UNIVERSITAT DE BARCELONA

ADVANCED MATHEMATICS
MASTER'S FINAL PROJECT

An Introduction to Stochastic Integration

Author:
Salim Boukfal Lazaar

Supervisor:
Xavier Bardina Simorra (UAB)
David Márquez Carreras (UB)

Facultat de Matemàtiques i Informàtica

June 28th, 2024

Abstract

The main purpose of this master's thesis is to continue and extend the study of the stochastic integral seen in the subject of Stochastic Calculus, providing (hopefully) an introductory text that will allow the average student of the subject (and to anyone who is already familiar with the stochastic integral with respect to the Brownian motion) to expand his knowledge.

To do so, in the second chapter we briefly review the construction of the Itô integral to then see how we exploit these ideas to generalize the construction to other processes, this is done following the construction provided in the third chapter [7]. Many of the presented results and notions of this part (and the following ones) have been completed with additional explanations and comparisons with the already seen objects in order to make the understanding of these clearer.

In the third chapter, we discuss the topic of stochastic integration with respect to random fields. We first treat the integral with respect to the space-time Gaussian white noise, following the construction presented in the first two chapters of [6], since it deals with objects which might be a bit more familiar to the intended audience as its construction uses the already studied Itô integral with respect to the Brownian motion. Before doing so, we introduce two crucial Gaussian processes (the isonormal process and the white noise), which generalize the Brownian motion and are crucial when it comes to define the stochastic integral with respect to the space-time white noise.

Next, and following the second chapter of [11], we introduce a wider class of random fields (which contains the ones already seen) that can be used as integrators and show how one constructs integrals with respect to such objects. During this process, we use the already studied Gaussian white noise as a canonical example that will serve us as a model to compare the new construction.

Finally, and to keep the reader entertained, in the Introduction and in Section 2.1 we pose a problem regarding the validity of the models that use the Brownian motion and the space-time Gaussian white noise as driving noises which is treated in Sections 2.3.4 and 3.2.4, providing results regarding the approximation in law of the stochastic integral with respect to the Brownian motion (Theorem 2.3.11, of which we could not find a statement nor a proof elsewhere) and the one with respect to the Brownian sheet (Theorem 3.2.2, which is a simplification of one of the results in [2]), respectively.

Resum

El principal objectiu d'aquesta tesi és el de continuar i estendre l'estudi de la integral estocàstica vist a l'assignatura de Càlcul Estocàstic, proporcionant (o almenys això esperem) un document introductori que permeti a l'estudiant mitjà d'aquesta assignatura (i a qualsevol altra persona que estigui familiaritzada amb la integral estocàstica respecte del moviment Brownià) expandir els seus coneixements.

Per tal d'aconseguir-ho, en el segon capítol revisem de manera breu la construcció de la integral estocàstica respecte del moviment Brownià per, després, veure com explotem les idees emprades per generalitzar la construcció a altres processos, tot seguint la construcció presentada al tercer capítol de [7]. Molts dels resultats d'aquesta part (i de la resta del treball) han sigut complementats amb explicacions addicionals i comparacions amb objectes ja coneguts amb la intenció de facilitar la comprensió d'aquests.

En el tercer capítol, tractem la integral estocàstica respecte camps aleatoris. En primer lloc, considerem la integral respecte el soroll blanc Gaussià en l'espai-temps, seguint els dos primers capítols de [6], ja que aquesta es construeix a partir d'objectes que poden ser més familiars com són el moviment Brownià i la integral d'Itô respecte d'aquest. Abans, però, haurem de presentar dos processos Gaussians que són crucials en la construcció presentada (el procés isonormal i el soroll blanc), que generalitzen el moviment Brownià.

A continuació, i seguint el segon capítol de [11], introduïm una classe més àmplia de camps aleatoris (que conté els ja estudiats) que podran ser usats com a integradors i veiem com es construeixen les corresponents integrals estocàstiques. Durant el procés, usem el ja estudiat soroll blanc Gaussià com a exemple canònic que ens servirà com a model per anar comparant la nova construcció.

Finalment, i per tal de mantenir al lector entretingut, a la Introducció i a la Secció 2.1 posem en dubte la validesa dels models que usen el moviment Brownià i el soroll blanc Gaussià com a pertorbacions aleatòries. El problema en qüestió l'acabem tractant a les Seccions 2.3.4 i 3.2.4, on donem resultats referents a l'aproximació en llei de les integrals estocàstiques respecte del moviment Brownià (Teorema 2.3.11, del qual no n'hem pogut trobar cap enunciat ni prova enlloc) i respecte del drap Brownià (Teorema 3.2.2, el qual és una simplificació d'un dels resultats presentats a [2]), respectivament.

