\mathrm{d}t = 0, \quad (1.75)$$

and after formally integrating by parts

$$\int_0^T \int u \nabla \varphi \, \mathrm{d}x\mathrm{d}t = 0. \quad (1.76)$$

We call a solution u of the Navier–Stokes equations *weak* if it verifies equations (1.74) and (1.76) for some test functions θ and φ.

Since no global existence results have been discovered for the strong solutions (their existence has only been proved for certain finite time intervals, not for all times), our knowledge of them is incipient at best. Therefore we will not define them here and instead refer the reader to the article [12, pp. 901–904], where the concept of strong solution is expounded.

Despite the difficulties in proving existence and uniqueness results and in finding solutions, some results have been obtained. In the two-dimensional case, it is known that existence and uniqueness are verified, which is also the case for the Euler equation. So in $\mathbb{R}^2$ the mathematical theory of the Navier–Stokes equations and of the Euler equation is fairly complete, unlike in $\mathbb{R}^3$. In the three-dimensional case, it is known that the Navier–Stokes equations always have a weak solution, but proving or disproving its uniqueness is an insurmountable problem in the present state of knowledge. In contrast to the Navier–Stokes equations, it has been proved that in $\mathbb{R}^3$ the Euler equation does not have unique weak solutions.

For a clear presentation of some of the extremely difficult open problems related to these equations see [10] and for more details and results [12].

1.4 Brief Note on the Importance and History of the Equations

Fluid mechanics models the motion of everything that is not in a solid state, and, in particular, the Navier–Stokes equations model the motions of viscous, incompressible fluids with uniform density past or through arbitrary objects. From this general description it is clear that the range of applications of the study of the Navier–Stokes equations is near limitless. We present here a few examples of where they are used: in modelling the fluid dynamics of everything from oceans and currents to the formation of

galaxies, nebulae and supernovae, and other astronomical phenomena; in haemodynamics, to understand the progression of cerebral aneurysms and the processes of thrombosis and arteriosclerosis; in meteorology, for predicting weather patterns and the behaviour of tornadoes; in aerodynamics, for the development of efficient wings for planes, propulsion systems for rockets, combustion engines, and for the design of a wide variety of aircraft and spacecraft; in particle physics and in acoustics.

We conclude this chapter with a brief historical sketch of the origins of the Navier–Stokes equations. Leonard Euler (1707–1783) published his equation of the motion of non-viscous fluids in 1757, equation that now bears his name, and in the early 1820s Claude-Louis Navier (1785–1836) and in 1831 Siméon Denis Poisson (1781–1840) derived for the first time the equations of fluid motion with friction, on the basis of considerations involving the action of intermolecular forces, generalizing the Euler equation. Adhémar Barré de Saint-Venant (1797–1886) derived the now called Navier–Stokes equations in 1843 under the sole assumption that the normal and tangential stresses are linear functions of the strain rates, that is, not following Navier's molecular approach, and identified the viscosity coefficient. George Gabriel Stokes (1819–1903) also derived the equations under the same assumption as Barré de Saint-Venant and published his result slightly later. Therefore the naming of the equations is yet another of the many cases of erroneous attribution of a result in the history of mathematics. For a detailed and very readable account of the history of these equations see [1, pp. 46–48, 88–92]. Another thorough historical survey of fluid mechanics is presented in [14, pp. 1–4] (despite the fact that this book is an authoritative and encyclopedic treatment of fluid mechanics, it should be noted that Barré de Saint-Venant's name is not mentioned once).

Chapter 2

The Wiener Process and Elements of Stochastic Analysis

2.1 From the Diffusion Equation to the Wiener Measure

The *diffusion equation* models the collective motion of particles in a material resulting from the random movement of each particle. Since one of the phenomena this equation models is the distribution of temperature in a bounded region, more specifically the diffusion of particles that are initially concentrated near the origin into the entire region, the equation is also known as the *heat equation*. It is important not to forget the distinction between heat and temperature: heat is the overall energy of the molecular motion, while temperature is the average energy of the molecular motion.

We start by presenting its derivation. Some concepts need to be defined for a more thorough presentation (we used [49, pp. 161, 213–215] as the source of these definitions).

Definition 2.1. The *heat capacity* of a body is the ratio of an infinitesimal amount of the heat transferred to the body to the corresponding change in the temperature of the body.

Definition 2.2. The *specific heat* c is the heat capacity of a unit mass of a homogeneous substance.

Definition 2.3. The *thermal conductivity* of the material, denoted by κ, corresponds to the amount of heat passing across unit area per unit time and per unit of the gradient of the temperature (see its formula in [49, pp. 213–214]).

We denote by D a region in $\mathbb{R}^n$, by $f(t,x)$ the temperature at the point x at time t (f can be interpreted more broadly as representing the density of some substance, such as a chemical concentration), and by $F(t,x)$ the flux of energy, or heat, through the boundary ∂D. The flux of energy $F(t,x)$ is defined as $F(t,x) := -\kappa\nabla f$, since we know by the Fourier law that heat flows from hot to cold regions at a rate proportional to the gradient of the temperature. By the conservation of energy we know that

$$\frac{\mathrm{d}}{\mathrm{d}t}\int_D c\rho f(t,x)\,\mathrm{d}x = -\int_{\partial D} F(t,x)\cdot n\,\mathrm{d}S, \tag{2.1}$$

i.e. the variation of the total amount of heat contained in D, which corresponds to the expression on the left-hand side of equation (2.1), is equal to the amount of heat leaving D through its boundary (in other words, the flow occurs from regions of higher to regions of lower concentration). Using successively the definition of $F(t,x)$ and the divergence theorem, the integral on the right-hand side of equation (2.1) becomes

$$\int\limits_{\partial D} \kappa \nabla f \cdot n \,\mathrm{d}S = \int\limits_{D} \kappa \nabla \cdot \nabla f \,\mathrm{d}x = \int\limits_{D} \kappa \Delta f \,\mathrm{d}x, \tag{2.2}$$

so we transform (2.1) into

$$\int\limits_{D} c\rho \partial_t f(t,x) \,\mathrm{d}x = \int\limits_{D} \kappa \Delta f \,\mathrm{d}x. \tag{2.3}$$

Note that in obtaining equation (2.3) from equation (2.1), we had to assume that $f(t,x)$ and $\partial_t f$ are continuous functions, and that

$$\int\limits_{-\infty}^{+\infty} |f(t,x)| \,\mathrm{d}x \qquad \text{and} \qquad \int\limits_{-\infty}^{+\infty} |\partial_t f| \,\mathrm{d}x \tag{2.4}$$

converge uniformly (the source of these provisos is [40, p. 420], and see also [19, s.v. 'Integral Calculus', p. 823]), otherwise the equality

$$\frac{\mathrm{d}}{\mathrm{d}t} \int\limits_{D} c\rho f(t,x) \,\mathrm{d}x = \int\limits_{D} c\rho \partial_t f(t,x) \,\mathrm{d}x \tag{2.5}$$

would be false.

Since the domain D if arbitrary, we obtain from equation (2.3) the *diffusion*, or *heat*, *equation*

$$\partial_t f(t,x) = \kappa \Delta f, \tag{2.6}$$

taking for simplicity $c = \rho = 1$. We could have also derived this equation by modelling the diffusion of a substance in a liquid, applying in the process the Fick law of diffusion. This derivation is very simple and can be seen in [40, pp. 14–15].

We present next an abbreviated derivation of the *fundamental solution* of the heat equation in one dimension, since that is the solution we need in what follows. Thus we will derive a solution of the second-order, linear, homogeneous partial differential equation in two variables

$$\frac{\partial f}{\partial t} = \kappa \frac{\partial^2 f}{\partial x^2}. \tag{2.7}$$

We assume that $|x| < \infty$, $t > 0$ and that $\partial_x f(t,x) \to 0$ as $x \to \pm\infty$, so integrating (2.7) we obtain

$$\frac{\mathrm{d}}{\mathrm{d}t} \int\limits_{-\infty}^{+\infty} f(t,x) \,\mathrm{d}x = 0, \tag{2.8}$$

that is,

$$\int_{-\infty}^{+\infty} f(t,x)\,\mathrm{d}x = \text{const}\,. \tag{2.9}$$

Since we want a normalized solution, we will consider the constant to be equal to one. Noticing that under the scaling $x \to ax$, $t \to a^2 t$, for a constant $a > 0$, the equation is unchanged, we look for a solution invariant under the dilation scaling (this nomenclature is from [8, p. 45]) $f(t,x) \to a^\alpha f(at, a^\beta x)$, i.e. we look for a solution of the form

$$f(t,x) = \frac{1}{t^\alpha} v\left(\frac{x}{t^\beta}\right), \tag{2.10}$$

for a smooth function v and constants α, β, setting $a = t^{-1}$ in the dilation scaling. Calculating the partial derivatives of (2.10), substituting them in (2.7) and setting $y = xt^{-\beta}$ we obtain the new equation

$$\alpha t^{-(1+\alpha)} v(y) + \beta t^{-(1+\alpha)} y \frac{\partial v}{\partial t} + \kappa t^{-(\alpha+2\beta)} \frac{\partial^2 v}{\partial x^2} = 0. \tag{2.11}$$

We want this equation to have only variable y, so we set $-(1+\alpha) = -(\alpha+2\beta)$, taking $\alpha = \beta = 1/2$, and the equation becomes

$$\frac{1}{2} v(y) + \frac{1}{2} y v'(y) + \kappa v''(y) = 0, \tag{2.12}$$

which is equivalent to

$$\frac{1}{2} \frac{\mathrm{d}}{\mathrm{d}y}(yv(y)) + \kappa v''(y) = 0. \tag{2.13}$$

Integrating this equation we obtain

$$\frac{1}{2} y v(y) + \kappa v'(y) = \text{const}\,. \tag{2.14}$$

If we set the constant to zero, then the solution of this equation is

$$v(y) = C \exp\left(-\frac{y^2}{4\kappa}\right) \tag{2.15}$$

for some constant C. Replacing y by its definition we obtain

$$v\left(\frac{x}{t^\beta}\right) = C \exp\left(-\frac{x^2}{4\kappa t}\right), \tag{2.16}$$

arriving thus at a solution of the form

$$f(t,x) = \frac{C}{\sqrt{t}} \exp\left(-\frac{x^2}{4\kappa t}\right). \tag{2.17}$$

To find the value of C we substitute this expression for $f(t,x)$ in (2.9) obtaining

$$\frac{C}{\sqrt{t}} \int_{-\infty}^{+\infty} \exp\left(-\frac{x^2}{4\kappa t}\right) \mathrm{d}x = 1, \tag{2.18}$$

and after changing the variable of integration we obtain

$$\frac{C}{\sqrt{t}}\sqrt{4\kappa t}\int_{-\infty}^{+\infty}\exp\left(-z^2\right)\mathrm{d}z = 1. \tag{2.19}$$

Using the following well-known calculation of the Gaussian integral

$$\begin{aligned}\int_{-\infty}^{+\infty}\exp\left(-x^2\right)\mathrm{d}x &= \left(\int_{-\infty}^{+\infty}\exp\left(-x^2\right)\mathrm{d}x\right)^{1/2}\left(\int_{-\infty}^{+\infty}\exp\left(-y^2\right)\mathrm{d}y\right)^{1/2}\\ &= \left(\int_0^{2\pi}\int_0^{+\infty} r\exp\left(-r^2\right)\mathrm{d}r\mathrm{d}\theta\right)^{1/2}\\ &= \sqrt{\pi},\end{aligned} \tag{2.20}$$

we obtain finally

$$C = \frac{1}{\sqrt{4\pi\kappa}}. \tag{2.21}$$

It is also clear from the calculation of the Gaussian integral that we have more generally

$$\int_{\mathbb{R}^n}\exp\left(-x^2\right)\mathrm{d}x = \prod_{i=1}^{n}\int_{-\infty}^{+\infty}\exp\left(-x_i^2\right)\mathrm{d}x_i = \pi^{n/2}. \tag{2.22}$$

Inserting the value of C we calculated in (2.21) into expression (2.17), we obtain the fundamental solution of the diffusion equation, which is

$$f(t,x) = \frac{1}{\sqrt{4\pi\kappa t}}\exp\left(-\frac{x^2}{4\kappa t}\right). \tag{2.23}$$

We conclude this section by presenting the Wiener measure, which is defined in terms of the fundamental solution of the diffusion equation, but before we do so it is important to mention that the Wiener measure was introduced with the very specific purpose of assigning a probability to the random path of a particle, with a view to giving a complete mathematical formulation of a phenomenon called *Brownian motion* (although later this phenomenon would prove to be an excellent model for many other physical phenomena). This phenomenon consists of the tiny irregular motions of small particles of pollen suspended in water and was extensively observed under the microscope and reported on by the botanist Robert Brown (1773–1858), hence its name. Brown misinterpreted the phenomenon by supposing that the small particles, which exhibited similar behaviour whether they came from organic or inorganic matter, were some kind of primitive molecule of living matter. He would have had to await the invention of the quantum theory of matter to understand that the haphazard movement of the particles was actually caused by the random impact of the water molecules. An extremely interesting and contemporary article describing Brown's original observations and others, even discussing which types of lenses and apparatuses can be used to perform the observations, is [30].

Let $C_0([0,1])$ be the vector space of continuous, real-valued functions $x(t)$ on $[0,1]$ that vanish at the

point zero. This space is in fact a metric space with the distance

$$\rho(x,y) = \max_{t\in[0,1]} |x(t) - y(t)|, \qquad x, y \in C_0([0,1]). \tag{2.24}$$

Definition 2.4. Let $A_1 =]a_1, b_1[, \ldots, A_n =]a_n, b_n[$ be Borel sets on the real line, i.e. open subsets of the real line, such that $x(t_i) \in A_i, i = 1, \ldots, n$. The set of functions

$$Q = \{x \in C_0([0,1])\colon x(t_i) \in A_i, 0 = t_0 < t_1 < \cdots \leq t_n = 1\} \tag{2.25}$$

is called a *quasi-interval* of $C_0([0,1])$, where a_i and b_i may be equal to $-\infty$ and $+\infty$, respectively, provided the $\leq$ symbol in (2.25) is substituted by $<$.

We understand an element $x(t)$ of the space $C_0([0,1])$ as denoting the path of a free particle, that is, of a particle not constrained by the influence of external forces, along the x-axis at time t. To give an idea of how erratic a path of such a particle is, consider the fact that a particle under normal conditions suffers 10^{21} collisions per second, each subtly altering its displacement (the source of this value is [47, p. 34]).

We define a 'measure' on $C_0([0,1])$ by the formula

$$\mu(Q) = \prod_{i=1}^{n} \int_{A_i} f(t_i - t_{i-1}, x_i - x_{i-1})\,\mathrm{d}x_i = \prod_{i=1}^{n} \int_{A_i} \frac{1}{\sqrt{2\pi(t_i - t_{i-1})}} \exp\left(-\frac{(x_i - x_{i-1})^2}{2(t_i - t_{i-1})}\right) \mathrm{d}x_i \tag{2.26}$$

where $x_i = x(t_i)$, each $f(t_i - t_{i-1}, x_i - x_{i-1})$, $i = 1, \ldots, n$, is a fundamental solution of the diffusion equation (2.7), taking $\kappa = 1/2$, and each integral

$$\int_{A_i} f(t_i - t_{i-1}, x_i - x_{i-1})\,\mathrm{d}x_i, \qquad i = 1, \ldots, n, \tag{2.27}$$

corresponds to the probability of finding a particle in the interval A_i after it has wandered from the origin. It is clear from these considerations that this 'measure' gives the probability of finding a particle in a certain path, since (2.26) represents the calculation of joint probabilities of a trajectory.

When this 'measure' is extended to a σ-additive probability measure on the Borel field of sets generated by the quasi-intervals, that is, on the Borel closure of the set of quasi-intervals, it is known as the *Wiener measure*, frequently denoted by μ_W (we used the definition given in [16, s.v. 'Wiener Measure']; the original definition can be seen in [43, p. 153]). In this case, the functions that belong to the space $C_0([0,1])$ we have been considering extend their domain to $\mathbb{R}_+$ (the extension to $\mathbb{R}$ is also possible) and we call then $C_0(\mathbb{R}_+)$ the *Wiener space*.

The rigorous construction of the Wiener measure and of the Wiener process was obtained by Norbert Wiener (1894–1964) in the article [43], and an abbreviated construction of the Wiener measure is also in [44]. It was one of the major mathematical results of the 20th century and laid the foundations for a mathematically rigorous theory of stochastic processes. We turn in the next section to the Wiener process, but before we do so, we would like to provide a short comment on the concept of mathematical

rigour.

The concept of mathematical rigour is one we mention throughout this work in quite a cavalier way (in truth, it seems to us that this is also the way the concept is used in general). Explaining it properly would take us far afield from the topics covered in this dissertation. Instead we refer the reader to [21], where the authors elucidate the evolution of this concept (and this is a key point: mathematical rigour depends on the philosophical positions of the community of practitioners of mathematics at a given period in history) and its ambiguity (on this latter point see especially [21, pp. 22–23]). Were we dealing in this work with problems closely related to the foundations of mathematics, such as set theory, we would include here a discussion on this topic. Since we are not, it suffices to say that when we mention that some proof is *rigorous*, we mean simply that it provides *correct and complete arguments*, to the extent that the ZFC system (i.e. the Zermelo–Fraenkel axioms of set theory plus the axiom of choice) and the rules of logic can be taken to be an adequate foundation for the edifice of mathematics, and therefore the basis for the notion of *correctness* in mathematics.

2.2 On the Construction of the Wiener Process

We will provide now a brief overview of the construction of the Wiener process, not only because it underpins the whole field of stochastic analysis, but also because of its intrinsic mathematical interest.

Definition 2.5. A continuous stochastic process W defined on a filtered probability space $(\Omega, \mathcal{F}, \mathcal{F}_t, P)$ is a *Wiener process* if it satisfies the following three properties:

1) $W(0) = 0$ a.s.

2) $W(t) - W(s) \sim N(0, t - s)$ for all $0 \leq s \leq t$.

3) The increments are independent of $\mathcal{F}_s$, that is, for all times $0 \leq s < t$, the random variables $W(t) - W(s)$ are independent of $\mathcal{F}_s$.

The Wiener process is the mathematical idealization of the Brownian motion phenomenon, that is, the continuous analogue of the random walk of a particle that is displaced, independently of the past, at discrete moments of time. It is also the most important stochastic process.

There are several methods for constructing the Wiener process, that is, of proving that this stochastic process exists mathematically. Two main approaches are usually considered and we shall refer to the first one as the *Kolmogorov construction* and to the second as the *Wiener–Lévy–Ciesielski construction*.

The Kolmogorov construction, named after Andrey Kolmogorov (1903–1987), involves setting a collection of random variables, attributing to each a finite-dimensional distribution and then using the Daniell–Kolmogorov theorem, also known as the Kolmogorov consistency theorem, to prove that the corresponding process exists, by constructing a probability measure (the Wiener measure) and a stochastic process on an appropriate measurable space (see the proof, for example, in [20, pp. 49–56], which includes the proof of Daniell–Kolmogorov theorem on pages 50–52).

The second construction, historically the first, involves considering a measurable space of functions, then defining a mapping from a certain probability space to this measurable space and finally deriving the corresponding finite-dimensional distributions. This construction employs random series, in the Wiener construction series with sine functions and in the Lévy construction series with Haar functions. Wiener's original construction was published in the aforementioned article [43], where all the steps of the construction described below are expounded, using results from three articles also published by the same author in 1920–1921 (all reprinted in [46, pp. 435–454]; we will not indicate their titles so as not to clutter the bibliography). A brief overview of the construction is also in [45, pp. 214–234], and in [47, pp. 15–24, 33–57]. This construction was simplified and generalized by Paul Lévy (1886–1971) in 1948 and by Zbigniew Ciesielski (1934–) in 1961 (for the references to their work see the bibliography of [20]).