Agraïments

Donat que aquest treball podria ser l'últim de l'estil (doncs no tinc la menor idea d'on aniré a parar l'any vinent) i que vaig oblidar per complet fer-ho en els dos TFGs, em sento amb el deure i la necessitat de dedicar un petit espai a aquelles persones que han fet d'aquesta travessia una meravellosa experiència. Sense més ni pus preàmbuls, doncs, anem a passar llista d'una vegada, no sigui el cas que quedi un document final massa llarg.

En primer lloc, agrair als meus familiars més propers, que a través del seu exemple m'han fet entendre la importància de l'esforç, la dedicació i la constància per aconseguir allò que hom es proposa, i, en particular, a les petites de la casa, la Maisa i la Noha, que des que van aparèixer han fet d'aquesta llar un lloc ple de joia. També agrair als xicots del parc: la Núria, l'Iván, en Borja, en Carlos i l'Aníbal, amb qui, en essència, ens dediquem a destrossar-nos a base de dominades i flexions (per a alguns pot semblar una tortura, cosa que no acabo d'entendre) i a la Sarah, que per algun motiu que desconec, però pel qual n'estic profundament agraït, segueix present a la meva vida, i espero que així segueixi.

Per un altre cantó, m'agradaria mencionar als meus companys de la doble titulació, especialment a en Nacho, la Míriam, en Jordi, l'Edu, la Clàudia i la Marina, que han fet que el pas per la universitat sigui considerablement més amè i m'han donat la motivació que em calia quan les ganes de llevar-me del llit pel matí eren pràcticament nul·les. També als meus fillols: la Jenni, en Salvador i l'Àngel, amb qui m'hagués encantat compartir classe d'igual manera que ho he fet amb la resta de companys. No oblidem tampoc als companys del màster, més concretament, a l'Òscar, en Joaquín, en Toni, en Guillem i en Guillermo, que, tot i haver-los conegut aquest any, m'ha donat la impressió que els coneixia de fa molt més. Sap greu que hagi durat tan poc aquest màster... o no, no sé fins quin punt m'agradaria, i us agradaria, allargar-lo.

En últim lloc, però no per això menys important, m'agradaria agrair als professors que m'han acompanyat durant tot aquest temps. Especialment a en Xavier, per la llibertat donada per poder fer el que un servidor vol, alhora que ha encaminat les idees proposades per tal que aquest treball no resultés en un desastre estrepitós.

M'omple d'orgull i satisfacció poder dir que he conegut a gent tan esplèndida, no crec que hi hagi paraules que facin justícia a tot allò que m'han aportat.

Contents

Chapter 1

Introduction

In the subject of Stochastic Calculus, one studies the stochastic integral with respect to the Brownian motion to give a rigorous definition of what is called a Stochastic Differential Equation (SDE), in this case driven by a Brownian motion. That is, we construct such objects to be able to make sense of expressions like

$$X_t = X_0 + \int_0^t \mu(s, X_s)ds + \int_0^t \sigma(s, X_s)dB_s, \tag{1.0.1}$$

which, in differential form, are written as

$$dX_t = \mu(t, X_t)dt + \sigma(t, X_t)dB_t, \tag{1.0.2}$$

where μ and σ are given functions satisfying some regularity conditions, $\{B_t \colon t \geq 0\}$ is a Brownian motion and $\{X_t \colon t \geq 0\}$ is a stochastic process determined by the previous SDE. In Chapter 2, Section 2.1, we will briefly see why are we interested in working with differential equations with random perturbations, why the differential "dB_t" does not make sense as a classical derivative, why the standard integration theory (Lebesgue-Stieltjes, for instance) does not work in this cases and why we really need to develop a new theory of integration.

It turns out that the ideas used to construct the integral with respect to the Brownian motion can be used to define stochastic integrals with respect to many other stochastic processes, which is done in Chapter 2, and hence, to define SDEs driven by more general stochastic processes. Moreover, as we will see in Chapter 3, the already constructed stochastic integrals can be used to define even more stochastic integrals, which will prove to be useful when it comes to define rigorously what is called a Stochastic Partial Differential Equation (SPDE).

On the other hand, and as discussed in Section 2.1, the choice of the Brownian motion as a driving noise (or the Gaussian white noise when it comes to SPDEs) is,

usually, an idealization or a simplification of a much more complicated situation. In other words, the random perturbations seen in SDEs and SPDEs are can be thought as simplifications, idealizations or limiting behaviours of the superposition of several factors that might be more complicated to study. Therefore, in order to make sure that the models are coherent and that simplified models lead to simplified solutions, an important task is to study under which conditions the solutions of such equations (SDEs and SPDEs) converge to the solution of the corresponding equation when the random perturbation is approximated by some family of stochastic processes.
To be a bit more specific, let $\mathcal{L}$ be an integro-differential operator, $\{\dot{W}_n\}_{n\in\mathbb{N}}$ be a sequence of random perturbations converging in some sense to a random perturbation $\dot{W}$ and $\{U_n(t,x)\}_{n\in\mathbb{N}}$ and $U(t,x)$ be, respectively, the solutions to the equations

$$\mathcal{L}U_n(t,x) = \sigma(t,x,U_n(t,x))\dot{W}_n(t,x), \quad \mathcal{L}U(t,x) = \sigma(t,x,U(t,x))\dot{W}(t,x).$$

Where t and x are, respectively, the time and space variables. Under which conditions is it true that the solutions U_n converge to U? In which sense is this convergence? Given that in most cases we are interested in the laws of the solutions, rather than, for instance, the exact form of their sample paths, we will focus our efforts on studying the convergence of the laws of the solutions of the SDEs and SPDEs or, in other words, on studying the weak convergence of the solutions.

However, before dealing with the convergence of the solutions of SDEs and SPDEs, we should first be concerned about the convergence of the noises involved, providing examples of processes $\{\dot{W}_n\}_n$ and $\dot{W}$. Therefore, in this work, apart from constructing the stochastic integral with respect to a certain class of noises, we will focus our efforts in proving some results regarding the weak approximation of such noises and integrals.

Chapter 2

Martingales as integrators

2.1 Motivation

As said in the introduction, one of the reasons to develop a theory of stochastic integration with respect to stochastic processes is to define rigorously expressions like (1.0.2). But why would someone want to work with such objects? Let us see an illustrative example where these objects turn out to be useful.

Particle in a fluid

Suppose that we have a particle submerged in a fluid (for simplicity, we will work in one dimension). Given the initial position and velocity of every single particle conforming the system, the trajectory of our particle as a function of time, X_t, can be simply found by using, for instance, Newton's laws, taking into account the several collisions, etc. However, even though simple, this way is unpractical from a computational point of view since the number of elements in the system is a little bit too high.

To simplify the computations, one can think that, at the end of the day, the effect of the several collisions translates into a friction force opposing the motion of the particle. In most cases, this friction depends, mainly, on the velocity of the particle, meaning that, by Newton's second law,

$$d\dot{X}_t = -\beta \dot{X}_t dt, \quad \beta := \frac{\zeta}{m} \tag{2.1.1}$$

where $\dot{X}_t = dX_t/dt$, $\zeta > 0$ is the friction coefficient and $m > 0$ is the mass of the particle we are studying. Given initial conditions $X_0, \dot{X}_0$, the solution to such ODE is given by

$$\dot{X}_t = \dot{X}_0 e^{-\beta t}, \quad X_t = X_0 + \frac{\dot{X}_0}{\beta}\left(1 - e^{-\beta t}\right).$$

The solution obtained, even though it describes several systems in a simple way, it may overlook particular details of the system (for instance, the frequency and intensity of the collisions of the particles) which cannot be summarized in the friction coefficient ζ.

One way to take into account some of these features is to introduce random perturbations in Eq.(2.1.1). To do so, we first discretize the time variable and suppose that our particle suffers a random perturbation on each step:

$$\dot{X}_{t+\Delta t} - \dot{X}_t = -\beta \dot{X}_t \Delta t + f \Delta V_t, \quad \Delta V_t = V_{t+\Delta t} - V_t. \tag{2.1.2}$$

Where V_t is the random perturbation introduced and f is some function modelling the intensity of the perturbation. The ΔV_t term accounts for the "random" displacements caused by the several collisions. If these collisions are not too strong (at the end of the day, they are caused by particles with low mass), the displacements will be "small" (of finite variance, for instance), meaning that ΔV_t accounts for the superposition of several small factors and hence, by the Central Limit Theorem, it is natural to assume that $\Delta V_t \sim \mathcal{N}(\tilde{\mu}, \tilde{\sigma}^2)$ (where $\tilde{\mu} \in \mathbb{R}$ and $\tilde{\sigma} > 0$ are some functions that might depend on the position of the particle, the time, etc.). Moreover, if we assume that the intensity of the collisions and their orientation is independent of the position of the particle X_t (this might be the case of an homogeneous medium), we can think that $f = f(t, \dot{X}_t)$ and similarly for $\tilde{\mu}$ and $\tilde{\sigma}$.

Another assumption we can make, for instance, is that the perturbations ΔV_t and $\Delta V_{t+\Delta t}$ are independent, which seems to be the case in a large variety of systems. Hence, Eq.(2.1.2) can be rewritten as

$$\dot{X}_{t+\Delta t} - \dot{X}_t = \mu(t, \dot{X}_t)\Delta t + \sigma(t, \dot{X}_t)\Delta B_t,$$

for some functions μ and σ and where $\{B_t \colon t \geq 0\}$ is a standard Brownian motion. Formally speaking, and taking the limit $\Delta t \to 0$, one arrives at equation (1.0.2) with $\dot{X}_t$ instead of X_t. Thus, we have converted the deterministic equation (2.1.1) into an equation modelling the sample paths of a stochastic process.