The original construction proceeds as follows: (a) the primary step is to prove by construction that a collection of random variables with the three properties given in definition 2.5 exists; (b) after showing that almost all sample paths of the Wiener process are Hölder continuous (see definition A.17 on page 71) of order less than $1/2$, which implies that the paths are continuous a.s. and nowhere differentiable (see the proof of these statements in [9, section 3.4]), the Wiener measure is defined; we can further note that the method used for constructing the measure on the functional space $C_0([0,1])$ consists of establishing a correspondence between the sample path functions of the process (a subclass of the functions of the entire space) and points on the unit interval (for more details on how to establish this correspondence see, for instance, [47, pp. 17–20, 37–40]); (c) since every continuous function f of period 2π has a Fourier series representation and the sample paths, which are random functions, are continuous, Wiener obtained a Fourier expansion, whose coefficients Z_k, $k \in \mathbb{N}$, are independent Gaussian random variables with zero mean; the process is then written as

$$W(t) = \sum_{k=0}^{\infty} Z_k \phi_k(t), \tag{2.28}$$

where

$$\phi_k(t) = \sin\Big[\frac{\pi}{2}(2k+1)t\Big], \tag{2.29}$$

and this series is proved to be uniformly convergent for every fixed t (we are using here the definition of ϕ_k found in [16, s.v. 'Wiener Process'], but others are possible). Some of the proofs of fundamental results used in this construction, such as the non-differentiability of sample paths or the Hölder continuity for almost every ω, were reworked and improved in the article [28], written jointly by Paley, Wiener and Zygmund. A more contemporary reworking of those proofs and also the proof of the unbounded variation of each time interval for almost every ω can be found, for example, in [9, section 3.4].

Lévy modified Wiener's original construction and substituted $\phi_k(t)$ as defined above with certain polygonal functions called Schauder functions, denoted by $s_k(t)$. This new construction of the process consists of performing interpolation using Schauder functions, which are constructed by integrating the

Haar functions $H_k(s)$ on $[0,1]$, i.e

$$s_k(t) = \int_0^t H_k(s)\,\mathrm{d}s, \qquad t \in [0,1], k \in \mathbb{N}. \tag{2.30}$$

The Haar functions are defined as follows: $H_0(s) := 1$ for $s \in [0,1]$,

$$H_1(s) := \begin{cases} 1 & \text{if } 0 \le t \le 1/2, \\ -1 & \text{if } 1/2 \le t \le 1, \end{cases} \tag{2.31}$$

and if $2^n \le k < 2^{n+1}, n \in \mathbb{N}$,

$$H_k(s) := \begin{cases} 2^{n/2} & \text{if } (k-2^n)/2^n \le t \le (k-2^n+1/2)/2^n, \\ -2^{n/2} & \text{if } (k-2^n+1/2)/2^n \le t \le (k-2^n+1)/2^n, \\ 0 & \text{otherwise.} \end{cases} \tag{2.32}$$

In the Lévy construction the Wiener process has the series representation

$$W(t) = \sum_{k=0}^{\infty} Z_k s_k(t). \tag{2.33}$$

The proof that this series converges uniformly to a continuous function, establishing the mathematical existence of this stochastic process, can be found in [20, pp. 57–58] and in [9, section 3.3]. The proof of the existence of the Wiener process can be seen in [20, pp. 58–59] and in [9, section 3.3].

Ciesielski generalized the Lévy construction by showing that the Wiener process can be represented in the form

$$W(t) = \sum_{k=0}^{\infty} Z_k \int_0^t \phi_k(s)\,\mathrm{d}s, \tag{2.34}$$

where $\{\phi_k\}_{k=0}^{\infty}$ is an arbitrary complete, orthonormal basis of the Hilbert space $L^2([0,1])$. Recall that $L^2([0,1])$ is the class of all square-integrable functions f on $[0,1]$ with respect to the Lebesgue measure, with the norm

$$\|f\|_2 = \left(\int_0^1 |f|^2\,\mathrm{d}x < \infty \right)^{1/2}. \tag{2.35}$$

We will not elaborate on this, but further note that the Haar functions used in the Lévy construction form such a base (the proof can be found in [9, section 3.3]).

2.3 Overview of Functional Integration

Unlike Lebesgue integration, where the domain of the integral is a region of n-dimensional space, *functional integration* consists of defining an integral which has its domain in a space of functions, a significantly more complicated mathematical object. There are at least two major approaches to functional

integration, one that arises in the vast fields of probability theory and stochastic analysis, and the other in the path integral approach to quantum mechanics and in quantum field theory. The two approaches are connected in deep and subtle ways.

The first approach was started first, before the field of stochastic analysis had even arisen. Percy John Daniell (1889–1946) introduced and developed, in a series of articles published in the period 1918–1920, the concept of the *Daniell integral*, sometimes also referred to as the *Daniell–Stone integral*. In contrast to the formulation of the Lebesgue integral, where the concept of a measure is axiomatic, and therefore requires the development of a measure theory before the integral can be defined, in the case of the Daniell integral it is the concept of an elementary integral which is axiomatically defined. In the context of the latter approach, *elementary functions* are bounded, real functions defined in an arbitrary set, forming a set we denote by T_0, and an *elementary integral* is a functional defined on T_0 which verifies linearity, non-negativity and continuity with respect to monotone convergence. It is not our goal to present the definition or the elementary properties of the Daniell–Stone integral, but simply to clarify how it differs from the Lebesgue integral, since the latter can be seen as a particular case of the former (in fact, under certain conditions, both integrals can be shown to be identical). For a very concise presentation of the Daniell–Stone integral see [19, s.v. 'Ordered Linear Spaces', pp. 1166–1167]. Adapting and extending the concept of the Daniell integral, Norbert Wiener defined a functional integral with respect to the Wiener measure, more specifically to a class of Brownian motion paths. This type of integral, called the *Wiener integral* is represented as

$$\int\limits_{C_0(I)} F(x)\,\mathrm{d}\mu_W, \tag{2.36}$$

where I is $[0,1]$, $\mathbb{R}_+$ or $\mathbb{R}$, and F is a functional defined on $C_0(I)$ that is measurable with respect to the Wiener measure μ_W (an article presenting the definition of the Wiener integral and reviewing some of the results pertaining to it, obtained in the first forty years after its introduction, is [23]). This concept of integral was later extended by Paley, Wiener and Zygmund, to the integral with respect to the Wiener process (based on earlier work by Wiener) which is defined as

$$\int\limits_0^1 g\,\mathrm{d}W := -\int\limits_0^1 g' W\,\mathrm{d}t, \tag{2.37}$$

where $g\colon [0,1] \to \mathbb{R}$ is a continuous, differentiable function (necessarily deterministic) belonging to $L^2([0,1])$ and verifying $g(0) = g(1) = 0$ (as above, the domain of each function in this space can be $\mathbb{R}_+$ or $\mathbb{R}$). Observe that this integral requires great mathematical skill to construct since the function $t \mapsto W(t)$ is nowhere differentiable, precluding any of the traditional approaches to the definition of the integral. In fact, notice that this integral is actually a random variable. For a few more details and results concerning this type of integral see [47, pp. 62–77] and [28], the latter being the article where this integral was first introduced. Further extensions of this integral are possible and we will have a lot more to say about them in the next section.

The second approach to functional integration was developed by Richard Feynman (1918–1988) with the goal of computing the probability that a certain quantum process will occur. This integral is defined over the same space as the Wiener integral, but is completely different: whereas the Wiener integral is based on a measure (like most rigorously defined integrals), the Feynman integral is not. For this reason its mathematical foundations are shaky and still a topic of very active research. It would take us far afield from the topics of this work to define the Feynman integral here, but the article [22] presents a very clear exposition of it and of its relation to the Wiener integral (on this latter point see also [29, pp. 6060–6061] and [4, pp. 4155–4159]). An excellent article synthesizing the development of functional integration after Feynman's lead is [4].

2.4 On the Possible Constructions of Stochastic Integrals

Our goal in this section is to present an overview of the constructions of the stochastic integrals. We make extensive use of one of these integrals, the Itô stochastic integral, in the remaining sections of this dissertation.

Concepts mentioned throughout this and the following sections of this chapter that are not defined here, are defined in the appendix. We therefore assume the reader is familiar with the concepts and notation presented in the appendix before tackling what follows.

We assume throughout this section that we have a filtered probability space $(\Omega, \mathcal{F}, \mathcal{F}_t, P)$.

Definition 2.6. A *martingale* with respect to $\{\mathcal{F}_t\}$ is a real-valued stochastic process M that verifies the following two conditions:

1) $\mathrm{E}[|M_t|] < \infty$ for all $t \in \mathbb{R}_+$.

2) $\mathrm{E}[M_t|\mathcal{F}_s] = M_s$ a.s. for all $0 \leq s \leq t$.

Intuitively, we can see $\mathrm{E}[M|\mathcal{F}_s]$ as giving an approximation of M, or more specifically, as reshaping M using the information provided by $\mathcal{F}_s$, which we interpret as the information about M that we are using. Therefore we can interpret the second condition of definition 2.6 as saying that the best prediction of unobserved future values of M is its last observation.

It is noteworthy that a real-valued Wiener process W is a martingale with continuous sample paths, since

$$\begin{aligned}
\mathrm{E}[W_s|\mathcal{F}_t] &= \mathrm{E}[W_s - W_t + W_t|\mathcal{F}_t] \\
&= \mathrm{E}[W_s - W_t|\mathcal{F}_t] + \mathrm{E}[W_t|\mathcal{F}_t] \\
&= \mathrm{E}[W_s - W_t] + W_t \\
&= W_t \qquad \text{a.s.,}
\end{aligned} \tag{2.38}$$

where we used the linearity of the expected value, the independence of the increments of this process and the fact that its expectation is zero.

Since martingales, and in particular the Wiener process, are not differentiable a.s. in time (unless they are constant), we cannot integrate with respect to martingales as we can with respect to functions of bounded variation. To arrive at the Itô stochastic integral we need a further extension of the functional integral in (2.37), one that is valid for a more general class of integrands. We start with a few preliminary definitions and results, and then make a brief presentation of the construction of the Itô integral and of more general integrals, following mainly [17, pp. 45–59], [20, pp. 129–148] and [39, pp. 43–54].

Definition 2.7. A *square-integrable* martingale is a right-continuous martingale M that verifies $\mathrm{E}[M_t^2] < \infty$ for every $t \in \mathbb{R}_+$.

We denote by $\mathcal{M}_2$ the space of square-integrable martingales that verify $M_0 = 0$ a.s. Further, if such processes are a.s. continuous we denote that space by $\mathcal{M}_2^{\mathsf{c}}$.

For $X, Y \in \mathcal{M}_2$, we set

$$|X|_t := \mathrm{E}[X_t^2]^{1/2} \qquad \text{and} \qquad |X| := \sum_{n=1}^{\infty} 2^{-n} \min\left\{|X|_n, 1\right\}, \tag{2.39}$$

which defines the metric $|X - Y|$. Observe that the definition of this metric (and of the metric (2.41) and (2.53) below) is based on the idea of the definition of the standard bounded metric $\bar{d}$, which is given by $\bar{d}(x, y) := \min\{d(x, y), 1\}$, and which induces the same topology as d.

Definition 2.8. Two elements $X, Y \in \mathcal{M}_2$ are said to be *equivalent* if the maps $t \mapsto X_t$ and $t \mapsto Y_t$ coincide a.s. In this case, we may write $X_t = Y_t$ a.s.

This definition, and also definitions 2.10 and 2.17 below, introduce equivalence relations.

Proposition 2.1. The space $\mathcal{M}_2$ is a complete metric space under the metric (2.39), and $\mathcal{M}_2^{\mathsf{c}}$ is a closed subspace of $\mathcal{M}_2$.

Proof. See [17, pp. 47–48] or [20, pp. 37–38]). QED

Definition 2.9. We denote by $\mathcal{L}_2$ the space of all real-valued, $\mathcal{F}_t$-adapted, measurable stochastic processes X such that for every $t \in \mathbb{R}_+$,

$$\|X\|_{2,t}^2 := \mathrm{E}\left[\int_0^t X^2 \,\mathrm{d}t\right] < \infty. \tag{2.40}$$

Definition 2.10. We say that $X, Y \in \mathcal{L}_2$ are *equivalent* if $\|X - Y\|_{2,t} = 0$ for every $t \in \mathbb{R}_+$, and in this case we write $X = Y$.

We may assume that any element of $\mathcal{L}_2$ is predictable, without loss of generality (the justification for this can be seen in [17, pp. 45–46]). After the fashion of the definition of the metric $|X - Y|$, we take $X, Y \in \mathcal{L}_2$ and if we set

$$\|X\|_2 := \sum_{n=1}^{\infty} 2^{-n} \min\left\{\|X\|_{2,n}, 1\right\}, \tag{2.41}$$

then $\|X - Y\|_2$ defines a metric, and $\mathcal{L}_2$ is complete under this metric (according to [17, p. 45] and [39, p. 44]).

Definition 2.11. A process $F \in \mathcal{L}_2$ is *simple* (this nomenclature is from [20, p. 132]) if there exists a strictly increasing sequence $\{t_n\}_{n=0}^{\infty}$ of real numbers with $t_0 = 0$ and $\lim_{n\to\infty} t_n = \infty$, and a sequence of random variables $\{f_n\}_{n=0}^{\infty}$ with $\sup_n \|f_n(\omega)\|_2 < \infty$, for every $\omega \in \Omega$, such that f_n is $\mathcal{F}(t_n)$-measurable for every $n \geq 0$ and

$$F_t(\omega) = \begin{cases} f_0(\omega) & \text{if } t = 0, \\ f_i(\omega) & \text{if } t \in]t_i, t_{i+1}], \qquad i = 0, 1, \ldots \end{cases} \tag{2.42}$$

Expression (2.42) can also be written as

$$F_t(\omega) = f_0(\omega)\mathbb{1}_{\{0\}}(t) + \sum_{i=0}^{\infty} f_i(\omega)\mathbb{1}_{]t_i, t_{i+1}]}(t), \qquad i = 0, 1, \ldots \tag{2.43}$$

We denote by $\mathcal{L}_0$ the class of all simple processes.

Proposition 2.2. The space $\mathcal{L}_0$ is dense in $\mathcal{L}_2$ with respect to the metric (2.41) .

Proof. See [17, pp. 46–47] or [39, p. 44]. QED

Let $F \in \mathcal{L}_0$ be defined by (2.43), and set

$$\begin{aligned} I(F) &:= \sum_{i=0}^{n-1} f_i(\omega)(W_{t_{i+1}}(\omega) - W_{t_i}(\omega)) + f_n(\omega)(W_t(\omega) - W_{t_n}(\omega)) \\ &= \sum_{i=0}^{\infty} f_i(W_{\min\{t,t_{i+1}\}} - W_{\min\{t,t_i\}}), \qquad t_n \leq t \leq t_{n+1}, \quad n = 0, 1, 2, \ldots \end{aligned} \tag{2.44}$$

We know that $|I(F)| = \|F\|_2$ (see [17, pp. 48–49] for details), so we let now $F \in \mathcal{L}_2$, defined as above. By proposition 2.2, we can find $F_n \in \mathcal{L}_0$ such that

$$\lim_{n\to\infty} \|F - F_n\|_2 = 0, \tag{2.45}$$

and since $|I(F_n) - I(F_m)| = \|F_n - F_m\|_2$, $I(F_n)$ is a Cauchy sequence that converges to a unique element $I(F) \in \mathcal{M}_2^c$ (by proposition 2.1). Therefore we make the following definition.

Definition 2.12. The *Itô stochastic integral* of $F \in \mathcal{L}_2$ with respect to the Wiener process is defined thus

$$\int_0^t F \, \mathrm{d}W := I(F)(t). \tag{2.46}$$

A very brief summary of the above procedure for defining the Itô integral can be found in [39, p. 49]. Observe that for a fixed t the integral is a random variable (in this case we still call it a stochastic integral though). The Itô integral was first rigorously formulated by Kiyosi Itô (1915–2008) in 1942. Later, further extensions of this definition were obtained, not only for wider classes of integrands, but also by defining stochastic integrals with respect to general continuous, local martingales. The above definition of the stochastic integral with respect to the Wiener process with integrands belonging to $\mathcal{L}_2$ can be adapted to accommodate integrands belonging to $\mathcal{L}_2^{\text{loc}}$, a class we define next.

Definition 2.13. The class of real-valued, $\mathcal{F}_t$-adapted, measurable processes X such that for every $t \in \mathbb{R}_+$,

$$\int_0^t X^2 \, \mathrm{d}t < \infty \qquad \text{a.s.}, \tag{2.47}$$

is denoted by $\mathcal{L}_2^{\mathsf{loc}}$.

For $X, Y \in \mathcal{L}_2^{\mathsf{loc}}$, we say that $X = Y$ if for every $t \in \mathbb{R}_+$,

$$\int_0^t |X - Y|^2 \, \mathrm{d}t = 0 \qquad \text{a.s.} \tag{2.48}$$

Definition 2.14. Let M be a stochastic process verifying $M_0 = 0$ a.s. If there exists a non-decreasing sequence $\{\tau_n \colon n \geq 1\}$ of stopping times of $\{\mathcal{F}_t\}$ such that $M_{\min\{t,\tau_n\}}$ is a martingale for each $n \geq 1$ and

$$\Pr\Big(\lim_{n\to\infty} \tau_n = \infty\Big) = 1, \tag{2.49}$$

then M is a *local martingale*.

We denote the space of local martingales and the space of continuous, local martingales by $\mathcal{M}^{\mathsf{loc}}$ and $\mathcal{M}^{\mathsf{c,loc}}$, respectively. Every martingale is a local martingale, but the converse is not true.

The definition of the stochastic integral of $F \in \mathcal{L}_2^{\mathsf{loc}}$ with respect to the Wiener process can be seen in [17, pp. 52–53] or in [39, pp. 47–48]. This integral, which is naturally denoted by same expression as (2.46), belongs to $\mathcal{M}_2^{\mathsf{c,loc}}$, the space of continuous, local, square-integrable martingales.

The last definition of the stochastic integral we will mention is that of the stochastic integral with respect to a martingale, i.e. the integral of the form

$$\int_0^t F \, \mathrm{d}M, \tag{2.50}$$

where M can belong to one of several spaces, such as $\mathcal{M}_2$, $\mathcal{M}_2^{\mathsf{c}}$ or $\mathcal{M}^{\mathsf{c,loc}}$. Note that, for instance, when trying to construct the stochastic integral with respect to $M \in \mathcal{M}_2^{\mathsf{c}}$, the integral cannot be defined for every $\omega \in \Omega$ separately, as is the case with the Lebesgue–Stieltjes integral, since M is not of bounded variation on any finite interval $[0, t]$ (the source of this observation is [20, pp. 129–130]). What makes the construction possible is that the martingale M has finite quadratic variation $\langle M \rangle$ (a concept defined below). We proceed by mentioning a few auxiliary, but important, results we will need.

Proposition 2.3. Let $M, N \in \mathcal{M}_2$.

1) There exists a natural, integrable, increasing process A such that $M_t^2 - A_t$ is a martingale, where A is uniquely determined. We denote A_t by $\langle M \rangle_t$, and call this last expression the *quadratic variation* of M.

2) There exists a process A, which is expressible as the difference of two natural, integrable, increasing processes such that $M_t N_t - A_t$ is a martingale, where A is uniquely determined. We denote A

by $\langle M, N\rangle$, and call this expression the *quadratic variational process* corresponding to M and N, or simply the *cross-variation* of M and N (note that $\langle M\rangle = \langle M, M\rangle$).

Proof. See [17, p. 53]. QED

We can also define the cross-variation process in another way.

Definition 2.15. Let $M, N \in \mathcal{M}_2^c$ and $\mathcal{P} = \{t_0, t_1, \ldots, t_m\}$ be a partition of $[0, t]$. Then

$$\langle M, N\rangle_t := \lim_{\|\mathcal{P}\|\to 0} \sum_{k=1}^{m} (X_{t_k} - X_{t_{k-1}})(Y_{t_k} - Y_{t_{k-1}}), \tag{2.51}$$

where $\|\mathcal{P}\| := \max_{1\leq k\leq n}(t_k - t_{k-1})$.

If $M, N \in \mathcal{M}_2^c$, then $\langle M, N\rangle$ is continuous. Several other properties and results concerning the cross-variation can be seen in [20, p. 31–32]. Another essential result tells us that for a d-dimensional Wiener process, $\langle W^i, W^j\rangle(t) = \delta_{ij}t$, for $i, j = 1, 2, \ldots, d$. From this result we conclude that the one-dimensional Wiener process has quadratic variation t at time t, and we deduce also from proposition 2.3 that $W_t^2 - t$ is a martingale.

Definition 2.16. Let $F \in \mathcal{L}_2$ be a real-valued, predictable process that verifies for every $t \in \mathbb{R}_+$,

$$(\|F\|_{2,t}^M)^2 := \mathrm{E}\int_0^t F^2 \,\mathrm{d}\langle M\rangle(s) < \infty. \tag{2.52}$$

We denote by $\mathcal{L}_2(M)$ the set of such processes.

Definition 2.17. We say that two measurable, adapted processes $X, Y \in \mathcal{L}_2(M)$ are *equivalent* if $X = Y$ a.s.