A first approach to study such equations is doing it for each $\omega \in \Omega$ (being Ω our sample space):

$$X_t(\omega) = \mu(t, X_t(\omega))dt + \sigma(t, X_t(\omega))dB_t(\omega)$$

transforming the equation into an ODE. Nevertheless, it is a very well-known fact that the sample paths of the Brownian motion are not of bounded variation and hence, that they are not differentiable. Meaning that the differential "$dB_t(\omega)$" is not well defined in the usual sense. In view of this fact, one might try to rewrite equation (1.0.2) in its integral form like in Eq.(1.0.1) and try to study the new equation for each ω:

$$X_t(\omega) = X_0(\omega) + \int_0^t \mu(s, X_s(\omega))ds + \int_0^t \sigma(s, X_s(\omega))dB_s(\omega).$$

However, the very same fact tells us that the integral $\int_0^t \sigma(s, X_s(\omega))dB_s(\omega)$ does not make sense as a Riemann-Stieltjes integral or a Lebesgue-Stieltjes one. So we do not have any classical tools to study equations (1.0.1) and (1.0.2) path by path. But we still have options!

Instead of defining the integral path by path as a limit of Riemann sums

$$\int_0^t \sigma(s, X_s(\omega))dB_s(\omega) = \lim_{n\to\infty} \sum_{j=1}^n \sigma\left(t_j^{(n)}, X_{t_j^{(n)}}(\omega)\right)\left(B_{t_j^{(n)}}(\omega) - B_{t_{j-1}^{(n)}}(\omega)\right),$$

we can exploit the ideas regarding the convergence of random variables to define the limit of these sums in a weaker sense, and this is what we will do.

As a final remark, observe that equation (1.0.2) has been obtained by assuming that the variance of the displacements is finite, among many other things. If any of these assumptions is not fulfilled, then it might not make any sense to consider SDEs driven by a Brownian motion and we would have to work with some other stochastic process. In this chapter, we will restrict ourselves to the case where the stochastic process is a continuous square integrable martingale and, at the end, we will mention some other constructions related to the ones seen here.

From now on, during this chapter we will work on a filtered probability space $\left(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t\in[0,T]}, \mathbb{P}\right)$ where Ω will be our sample space, $\mathcal{F}$ a σ-field on Ω, $\{\mathcal{F}_t\}_{t\in[0,T]}$ a filtration (being $T > 0$ is some fixed time horizon) and $\mathbb{P}$ is a probability measure. We will assume as well that the filtration satisfies the usual conditions: it is right-continuous and complete or, in other words,

$$\mathcal{F}_t = \bigcap_{s>t} \mathcal{F}_s$$

for any $t \in [0, T]$ and $\mathcal{F}_0$ contains all $\mathbb{P}$-null sets and the ones contained in such sets.

Usually, we will require the filtration to be the one generated by the integrator (the process with respect to which we will integrate) completed so that the usual conditions are fulfilled or that the integrator is adapted to the given filtration.

2.2 Integrals with respect to the Brownian motion, an overview

Before we start talking about integration with respect to martingales and other processes, we first recall the main ideas and results seen in the subject of Stochastic Calculus regarding the construction of the (L^2) integral with respect to the Brownian motion.

In this section we will assume that the filtration $\{\mathcal{F}_t\}_t$ is the one generated by the Brownian motion completed so that the usual conditions are fulfilled. This in particular implies that increments of the form $B_t - B_s$ are independent of $\mathcal{F}_r$ for $0 \leq r \leq s < t < \infty$.

To talk about integrals of the form $\int_0^T f_t dB_t$ where $f = \{f_t \colon 0 \leq t \leq T\}$ (being $T > 0$ some fixed time) is some suitable stochastic process, we first define the class of functions for which the integral will be defined.

Definition 2.2.1. *Let $\nu = \nu[0,T]$ be the class of functions $f \colon [0,T] \times \Omega \to \mathbb{R}$, $(t,\omega) \mapsto f(t,\omega)$ such that*

(i) $(t,\omega) \mapsto f(t,\omega)$ is $\mathcal{B} \times \mathcal{F}$-measurable (where $\mathcal{B}$ is the Borel σ-field on $[0,T]$).

(ii) For each $t \in [0,T]$, $\omega \mapsto f(t,\omega)$ is $\mathcal{F}_t$-measurable (the process f is $\{\mathcal{F}_t\}_t$-adapted).

(iii) $\mathbb{E}\left[\int_0^T f^2(t,\omega)dt\right] < \infty$, where $f^2(t,\omega) = f(t,\omega) \cdot f(t,\omega)$ and $\mathbb{E}$ is the expected value with respect to the probability measure $\mathbb{P}$.

Observe that the space ν can be identified with the space of functions in $L^2\left([0,T] \times \Omega, \mathcal{B} \times \mathcal{F}, d\mathbb{P}dt\right)$ (here dt denotes the Lebesgue measure on $[0,T]$) which are adapted to the filtration $\{\mathcal{F}_t\}_t$. In fact, one can show that ν is closed with respect to the norm $||f||^2 = \mathbb{E}\left[\int_0^T f^2(t,\omega)dt\right]$, $f \in \nu$. So ν is a Hilbert space with respect to the inner product $(f,g) = \mathbb{E}\left[\int_0^T f(t,\omega)g(t,\omega)dt\right]$ (where functions $f, g \in \nu$ such that $||f - g|| = 0$ are identified, of course).

The procedure now is quite straightforward. First we consider the subset of elementary processes $\mathcal{E} \subset \nu$ of the form

$$\phi_t(\omega) = \phi(t,\omega) = \sum_{j=0}^{n-1} e_j(\omega)\mathbb{I}_{[t_j,t_{j+1})}(t),$$

where $0 = t_0 < ... < t_n = T$ constitutes a partition of the interval $[0,T]$ and the functions e_j are $\mathcal{F}_{t_j}$-measurable and bounded (here $\mathbb{I}$ denotes the indicator function).

For these processes, the definition of stochastic integral is given by

$$\left(\int_0^T \phi_t dB_t\right)(\omega) = \sum_{j=0}^{n-1} e_j(\omega)\left(B_{t_{j+1}}(\omega) - B_{t_j}(\omega)\right).$$

One then sees that the stochastic integral (also known as Itô integral) acts linearly on $\mathcal{E}$ and that any function in ν can be approximated by elementary processes with respect to the norm $||\cdot||$. That is, $\mathcal{E}$ is dense in ν. Finally, one defines the Itô integral with respect to the Brownian motion of any function $f \in \nu$ as the $L^2(\Omega)$ limit of Itô integrals of elementary processes that approximate f. To justify the existence of such limit, one uses the following lemma, which is crucial in such task:

Lemma 2.2.1 (The Itô isometry). *Let $\phi \in \mathcal{E}$ be an elementary process. Then*

$$\mathbb{E}\left[\left(\int_0^T \phi_t dB_t\right)^2\right] = \mathbb{E}\left[\int_0^T \phi_t^2 dt\right]. \tag{2.2.1}$$

To understand what will be going on in the following sections, let us see the proof of this result.

Proof. Let us write $\phi_t = \sum_{j=0}^{n-1} e_j \mathbb{I}_{[t_j,t_{j+1})}(t)$ and $\Delta B_j := B_{t_{j+1}} - B_{t_j}$. Then we have that

$$\mathbb{E}\left[\left(\int_0^T \phi_t dB_t\right)^2\right] = \sum_{j=0}^{n-1} \mathbb{E}[e_j^2(\Delta B_j)^2] + 2\sum_{0\le i<j\le n-1} \mathbb{E}[e_i e_j \Delta B_i \Delta B_j].$$

Now, given that e_j is $\mathcal{F}_{t_j}$-measurable and that ΔB_j is independent of $\mathcal{F}_{t_j}$, we have that, by law of total expectation

$$\begin{aligned}
\mathbb{E}[e_j^2(\Delta B_j)^2] &= \mathbb{E}\left[\mathbb{E}\left[e_j^2(\Delta B_j)^2 \middle| \mathcal{F}_{t_j}\right]\right] \\
&= \mathbb{E}\left[e_j^2 \mathbb{E}\left[(\Delta B_j)^2 \middle| \mathcal{F}_{t_j}\right]\right] \\
&= \mathbb{E}\left[e_j^2 \mathbb{E}\left[(\Delta B_j)^2\right]\right] \\
&= \mathbb{E}\left[e_j^2\right](t_{j+1} - t_j).
\end{aligned}$$

Similar reasons lead to

$$\begin{aligned}
\mathbb{E}[e_i e_j \Delta B_i \Delta B_j] &= \mathbb{E}\left[\mathbb{E}\left[e_i e_j \Delta B_i \Delta B_j \middle| \mathcal{F}_{t_j}\right]\right] \\
&= \mathbb{E}\left[e_i e_j \Delta B_i \mathbb{E}\left[\Delta B_j \middle| \mathcal{F}_{t_j}\right]\right] \\
&= \mathbb{E}\left[e_i e_j \Delta B_i \mathbb{E}\left[\Delta B_j\right]\right] \\
&= 0
\end{aligned}$$

for $0 \leq i < j \leq n-1$. Hence,

$$\mathbb{E}\left[\left(\int_0^T \phi_t dB_t\right)^2\right] = \sum_{j=0}^{n-1} \mathbb{E}\left[e_j^2\right](t_{j+1}-t_j).$$