Let us set

$$\|F\|_2^M := \sum_{n=1}^{\infty} 2^{-n} \min\left\{\|F\|_{2,n}^M, 1\right\}. \tag{2.53}$$

$\mathcal{L}_0$ is dense in $\mathcal{L}_2(M)$ with respect to the metric (2.53), according to [17, p. 54] and [39, p. 51].

The stochastic integral of a process $F \in \mathcal{L}_2(M)$ with respect to $M \in \mathcal{M}_2$ is itself a member of $\mathcal{M}_2$, and is constructed in a very similar way to the Itô stochastic integral. More specifically, we let $F \in \mathcal{L}_0$ be defined by (2.43), then we set $I^M(F)$ as in (2.44), substituting in that definition the Wiener process W by a martingale M, and finally we extend $F \in \mathcal{L}_0$ to $F \in \mathcal{L}_2(M)$ (see the complete details in [17, pp. 54–55] and in [39, pp. 51–52]).

If the integral is defined with respect to $M \in \mathcal{M}_2^c$, another possibility, then the integral also belongs to $M \in \mathcal{M}_2^c$. This integral, and by implication all the narrower definitions mentioned above, verifies some important properties.

Proposition 2.4. If $F, G \in \mathcal{L}_2(M)$, then the stochastic integral with respect to $M \in \mathcal{M}_2$ verifies the following properties, for $t > s \geq 0$:

1) $\int_0^t (\alpha F(s) + \beta G(s))\,\mathrm{d}M(s) = \alpha \int_0^t F(s)\,\mathrm{d}M(s) + \beta \int_0^t G(s)\,\mathrm{d}M(s)$, for $\alpha, \beta \in \mathbb{R}$, a.s.

2) $\int_0^0 F(s)\,\mathrm{d}M(s) = 0$ a.s.

3) $\mathrm{E}\left[\int_0^t F(s)\,\mathrm{d}M(u) - \int_0^s F(s)\,\mathrm{d}M(u) \,\middle|\, \mathcal{F}_s\right] = 0$ a.s.

4) $\mathrm{E}\left[\left(\int_0^t F(u)\,\mathrm{d}M(u) - \int_0^s F(s)\,\mathrm{d}M(u)\right)^2 \,\middle|\, \mathcal{F}_s\right] = \mathrm{E}\left[\int_s^t F^2(u)\,\mathrm{d}\langle M\rangle(u) \,\middle|\, \mathcal{F}_s\right]$ a.s.

Proof. See [20, pp. 137–139]. QED

Proposition 2.5. If $M, N \in \mathcal{L}_2$, $F \in \mathcal{L}_2(M)$ and $G \in \mathcal{L}_2(N)$, then

$$\left\langle \int_0^t F(u)\,\mathrm{d}M(u), \int_0^t G(u)\,\mathrm{d}N(u) \right\rangle = \int_0^t (FG)(u)\,\mathrm{d}\langle M, N\rangle(u) \quad \text{a.s.}, \tag{2.54}$$

which we call the *cross-variation formula* (according to the nomenclature of [20, p. 142]).

Proof. See [20, pp. 142–145]. QED

Propositions 2.4 and 2.5 are valid for a further extension of the stochastic integral, this time to a stochastic integral with respect to processes in the space of local, square-integrable martingales $\mathcal{M}_2^{\mathsf{loc}}$. For more details on this extension see [17, p. 57]. The integral can also be defined with respect to processes $M \in \mathcal{M}^{\mathsf{c,loc}}$, replacing the condition in definition 2.16 with the condition

$$\Pr\left(\int_0^t F^2\,\mathrm{d}\langle M\rangle(s) < \infty\right) = 1. \tag{2.55}$$

We will not elaborate further on this, and instead refer the reader to [20, pp. 130–131,145–147] for details. We mention only that with this definition of the integral, only properties 1) and 2) from proposition 2.4 are valid (proposition 2.5 is also valid).

2.5 Semimartingales, the Chain Rule and Stochastic Differential Equations

In this section we will motivate and define the concept of semimartingale, define the change-of-variable formula, also known as the Itô formula, and end by explaining what a stochastic differential equation is. The conditions for the existence and uniqueness of solutions of a SDE are also presented.

Like in the previous section, we assume that we have a filtered probability space $(\Omega, \mathcal{F}, \mathcal{F}_t, P)$.

Definition 2.18. A stochastic process is said to be of the *Markovian type* if an infinitesimal change on the process depends only on the present state of the process, that is, if the process is independent of its past history.

We deal exclusively with Markovian-type processes in this work, but bear in mind that the stochastic differential equations defined in this section can also be defined for processes of the non-Markovian type, although the latter are less common (according to [19, s.v. 'Stochastic Differential Equations', p. 1536]).

Definition 2.19. We call a Markovian-type process X a *diffusion process* if the sample function $X_t(\omega)$ is continuous in t with probability one until a certain random time, called the terminal time, after which $X_t(\omega)$ stays at the terminal point.

The Wiener process is an example of a diffusion process. Since the above informal, intuitive definitions of Markovian process and diffusion process suffice for our purposes, we refrain from giving their technical definitions. For those definitions and more details see [19, s.v. 'Markov Processes', pp. 971–973] and [35, pp. 110–111] on the Markovian processes and diffusion processes, respectively.

Definition 2.20. An *Itô process* is a real-valued stochastic process X such that

$$X(t) = X(0) + \int_0^t F\,\mathrm{d}t + \int_0^t G\,\mathrm{d}W \qquad \text{a.s.} \tag{2.56}$$

where F and G are some adapted processes belonging to one of the martingale spaces previously defined, provided that the integrals are well-defined.

The first integral on the right-hand side of (2.56) is frequently called the *mean motion*, and the second integral the *fluctuation*, or *noise*. We write, formally,

$$X(t) - X(0) = \int_0^t \mathrm{d}X(s) \qquad \text{a.s.} \tag{2.57}$$

and call $\mathrm{d}X$, $\mathrm{d}X_s$ or $\mathrm{d}X(s)$, the *stochastic differential* of the process X. Therefore the Itô process X has the stochastic differential

$$\mathrm{d}X = F\,\mathrm{d}t + G\,\mathrm{d}W. \tag{2.58}$$

Note that the differential symbols, such as $\mathrm{d}W$, have no intrinsic mathematical meaning, and are merely an abbreviation for the corresponding integral expressions.

Stochastic processes of the form (2.56) are but a particular case of a mathematical object we will later make extensive use of — the semimartingale.

Definition 2.21. A *continuous semimartingale* X is an adapted stochastic process which has the unique decomposition

$$X_t = X_0 + M_t + V_t \qquad \text{a.s.}, \tag{2.59}$$

where X_0 is a random variable, $M \in \mathcal{M}^{\mathsf{c,loc}}$ and V_t is a right-continuous, $\mathcal{F}_t$-adapted, integrable process

(verifying $V_0 = 0$) such that $t \mapsto V_t$ is of bounded variation a.s. We call M the *martingale part* of the decomposition and V the *drift part*.

Proposition 2.6. A continuous semimartingale X with the decomposition (2.59) has a finite quadratic variation and verifies $\langle X, X\rangle = \langle M, M\rangle$.

Proof. See [33, p. 128]. QED

Of fundamental importance are the *change-of-variable*, or *Itô*, *formula* and the *Itô product rule*, which we define next.

Theorem 2.1 (Multidimensional Itô Formula)**.** *Let $f\colon \mathbb{R}_+ \times \mathbb{R}^d \to \mathbb{R}$ be a function of class $C^{1,2}$, and let X be a multidimensional continuous semimartingale, where in that case we write the martingale part of its decomposition as the vector $(M_t^1, \ldots, M_t^d)$, where $M_t^i \in \mathcal{M}^{c,loc}$, and the drift part as the vector $(V_t^1, \ldots, V_t^d)$, V_t^i being an adapted process of bounded variation verifying $V_0^i = 0$, $i = 1, \ldots, d$. We call*

$$\begin{aligned} \mathrm{d}f(t, X_t) &= \frac{\partial}{\partial t} f(t, X_t)\,\mathrm{d}t + \sum_{i=1}^{d} \frac{\partial}{\partial x_i} f(t, X_t)\,\mathrm{d}X_t^i + \frac{1}{2}\sum_{i=1}^{d}\sum_{j=1}^{d} \frac{\partial^2}{\partial x_i \partial x_j} f(t, X_t)\,\mathrm{d}\langle M^i, M^j\rangle_t \qquad (2.60) \\ &= \frac{\partial}{\partial t} f(t, X_t)\,\mathrm{d}t + \sum_{i=1}^{d} \frac{\partial}{\partial x_i} f(t, X_t)\,\mathrm{d}V_t^i + \sum_{i=1}^{d} \frac{\partial}{\partial x_i} f(t, X_t)\,\mathrm{d}M_t^i \\ &\quad + \frac{1}{2}\sum_{i=1}^{d}\sum_{j=1}^{d} \frac{\partial^2}{\partial x_i \partial x_j} f(t, X_t)\,\mathrm{d}\langle M^i, M^j\rangle_t, \qquad (2.61) \end{aligned}$$

the multidimensional Itô formula (in differential form). This formula is also sometimes called the chain rule for stochastic calculus.

Proof. For the one-dimensional case see [20, p. 150–153]. QED

This formula is a generalization of the original result proved by Itô in 1944 for the one-dimensional case and where X is the Wiener process.

Since the stochastic differentials have no mathematical meaning, as we mentioned before, expression (2.61) is actually an abbreviation for the equation

$$\begin{aligned} f(t, X_t) &= \int_0^t \frac{\partial}{\partial t} f(s, X_s)\,\mathrm{d}s + \sum_{i=1}^{d} \int_0^t \frac{\partial}{\partial x_i} f(s, X_s)\,\mathrm{d}V_s^i + \sum_{i=1}^{d} \int_0^t \frac{\partial}{\partial x_i} f(s, X_s)\,\mathrm{d}M_s^i \\ &\quad + \frac{1}{2}\sum_{i=1}^{d}\sum_{j=1}^{d} \int_0^t \frac{\partial^2}{\partial x_i \partial x_j} f(s, X_s)\,\mathrm{d}\langle M^i, M^j\rangle_s \qquad \text{a.s.} \qquad (2.62) \end{aligned}$$

Observe that only the third integral on the right-hand side of (2.62) is a stochastic integral (belonging to $\mathcal{M}^{c,loc}$). All the other integrals are Lebesgue–Stieltjes integrals. Note further that $f(t, X_t)$ is a continuous semimartingale belonging to $\mathcal{M}^{c,loc}$.

By proposition 2.6 we know that equation (2.60) can also be written as

$$\mathrm{d}f(t, X_t) = \frac{\partial}{\partial t} f(t, X_t)\,\mathrm{d}t + \sum_{i=1}^{d} \frac{\partial}{\partial x_i} f(t, X_t)\,\mathrm{d}X_t^i + \frac{1}{2}\sum_{i=1}^{d}\sum_{j=1}^{d} \frac{\partial^2}{\partial x_i \partial x_j} f(t, X_t)\,\mathrm{d}\langle X^i, X^j\rangle_t \qquad (2.63)$$

We will deal frequently with the particular case where $\mathrm{d}M_t = \mathrm{d}W_t$ and $\mathrm{d}V_t = \mathrm{d}t$, that is, we will consider the continuous semimartingale $\mathrm{d}X = \mathrm{d}W + \mathrm{d}t$, which corresponds to an Itô process. In this case, we simplify the last term on the right-hand side of equation (2.63) using the formal multiplication rules

$$(\mathrm{d}t)^2 = 0, \tag{2.64}$$

$$\mathrm{d}t\,\mathrm{d}W^i = 0, \tag{2.65}$$

$$\mathrm{d}W^i\,\mathrm{d}W^j = \delta_{ij}\,\mathrm{d}t, \tag{2.66}$$

for $i, j = 1, \ldots, d$, and obtain the multidimensional Itô formula in the (differential) form

$$\begin{aligned}\mathrm{d}f &= \frac{\partial f}{\partial t}\,\mathrm{d}t + \sum_{i=1}^{n}\frac{\partial f}{\partial x_i}\,\mathrm{d}X^i + \frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{n}\frac{\partial^2 f}{\partial x_i \partial x_j}\,\mathrm{d}t \\ &= \left(\frac{\partial f}{\partial t} + \sum_{i=1}^{n}\frac{\partial f}{\partial x_i} + \frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{n}\frac{\partial^2 f}{\partial x_i \partial x_j}\right)\mathrm{d}t + \sum_{i=1}^{n}\frac{\partial f}{\partial x_i}\,\mathrm{d}W.\end{aligned} \tag{2.67}$$

Since the process X has continuous sample paths a.s., the functions $t \mapsto \partial f/\partial t$, $t \mapsto \partial f/\partial x_i$ and $t \mapsto \partial^2 f/\partial x_i \partial x_j$ are continuous for almost every ω, which ensures that when we write (2.67) in integral form, all integrals are well-defined.

Proposition 2.7. Consider the semimartingales X_t and Y_t, defined as

$$X_t = X_0 + M_t + V_t \tag{2.68}$$

and

$$Y_t = Y_0 + N_t + U_t, \tag{2.69}$$

where $M, N \in \mathcal{M}^{\mathsf{c,loc}}$, and V and U are adapted, continuous processes of bounded variation verifying $V_0 = U_0 = 0$ a.s. Then

$$\int_0^t X_s\,\mathrm{d}Y_s = X_tY_t - X_0Y_0 - \int_0^t Y_s\,\mathrm{d}X_s - \langle M, N\rangle_t, \qquad \text{a.s.}, \tag{2.70}$$

and we call this equation the *integration-by-parts formula*.

Proof. See [35, pp. 59–60]. QED

Observe that this formula would be the integration-by-parts formula of ordinary calculus if it were not for the presence of the last term on the right-hand side.

If we write the integration-by-parts formula in differential form and rearrange the terms we obtain

$$\mathrm{d}(X_tY_t) = X_t\,\mathrm{d}Y_t + Y_t\,\mathrm{d}X_t + \mathrm{d}\langle M, N\rangle_t, \tag{2.71}$$

and we call this equation the *Itô product formula*.

We proceed now to the definition of stochastic differential equation with respect to the Wiener process, mentioning also the essential theorem which states the conditions for the existence and uniqueness of solutions of such an equation.

Definition 2.22. Let $b_i(t,x), B_{ij}(t,x)\colon \mathbb{R}_+ \times \mathbb{R}^d \to \mathbb{R}$, for $1 \leq i \leq d$ and $1 \leq j \leq r$, be Borel-measurable functions. We call $b(t,x) = (b_1(t,x), \ldots, b_d(t,x))$ the drift vector and $B(t,x) = [B_{ij}]$ the dispersion matrix. Let also X be a progressively measurable stochastic process. With these assumptions we call

$$\mathrm{d}X_t = b(t, X_t)\,\mathrm{d}t + B(t, X_t)\,\mathrm{d}W_t \tag{2.72}$$

a *stochastic differential equation* (SDE) and X_0 its *initial value*.

This equation is written in terms of its components as

$$\mathrm{d}X_t^i = b_i(t, X_t)\,\mathrm{d}t + \sum_{j=1}^{r} B_{ij}(t, X_t)\,\mathrm{d}W_t^j, \qquad 1 \leq i \leq d, \tag{2.73}$$

where W is an r-dimensional Wiener process.

We can interpret a SDE as being an ordinary differential equation with an extra term, which accounts for the random perturbations of the trajectory of its solution. These random perturbations are accounted for in the term with the differential $\mathrm{d}W$, which is therefore called the *random term*, or *diffusion term*.

Definition 2.23. We say that X is a *solution* of the SDE (2.72) if the following integral equation

$$X_t = X_0 + \int_0^t b(s, X_s)\,\mathrm{d}s + \int_0^t B(s, X_s)\,\mathrm{d}W_s \tag{2.74}$$

holds with probability one, that is, if it holds a.s. for all times $t \in \mathbb{R}_+$.

Theorem 2.2 (Existence and uniqueness of solution of a SDE)**.** *Let b and B be the measurable functions of definition 2.22 satisfying, for all $t \in \mathbb{R}_+$ and $x, y \in \mathbb{R}^d$, the following Lipschitz conditions, for some constant K:*

1) $\sup_t \sup_x |b(t,x)| \leq K$.

2) $\sup_t \sup_x \|B(t,x)\| \leq K$.

3) $|b(t,x) - b(t,y)| + \|B(t,x) - B(t,y)\| \leq K|x-y|$.

Let also X_0 be a random variable verifying $\mathrm{E}[|X_0|^2] < \infty$. Then there exists a unique solution $X \in \mathcal{L}_2$ of the stochastic differential equation (2.72).

Proof. See [41, pp. 124–126]. QED

The three conditions in theorem 2.2 ensure that the approach of finding a numerical solution using an iteration scheme is solid, since they imply that the iteration scheme converges to a limiting process X, which we call the solution of the SDE (this is the method used in the proof of the theorem). In particular,

conditions 1) and 2) tell us that the functions b and B are Lipschitz continuous. Condition 3), which ensures that there is linear growth in x, is actually a consequence of the previous two conditions, and can be written in the following alternative way:

$$|b(t,x)| + \|B(t,x)\| \leq K(1+|x|). \tag{2.75}$$

Recall that in the case of ordinary differential equations, given sufficient boundary and initial conditions, the solution of the equation is unique, whereas when we say that a certain stochastic process X solves a SDE, this solution is itself nondeterministic. We can attempt to grasp the complexity of finding a solution to a stochastic differential equation by noting that a SDE induces a transformation of the Wiener space $C_0(\mathbb{R}_+)$ into itself, that is, a transformation of each sample path of the Wiener process into the corresponding sample path of the solution. Despite this difficulty, explicit formulas for solutions do exist, namely solutions of linear stochastic differential equations. For the definition of this type of SDE and a brief presentation of some formulas for its solution see [9, section 5.4].

Definition 2.24. A solution X of the SDE (2.72) is called *weak* if it verifies

$$\Pr\left(\int_0^t \left(|b_i(s,X_s)| + B_{ij}^2(s,X_s)\right)\,\mathrm{d}s < \infty\right) = 1, \tag{2.76}$$

for every $1 \leq i \leq d$, $1 \leq j \leq r$ and $t \in \mathbb{R}_+$.

The concept of strong solution also exists, but since we will have no use for it in this dissertation, we refrain from presenting its definition (it can be seen, for example, in [20, p. 285]).

2.6 The Skorokhod Problem and Stochastic Differential Equations with Reflecting Boundaries

We move now to two intimately connected topics — the Skorokhod problem and stochastic differential equations with reflection.

Let us assume that there are particles located inside an open set $D \subset \mathbb{R}^d$, the boundary of which is impermeable. The Skorokhod problem is a mathematical formalization of the problem of finding the position of the particles after they have hit the boundary and been reflected. This reflection can be normal or oblique, but we will deal only with the former case in this work (the latter case being significantly more complicated). We start by giving the most general definition of this problem, named after Anatoliy Skorokhod (1930–2011), who defined it for the first time in the early 1960s, mentioning next its physical interpretation.

Definition 2.25. Let us denote by K_x be the set of reflecting directions for $x \in \partial D$. We assume that this set is non-empty and that $|v| = 1$, for any $v \in K_x$ (which implies that v is a reflecting vector field). Let $f\colon C(\mathbb{R}_+) \to \mathbb{R}^d$ be a known function that verifies $f(0) \in \overline{D}$. We always assume in this

definition that $t \in \mathbb{R}_+$. Following [32, pp. 24–25], we call a pair of continuous functions (g, φ), where $g, \varphi \colon \mathbb{R}_+ \to \mathbb{R}_+ \times \mathbb{R}^d$, a solution of the *Skorokhod problem* for the triple (D, K, f) if the following three conditions are met, assuming throughout that $g(t) \in \overline{D}$, that l is a non-decreasing function verifying $l(0) = 0$, and that $v(g(t)) \in K_{g(t)}$ if $g(t) \in \partial D$:

1) $g(t) = f(t) + \varphi(t)$, where $\varphi(t) := \int_0^t v(g(s))\,\mathrm{d}l(s)$.

2) $\int_0^t |v(g(s))|\,\mathrm{d}l(s) < \infty$.

3) $l(t) = \int_0^t \mathbb{1}_{\partial D}(g(s))\,\mathrm{d}l(s)$.

We interpret $g(t)$ as indicating the position of a particle at the instant t, and $f(t)$ as being a trajectory of a Wiener process. The moment t when a particle hits the boundary corresponds to the position $g(t) = 0$. Condition 3) signifies that l does not increase when $g(s) \notin \partial D$, that is, that l may only increase at those instants s when the particle is hitting the boundary. Observe that if the set K_x contains only one element $v(x)$, then we can interpret condition 1) as indicating reflection along a vector field v.

The existence and uniqueness proof of the solution to the Skorokhod problem can be seen in [35, p. 118], and an alternative uniqueness proof is in [32, pp. 2–3].

We make the identification

$$g(t) = \Gamma f(t) = f(t) + l(t), \tag{2.77}$$

where

$$l(t) = -\min\left\{\min_{s\in[0,t]} f(s), 0\right\}, \tag{2.78}$$

calling the continuous map $\Gamma \colon C(\mathbb{R}_+, \mathbb{R}^d) \to C(\mathbb{R}_+, \mathbb{R}^d)$ the *Skorokhod map*. For the specific domains we will mention in this work (essentially $D = \mathbb{R}_+$ and $D = \mathbb{R}^d_+$, as we will see), and always considering normal reflections, the proposition that follows is valid.

Proposition 2.8. Let $f_1,\ f_2 \in C([0, T], \mathbb{R}^d)$. Then the following two inequalities hold, where $g_i = \Gamma f_i$, $l_i = f_i - g_i$, $i = 1, 2$, and $\|f\| := \max_{t\in[0,T]} |f(t)|$ (the source of these inequalities is [32, p. 4]):

1) $\|g_1 - g_2\| \leq 2\|f_1 - f_2\|$.

2) $|l_1 - l_2\| \leq \|f_1 - f_2\|$.

Since Γf is continuous, writing inequality 1) of proposition 2.8 as

$$\|\Gamma f\| \leq 2\|f\|, \tag{2.79}$$

it is clear that Γ is a Lipschitz-continuous map. It is important to bear in mind that property (2.79) of the Skorokhod map may fail for some specific problems. An example of the failed attempt to prove the Lipschitz continuity of Γ, assuming oblique reflections in the domain $D = \mathbb{R}^d_+$, can be seen in [32, pp. 28–29].

Our ultimate goal is to study stochastic differential equations, which correspond to diffusion processes, in a domain with a reflecting boundary. More specifically, we want to construct a stochastic process $\xi(t)$ that continuously reflects when it hits the boundary of a multidimensional domain D. As in the above description of the Skorokhod problem, we always assume below that $t \in \mathbb{R}_+$. The form of the SDE with reflection at the boundary that we will consider is the following:

$$\mathrm{d}\xi(t) = a(t, \xi(t))\,\mathrm{d}t + \sum_k b_k(t, \xi(t))\,\mathrm{d}W_k(t) + v(\xi(t))\,\mathrm{d}l(t), \tag{2.80}$$

where $a = a(t,x)\colon \mathbb{R}_+ \times \mathbb{R}^d \to \mathbb{R}^d$ and $b_k = b_k(t,x)\colon \mathbb{R}_+ \times \mathbb{R}^d \to \mathbb{R}^d$ are measurable functions, $\xi(0) = \xi_0$, where $\xi_0 \in \overline{D}$ is an $\mathcal{F}_0$-measurable random variable, and all the conditions of the Skorokhod problem are satisfied taking $g(t) = \xi(t)$ and

$$f(t) = \xi_0 + \int_0^t a(s, \xi(s))\,\mathrm{d}s + \sum_k \int_0^t b_k(s, \xi(s))\,\mathrm{d}W_k(s) \qquad \text{a.s.}, \tag{2.81}$$

assuming all the integrals in this equation are well defined.

Definition 2.26. We say that (2.80) has a *weak solution* if there exists a pair $(\xi(t), \varphi(t))$ of continuous $\mathcal{F}_t$-adapted processes that verify all the above conditions.

Theorem 2.3. *Let the functions $a = a(t,x)$ and $b_k = b_k(t,x)$ in equation* (2.80) *be continuous in (t,x) and bounded, i.e. let the following condition be satisfied, for some constant K:*

$$\sup_t \sup_x \left(|a(t,x)| + \sum_k |b_k(t,x)| \right) \leq K. \tag{2.82}$$

Then there exists a weak solution to equation (2.80)*.*

Proof. See [32, pp. 29–33]. QED

As is the case for stochastic differential equations without reflection, equations with reflection can also have strong solutions. This is a concept we will not need, and therefore we will not define it here (see more details on this concept in [36, pp. 456–459]).

Observe that taking f as in equation (2.81), the process (2.80) is a solution of the Skorokhod problem. If we make the identification $\xi(t) = \Gamma f(t)$ as we did in equation (2.77), then equation (2.81) becomes

$$f(t) = \xi_0 + \int_0^t a(s, \Gamma f(s))\,\mathrm{d}s + \sum_k \int_0^t b_k(s, \Gamma f(s))\,\mathrm{d}W_k(s) \qquad \text{a.s.}, \tag{2.83}$$

and if $f(t)$ is a solution of this equation, then $\xi(t) = \Gamma f(t)$ is a solution of (2.80).

Recall that in theorem 2.2 we presented the Lipschitz conditions that ensure the existence of a unique solution of a SDE. Those conditions are essentially the ones we need to ensure the existence and uniqueness of solutions of equation (2.83). Therefore assuming that the functions a and b_k satisfy

the global Lipschitz condition in x, namely

$$|a(t,x_1)-a(t,x_2)|+\sum_k |b_k(t,x_1)-b_k(t,x_2)|\leq K|x_1-x_2|, \tag{2.84}$$

for x_1, $x_2\in\mathbb{R}^d_+$ and some constant K, and the linear growth condition in x, which is

$$|a(t,x)|+\sum_k |b_k(t,x)|\leq K(1+|x|), \tag{2.85}$$

for $x\in\mathbb{R}^d_+$ and some constant K, we know that equation (2.83) has a unique solution, provided the Skorokhod map is Lipschitz-continuous.

The definition presented above of the Skorokhod problem is the most general one, and it is instructive to indicate how that abstract definition in any domain is modified when applied to some specific domains. We mention also in each case the associated reflecting SDE, which is naturally a modified version of equation (2.80). The Skorokhod problem and its associated reflecting SDE are presented in [32, pp. 1–2, 5] for the positive half-line $\mathbb{R}_+$, and in [32, pp. 21–23, 27–28] for the half-space $\mathbb{R}^d_+=\mathbb{R}^{d-1}\times\mathbb{R}_+$, in the latter case for normal reflections (see [32, 27–29] for a brief description of this problem with oblique reflections). For both these cases, from theorem 2.3 we conclude that the associated SDE with reflection has a unique solution.

For the case of the domain $D=\mathbb{R}_+$, we have simply

$$\int_0^t v(g(s))\,\mathrm{d}l(s)=l(t), \tag{2.86}$$

so condition 1) of definition 2.25 becomes $g(t)=f(t)+l(t)$. For this problem, the associated reflecting SDE is (2.80) with $k=1$, i.e.

$$\mathrm{d}\xi(t)=a(t,\xi(t))\,\mathrm{d}t+b(t,\xi(t))\,\mathrm{d}W(t)+\mathrm{d}l(t), \tag{2.87}$$

with reflection at zero. The proof of the existence of an unique solution to this equation can be seen in [32, pp. 6–7].

For the domain $D=\mathbb{R}^d_+$, assuming we have normal reflections at the boundary $\partial\mathbb{R}^d_+=\mathbb{R}^{d-1}_+\times\{0\}$, we have $f=(f_1,\ldots,f_d)$ and $g=(g_1,\ldots,g_d)$, where $f,g\in C(\mathbb{R}_+,\mathbb{R}^d)$. Thus $g(t)=f(t)+nl(t)$, where $n=(0,\ldots,0,1)$, that is, in terms of coordinates we have

$$\begin{aligned} g_1(t)&=f_1(t),\\ &\vdots\\ g_{d-1}(t)&=f_{d-1}(t),\\ g_d(t)&=f_d(t)+l(t). \end{aligned} \tag{2.88}$$

Therefore the pair (g_d,l) is the exact solution of the above one-dimensional Skorokhod problem for f_d.

In this case we have a multidimensional Skorokhod map defined as

$$\Gamma f(t) = \Gamma(f_1, \ldots, f_d)(t) = (f_1(t), \ldots, f_{d-1}(t), \Gamma f_d(t)). \tag{2.89}$$

As is the case for the unidimensional Skorokhod map, the multidimensional version is continuous and verifies the Lipschitz condition (2.79). The multidimensional reflecting SDE in $\mathbb{R}^d_+$ with normal reflection at the boundary is now

$$\mathrm{d}\xi(t) = a(t, \xi(t))\,\mathrm{d}t + \sum_k b_k(t, \xi(t))\,\mathrm{d}W_k(t) + n\,\mathrm{d}l(t). \tag{2.90}$$

Observe that when writing equation (2.90) in coordinate form, the differential of $l(t)$ is only present in the SDE corresponding to the coordinate d, that is, in the equation

$$\mathrm{d}\xi_d(t) = a_d(t, \xi(t))\,\mathrm{d}t + \sum_k b_{k,d}(t, \xi(t))\,\mathrm{d}W_k(t) + \mathrm{d}l(t), \tag{2.91}$$

since in all the other coordinates we have $n = 0$.

If we write equation (2.90) in integral form, take $f(t)$ to be defined by equation (2.81), and make the identification $\xi(t) = \Gamma f(t)$, then by the Lipschitz-continuity of the Skorokhod map and by theorem 2.3, we conclude that equation (2.90) has a unique solution.

2.7 The Local Time of a Continuous Semimartingale

In this section we define and make a few remarks on the concept of local time, using as main references [41, pp. 117–120], [17, pp. 113–116], [20, pp. 201–209] and [35, pp. 95–97].

We identify the function $l(t)$ in the definition of the Skorokhod problem and in the definition of a SDE with reflection, with the local time $L_t(x)$ defined below. At the end of this section we use this concept to demonstrate that $l(t)$ is a local time at the point zero of the process $\xi(t)$ on the domain $\mathbb{R}_+$.

Definition 2.27. The *local time*, or the *sojourn time density*, of the Wiener process W_t is a family of non-negative random variables $\{L_t(x, \omega) \colon (t, x) \in \mathbb{R}_+ \times \mathbb{R}, \omega \in \Omega\}$ that verify the following two conditions with probability one:

1) The map $(t, x) \longmapsto L_t(x, \omega)$ is continuous for almost all $\omega \in \Omega$.

2) For every Borel subset A of $\mathbb{R}$ and $t \in \mathbb{R}_+$,

$$\int_0^t \mathbb{1}_A(W_s)\,\mathrm{d}s = 2\int_A L_t(x, \omega)\,\mathrm{d}x, \tag{2.92}$$

where the integral on the left-hand side of this equation corresponds to the Lebesgue measure of the set $\{0 \leq s \leq t \colon W_s \in A\}$, and is usually called the *occupation time* of A by the Brownian path up to time t.

The concept of local time corresponds to the measurement of the amount of time spent by a path of the Wiener process in the neighbourhood of a point $x \in \mathbb{R}$, and was introduced by Paul Lévy in 1948 (the publication where this concept was first defined is included in the bibliography of [17]). According to [41, p. 117], the first rigorous proof of the existence of local time was published in 1958 by Hale Trotter (1931–). The article with this proof is listed in the bibliography of [20]. It is interesting to note that, according to [33, p. 226], the family $\{L_t(x)\}$ may be chosen such that the map $(t, x) \longmapsto L_t(x)$ is a.s. Hölder continuous of order γ, for every $\gamma < 1/2$ (as is the case for the Wiener process).

Equation (2.92) is actually a particular case of the equation in the proposition that follows.

Proposition 2.9. Let $f\colon \mathbb{R} \to \mathbb{R}_+$ be a Borel-measurable function. Then

$$\int_0^t f(W_s)\,\mathrm{d}s = 2\int_{-\infty}^{+\infty} f(x)L_t(x,\omega)\,\mathrm{d}x \qquad \text{a.s.} \tag{2.93}$$

Proof. See [17, p. 116]. QED

The existence of this family $\{L_t(x,\omega)\}$ of random variables can be proved by showing that the right-hand side of the Tanaka formula (equation (2.101) below) leads to a family of random variables verifying conditions 1) and 2) of definition 2.27 (see the proof in [20, pp. 207–209]). This family, which is unique (according to [17, p. 113]), is given by the following limit, for $x \in \mathbb{R}$:

$$L_t(x) = \lim_{\varepsilon \to 0} \frac{1}{4\varepsilon} \int_0^t \mathbb{1}_{]x-\varepsilon, x+\varepsilon[}(W_s)\,\mathrm{d}s \qquad \text{a.s.,} \tag{2.94}$$

which is finite and always exists. The integral on the right-hand side of (2.94) corresponds to the Lebesgue measure of the set $\{0 \leq s \leq t\colon |W_s - x| \leq \varepsilon\}$. The introduction of this definition was motivated by the discovery that, for every $x \in \mathbb{R}$, the set $\{t \in \mathbb{R}_+\colon W_t(\omega) = x\}$ has Lebesgue measure zero (the proof can be found in [20, pp. 105–106]), which gives us no information about the time spent by a path in the neighbourhood of x. Although we have been discussing specifically the concept of the local time of the Wiener process, it is also possible to define the local time of any continuous semimartingale X. In that case we substitute $\mathrm{d}s$ in equation (2.94) by $\mathrm{d}\langle X, X\rangle_s$.

Proposition 2.10. Let f be a Borel-measurable function. If X is a continuous semimartingale, then the *occupation-time equation* is

$$\int_0^t f(X_t)\,\mathrm{d}\langle X\rangle_s = 2\int_{-\infty}^{+\infty} f(x)L_t(x,\omega)\,\mathrm{d}x \qquad \text{a.s.} \tag{2.95}$$

Proof. See [39, pp. 126–127]. QED

Equation (2.93) is clearly a particular case of this equation, since we know that the quadratic variation of the Wiener process verifies $\langle X, X\rangle_s = s$, denoting the Wiener process by X.

We want now to derive another representation for local time, called the Tanaka formula, but before we do so we take a brief detour and make a few observations regarding the delta function, not only because

it is used in the deduction of that formula, but also because of its importance and ubiquitousness. The Dirac delta function δ, which is actually something of a misnomer, since it is not a function at all, at least not in the traditional sense (we will not go into its rigorous definition here, since it requires concepts from Schwartz's theory of distributions, which is not the subject of this thesis), satisfies two essential conditions:

$$\delta(x) = \begin{cases} +\infty & \text{if } x = 0, \\ 0 & \text{if } x \neq 0, \end{cases} \qquad \text{and} \qquad \int_{-\infty}^{+\infty} \delta(x)\,\mathrm{d}x = 1. \tag{2.96}$$

The following fundamental equality, sometimes also taken as the definition of the delta function,

$$\int_{-\infty}^{+\infty} \delta(x-a)\varphi(x)\,\mathrm{d}x = \varphi(a), \tag{2.97}$$

is valid for a continuous function φ (again we would have to go into the theory of distributions to justify this). The Dirac delta function is useful for describing the spatial density of a physical magnitude which is concentrated at a point in space. Despite bearing the name of Paul Dirac (1902–1984), who used it in his publications since the 1920s, the definition (2.97) was actually anticipated by Augustin-Louis Cauchy (1789–1857) a century earlier (in [2, pp. 890–891] is presented a formula used by Cauchy which is essentially equation (2.97)). It is quite striking to realize that a fairly circumscribed concept had such a wide-raging influence. For instance, the avoidance of its use was one of the main motivations for the pioneering mathematically rigorous construction of quantum mechanics in terms of operators in Hilbert space, obtained by John von Neumann (1903–1957) and published in book form in 1932. In this major work, the author calls expressions such as (2.97) 'fiction' ([42, p. 14]) and functions such as the delta function and its derivative 'unreal objects' ([42, p. 15]). In the 1950s Laurent Schwartz (1915–2002) gave a rigorous definition of the delta function in the context of his theory of distributions, which in addition to having had an enormous influence on many branches of mathematics, would later be used to provide the underpinnings of axiomatic quantum field theory, one of the formulations of this very important physical theory, which models some of the fundamental interactions of matter.

Returning to our original purpose, which is to derive the Tanaka formula, we fix a number $a \in \mathbb{R}$ and take $f(x)$ in equation (2.93) to be the Dirac delta function evaluated at $x - a$, obtaining

$$\int_0^t \delta(W_s(\omega) - a)\,\mathrm{d}s = 2L_t(a,\omega), \tag{2.98}$$

using equality (2.97). The integral on the left-hand side of this equation is merely formal, but we want to offer a reasonable interpretation of it. This interpretation is in essence the Tanaka formula, as we will see, since that formula presents us with a representation of $L_t(a)$ in terms of expressions that are perfectly well-defined.

If we consider the function $u(x) = |x|$, then $u'(x) = \operatorname{sgn}(x)$ and $u''(x) = 2\delta(x)$, once again in the context of the theory of distributions. Therefore if we take u to be the non-decreasing, continuously differentiable function $u(x) = (x-a)^+$ on $\mathbb{R}\backslash\{a\}$, then $u''(x) = \delta(x-a)$, and so applying the Itô formula

in one dimension with $X(\cdot) = W(\cdot)$, we obtain

$$\mathrm{d}u = u'(x)\,\mathrm{d}W + \frac{1}{2}u''(x)\,\mathrm{d}t = \mathbb{1}_{]a,+\infty[}(W)\,\mathrm{d}W + \frac{1}{2}\delta(W-a)\,\mathrm{d}t, \tag{2.99}$$

or in integral form

$$(W_t - a)^+ - (W_0 - a)^+ = \int_0^t \mathbb{1}_{]a,+\infty[}(W_s)\,\mathrm{d}W_s + \frac{1}{2}\int_0^t \delta(W_s - a)\,\mathrm{d}s \qquad \text{a.s.} \tag{2.100}$$

Substituting equation (2.98) in this last equation we obtain the *Tanaka formula*

$$L_t(a) = (W_t - a)^+ - (W_0 - a)^+ - \int_0^t \mathbb{1}_{]a,+\infty[}(W_s)\,\mathrm{d}W_s \qquad \text{a.s.} \tag{2.101}$$

Since we have $f = f^+ - f^-$ and $|f| = f^+ + f^-$, by symmetry we can also write the Tanaka formula as

$$L_t(a) = (W_t - a)^- - (W_0 - a)^- + \int_0^t \mathbb{1}_{]-\infty,a]}(W_s)\,\mathrm{d}W_s \qquad \text{a.s.} \tag{2.102}$$

Adding these two representations of the Tanaka formula, and using the following result (taken from [20, p. 205])

$$\Pr\left(\int_0^t \mathbb{1}_{\{a\}}(W_s)\,\mathrm{d}W_s = 0\right) = 1, \qquad t \geq 0, \tag{2.103}$$

we obtain a third representation of the Tanaka formula, which is

$$2L_t(a) = |W_t - a| - |W_0 - a| - \int_0^t \mathrm{sgn}(W_s - a)\,\mathrm{d}W_s \qquad \text{a.s.} \tag{2.104}$$

Note that

$$M_t = \int_0^t \mathrm{sgn}(W_s)\,\mathrm{d}W_s \tag{2.105}$$

is also a Wiener process, since from its expression in differential form $\mathrm{d}M_t = \mathrm{sgn}(W_t)\,\mathrm{d}W_t$, we conclude that $(\mathrm{d}M_t)^2 = (\mathrm{d}W_t)^2 = \mathrm{d}t$.