On the other hand,

$$\begin{aligned}
\mathbb{E}\left[\int_0^T \phi_t^2 dt\right] &= \mathbb{E}\left[\int_0^T \left(\sum_{j=0}^{n-1} e_j \mathbb{I}_{[t_j,t_{j+1})}(t)\right)^2 dt\right] \\
&= \mathbb{E}\left[\int_0^T \sum_{j=0}^{n-1} e_j^2 \mathbb{I}_{[t_j,t_{j+1})}(t) dt\right] \\
&= \mathbb{E}\left[\sum_{j=0}^{n-1} e_j^2 (t_{j+1}-t_j)\right] \\
&= \sum_{j=0}^{n-1} \mathbb{E}\left[e_j^2\right](t_{j+1}-t_j). \qquad (2.2.2)
\end{aligned}$$

Proving the desired result. □

One of the key points in this last proof is that the expected values on the double sum cancel out. In some sense, this means that there is some sort of $L^2(\Omega)$ orthogonality between the increments of the integrator, $\mathbb{E}[\Delta B_i \Delta B_j] = 0$. Further, notice that this orthogonality has been proven using the fact that the conditional expectation $\mathbb{E}\left[\Delta B_j \middle| \mathcal{F}_{t_j}\right]$ vanishes, which is, in essence, the condition for the integrator to be a martingale.

Observe as well that both sides of equation (2.2.1) lead to (2.2.2), which depends linearly on the increment of time. This is no coincidence! As we will see in the near future, even though the Brownian motion (and, in general, non-constant continuous martingales) are of unbounded variation, they possess quadratic variation (a concept to be defined), meaning that, roughly speaking, the squared increments of such processes behave like linear increments of time. For the case of the Brownian motion, the quadratic variation on an interval $[t_j, t_{j+1})$ is exactly $t_{j+1} - t_j$

We will exploit these observations in the following section, where we will adapt the ideas seen in this section to construct the integral with respect to continuous square integrable martingales.

2.3 Integrals with respect to continuous square integrable martingales

One might think that the Brownian motion is a special stochastic process in the sense that there are not many processes with the property that its sample paths are not of bounded variation, but nothing could be further from reality!

Theorem 2.3.1. *If $M = \{M_t\colon t \geq 0\}$ is a continuous martingale of bounded variation, then*

$$\mathbb{P}\left\{\omega \in \Omega\colon M_t(\omega) = M_0(\omega) \textit{ for all } t \geq 0\right\} = 1.$$

One can find a proof of this result in page 240, Theorem. 3.1.1 of [8].

Thus, in general, any non-constant continuous martingale will have unbounded variation, meaning that, as in the case of the Brownian motion, integrals of the form $\int_0^T f_t dM_t$ cannot be defined path-wise. But we can make use of similar arguments to the ones seen in the previous section to define such objects. Before dealing with such task, we first talk about a process of utmost importance related to continuous square integrable martingales.

2.3.1 Quadratic variation

As already mentioned in a couple times, non-constant continuous martingales do not have bounded variation. Recall that the total variation of a function $f\colon [a,b] \subset \mathbb{R} \to \mathbb{R}$ on $[a,b]$, $V_a^b(f)$, is defined as

$$V_a^b(f) \coloneqq \sup_{P \in \mathcal{P}} \sum_{j=0}^{n_P - 1} |f(x_{j+1}) - f(x_j)|.$$

Where $\mathcal{P}$ is the set of partitions of $[a,b]$ of the form $P = \{a = x_0 < \ldots < x_{n_P} = b\}$ with $n_P \in \mathbb{N}$. Recall as well that a function of bounded variation is differentiable almost everywhere and that any differentiable function is of bounded variation.

To begin with, let us consider the case of a standard Brownian motion $B = \{B_t\colon t \geq 0\}$. As mentioned before, its sample paths are not of bounded variation. Roughly speaking, and to gain some intuition, this can be though as a consequence of the fact that, since $\mathbb{E}\left[(B_{t+\Delta t} - B_t)^2\right] = \Delta t$, for Δt small enough (and in quadratic mean),

$$\frac{B_{t+\Delta t} - B_t}{\Delta t} \approx \frac{\sqrt{\Delta t}}{\Delta t} = \frac{1}{\sqrt{\Delta t}}$$

which diverges as Δt approaches 0 and hence, it cannot have bounded variation.

However, the fact that $\mathbb{E}\left[(B_{t+\Delta t} - B_t)^2\right] = \Delta t$ (and using the independence of the increments of the Brownian motion) lead to the following result:

Lemma 2.3.1. *Let $B = \{B_t \colon t \geq 0\}$ be a standard Brownian motion and let $0 \leq a < b < \infty$, then*

$$L^2(\Omega) - \lim_{|P|\to 0} \sum_{j=0}^{n-1} \left(B_{t_{j+1}} - B_{t_j}\right)^2 = b - a \tag{2.3.1}$$

where P ranges over all partitions of $[a, b]$ of the form $\{a = t_0 < \ldots < t_n = b\}$ and where $|P| = \sup_{0\leq j\leq n-1} |t_{j+1} - t_j|$ is the mesh of the partition P.