We return now to the case of the Skorokhod problem on the domain $D = \mathbb{R}_+$, and we demonstrate that $l(t)$ is a local time at the point zero of the process $\xi(t)$ (adapting the presentation in [32, pp. 10–12]). If $\xi(t)$ is a solution of equation (2.87), then $\xi(t) = \xi^+(t) = (\xi(t) - 0)^+$. Therefore by the Tanaka formula in the representation (2.101) we obtain

$$\begin{aligned} L_t(0) &= (\xi(t) - 0)^- (\xi(0) - 0)^- \int_0^t \mathbb{1}_{]0,+\infty[}(\xi(s))\,\mathrm{d}\xi(s) \qquad \text{a.s.} \\ &= \xi(t) - \xi(0) - \int_0^t \mathbb{1}_{]0,+\infty[}(\xi(s))a(s,\xi(s))\,\mathrm{d}s - \int_0^t \mathbb{1}_{]0,+\infty[}(\xi(s))b(s,\xi(s))\,\mathrm{d}W(s) \end{aligned}$$

$$- \int_0^t \mathbb{1}_{]0,+\infty[}(\xi(s))\,\mathrm{d}l(s) \qquad \text{a.s.} \tag{2.106}$$

$$= \xi(t) - \xi(0) - \int_0^t a(s,\xi(s))\,\mathrm{d}s + \int_0^t \mathbb{1}_{\{0\}}(\xi(s))a(s,0)\,\mathrm{d}s - \int_0^t b(s,\xi(s))\,\mathrm{d}W(s)$$

$$+ \int_0^t \mathbb{1}_{\{0\}}(\xi(s))b(s,0)\,\mathrm{d}W(s) \qquad \text{a.s.}$$

$$= l(t) + \int_0^t \mathbb{1}_{\{0\}}(\xi(s))a(s,0)\,\mathrm{d}s + \int_0^t \mathbb{1}_{\{0\}}(\xi(s))b(s,0)\,\mathrm{d}W(s) \qquad \text{a.s.} \tag{2.107}$$

Observe that the last integral on the right-hand side of equation (2.106) is zero due to condition 3) in definition 2.25, and that the term $l(t)$ in equation (2.107) is obtained from the previous equality by writing equation (2.87), in integral form, in terms of $l(t)$. Using now the occupation-time equation (2.95) with $f = \mathbb{1}_{\{0\}}$, we obtain

$$\int_0^t \mathbb{1}_{\{0\}}(\xi(s))\,\mathrm{d}\langle\xi\rangle_s = 2\int_{-\infty}^{+\infty} \mathbb{1}_{\{0\}}(x)L_t(x)\,\mathrm{d}x = 0 \qquad \text{a.s.,} \tag{2.108}$$

and since $\langle\xi\rangle_s = \langle W\rangle_s$ by proposition 2.6 (we also know that $\langle W\rangle_s = s$), then

$$\int_0^t \mathbb{1}_{\{0\}}(\xi(s))b^2(s,0)\,\mathrm{d}\langle W\rangle_s = 0. \tag{2.109}$$

Assuming that $b(s,0) \neq 0$ almost everywhere, we conclude that

$$\Pr\left(\int_0^t \mathbb{1}_{\{0\}}(\xi(s))\,\mathrm{d}s = 0\right) = 1 \tag{2.110}$$

and that the two integrals in equation (2.107) are zero.

So we have $L_t(0) = l(t)$ a.s., for $t \in \mathbb{R}_+$. This result tells us that the amount of time spent by process ξ at $\xi(s) = 0$ is zero.

Chapter 3

Derivation of the Navier–Stokes Equations from a Stochastic Variational Principle

3.1 Introduction

We understand by a *variational principle*, a principle which describes the conditions under which certain quantities attain extreme values. The variational principle is regarded as a postulate that furnishes the foundations of a physical theory and its most powerful attribute is its generality. Examples of physical theories where variational principles are applied include classical mechanics and quantum mechanics, electromagnetism, general relativity and field theory, in other words, almost all the established physical theories.

Definition 3.1. The *action functional* S is the integral of the *Lagrange function*, or *Lagrangian*, L of the physical system with respect to the time, that is,

$$S = \int_{t_1}^{t_2} L \, \mathrm{d}t. \tag{3.1}$$

Remember that L is defined as $L := T - V$, where T is the kinetic energy of the system, and V is the potential energy due to the external or internal forces.

The *principle of least action*, also known as the *principle of stationary action*, postulates that particles always follow the path of least action. By minimizing the variation of the action, denoted by δS, we can determine the path, among the infinite number of possible paths, that is actually followed (since this path represents a minimum). This is done by describing the unknown action functional as a minimal term and a variation, i.e.

$$S \longrightarrow S + \delta S. \tag{3.2}$$

Therefore the path followed by the particles is the one for which $\delta S = 0$, provided L was adequately chosen. Computing the variation δS leads to the equations of motion for the system.

The principle of least action was first formulated for general cases by Joseph-Louis Lagrange (1736–1813), building on the pioneering work of Pierre de Fermat (1601–1665), Pierre-Louis Maupertuis (1698–1759) and Euler. See an account of these and later developments in [50, pp. 19–32].

The significance of the variational principle approach is that most physical theories may be formulated by specifying a suitable Lagrangian, independently of the choice of a coordinate system. For a clear and concise discussion of the Lagrangian formalism and more, see [31, pp. 471–492]. The author of that work has some distinctive views, and especially pertinent to the above comments on the generality of the Lagrangian approach is the author's criticism of this very same generality (see [31, p. 291]). A very clear exposition on variational principles, covering both their historical and philosophical facets, is [50].

Since they are used below, we briefly define what Lagrange multipliers are. The idea behind Lagrange multipliers is to replace the Lagrange function L by a modified Lagrange function $\mathcal{L}$ of the form

$$\mathcal{L} = L + \sum_k \lambda_k(t) N_k, \tag{3.3}$$

where N_k are smooth functions called *constraints* and λ_k are functions called *Lagrange multipliers*, with the goal of determining the extreme values of S, that is, of minimizing it, subject to certain constraints.

We will proceed in the next section by providing an expanded exposition of the results presented in [5, pp. 2–5]. The objective is to derive the Navier–Stokes equations from a stochastic variational principle with a Lagrange multiplier expressed in terms of the pressure. This is one of the approaches to the broader problem of deriving these equations from variational principles. Several approaches have been tried and the bibliography of [5] includes the references to some of the most important articles where this problem has been tackled. We provide a few observations on the early attempts before we proceed. It has been shown that the Euler equation can be derived from variational principles, but doing the same for the Navier–Stokes equations has proved to be much more challenging. This is due to the fact that these equations take into account the viscosity of a fluid, which corresponds to the dissipative term in the latter equations (recall that the Euler equations do not have this term, making them a conservative system). According to [48, p. 133], the way to overcome the difficulty of deriving the Navier–Stokes equations is to extend the notions of the traditional calculus of variations, which was done for the first time in [18]. It involved combining concepts from the traditional calculus of variations and stochastic analysis, since the viscosity term of the equations can be seen as introducing randomness into the model (and stochastic analysis is the field of mathematics that deals the most with this multifaceted phenomenon). More precisely, it involved deriving the Navier–Stokes equations from the Euler equation while taking into account the fluctuations of Brownian motion. In the article [27], an approach to the stochastic calculus of variations in terms of the sample paths of the Wiener process was developed, building on an earlier extension of the classical calculus of variations to stochastic processes. Due to limitations in the method presented in this article (see the explanation in [48, p. 134]), one of its authors published a different attempt at deriving the Navier–Stokes equations via a stochastic calculus of variations in [48]. Further

details and references on the more recent approaches used to tackle this problem can be found in [5, pp. 1–2].

3.2 Derivation of the Equations

We assume throughout this section that we have a filtered probability space $(\Omega, \mathcal{F}, \mathcal{F}_t, P)$. Let $\mathbb{T}$ be a flat torus of dimension d, which we identify with $[0, 2\pi]^d$ (corresponding to periodic boundary conditions), and let ξ be a semimartingale defined, in differential form, as

$$\mathrm{d}\xi_t(x, \omega) = \mathrm{d}M_t(x, \omega) + \mathrm{D}_t\xi_t(x, \omega)\,\mathrm{d}t, \qquad x \in \mathbb{T}, \tag{3.4}$$

and such that $\xi_0(x) = x$. We call $\mathrm{d}M_t$ the martingale part of the decomposition of ξ_t and $\mathrm{D}_t\xi_t$ the drift. From now on we omit the parameter $\omega \in \Omega$. As we mentioned in section 2.5, equation (3.4) is actually an abbreviation for the integral expression

$$\xi_t(x) = \xi_0(x) + M_t(x) + \int_0^t \mathrm{D}_s\xi_s(x)\,\mathrm{d}s \qquad \text{a.s.} \tag{3.5}$$

Remember that since martingales are not differentiable a.s. in time, we cannot compute their derivative with respect to time. Nevertheless, it is possible to define a time derivative after conditional expectation of the paths.

Definition 3.2. The *generalized derivative*, or *mean forward derivative (of Nelson)*, D_t of a martingale $F\colon [0, T] \times \mathbb{T} \to \mathbb{R}$ is defined thus

$$\mathrm{D}_t F(t, \xi_t(x)) := \lim_{\varepsilon \to 0} \frac{1}{\varepsilon} \mathrm{E}\left[F(t + \varepsilon, \xi_{t+\varepsilon}(x)) - F(t, \xi_t(x)) | \mathcal{F}_t\right] \qquad \text{a.s.} \tag{3.6}$$

when the limit exists.

Note that the generalized derivative is also well-defined for more general domains, namely for the domain $[0, \infty[\times M$, where M is a Hausdorff, second-countable, differentiable n-manifold (the source of this observation is [6, pp. 194, 201]). The mean backward derivative can also be defined by taking the conditional expectation of $F(t, \xi_t(x)) - F(t - \varepsilon, \xi_{t-\varepsilon}(x))$ in definition 3.2, but we will not use this definition in what follows.

Observe that when $F(t, x) = x$, the definition of generalized derivative becomes

$$\mathrm{D}_t\xi_t(x) = \lim_{\varepsilon \to 0} \frac{1}{\varepsilon} \mathrm{E}\left[\xi_{t+\varepsilon}(x) - \xi_t(x) | \mathcal{F}_t\right], \tag{3.7}$$

that is, it corresponds to the semimartingale's drift.

Consider now the following SDE

$$\mathrm{d}g_t(x) = \sqrt{2\nu}\,\mathrm{d}W_t + u(t, g_t(x))\,\mathrm{d}t, \qquad x \in \mathbb{T}, \tag{3.8}$$

with $g_0(x) = x$, and where $g_t(\cdot)$ are diffeomorphisms, which represent diffusions on the torus. Notice that this SDE is written in integral form and in terms of its components thus

$$g_t^i(x) = g_0^i(x) + \sqrt{2\nu} \int_0^t \mathrm{d}W_s^i + \int_0^t u^i(s, g_s(x))\,\mathrm{d}s, \qquad i = 1, \ldots, d \tag{3.9}$$

which holds almost surely for each i.

As above, for the case of $F(t,x) = x$, we have $\mathrm{D}_t g_t(x) = u(t, g_t(x))$. For the case of $F(t,x) = u(t,x)$, using the Itô formula in d dimensions we obtain the stochastic differential of $\mathrm{D}_t g_t(x)$, which is

$$\begin{aligned}
\mathrm{d}u^i(t, g_t(x)) &= \partial_t u^i\,\mathrm{d}t + \sum_{j=1}^d \partial_j u^i\,\mathrm{d}g_t^j(x) + \frac{1}{2}\sum_{i=1}^d\sum_{j=1}^d \partial_i\partial_j u^i 2\nu\,\mathrm{d}t \\
&= \partial_t u^i\,\mathrm{d}t + \sum_{j=1}^d \partial_j u^i \left(\sqrt{2\nu}\,\mathrm{d}W_t^j + u^j(t, g_t(x))\,\mathrm{d}t\right) + \nu\sum_{i=1}^d\sum_{j=1}^d \partial_i\partial_j u^i\,\mathrm{d}t \\
&= \sqrt{2\nu}\sum_{j=1}^d \partial_j u^i\,\mathrm{d}W_t^j + \left(\partial_t u^i + \sum_{j=1}^d u^j\partial_j u^i + \nu\sum_{i=1}^d\sum_{j=1}^d \partial_i\partial_j u^i\right)\mathrm{d}t \\
&= \sqrt{2\nu}\sum_{j=1}^d \partial_j u^i\,\mathrm{d}W_t^j + (\partial_t u^i + u^i\cdot\nabla u^i + \nu\Delta u^i)\,\mathrm{d}t,
\end{aligned} \tag{3.10}$$

so the second generalized derivative of $g_t(x)$ is

$$\mathrm{D}_t\mathrm{D}_t g_t^i(x) = \mathrm{D}_t u^i(t, g_t(x)) = (\partial_t u^i + u^i\cdot\nabla u^i + \nu\Delta u^i)(t, g_t(x)), \qquad i = 1, \ldots, d. \tag{3.11}$$

This equation can be written in the abbreviated form

$$\mathrm{D}_t u(t, g_t(x)) = (\partial_t u + u\cdot\nabla u + \nu\Delta u)(t, g_t(x)). \tag{3.12}$$

We now denote by $\mathbb{H}$ a linear subspace dense in $L^2([0,T]\times\mathbb{T})$, by E the expectation with respect to the underlying probability measure P, and we let $p \in \mathbb{H}$ be a function. We define the following action functional

$$\begin{aligned}
S(g_t, p) &:= \frac{1}{2}\,\mathrm{E}\int_0^T\int |\mathrm{D}_t g_t(x)|^2\,\mathrm{d}x\mathrm{d}t + \mathrm{E}\int_0^T\int p(t, g_t(x))(\det\nabla g_t(x) - 1)\,\mathrm{d}x\mathrm{d}t \\
&:= S^1(g_t) + S^2(g_t, p),
\end{aligned} \tag{3.13}$$

where $S^1(g_t)$ can be interpreted as corresponding to an action integral of the Lagrangian composed solely of the kinetic energy and $S^2(g_t, p)$ as corresponding to the part of a modified Lagrange function which includes the Lagrange multiplier expressed in terms of the pressure p. In fact, since $\mathrm{D}_t g_t(x) = u(t, g_t(x))$, we can interpret $S^1(g_t)$ as corresponding to the total energy functional $\mathcal{E}$ of the fluid, which is

written in terms of the d components of the drift velocity field u as

$$\mathcal{E} = \frac{1}{2}\sum_{i=1}^{d}\int_0^T\int (u^i(t, g_t(x)))^2 \,\mathrm{d}x\mathrm{d}t. \tag{3.14}$$

It is important to mention that a solution u of the Navier–Stokes equation is only considered to be physically reasonable if it satisfies

$$\int |u(t,x)|^2 \,\mathrm{d}x < \infty, \qquad t \in \mathbb{R}_+, \tag{3.15}$$

i.e. the energy cannot be infinite. Recalling the definition of the modified Lagrange function (3.3), we see that $p(t, g_t(x))$ corresponds to the Lagrange multiplier and $\det \nabla g_t(x) - 1$ to the restrictions. The origin of the factor $\det \nabla g_t(x) - 1$ in the definition of $S^2(g_t, p)$ warrants a brief explanation. It is a well-known fact that when we make a change of variables in an integral we have

$$\int f(g_t(x)) \,\mathrm{d}x = \int f(x) \det \nabla g_t(x) \,\mathrm{d}x, \tag{3.16}$$

where ∇g_t is the Jacobian matrix. If we set $\det \nabla g_t(x) = 1$, then the Lebesgue measure is conserved by the transformation $g_t(x)$ for all t. If we calculate the derivative of (3.16) with respect to t, we obtain

$$\int f \nabla \cdot u \,\mathrm{d}x = 0, \tag{3.17}$$

for any f, which implies that $\nabla \cdot u = 0$ (this is the condition for the incompressibility of a fluid, obtained in equation (1.32)).

We will consider the following functional variations of g_t and p:

$$g_t(\cdot) \longrightarrow g_t^\varepsilon(\cdot) = g_t(\cdot) + \varepsilon h(t, g_t(\cdot)), \tag{3.18}$$

$$p(t,\cdot) \longrightarrow p^\varepsilon(t,\cdot) = p(t,\cdot) + \varepsilon\varphi(t, g_t(\cdot)), \tag{3.19}$$

where $h(t,x)$ and $\varphi(t,x)$ are smooth functions in $x, \varphi \in \mathbb{H}$ such that $h(T,\cdot) = h(0,\cdot) = 0$. We assume also that we have periodic boundary conditions with period 2π for all x.

Theorem 3.1. *The diffusions $g_t(\cdot)$ and a function $p \in \mathbb{H}$ verify $\delta S(g_t, p) = 0$, i.e. they are critical to the action functional S, if and only if the drift $u(t,\cdot)$ of $g_t(\cdot)$ satisfies the Navier–Stokes equations* (1.63) *and* (1.64)*, for $x \in \mathbb{T}$ and $t \in [0,T]$, taking $\rho = 1$ and $f = 0$.*

Proof. We start by assuming that u satisfies the Navier–Stokes equations and we prove that $\delta S = 0$. To compute the variations of $S^1(g_t)$ and $S^2(g_t, p)$ we need some additional results. These are presented before the computation of each variation.

If X_t and Y_t are two martingales, then using the Itô product formula (2.71) we know that

$$\mathrm{d}(X_tY_t) = X_t \,\mathrm{d}Y_t + Y_t \,\mathrm{d}X_t + \mathrm{d}X_t \,\mathrm{d}Y_t. \tag{3.20}$$

The last term on the right-hand side of this equality is known as the *Itô contraction*. So we have

$$\mathrm{d}(\mathrm{D}_t g_t \cdot h(t, g_t)) = \mathrm{d}\mathrm{D}_t g_t \cdot h(t, g_t) + \mathrm{D}_t g_t \cdot \mathrm{d}h(t, g_t) + \mathrm{d}\mathrm{D}_t g_t \cdot \mathrm{d}h(t, g_t) \tag{3.21}$$

and

$$\mathrm{D}_t(\mathrm{D}_t g_t \cdot h(t, g_t)) = \mathrm{D}_t\mathrm{D}_t g_t \cdot h(t, g_t) + \mathrm{D}_t g_t \cdot \mathrm{D}_t h(t, g_t) + \mathrm{d}\mathrm{D}_t g_t \cdot \mathrm{d}h(t, g_t). \tag{3.22}$$

Denoting by $\langle \cdot, \cdot \rangle$ the scalar product in $L^2(\mathbb{T})$, from the previous two equations we deduce that

$$\mathrm{d}\langle \mathrm{D}_t g_t, h \rangle = \langle \mathrm{d}\mathrm{D}_t g_t, h \rangle + \langle \mathrm{D}_t g_t, \mathrm{d}h \rangle + \langle \mathrm{d}\mathrm{D}_t g_t, \mathrm{d}h \rangle \tag{3.23}$$

and that

$$\mathrm{D}_t\langle \mathrm{D}_t g_t, h \rangle = \langle \mathrm{D}_t\mathrm{D}_t g_t, h \rangle + \langle \mathrm{D}_t g_t, \mathrm{D}_t h \rangle + \langle \mathrm{d}\mathrm{D}_t g_t, \mathrm{d}h \rangle. \tag{3.24}$$

By equation (3.10) we know that

$$\mathrm{d}\mathrm{D}_t g_t(x) = \sqrt{2\nu}\nabla u \, \mathrm{d}W_t + (\partial_t u + u \cdot \nabla u + \nu \Delta u) \, \mathrm{d}t, \tag{3.25}$$

and using again the Itô formula we obtain

$$\begin{aligned}
\mathrm{d}h^i(t, g_t(x)) &= \partial_t h^i \, \mathrm{d}t + \sum_{j=1}^{d} \partial_j h^i \, \mathrm{d}g_t^j(x) + \frac{1}{2} \sum_{i=1}^{d} \sum_{j=1}^{d} \partial_i \partial_j h^i 2\nu \, \mathrm{d}t \\
&= \partial_t h^i \, \mathrm{d}t + \sum_{j=1}^{d} \partial_j h^i \left(\sqrt{2\nu} \, \mathrm{d}W_t^j + u^j(t, g_t(x)) \, \mathrm{d}t \right) + \nu \sum_{i=1}^{d} \sum_{j=1}^{d} \partial_i \partial_j h^i \, \mathrm{d}t \\
&= \sqrt{2\nu} \sum_{j=1}^{d} \partial_j h^i \, \mathrm{d}W_t^j + \left(\partial_t h^i + \sum_{j=1}^{d} u^j \partial_j h^i + \nu \sum_{i=1}^{d} \sum_{j=1}^{d} \partial_i \partial_j h^i \right) \mathrm{d}t,
\end{aligned} \tag{3.26}$$

for $i = 1, \ldots, d$, that is,

$$\mathrm{d}h(t, g_t(x)) = \sqrt{2\nu}\nabla h \, \mathrm{d}W_t + (\partial_t h + u \cdot \nabla h + \nu \Delta h)\mathrm{d}t. \tag{3.27}$$

Therefore multiplying equations (3.25) and (3.27) gives us the Itô contraction of $\mathrm{D}_t g_t(x)$ and $\mathrm{d}h(t, g_t(x))$, which is

$$\mathrm{d}\mathrm{D}_t g_t(x) \cdot \mathrm{d}h(t, g_t(x)) = 2\nu(\nabla u \cdot \nabla h)(t, g_t(x)) \, \mathrm{d}t. \tag{3.28}$$

For this calculation we used the formal multiplication rules (2.64)–(2.66), which cause all terms of the product, except one, to vanish.

If we let the variation of the action functional S be denoted by

$$\delta S(g_t, p) = \frac{\mathrm{d}}{\mathrm{d}\varepsilon}\bigg|_{\varepsilon=0} S(g_t^\varepsilon, p^\varepsilon), \tag{3.29}$$

then the variation of $S^1(g_t)$ is

$$\begin{aligned}
\delta S^1(g_t) &= \frac{\mathrm{d}}{\mathrm{d}\varepsilon}\Big|_{\varepsilon=0} \frac{1}{2}\,\mathrm{E}\int_0^T\int |\mathrm{D}_t g_t^\varepsilon(x)|^2\,\mathrm{d}x\mathrm{d}t \\
&= \mathrm{E}\int_0^T\int \mathrm{D}_t g_t(x)\frac{\mathrm{d}}{\mathrm{d}\varepsilon}\Big|_{\varepsilon=0}(\mathrm{D}_t g_t(x) + \varepsilon \mathrm{D}_t h(t, g_t(x)))\,\mathrm{d}x\mathrm{d}t \\
&= \mathrm{E}\int_0^T\int \mathrm{D}_t g_t(x)\cdot \mathrm{D}_t h(t, g_t(x))\,\mathrm{d}x\mathrm{d}t \\
&= \mathrm{E}\int_0^T\int \mathrm{D}_t\langle \mathrm{D}_t g_t, h\rangle\,\mathrm{d}x\mathrm{d}t - \mathrm{E}\int_0^T\int \mathrm{D}_t\mathrm{D}_t g_t(x)\cdot h(t, g_t(x))\,\mathrm{d}x\mathrm{d}t \\
&\quad - \mathrm{E}\int_0^T\int \mathrm{d}\mathrm{D}_t g_t(x)\cdot \mathrm{d}h(t, g_t(x))\,\mathrm{d}x\mathrm{d}t \qquad (3.30) \\
&= -\mathrm{E}\int_0^T\int \mathrm{D}_t\mathrm{D}_t g_t(x)\cdot h(t, g_t(x))\,\mathrm{d}x\mathrm{d}t - 2\nu\,\mathrm{E}\int_0^T\int (\nabla u\cdot\nabla h)(t, g_t(x))\,\mathrm{d}x\mathrm{d}t \qquad (3.31) \\
&= -\mathrm{E}\int_0^T\int ((\partial_t u + u\cdot\nabla u + \nu\Delta u)\cdot h)(t, g_t(x))\,\mathrm{d}x\mathrm{d}t \\
&\quad - 2\nu\,\mathrm{E}\left(\int \nabla u\cdot h|_{t=0}^{t=T}\,\mathrm{d}x - \int_0^T\int (\Delta u\cdot h)\,\mathrm{d}x\mathrm{d}t\right) \qquad (3.32) \\
&= -\mathrm{E}\int_0^T\int ((\partial_t u + u\cdot\nabla u - \nu\Delta u)\cdot h)(t, g_t(x))\,\mathrm{d}x\mathrm{d}t. \qquad (3.33)
\end{aligned}$$

For obtaining equality (3.30) we used (3.24) and for obtaining (3.31) we used the following computation

$$\mathrm{E}\int_0^T\int \mathrm{D}_t\langle \mathrm{D}_t g_t, h\rangle\,\mathrm{d}x\mathrm{d}t = \mathrm{E}\int \langle D_t g_t, h\rangle|_{t=0}^{t=T}\,\mathrm{d}x = 0, \qquad (3.34)$$

which is valid since we are assuming that $h(T,\cdot) = h(0,\cdot) = 0$, and also equation (3.28). Finally, for obtaining (3.32) we used equation (3.12) and integration by parts.

Before we compute the variation of $S^2(g_t, p)$ we need a few auxiliary results. The article [37, pp. 868–870] includes several of the computations we present below but using a different notation. Regarding notation, it should be noted that if $A_{m\times n} = [A_{ij}]$ and $B_{n\times m} = [B_{ji}]$ are two matrices, then the trace of AB is defined as

$$\mathrm{tr}(AB) = \sum_{i=1}^m\sum_{j=1}^n A_{ij}B_{ji}. \qquad (3.35)$$

The trace of a single matrix A is the sum of the elements that verify $i = j$. Using the notation of summation over repeated indices, we can write

$$\mathrm{tr}(A) = A_{ii}. \qquad (3.36)$$

As mentioned before, ∇g_t is the Jacobian matrix of g_t, and we frequently use the notation $(\nabla g_t)_{ij} = \partial_j g_t^i$. It is clear that we have the equality

$$(\nabla g_t)^{-1}_{ij}(\nabla g_t)_{jk} = \delta_{ik}, \tag{3.37}$$

so differentiating it and summing with respect to i, we obtain

$$\sum_i \left(\partial_i(\nabla g_t)^{-1}_{ij}(\nabla g_t)_{jk} + (\nabla g_t)^{-1}_{ij}\partial_i(\nabla g_t)_{jk}\right) = 0, \tag{3.38}$$

which is equivalent to

$$\sum_i \left(\partial_i(\nabla g_t)^{-1}_{ij}\partial_k g_t^j + (\nabla g_t)^{-1}_{ij}\partial_i\partial_k g_t^j\right) = 0, \tag{3.39}$$

if and only if

$$\sum_i \partial_i(\nabla g_t)^{-1}_{ik} = -\sum_i((\nabla g_t)^{-1}_{ij}\partial_k\partial_i g_t^j)(\nabla g_t)^{-1}_{jk} = 0, \tag{3.40}$$

where the last equality is justified by result (3.43), which we obtain below. That is, we have the equality

$$\sum_i \partial_i(\nabla g_t)^{-1}_{ij} = 0. \tag{3.41}$$

It is a well-known result from linear algebra, sometimes called the Jacobi formula, that if A is an invertible matrix then

$$\frac{\mathrm{d}}{\mathrm{d}t}\det A(t) = \det A(t)\,\mathrm{tr}\left(A(t)^{-1}\frac{\mathrm{d}}{\mathrm{d}t}A(t)\right). \tag{3.42}$$

Using this result to calculate the derivative of $\det \nabla g_t(x) = 1$ we obtain

$$\begin{aligned}\partial_k \det(\nabla g_t) &= \det(\nabla g_t)\,\mathrm{tr}\left((\nabla g_t)^{-1}\partial_k(\nabla g_t)\right)\\ &= \mathrm{tr}\left((\nabla g_t)^{-1}_{ij}\partial_k\partial_i g_t^j\right)\\ &= \sum_i (\nabla g_t)^{-1}_{ij}\partial_k\partial_i g_t^j = 0.\end{aligned} \tag{3.43}$$

Using the Jacobi formula again to calculate the derivative of $\det \nabla g_t^\varepsilon(x)$ we obtain

$$\begin{aligned}\left.\frac{\mathrm{d}}{\mathrm{d}\varepsilon}\right|_{\varepsilon=0}\det \nabla g_t^\varepsilon(x) &= \det(\nabla g_t(x))\,\mathrm{tr}\left((\nabla g_t(x))^{-1}\left.\frac{\mathrm{d}}{\mathrm{d}\varepsilon}\right|_{\varepsilon=0}\nabla g_t^\varepsilon(x)\right)\\ &= \det(\nabla g_t(x))\,\mathrm{tr}\left((\nabla g_t(x))^{-1}\nabla h(t, g_t(x))\right)\\ &= \det(\nabla g_t(x))(\nabla g_t(x))^{-1}_{ij}\partial_i h^j(t, g_t).\end{aligned} \tag{3.44}$$

Is is clear that

$$\begin{aligned}\partial_i\left(p(t, g_t)(\nabla g_t)^{-1}_{ij}h^j(t, g_t)\right) &= \partial_i(p(t, g_t))(\nabla g_t)^{-1}_{ij}h^j(t, g_t) + p(t, g_t)\partial_i((\nabla g_t)^{-1}_{ij})h^j(t, g_t)\\ &\quad + p(t, g_t)(\nabla g_t)^{-1}_{ij}\partial_i(h^j(t, g_t)),\end{aligned} \tag{3.45}$$

so integrating the left-hand side this equation we obtain

$$\int_0^T \int \partial_i \left(p(t, g_t(x))(\nabla g_t(x))^{-1}_{ij} h^j(t, g_t(x)) \right) \mathrm{d}x\mathrm{d}t = \int_0^T p(t, g_t(x))(\nabla g_t(x))^{-1}_{ij} h^j(t, g_t(x))\Big|_{x=0}^{x=2\pi} \mathrm{d}t = 0, \tag{3.46}$$

because we are considering periodic boundary conditions in x. By equation (3.41), the second term on the right-hand side of (3.45) vanishes. Therefore the integration of equation (3.45) produces

$$\int_0^T \int p(t, g_t)(\nabla g_t)^{-1}_{ij} \partial_i (h^j(t, g_t)) \det \nabla g_t(x) \, \mathrm{d}x\mathrm{d}t = -\int_0^T \int \partial_i (p(t, g_t))(\nabla g_t)^{-1}_{ij} h^j(t, g_t) \det \nabla g_t(x) \, \mathrm{d}x\mathrm{d}t. \tag{3.47}$$

We can finally compute the variation of $S^2(g_t, p)$, which is

$$\begin{aligned}
\delta S^2(g_t, p) &= \frac{\mathrm{d}}{\mathrm{d}\varepsilon}\Big|_{\varepsilon=0} \mathrm{E} \int_0^T \int p^\varepsilon(t, g_t^\varepsilon(x))(\det \nabla g_t^\varepsilon(x) - 1) \, \mathrm{d}x\mathrm{d}t \\
&= \frac{\mathrm{d}}{\mathrm{d}\varepsilon}\Big|_{\varepsilon=0} \mathrm{E} \int_0^T \int (p(t, g_t + \varepsilon h) + \varepsilon\varphi(t, g_t + \varepsilon h))(\det \nabla g_t(x)^\varepsilon - 1) \, \mathrm{d}x\mathrm{d}t \\
&= \mathrm{E} \int_0^T \int \frac{\mathrm{d}}{\mathrm{d}\varepsilon}\Big|_{\varepsilon=0} p(t, g_t + \varepsilon h)(\det \nabla g_t^\varepsilon(x) - 1) \, \mathrm{d}x\mathrm{d}t \\
&\quad + \mathrm{E} \int_0^T \int \frac{\mathrm{d}}{\mathrm{d}\varepsilon}\Big|_{\varepsilon=0} \varepsilon\varphi(t, g_t + \varepsilon h)(\det \nabla g_t^\varepsilon(x) - 1) \, \mathrm{d}x\mathrm{d}t \\
&\quad + \mathrm{E} \int_0^T \int p(t, g_t(x)) \frac{\mathrm{d}}{\mathrm{d}\varepsilon}\Big|_{\varepsilon=0} \det \nabla (g_t(x)^\varepsilon - 1) \, \mathrm{d}x\mathrm{d}t \\
&= \mathrm{E} \int_0^T \int (\nabla p \cdot h)(\det \nabla g_t(x) - 1) \, \mathrm{d}x\mathrm{d}t \\
&\quad + \mathrm{E} \int_0^T \int \varphi(t, g_t(x))(\det \nabla g_t(x) - 1) \, \mathrm{d}x\mathrm{d}t \\
&\quad + \mathrm{E} \int_0^T \int p(t, g_t(x)) \frac{\mathrm{d}}{\mathrm{d}\varepsilon}\Big|_{\varepsilon=0} \det \nabla (g_t(x) + \varepsilon h(t, g_t(x))) \, \mathrm{d}x\mathrm{d}t \qquad (3.48) \\
&= \mathrm{E} \int_0^T \int p(t, g_t(x)) \frac{\mathrm{d}}{\mathrm{d}\varepsilon}\Big|_{\varepsilon=0} \det \nabla (g_t(x) + \varepsilon h(t, g_t(x))) \, \mathrm{d}x\mathrm{d}t \qquad (3.49) \\
&= \mathrm{E} \int_0^T \int p(t, g_t(x))(\nabla g_t(x))^{-1}_{ij} \partial_i (h^j(t, g_t(x))) \det \nabla g_t(x) \, \mathrm{d}x\mathrm{d}t \qquad (3.50) \\
&= -\mathrm{E} \int_0^T \int \partial_i (p(t, g_t))(\nabla g_t(x))^{-1}_{ij} h^j(t, g_t) \det \nabla g_t(x) \, \mathrm{d}x\mathrm{d}t \qquad (3.51) \\
&= -\mathrm{E} \int_0^T \int \nabla p(t, g_t) \cdot h(t, g_t) \, \mathrm{d}x\mathrm{d}t. \qquad (3.52)
\end{aligned}$$

We now justify equalities (3.49), (3.50) and (3.51). Since φ is arbitrary we conclude from the second expectation in (3.48) that the critical points of the action verify $\det \nabla g_t(x) - 1 = 0$, so this expectation is zero, which also implies that the first expectation of the same equality is zero. This justifies equality (3.49). Equality (3.50) is justified by computation (3.44), and equality (3.51) by equation (3.47).

Adding finally equations (3.33) and (3.52) we obtain

$$\delta S = \delta S^1 + \delta S^2 = -\,\mathrm{E} \int_0^T \int ((\partial_t u + u \cdot \nabla u - \nu \Delta u + \nabla p) \cdot h)(t, g_t(x))\, \mathrm{d}x \mathrm{d}t = 0, \tag{3.53}$$

since part of the integrand is the Navier–Stokes equation, which we assumed u verifies. The other implication, which states that $\delta S = 0$ implies that u satisfies the Navier–Stokes equation, follows immediately.

QED

3.3 Diffusions on the Multidimensional Half-Space With Reflection at the Boundary

In this last section of this work, our goal is to study diffusions of the type of equation (3.8) on a d-dimensional domain D with normal reflection at the boundary. As we explained in section 2.6, when we introduced the Skorokhod problem and stochastic differential equations with reflection, due to the difficulty of the study of these mathematical objects when we consider non-regular domains (i.e. domains with non-smooth boundaries) or oblique reflections, we will not deal with such cases in this dissertation.

In the previous section we assumed that the Navier–Stokes equations were defined in the domain $[0, T] \times \mathbb{T}$, where $\mathbb{T}$ is the d-dimensional flat torus. In this section we consider the Navier–Stokes equations with Neumann boundary conditions in the domain $[0, T] \times \mathbb{R}^d_+$, or more specifically, we consider the system (1.72), taking $\rho = 1$ and $f = 0$. We will see how the results expounded in section 3.2 change when we consider the domain $D = \mathbb{R}^d_+$, and take the SDE (3.8) and write it in the form of the SDE (2.80) with normal reflections at the boundary $\partial \mathbb{R}^d_+ = \mathbb{R}^{d-1}_+ \times \{0\}$. The former equation becomes

$$\mathrm{d}g_t(x) = u(t, g_t(x))\, \mathrm{d}t + \sqrt{2\nu} \sum_{k=1}^{d} \mathrm{d}W_k(t) + n\, \mathrm{d}l(t), \tag{3.54}$$

where $g_t(\cdot)$ are diffeomorphisms (representing diffusions on the half-space) such that $g_0(x) = x$, $n = (0, \ldots, 0, 1)$, and l is a non-decreasing function verifying $l(0) = 0$ and the other conditions given in definition 2.25.

Recall that according to definition 2.26, a pair of processes $(g_t(x), l(t))$ is a weak solution of equation (3.54) if u is measurable and all the conditions of the Skorokhod problem are a.s. satisfied for $t \in \mathbb{R}_+$. Those conditions are the following: $g_t(x) \in \mathbb{R}^d_+$, l is a non-decreasing function such that $l(0) = 0$ and

$$l(t) = \int_0^t \mathbb{1}_{\partial \mathbb{R}^d_+}(g_s(x))\, \mathrm{d}l(s), \tag{3.55}$$

and

$$g_t(x) = g_0(x) + \int_0^t u(s, g_s(x))\,\mathrm{d}s + \sqrt{2\nu}\sum_{k=1}^{d}\int_0^t \mathrm{d}W_k(s) + \int_0^t n\,\mathrm{d}l(s) \qquad \text{a.s.,} \tag{3.56}$$

assuming all the integrals are well-defined. We write equation (3.56) in terms of its d coordinates as

$$\begin{aligned}
g_t^1(x) &= g_0^1(x) + \int_0^t u^1(s, g_s(x))\,\mathrm{d}s + \sqrt{2\nu}\sum_{k=1}^{d}\int_0^t \mathrm{d}W_k(s) \qquad \text{a.s.} \\
&\vdots \\
g_t^{d-1}(x) &= g_0^{d-1}(x) + \int_0^t u^{d-1}(s, g_s(x))\,\mathrm{d}s + \sqrt{2\nu}\sum_{k=1}^{d}\int_0^t \mathrm{d}W_k(s) \qquad \text{a.s.} \\
g_t^{d}(x) &= g_0^{d}(x) + \int_0^t u^{d}(s, g_s(x))\,\mathrm{d}s + \sqrt{2\nu}\sum_{k=1}^{d}\int_0^t \mathrm{d}W_k(s) + l(t) \qquad \text{a.s.}
\end{aligned} \tag{3.57}$$

We know that a solution of (3.54) is also a solution of the Skorokhod problem (definition 2.25) if we write, in terms of the notation used in that definition,

$$f(t) = g_0(x) + \int_0^t u(s, g_s(x))\,\mathrm{d}s + \sqrt{2\nu}\sum_{k=1}^{d}\int_0^t \mathrm{d}W_k(s) \qquad \text{a.s.,} \tag{3.58}$$

thus replacing condition (3.56) with the condition

$$g_t(x) = f(t) + nl(t). \tag{3.59}$$

Using the multidimensional Skorokhod map given by equation (2.89), we write

$$g_t(x) = (\Gamma f)(t) = (f^1(t), \ldots, f^{d-1}(t), (\Gamma f^d)(t)), \tag{3.60}$$

where $g_t^d(x) = (\Gamma f^d)(t)$ is the unidimensional Skorokhod map. If we write equation (3.58) in terms of the Skorokhod map, then we obtain

$$f(t) = g_0(x) + \int_0^t u(s, (\Gamma f)(s))\,\mathrm{d}s + \sqrt{2\nu}\sum_{k=1}^{d}\int_0^t \mathrm{d}W_k(s) \qquad \text{a.s.} \tag{3.61}$$

Since Γ is Lipschitz continuous, applying theorem 2.3 we conclude that equation (3.56) has a unique solution, and so does (3.54).

Proceeding now as in the previous section, we calculate the generalized derivative of (3.54) obtaining

$$\begin{aligned}
\mathrm{D}_t g_t(x) &= \lim_{\varepsilon\to 0}\frac{1}{\varepsilon}\,\mathrm{E}\left[g_{t+\varepsilon}(x) - g_t(x)|\mathcal{F}_t\right] \\
&= \lim_{\varepsilon\to 0}\frac{1}{\varepsilon}\,\mathrm{E}\left[\int_t^{t+\varepsilon} u(s, g_s(x))\,\mathrm{d}s + \sqrt{2\nu}\sum_{k=1}^{d}\int_t^{t+\varepsilon}\mathrm{d}W_k(s) + \int_t^{t+\varepsilon} n\,\mathrm{d}l(s)\,\middle|\,\mathcal{F}_s\right]
\end{aligned}$$

$$= \lim_{\varepsilon\to 0}\frac{1}{\varepsilon}\,\mathrm{E}\left[\int_t^{t+\varepsilon} u(s,g_s(x))\,\mathrm{d}s \;\middle|\; \mathcal{F}_s\right] + \sqrt{2\nu}\sum_{k=1}^{d}\lim_{\varepsilon\to 0}\frac{1}{\varepsilon}\,\mathrm{E}\left[\int_t^{t+\varepsilon}\mathrm{d}W_k(s) \;\middle|\; \mathcal{F}_s\right]$$
$$+ n\lim_{\varepsilon\to 0}\frac{1}{\varepsilon}\,\mathrm{E}\left[\int_t^{t+\varepsilon}\mathrm{d}l(s) \;\middle|\; \mathcal{F}_s\right] \tag{3.62}$$
$$= u(t,g_t(x)) + nl(t) \qquad \text{a.s.}, \tag{3.63}$$

where the second limit on the right-hand side of equation (3.62) is zero due to equation 3) in proposition 2.4 (see p. 36). Using the Itô formula in d dimensions we obtain the stochastic differential of $\mathrm{D}_t g_t(x)$, which is

$$\begin{aligned}
\mathrm{d}(u^i(t,g_t(x)) + nl(t)) &= \partial_t(u^i + nl(t))\,\mathrm{d}t + \sum_{j=1}^{d}\partial_j(u^i + nl(t))\,\mathrm{d}g_t^j(x) + \frac{1}{2}\sum_{i=1}^{d}\sum_{j=1}^{d}\partial_i\partial_j(u^i + nl(t))2\nu\,\mathrm{d}t \\
&= \left(\partial_t u^i + n\partial_t l(t) + \nu\sum_{i=1}^{d}\sum_{j=1}^{d}\partial_i\partial_j u^i\right)\mathrm{d}t \\
&\quad + \sum_{j=1}^{d}\partial_j u^i\left(u^j(t,g_t(x))\,\mathrm{d}t + \sqrt{2\nu}\sum_{k=1}^{d}\mathrm{d}W_k^j(t) + n\,\mathrm{d}l(t)\right) \\
&= \left(\partial_t u^i + n\partial_t l(t) + \sum_{j=1}^{d}u^j\partial_j u^i + \nu\sum_{i=1}^{d}\sum_{j=1}^{d}\partial_i\partial_j u^i\right)\mathrm{d}t + \sqrt{2\nu}\sum_{j=1}^{d}\sum_{k=1}^{d}\partial_j u^i\mathrm{d}W_k^j(t) \\
&\quad + n\sum_{j=1}^{d}\partial_j u^i\,\mathrm{d}l(t) \\
&= (\partial_t u^i + n\partial_t l(t) + u^i\cdot\nabla u^i + \nu\Delta u^i)\,\mathrm{d}t + \sqrt{2\nu}\sum_{k=1}^{d}\nabla u\,\mathrm{d}W_k(t) + n\nabla u\,\mathrm{d}l(t),
\end{aligned} \tag{3.64}$$

for $i = 1,\ldots,d$. We can simplify this result by observing that $\partial_t l(t) = 0$, since $l(t)$ is constant on each interval in the complement of the closed set $\{t\in\mathbb{R}_+ \colon W_t(x) = x\}$, a set we had already encountered on page 47 of this dissertation (the source of this result is [20, p. 202]). Therefore we conclude that the second generalized derivative of $g_t(x)$ is

$$\mathrm{D}_t\mathrm{D}_t g_t(x) = (\partial_t u + u\cdot\nabla u + \nu\Delta u)(t,g_t(x)) + n\nabla u\,l(t). \tag{3.65}$$

We consider again the definition of the action functional (3.13) and the functional variations of g_t and p in (3.18) and (3.19), respectively, in whose definitions the smooth functions h and φ are used. We assume, as before, that $h(T,\cdot) = h(0,\cdot) = 0$.