In particular, this lemma implies that the limit in (2.3.1) holds in probability.

One can easily check that the process $B^2 = \{B_t^2 \colon t \geq 0\}$ is not a martingale, however, the process $\{B_t^2 - t \colon t \geq 0\}$ is so. The point in making this observation is that, in order to transform B^2 into a martingale, we have to subtract from it its quadratic variation (starting at 0).

These ideas can be applied to any stochastic process, however, only a limited class of stochastic processes (which contains the continuous square integrable martingales) will turn out to have similar properties to the ones presented by the Brownian motion.

Theorem 2.3.2. *Let $M = \{M_t \colon t \geq 0\}$ be a continuous square integrable martingale with respect to a filtration $\{\mathcal{F}_t\}_t$ satisfying the usual conditions. Then there exists a unique non-decreasing, continuous adapted process $\langle M\rangle = \{\langle M_t\rangle \colon t \geq 0\}$ such that:*

(i) $\langle M\rangle_0 = 0$ *a.s.; and*

(ii) $t \mapsto M_t^2 - \langle M\rangle_t$ *is a martingale.*

Moreover,

$$\langle M\rangle_t = \lim_{|P|\to 0} \sum_{j=0}^{n-1} \left(M_{t_{j+1}} - M_{t_j}\right)^2 \tag{2.3.2}$$

where P ranges over all partitions of $[0, t]$ and the convergence holds in probability.

For a proof of this result, see Theorem 3.2.1, page 242, of [8].

The process $\langle M\rangle$ in the previous theorem is the quadratic variation of M and it uniquely determines the martingale M among all the continuous square integrable martingales.

One then can define the covariation of two continuous square integrable martingales M and N by polarization:

$$\langle M, N\rangle \coloneqq \frac{1}{4} \left(\langle M + N\rangle - \langle M - N\rangle\right).$$

It can be defined, as well, as the following limit in probability:

$$\langle M, N\rangle_t = \lim_{|P|\to 0} \sum_{j=0}^{n-1} \left(M_{t_{j+1}} - M_{t_j}\right) \left(N_{t_{j+1}} - N_{t_j}\right). \tag{2.3.3}$$

The covariation process is adapted, continuous and starts at zero, but it is not necessarily non-decreasing. It also satisfies that $\{M_tN_t - \langle M, N\rangle_t \colon t \geq 0\}$ is a martingale. In particular, one has that

$$\mathbb{E}\left[M_tN_t - \langle M, N\rangle_t\right] = \mathbb{E}\left[M_0N_0 - \langle M, N\rangle_0\right] = \mathbb{E}\left[M_0N_0\right] \tag{2.3.4}$$

It is easy to check that the covariation is linear on each of its arguments ($\langle \cdot, \cdot \rangle$ is bilinear), symmetric and that $\langle M, M\rangle = \langle M\rangle$.

In general, the quadratic variation or the covariation will be stochastic processes. The fact that the quadratic variation of the Brownian motion is non-random is a special feature of this process.

We point out that these results can be extended to the case of continuous local martingales and continuous semimartingales (see Definitions 2.4.1 and 2.4.2) and to càdlàg local martingales and semimartingales. However, the process $\{M_tN_t - \langle M, N\rangle_t \colon t \geq 0\}$ is no longer a martingale, but a local martingale or a semimartingale.

As a final remark, the notation used here to denote the quadratic variation coincides with the one usually used to denote the compensator in the Doob-Meyer decomposition. In general, both quantities are not equal, but given that the sample paths of the processes considered here are continuous they do agree almost surely, justifying the use of such notation.

With this we can start constructing the desired stochastic integral.

2.3.2 Construction of the stochastic integral

The idea now is to replicate the construction seen in Sec.2.2 (for this part, we have mainly followed Chapter 3 of [7]). To do so, we first start by observing that the third condition in Definition 2.2.1 can be rewritten as follows:

$$\mathbb{E}\left[\int_0^T X^2(t,\omega)d\langle B\rangle_t\right] < \infty.$$

A subtle change has been introduced, we have used that $dt = d\langle B\rangle_t$. This change, even though it might seem simple, allows us to naturally introduce the class of integrable functions with respect to a given continuous square integrable martingale $M = \{M_t \colon t \geq 0\}$. Unless stated otherwise, we shall assume as well that $M_0 = 0$.

Definition 2.3.1. *Let $\nu_M = \nu_M[0,T]$ be the class of functions $X \colon [0,T] \times \Omega \to \mathbb{R}$, $(t,\omega) \mapsto X(t,\omega)$ such that*

(i) $(t,\omega) \mapsto X(t,\omega)$ is $\mathcal{B} \times \mathcal{F}$-measurable (where $\mathcal{B}$ is the Borel σ-field on $[0,T]$).