Using the Itô formula we compute the stochastic differential of h, as we did in equation (3.26), obtaining

$$\begin{aligned}
\mathrm{d}h^i(t,g_t(x)) &= \partial_t h^i\,\mathrm{d}t + \sum_{j=1}^{d}\partial_j h^i\,\mathrm{d}g_t^j(x) + \frac{1}{2}\sum_{i=1}^{d}\sum_{j=1}^{d}\partial_i\partial_j h^i 2\nu\,\mathrm{d}t \\
&= \partial_t h^i\,\mathrm{d}t + \sum_{j=1}^{d}\partial_j h^i\left(u^j(t,g_t(x))\,\mathrm{d}t + \sqrt{2\nu}\sum_{k=1}^{d}\mathrm{d}W_k^j(t) + n\,\mathrm{d}l(t)\right) + \nu\sum_{i=1}^{d}\sum_{j=1}^{d}\partial_i\partial_j h^i\,\mathrm{d}t
\end{aligned}$$

$$= \sqrt{2\nu} \sum_{j=1}^{d} \sum_{k=1}^{d} \partial_j h^i \, \mathrm{d}W_k^j(t) + \left(\partial_t h^i + \sum_{j=1}^{d} u^j \partial_j h^i + \nu \sum_{i=1}^{d} \sum_{j=1}^{d} \partial_i \partial_j h^i \right) \mathrm{d}t$$
$$+ n \sum_{j=1}^{d} \partial_j h^i \, \mathrm{d}l(t), \tag{3.66}$$

for $i = 1, \ldots, d$, that is,

$$\mathrm{d}h(t, g_t(x)) = \sqrt{2\nu} \sum_{k=1}^{d} \nabla h \, \mathrm{d}W_k(t) + (\partial_t h + u \cdot \nabla h + \nu \Delta h)\mathrm{d}t + n \nabla h \, \mathrm{d}l(t). \tag{3.67}$$

The Itô contraction of $\mathrm{d}\mathrm{D}_t g_t(x)$ and $\mathrm{d}h(t, g_t(x))$ is equal to the Itô contraction we obtained in equation (3.28), since $\mathrm{d}t \, \mathrm{d}l(t) = 0$ and $\mathrm{d}l(t) \, \mathrm{d}l(t) = 0$. We can justify these contractions by observing that when we apply the Itô formula to the function $f(x) = x^2$, we obtain the same result, regardless of the presence of the term $l(t)$ (corresponding to the local time). This result is used, for instance, in the proof of the existence and uniqueness of solutions of the reflecting SDE (2.87), which is implied by the validity of the Lipschitz conditions (2.84) and (2.85). See [32, pp. 6–7] for the proof of this result (the application of the Itô formula to the said function is on page 7).

The above observations on the calculus of the differential of the local time justify that we make a slight detour from the main thrust of this section, and briefly mention the problem of integration with respect to local time. Local time depends on a time and on a space variable, and integration can be defined with respect to the two variables or with respect to each one individually (keeping the other fixed). Although we had not mentioned this fact before, it was clear from the notation used for the local time that the integrals with respect to it that made an appearance in this dissertation were only with respect to the variable t. In [7], the following adaptation of the Itô formula is established:

$$f(t, W_t) = f(0, W_0) + \int_0^t \frac{\partial f}{\partial t}(s, W_s) \, \mathrm{d}s + \int_0^t \frac{\partial f}{\partial x}(s, W_s) \, \mathrm{d}s - \frac{1}{2} \int_0^t \int_{\mathbb{R}} \frac{\partial f}{\partial x}(s, W_s) \, \mathrm{d}L_s(x), \tag{3.68}$$

where W_t is the Wiener process, $L_t(x)$ is the local time of the Wiener process and f is a measurable function defined on $\mathbb{R}_+ \times \mathbb{R}$, whose derivatives are continuous. The differential $\mathrm{d}L_t(x)$ refers to an area integration with respect to the map $(t, x) \mapsto L_t(x, \omega)$, which we know to be continuous according to the definition of local time (see page 46). The last integral on the right-hand side of equation (3.68) is sometimes called in the literature the *local time-space integral*, and if its integrand H is a function of class C^2, then the following integration-by-parts formula is valid:

$$\int_0^t \int_{\mathbb{R}} H(s, x) \, \mathrm{d}L_s(x) = \int_{\mathbb{R}} H(t, x) \, \mathrm{d}L_s(x) - \int_0^t \left(\int_{\mathbb{R}} \partial_t H(s, x) \, \mathrm{d}L_s(x) \right) \mathrm{d}s, \tag{3.69}$$

where both integrals with respect to the local time on the right-hand side are integrals with respect to the x variable only (the proof can be found in [11, p. 187]). Formula (3.68) is valid not only for the Wiener process, but for any continuous semimartingale (according to [11, pp. 181–182]), and can be used to

prove several other extensions of the Itô formula (a comprehensive list can be seen in [11, pp. 180–182], and the proofs of some of the formulas can be found in [11, pp. 189–192]). Observe that equation (3.67) is a consequence of equation (3.68) in the d-dimensional form.

We continue now the adaptation of the results of the previous section. Using equation (3.65), and omitting the previously calculated steps on page 57, we obtain the new computation of the variation of $S^1(g_t)$, which is

$$\delta S^1(g_t) = -\operatorname{E} \int_0^T \int [(\partial_t u + u \cdot \nabla u - \nu \Delta u) \cdot h + n \nabla u \, l(t) \cdot h] \, \mathrm{d}x \mathrm{d}t. \tag{3.70}$$

This variation is equal to the variation (3.33) for the first $d-1$ coordinates of g_t (for the coordinate d, we take $n = 1$ in the equation). The variation of $S^2(g_t, p)$ is the one obtained in computation (3.52), provided we make the same assumptions we made then. Therefore we obtain

$$\delta S = \delta S^1 + \delta S^2 = -\operatorname{E} \int_0^T \int [(\partial_t u + u \cdot \nabla u - \nu \Delta u + \nabla p) \cdot h + n \nabla u \, l(t) \cdot h] \, \mathrm{d}x \mathrm{d}t. \tag{3.71}$$

This variation would be zero (as in theorem 3.1) if and only if u satisfied the Navier–Stokes equations and

$$\nabla u \cdot \boldsymbol{n} \, l(t) = 0 \qquad \text{on } [0,T] \times \partial \mathbb{R}^d_+. \tag{3.72}$$

Since we are assuming that u verifies the Navier–Stokes equations with the Neumann boundary conditions, which tells us, in particular, that

$$\nabla u \cdot \boldsymbol{n} = 0 \qquad \text{on } [0,T] \times \partial \mathbb{R}^d_+, \tag{3.73}$$

we conclude that variation (3.71) is indeed zero. From these results we deduce that theorem 3.1 is valid when we consider diffusions with reflection on the domain $\mathbb{R}^d_+$, provided that we assume that the Navier–Stokes equations satisfy the Neumann boundary conditions.

Bibliography

[1] J. D. Anderson, Jr. *A History of Aerodynamics and Its Impact on Flying Machines*. Cambridge University Press, 1997.

[2] J. Bair, P. Błaszczyk, R. Ely, et al. Is mathematical history written by the victors? *Notices of the American Mathematical Society*, 60(7):886–904, 2013.

[3] G. Batchelor. *An Introduction to Fluid Dynamics*. Oxford University Press, 1967.

[4] P. Cartier and C. DeWitt-Morette. Functional integration. *J. Math. Phys.*, 41(6):4154–4187, 2000.

[5] A. B. Cruzeiro. Navier–Stokes and stochastic Navier–Stokes equations via Lagrange multipliers. To appear in *J. Geom. Mechanics*.

[6] T. G. Dankel, Jr. Mechanics on manifolds and the incorporation of spin into Nelson's stochastic mechanics. *Arch. Rational Mech. Anal.*, 37:192–221, 1970.

[7] N. Eisenbaum. Integration with respect to local time. *Potential Anal.*, 13:303–328, 2000.

[8] L. C. Evans. *Partial Differential Equations*. American Mathematical Society, 1998.

[9] L. C. Evans. *An Introduction to Stochastic Differential Equations*. American Mathematical Society, 2013.

[10] C. L. Fefferman. Existence and smoothness of the Navier–Stokes equation. In J. Carlson, A. Jaffe, and A. Wiles, editors, *The Millennium Prize Problems*. American Mathematical Society, 2006.

[11] R. Ghomrasni and G. Peskir. Local time-space calculus and extensions of Itô's formula. In J. Hoffmann-Jörgensen, J. Wellner, and M. Marcus, editors, *High Dimensional Probability III*. Birkhäuser Basel, 2003.

[12] Y. Giga. Weak and strong solutions of the Navier–Stokes initial value problem. *Publ. Res. Inst. Math. Sci.*, 19:887–910, 1983.

[13] L. Godinho and J. Natário. *An Introduction to Riemannian Geometry: With Applications to Mechanics and Relativity*. Springer-Verlag, 2014.

[14] R. A. Granger. *Fluid Mechanics*. Dover Publications, Inc., 1995.

[15] N. G. Hadjiconstantinou. The limits of Navier–Stokes theory and kinetic extensions for describing small-scale gaseous hydrodynamics. *Phys. Fluids*, 18:111301–1–111301–19, 2006.

[16] M. Hazewinkel, editor. *Encyclopaedia of Mathematics*. Kluwer Academic Publishers, 1993.

[17] N. Ikeda and S. Watanabe. *Stochastic Differential Equations and Diffusion Processes*. North-Holland Publishing Company, second edition, 1989.

[18] A. Inoue and T. Funaki. On a new derivation of the Navier–Stokes equation. *Commun. Math. Phys.*, 65:83–90, 1979.

[19] K. Itô, editor. *Encyclopedic Dictionary of Mathematics*. The MIT Press, second English edition, 1993.

[20] I. Karatzas and S. E. Shreve. *Brownian Motion and Stochastic Calculus*. Springer-Verlag, second edition, 1991.

[21] K. U. Katz and M. G. Katz. A Burgessian critique of nominalistic tendencies in contemporary mathematics and its historiography. *Found. Sci.*, 17:51–89, 2012.

[22] J. Keller and D. McLaughlin. The Feynman integral. *The American Mathematical Monthly*, 82(5):451–465, 1975.

[23] I. M. Koval'chik. The Wiener integral. *Russ. Math. Surv.*, 18:97–134, 1963.

[24] L. Landau and E. Lifshitz. *Fluid Mechanics*. Number 6 in Course of Theoretical Physics. Pergamon Press, third revised English edition, 1959.

[25] G. Łukaszewicz and P. Kalita. *Navier–Stokes Equations: An Introduction with Applications*. Springer-Verlag, 2016.

[26] J. R. Munkres. *Topology*. Prentice-Hall, Inc., second edition, 2000.

[27] T. Nakagomi, K. Yasue, and J. Zambrini. Stochastic variational derivations of the Navier–Stokes equation. *Lett. Math. Phys.*, 5:545–552, 1981.

[28] R. Paley, N. Wiener, and A. Zygmund. Notes on random functions. *Math. Z.*, 37:647–668, 1933. Reprinted in [46, pp. 536–557].

[29] M. Pavon. Stochastic mechanics and the Feynman integral. *J. Math. Phys.*, 41(9):6060–6078, 2000.

[30] P. Pearle, B. Collett, K. Bart, et al. What Brown saw and you can too. *Am. J. Phys.*, 78(12):1278–1289, 2010.

[31] R. Penrose. *The Road to Reality: A Complete Guide to the Laws of the Universe*. Jonathan Cape, 2004.

[32] A. Pilipenko. *An Introduction to Stochastic Differential Equations with Reflection*. Potsdam University Press, 2014.

[33] D. Revuz and M. Yor. *Continuous Martingales and Brownian Motion*. Springer-Verlag, 1999.

[34] M. Ricou. Medida e Integração. Unpublished book draft, Departamento de Matemática, Instituto Superior Técnico, 2013.

[35] L. Rogers and D. Williams. *Itô Calculus*, volume 2 of *Diffusions, Markov Processes, and Martingales*. Cambridge University Press, second edition, 2000.

[36] Y. Saisho. Stochastic differential equations for multi-dimensional domain with reflecting boundary. *Probab. Th. Rel. Fields*, 74:455–477, 1987.

[37] R. Shankar. Symmetries and conservation laws of the Euler equations in Lagrangian coordinates. *J. Math. Anal. Appl.*, 447:867–881, 2017.

[38] G. Simmons. *Introduction to Topology and Modern Analysis*. McGraw-Hill Book Company, Inc., and Kōgakusha Co., Ltd., international student edition, 1963.

[39] R. Situ. *Theory of Stochastic Differential Equations with Jumps and Applications*. Springer-Verlag, 2005.

[40] W. A. Strauss. *Partial Differential Equations: An Introduction*. John Wiley & Sons, Inc., 2008.

[41] D. W. Stroock and S. Varadhan. *Multidimensional Diffusion Processes*. Springer-Verlag, 1997.

[42] J. von Neumann. *Mathematische Grundlagen der Quantenmechanik*. Springer-Verlag, second edition, 1996. Reprint of the 1932 edition.

[43] N. Wiener. Differential-space. *J. Math. and Phys. M.I.T.*, 2:131–174, 1923. Reprinted in [46, pp. 455–498].

[44] N. Wiener. The average value of a functional. *Proc. London Math. Soc.*, 22:454–467, 1924. Reprinted in [46, pp. 499–512].

[45] N. Wiener. Generalized harmonic analysis. *Acta Math.*, 55:117–258, 1930.

[46] N. Wiener. *Collected Works*, volume I. The M.I.T. Press, 1976.

[47] N. Wiener, A. Siegel, B. Rankin, and W. T. Martin. *Differential Space, Quantum Systems, and Prediction*. The M.I.T. Press, 1966.

[48] K. Yasue. A variational principle for the Navier–Stokes equation. *J. Funct. Anal.*, 51:133–141, 1983.

[49] B. Yavorsky and A. Detlaf. *Handbook of Physics*. Mir Publishers, 1975.

[50] W. Yourgrau and S. Mandelstam. *Variational Principles in Dynamics and Quantum Theory*. Pitman Publishing Corporation, second edition, 1960.

Appendix A

A.1 Some Notions of Measure Theory

Definition A.1. Let X be a set. A *σ-algebra* is a non-empty collection $\mathcal{M}$ of subsets of X such that $X \in \mathcal{M}$, and that $\mathcal{M}$ is closed for the set-theoretic operations of complementation and countable union.

Definition A.2. A *measure* is a function $\mu\colon \mathcal{M} \to Y$, for $Y = \mathbb{R}$ (or some subset of $\mathbb{R}$) or $Y = \mathbb{C}$, such that $\mu(\emptyset) = 0$ and for any disjoint sets $A_n \in \mathcal{M}$,

$$\mu\left(\bigcup_{n=1}^{\infty} A_n \in \mathcal{M}\right) = \sum_{n=1}^{\infty} \mu(A_n). \tag{A.1}$$

A function that verifies this equality is called a *σ-additive* function.

Definition A.3. We call the tuple $(X, \mathcal{M})$ a *measurable space*, and the triple $(X, \mathcal{M}, \mu)$ a *measure space*.

Definition A.4. A *probability measure* is a measure that verifies $\mu(X) = 1$.

Definition A.5. A *probability space* is a measure space, denoted henceforth by $(\Omega, \mathcal{F}, P)$, where P is a probability measure. We call also $(\Omega, \mathcal{F})$ the *sample space*, Ω the *state space*, each $\omega \in \Omega$ a *sample point*, and each $A \in \mathcal{F}$ an *event*.

Definition A.6. The *probability* of an event A corresponds to $P(A)$, which we sometimes denote by $\Pr(A)$.

Proposition A.1. If $\{\mathcal{M}_\alpha\colon \alpha \in J\}$ is a non-empty set of σ-algebras in X (J is the index set), then the collection

$$\mathcal{M} = \bigcap_{\alpha \in J} \mathcal{M}_\alpha \tag{A.2}$$

is a σ-algebra in X.

Proof. See [34, p. 113]. QED

If C is a collection of subsets of X, then, by this proposition, the intersection of all σ-algebras that contain C is the smallest σ-algebra in X that contains C. This fact motivates the following definition.

Definition A.7. If C is a collection of subsets of X, then we call the intersection of all σ-algebras in X that contain the collection C the *σ-algebra generated by C*.

Definition A.8. The *Borel σ-algebra*, which we denote by $\mathcal{B}(\mathbb{R}^N)$, is the σ-algebra generated by the open sets of $\mathbb{R}^N$. The sets belonging to $\mathcal{B}(\mathbb{R}^N)$ are called *Borel sets* or *Borel-measurable sets*.

Observe that any open (or closed) set is Borel-measurable. As an interesting aside, note that the cardinality of $\mathcal{B}(\mathbb{R}^N)$, denoted by $|\mathcal{B}(\mathbb{R}^N)|$, is the same as the cardinality of the continuum, that is, $2^{\aleph_0}$, where $\aleph_0$ is the cardinality of $\mathbb{N}$. Note also that the cardinality of the power set of $\mathbb{R}^N$, denoted by $|\mathcal{P}(\mathbb{R}^N)|$, which has actually the same cardinality as that of $\mathcal{P}(\mathbb{R})$ (since $|\mathbb{R}| = |\mathbb{R}^N|$), is strictly larger than the cardinality of the continuum.

Definition A.9. If $\mathcal{M} = \mathcal{B}(\mathbb{R}^N)$, then μ is a *Borel measure*.

Definition A.10. Let $(X, \mathcal{M})$ and $(Y, \mathcal{N})$ be measurable spaces. The *product of σ-algebras $\mathcal{M}$* and $\mathcal{N}$, which we denote by $\mathcal{M} \otimes \mathcal{N}$, is the σ-algebra generated in $X \times Y$ by the sets $A \times B$, where $A \in \mathcal{M}$ and $B \in \mathcal{N}$.

Definition A.11. Let $(X, \mathcal{M})$ and $(Y, \mathcal{N})$ be measurable spaces. The map $f\colon X \to Y$ is *measurable* if $f^{-1}(B) \in \mathcal{M}$ for every $B \in \mathcal{N}$.

Definition A.12. Let $E \subseteq S \subseteq X$ and $f\colon X \to \overline{\mathbb{R}}$. We call f *$\mathcal{M}$-measurable* in E if and only if

$$\{(x, y) \in X \times \mathbb{R}\colon x \in E, 0 < y < f(x), 0 > y > f(x)\} \in \mu \otimes \mathcal{B}(\mathbb{R}). \tag{A.3}$$

We say that f is a *Borel-measurable function* in a set E if and only if the set in expression (A.3) is Borel-measurable in $\mathbb{R}^{N+1}$.

Definition A.13. Let U be an *N-rectangle*, i.e. let it be defined as $U = I_1 \times I_2 \times \cdots \times I_N$, where $I_i =]a_i, b_i[$ are intervals in $\mathbb{R}$ (each I_i can also be closed, or neither open nor closed). We define the *length*, or *content*, of I_i as $c_1(I_i) = b_i - a_i$, so the content of U is $c_N(U) = c_1(I_1) \times c_1(I_2) \times \cdots \times c_1(I_N)$.

We denote by $\mathcal{E}(\mathbb{R}^N)$ the class of sets that are countable unions of sets which in turn are finite unions of limited rectangles.

Definition A.14. The *exterior Lebesgue measure* in $\mathbb{R}^N$ is the function $m^*\colon \mathcal{P}(\mathbb{R}^N) \to \mathbb{R}_+$ given by

$$m^*(E) = \inf\{c_N(U)\colon E \subseteq U, U \in \mathcal{E}(\mathbb{R}^N)\}. \tag{A.4}$$

Definition A.15. A set $E \subseteq \mathbb{R}^N$ is *Lebesgue-measurable* in $\mathbb{R}^N$ if and only if

$$m^*(F) = m^*(F \cap E) + m^*(F \setminus E), \tag{A.5}$$

for any $F \subseteq \mathbb{R}^N$.

We denote by $\mathcal{L}(\mathbb{R}^N)$ the class of the Lebesgue-measurable sets in $\mathbb{R}^N$.

Definition A.16. The *Lebesgue measure* $m\colon \mathcal{L}(\mathbb{R}^N) \to \mathbb{R}_+$ is the restriction of m^* to $\mathcal{L}(\mathbb{R}^N)$.

Definition A.17. We say that a function $f\colon [0,T] \to \mathbb{R}$ is *Hölder continuous* at the point $s \in [0,T]$ with exponent $\gamma \in]0,1]$ if there exists a constant c such that for all $t \in [0,T]$,

$$|f(t) - f(s)| \leq c|t-s|^\gamma. \tag{A.6}$$

A.2 A Brief Note on the Lebesgue–Stieltjes Integral

The Lebesgue–Stieltjes integral is used in this work so we write a few words about it and about the broader problematic of the development of classical integration theory. Recall that the *Riemann integral* of a bounded function $f\colon [a,b] \to \mathbb{R}$ is defined, when the limit exists, as

$$\int_a^b f(x)\,\mathrm{d}x := \lim_{\|\mathcal{P}\|\to 0} \sum_{k=1}^{n} f(x_k^*)(x_k - x_{k-1}), \qquad x_{k-1} \leq x_k^* \leq x_k, \tag{A.7}$$

where $\mathcal{P}$ is a partition of $[a,b]$ and $\|\mathcal{P}\| := \max_{1\leq k\leq n}(t_k - t_{k-1})$. Thomas Stieltjes (1856–1894) generalized this concept of integral by defining an integral with respect to an arbitrary real-valued, bounded function F. This integral is called the *Riemann-Stieltjes integral*, or simply the *Stieltjes integral*, and is defined as follows:

$$\int_a^b f(x)\,\mathrm{d}F(x) := \lim_{\|\mathcal{P}\|\to 0} \sum_{k=1}^{n} f(x_k^*)(F(x_k) - F(x_{k-1})), \qquad x_{k-1} \leq x_k^* \leq x_k, \tag{A.8}$$

when the limit exists. Observe that the Riemann integral is obtained from the Riemann–Stieltjes integral by choosing $F(x) = x$.

Definition A.18. If $f\colon S \subseteq \mathbb{R} \to \mathbb{R}$ and $I \subseteq S$, then we denote by $\mathcal{V}_f(I)$ the *total variation* of f in I and define it as

$$\mathcal{V}_f(I) := \sup\left\{ \sum_{k=1}^{n} |f(x_k) - f(x_{k-1})|\colon x_0 < x_1 < \cdots < x_n, x_k \in I \right\}. \tag{A.9}$$

Definition A.19. Let f be a function and I and internal, both as in definition A.18. We say that f is of *bounded variation* in I if $\mathcal{V}_f(I) < +\infty$.

Definition A.20. Let f be a function and I and internal, both as in definition A.18. We say that f is of *finite variation* if it is of bounded variation on each bounded interval I.

The Riemann-Stieltjes integral is defined when f is continuous and F is of bounded variation, or when f is of bounded variation (not necessarily continuous) and F is continuous (not necessarily of bounded variation). The source of these statements is [19, s.v. 'Curvilinear Integrals and Surface Integrals', p. 354].

Definition A.21. If μ is a real measure in $(X, \mathcal{M})$ and $E \in \mathcal{M}$, then

$$\mu^+(E) := \sup\{\mu(A \cap E)\colon A \in \mathcal{M}\} \tag{A.10}$$

is the *positive variation* of μ, and

$$\mu^-(E) := -\inf\{\mu(A \cap E) \colon A \in \mathcal{M}\} \tag{A.11}$$

is the *negative variation* of μ.

Definition A.22. If μ is a measure defined in $\mathcal{M}$ and $S \in \mathcal{M}$, then μ is *concentrated* in S if and only if $\mu(E) = \mu(E \cap S)$ for all $E \in \mathcal{M}$.

Definition A.23. Let μ^+ and μ^- be two positive, finite measures concentrated in disjoint sets. If $\mu = \mu^+ - \mu^-$, then we call this the unique *Jordan decomposition* of μ.

Definition A.24. Let $\mu = \mu^+ - \mu^-$ be the Jordan decomposition of a real, positive measure, and let also $X = \mathbb{R}^N$. The *Lebesgue–Stieltjes integral* is

$$\int_X f \,\mathrm{d}\mu := \int_X f \,\mathrm{d}\mu^+ - \int_X f \,\mathrm{d}\mu^-, \tag{A.12}$$

if the integrals on the right-hand side of this equation exist.

The Lebesgue–Stieltjes integral is a generalization of the Lebesgue integral, which is defined only with respect to the Lebesgue measure. Therefore if μ is the Lebesgue measure, then the Lebesgue–Stieltjes integral is equivalent to the Lebesgue integral. Since the Lebesgue integral is also an extension of the Riemann integral, which makes any Riemann integral a Lebesgue integral, we can also see the Lebesgue–Stieltjes integral as a generalization of the Riemann-Stieltjes integral. Incidently, it is very important to be aware of the reason why the Lebesgue integral is technically superior to the Riemann integral, which causes it to be extensively used all across modern mathematics, such as in probability theory or in the theory of partial differential equations. It is mainly due to the result (usually called the *Lebesgue convergence theorem*) that tells us that

$$\lim_{n\to\infty} \int_a^b f_n \,\mathrm{d}m = \int_a^b \lim_{n\to\infty} f_n \,\mathrm{d}m, \tag{A.13}$$

where m is the Lebesgue measure, which holds under fairly non-restrictive assumptions (which we will not mention here, but can be seen, along with the proof, in [34, pp. 168–169]). This theorem is not valid for the Riemann integral.

Depending on the properties of F, several results are known. For instance, if $F(x)$ is a strictly monotone, increasing, continuous function and $G(y)$ is its inverse function, then

$$\int_a^b f(x) \,\mathrm{d}F(x) = \int_{F(a)}^{F(b)} f(G(y)) \,\mathrm{d}y, \tag{A.14}$$

where the left-hand side of the equality is a Lebesgue–Stieltjes integral and the right-hand side is a Lebesgue integral.

If $F(x)$ is absolutely continuous or differentiable, then

$$\int_a^b f(x)\,\mathrm{d}F(x) = \int_a^b f(x)F'(x)\,\mathrm{d}x, \tag{A.15}$$

where the integral on the right-hand side of the equality is a Lebesgue integral.

A.3 Stochastic Processes, Filtered Probability Spaces, Convergence and Conditional Expectation

Definition A.25. A real-valued *random variable* X is a function $X\colon \Omega \to \mathbb{R}$ that is $\mathcal{F}$-measurable.

We rarely write $X(\omega)$ to denote the image of $\omega \in \Omega$, and we usually refer to the random variable simply as X, not displaying the dependence on the sample point ω.

Definition A.26. A collection of random variables $X = \{X(t,\omega)\colon t \in \mathbb{R}_+\}$ is a *stochastic process*. We interpret the parameter t as indicating time.

We often write $X_t(\omega)$ instead of $X(t,\omega)$, and usually abbreviate this by $X(t)$, X_t or X. Notice then that X may denote a random variable or a stochastic process, but we always dispel this ambiguity by indicating which mathematical object the symbol is denoting when we are using it.

For a fixed point $\omega \in \Omega$ the map $t \mapsto X(t,\omega)$ is the corresponding *sample path*, that is, an observation of the value of the various random variables that belong to the stochastic process as time evolves. A stochastic process therefore provides a mathematical model for a random experiment observed continuously in time, such as the trajectory of a particle subjected to random disturbances. Observe that one sample path is one realization of a stochastic process, while one realization of a random variable is just one value $X(\omega)$, obtained by the map $\omega \mapsto X(t,\omega)$ for a fixed t.

Definition A.27. The *indicator function* $\mathbb{1}$ of f in the set A is

$$\mathbb{1}_A(f) = \begin{cases} 1 & \text{if } f \in A, \\ 0 & \text{if } f \notin A. \end{cases} \tag{A.16}$$

The mathematical object f in the above definition need not be a function. If we let $(\Omega, \mathcal{F})$ be a sample space, and take $A \in \mathcal{F}$ and $\omega \in \Omega$, then the indicator function

$$\mathbb{1}_A(\omega) = \begin{cases} 1 & \text{if } \omega \in A, \\ 0 & \text{if } \omega \notin A, \end{cases} \tag{A.17}$$

is actually a random variable.

We proceed by defining two types of convergence of sequences of random variables in probability spaces, which are used throughout this work.

Definition A.28. Let $\{X_n\}_{n=1}^{\infty}$ be a sequence of random variables, and let X be another random variable. We say that $\{X_n\}$ is *convergent in probability* to a random variable X if and only if

$$\lim_{n\to\infty} \Pr(|X_n - X| > \varepsilon) = 0, \tag{A.18}$$

for every $\varepsilon > 0$. In this case, we call X the *probability limit* of the sequence, and convergence is indicated by writing $X_n \xrightarrow{P} X$, or

$$\operatorname*{plim}_{n\to\infty} X_n = X. \tag{A.19}$$

Definition A.29. We say that a sequence of random variables $\{X_n\}_{n=1}^{\infty}$ *converges almost surely*, or *almost everywhere*, to the random variable X, which we often abbreviate by saying that it converges *a.s.*, and write $X_n \xrightarrow{a.s.} X$, if

$$\Pr(\{\omega\colon X_n(\omega) \longrightarrow X(\omega)\}) = 1, \tag{A.20}$$

i.e. the events for which X_n does not converge to X have probability zero.

In general, a property which is true except for an event of probability zero is also said to hold a.s. In the context of measure theory, we say that a statement holds for almost every a if the statement holds for all a except for those values of a belonging to a countable set. It is a well-known result that almost sure convergence implies convergence in probability.

Definition A.30. Let $(\Omega, \mathcal{F}, P)$ be a probability space. We call the family of σ-algebras $\{\mathcal{F}_t\colon t \geq 0\} \subseteq \mathcal{F}$ a *filtration* if $\mathcal{F}_s \subset \mathcal{F}_t$ for $0 \leq s < t < \infty$.

Informally, we can think of $\mathcal{F}_t$ as containing all the information on the process available at time t, i.e. a filtration describes the history of a process. Given a set A, we can interpret $A \in \mathcal{F}_t$ as saying that an observer of the process knowns whether A has occurred or not by time t.

Definition A.31. We call the quadruple $(\Omega, \mathcal{F}, \mathcal{F}_t, P)$ a *filtered probability space* when we have a fixed probability space with an associated filtration.

Definition A.32. A stochastic process X is said to be *adapted to the filtration* $\{\mathcal{F}_t\}$, or *non-anticipating*, if X_t is an $\mathcal{F}_t$-measurable random variable for each $t \in \mathbb{R}_+$, that is, if $X_t \in \mathcal{F}_t$. We frequently abbreviate this by saying that the process X is $\mathcal{F}_t$-adapted.

The process is $\mathcal{F}_t$-adapted if X_t is completely determined by the information available up to time t and there is no need to look into the future.

Definition A.33. A stochastic process X is *measurable* if $\{(t,\omega)\colon X_t(\omega) \in A\} \in \mathcal{B}(\mathbb{R}_+) \otimes \mathcal{F}$, for every $A \in \mathcal{B}(\mathbb{R}^N)$.

Definition A.34. Given a stochastic process X, the smallest σ-algebra with respect to which X_s is measurable for every $s \in [0,t]$, is the filtration generated by the process itself, which we denote by $\mathcal{F}_t^X$.

When we write $A \in \mathcal{F}_t^X$, we mean that an observer of X knows whether A has occurred or not by time t. Notice that every process X is $\mathcal{F}_t^X$-adapted.

Definition A.35. A filtration $\{\mathcal{F}_t\}$ is *right-continuous* if

$$\mathcal{F}_t = \mathcal{F}_{t^+} := \bigcap_{\varepsilon>0} \mathcal{F}_{t+\varepsilon}, \qquad t \in \mathbb{R}_+, \tag{A.21}$$

that is, if $\mathcal{F}_t$ is the σ-algebra of events immediately after $t \in \mathbb{R}_+$. Similarly, we say that the filtration if *left-continuous* if $\mathcal{F}_t = \mathcal{F}_{t^-}$ is the σ-algebra of events strictly prior to $t > 0$.

We can interpret the right-continuity of $\mathcal{F}_t^X$ as saying that if X_s has been observed for $0 \leq s \leq t$, then nothing more can be learned by looking infinitesimally far into the future.

Definition A.36. A continuous stochastic process X is *predictable* if it is measurable with respect to the σ-algebra generated by all left-continuous adapted processes.

Intuitively, we say that a process if predictable if its value is knowable at an earlier time. If a process if left-continuous, then it is predictable.

Definition A.37. A process X is said to be *progressively measurable* with respect to the filtration $\{\mathcal{F}_t\}$ if $\{(s,\omega)\colon 0 \leq s \leq t, \omega \in \Omega, X_s(\omega) \in \mathcal{B}(\mathbb{R}^d)\} \in \mathcal{B}([0,t]) \otimes \mathcal{F}_t$, i.e. if the map $(s,\omega) \mapsto X_s(\omega)$ is $\mathcal{B}([0,t]) \otimes \mathcal{F}_t$-measurable, for every time t.

Note that any progressively measurable process is adapted, and that any right-continuous, adapted process is progressively measurable (a proof of the second assertion can be seen in [41, p. 20]).

The concept of the progressive measurability of X indicates that the process is non-anticipating and measurable in the variables t and ω together.

Definition A.38. A *stopping time* is a random variable $\tau\colon \Omega \to \mathbb{R}$ such that $\{\omega \in \Omega\colon \tau(\omega) \leq t\} \in \mathcal{F}_t$ for all $t \in \mathbb{R}_+$, where τ may take the value $+\infty$.

We interpret $\tau(\omega)$ as corresponding to the instant at which a certain phenomenon occurs for the first time. The notion of stopping time is a special case of the concept of *random time*, which is simply an $\mathcal{F}_t$-measurable random variable with values in $\mathbb{R}_+$.

Definition A.39. An *integrable process* A is a process that verifies $\mathrm{E}[A_t] < \infty$, for every $t \in \mathbb{R}_+$.

Definition A.40. An *increasing* process A is a process that verifies the following three conditions:

1) A is $\mathcal{F}_t$-adapted.

2) $A_0 = 0$ a.s.

3) The map $t \mapsto A_t$ is right-continuous and increasing a.s. (which implies, given the previous condition, that $A_t \in \mathbb{R}_+$ a.s.).

Definition A.41. An integrable, increasing process A is *natural* if for every bounded right-continuous martingale M,

$$\mathrm{E}\int_0^t M_s \,\mathrm{d}A_s = \mathrm{E}\int_0^t M_{-s} \,\mathrm{d}A_s, \qquad t \in \mathbb{R}_+. \tag{A.22}$$

The above equation can be written in the equivalent form

$$\mathrm{E}[M_t A_t] = \mathrm{E} \int_0^t M_{-s} \, \mathrm{d}A_s, \tag{A.23}$$

and the proof of this fact can be seen in [20, pp. 23–24]. If the map $t \mapsto A_t$ is continuous a.s., then an integrable, increasing process is natural. Further, an integrable increasing process is natural if and only if it is predictable (according to [17, p. 35]).

It should be noted that if A is an increasing process and X is a measurable process, then if we fix $\omega \in \Omega$, the sample path $t \mapsto X(t, \omega)$ is a measurable function from $\mathbb{R}_+$ into $\mathbb{R}$, which implies that the Lebesgue–Stieltjes integrals

$$\int_0^t X_s^{\pm}(\omega) \, \mathrm{d}A_s(\omega) \tag{A.24}$$

are well-defined (the source of these observations is [20, p. 23]).

Definition A.42. We call the *expected value* of a random variable X, which we denote by $\mathrm{E}(X)$, the Lebesgue integral of X with respect to the probability measure P, i.e.

$$\mathrm{E}(X) = \int_\Omega X \, \mathrm{d}P := \int_\Omega X^+ \, \mathrm{d}P - \int_\Omega X^- \, \mathrm{d}P, \tag{A.25}$$

provided that at least one of the integrals on the right-hand side of the equality is finite.

Definition A.43. A random variable X is *integrable* if

$$\int_\Omega |X| \, \mathrm{d}P < \infty. \tag{A.26}$$

Let Y be a random variable that is a function of X, i.e. $Y = g(X)$, where $g \colon \mathbb{R}^n \to \mathbb{R}$ is an integrable function. Then Y is $\mathcal{F}(X)$-measurable, where $\mathcal{F}(X) = \{X^{-1}(B) \colon B \in \mathcal{B}\}$ is a σ-algebra (the σ-algebra generated by X), and

$$E(Y) = \int_{\mathbb{R}^n} g(x) f(x) \, \mathrm{d}x. \tag{A.27}$$

In particular,

$$E(X) = \int_{\mathbb{R}^n} x f(x) \, \mathrm{d}x. \tag{A.28}$$

We want to define now the expected value of a random variable X given another random variable Y, that is, given a point $\omega \in \Omega$ and the value $Y(\omega)$, we want to obtain the expected value of $X(\omega)$.

Definition A.44. Let X and Y be two random variables. We define the *conditional expected value* of X given Y, and denote it by $\mathrm{E}(X|Y)$, as any random variable $\mathrm{E}(X|Y)$ that verifies

$$\int_A X \, \mathrm{d}P = \int_A \mathrm{E}(X|Y) \, \mathrm{d}P \tag{A.29}$$

for all $A \in \mathcal{F}(Y)$.

Definition A.45. Supposing that $\mathcal{G}$ is a σ-algebra such that $\mathcal{G} \subseteq \mathcal{F}$ and X is an integrable random variable, we define $\mathrm{E}(X|\mathcal{G})$ to be any random variable on Ω such that

$$\int_A X \,\mathrm{d}P = \int_A \mathrm{E}(X|\mathcal{G}) \,\mathrm{d}P \tag{A.30}$$

for all $A \in \mathcal{G}$, and such that $\mathrm{E}(X|\mathcal{G})$ is $\mathcal{G}$-measurable.

We can interpret the random variable $\mathrm{E}(X|\mathcal{G})$ as providing an estimate of X (its expectation, in fact) using the information available in $\mathcal{G}$. By requiring that $\mathrm{E}(X|\mathcal{G})$ is $\mathcal{G}$-measurable, we are saying that $\mathrm{E}(X|\mathcal{G})$ must be constructed from the information available in $\mathcal{G}$. Equation (A.30) forces the estimate we obtain to be consistent with X concerning integration over events in $\mathcal{G}$. The three properties of conditional expectation that we will need are listed in the following proposition.

Proposition A.2. Let X and Y be integrable random variables, and let $\mathcal{G}$ be a σ-algebra. Then

1) $\mathrm{E}(aX + bY|\mathcal{G}) = a\,\mathrm{E}(X|\mathcal{G}) + b\,\mathrm{E}(Y|\mathcal{G})$, for $a, b \in \mathbb{R}$, a.s.

2) If X is $\mathcal{G}$-measurable, then $\mathrm{E}(X|\mathcal{G}) = X$ a.s.

3) If X is independent of $\mathcal{G}$, then $\mathrm{E}(X|\mathcal{G}) = \mathrm{E}(X)$ a.s.

A final note regarding notation: when writing the non-conditional expectation of an expression, we sometimes omit the square brackets around the expression.