(ii) For each $t \in [0,T]$, $\omega \mapsto X(t,\omega)$ is $\mathcal{F}_t$-measurable (the process X is $\{\mathcal{F}_t\}_t$-adapted).

(iii) $\mathbb{E}\left[\int_0^T X^2(t,\omega)d\langle M\rangle_t\right] < \infty$*, where* $X^2(t,\omega) = X(t,\omega)\cdot X(t,\omega)$ *and* $\mathbb{E}$ *is the expected value with respect to the probability measure* $\mathbb{P}$*.*

One may argue that, again, we need to define the integral $\int_0^T X^2(t,\omega)d\langle M\rangle_t$ since the quantity $\langle M\rangle_t$ is, in general, random. However, in this case there is no need to develop a new theory of integration. Indeed, recall that, even though $\langle M\rangle_t$ is random, for almost all $\omega\in\Omega$, the function $t\mapsto\langle M\rangle_t(\omega)$ is non-decreasing and hence, of bounded variation, meaning that the integral $\int_0^T X^2(t,\omega)d\langle M\rangle_t$ can be defined path-wise as a Lebesgue-Stieltjes integral.

The space ν_M can be identified with the space of functions in $L^2\left([0,T]\times\Omega,\mathcal{B}\times\mathcal{F},d\mu_M\right)$ which are adapted to the filtration $\{\mathcal{F}_t\}_t$. Here μ_M is the measure defined by

$$\mu_M(A) \coloneqq \mathbb{E}\left[\int_0^T \mathbb{I}_A(t,\omega)d\langle M\rangle_t(\omega)\right], \quad A\in\mathcal{B}\times\mathcal{F}.$$

In ν_M (and in $L^2\left([0,T]\times\Omega,\mathcal{B}\times\mathcal{F},d\mu_M\right)$), we consider the norm

$$||X||_M^2 = \mathbb{E}\left[\int_0^T X^2(t,\omega)d\langle M\rangle_t\right].$$

Again, we identify functions $X,Y\in\nu_M$ such that $||X-Y||_M = 0$. With all this in mind, one can show that ν_M is a real Hilbert space with respect to the inner product $(X,Y)_M \coloneqq \mathbb{E}\left[\int_0^T X(t,\omega)Y(t,\omega)d\langle M\rangle_t\right]$.

Theorem 2.3.3. *The space* ν_M *with the norm* $||\cdot||_M$ *is closed.*

Proof. Suppose we have a sequence of processes $\{X^{(n)}\}_{n\in\mathbb{N}}\subset\nu_M$ that converges, in norm $||\cdot||_M$ to some process X, We have to show that $X\in\nu_M$.

Given that $L^2\left([0,T]\times\Omega,\mathcal{B}\times\mathcal{F},d\mu_M\right)$ is a Hilbert space and that is a subspace of the latter, we already have that X is $\mathcal{B}\times\mathcal{F}$-measurable and that $||X||_M<\infty$. The only thing left to show is that it is adapted to the filtration $\{\mathcal{F}_t\}_t$ (or that it has a representative which is so).

Up to a subsequence, we have that

$$\mu_M\left\{(t,\omega)\in[0,T]\times\Omega\colon \lim_{n\to\infty} X_t^{(n)}(\omega)\neq X_t(\omega)\right\} = 0. \tag{2.3.5}$$

Now consider the process Y defined by

$$Y_t(\omega) = \begin{cases} \lim_{n\to\infty} X_t^{(n)}(\omega), & (t,\omega)\in A \\ 0, & (t,\omega)\notin A \end{cases}$$

where
$$A := \left\{(t,\omega) \in [0,T] \times \Omega \colon \lim_{n\to\infty} X_t^{(n)}(\omega) \text{ exists in } \mathbb{R}\right\}$$
(recall that $X^{(n)}$ might be denoting a subsequence of our original sequence). Then, by the μ_M-a.e. uniqueness of the limit and (2.3.5), Y is a representative of X in $L^2([0,T] \times \Omega, \mathcal{B} \times \mathcal{F}, d\mu_M)$ (that is, $||X - Y||_M = 0$) such that, for each $t \in [0,T]$, Y_t is the limit of $\mathcal{F}_t$-measurable random variables and hence, $\mathcal{F}_t$-measurable itself. □

We shall now see that any process $f \in \nu_M$ can be approximated in norm $||\cdot||_M$ by elementary processes (the same definition as before for the subset $\mathcal{E}$ works).

Theorem 2.3.4. *Let $\{A_t \colon t \geq 0\}$ be a continuous, non-decreasing process starting at 0 adapted to the filtration $\{\mathcal{F}_t\}_t$ (to which the martingale M is adapted). If $X = \{X_t, \colon t \geq 0\}$ is a progressively measurable process (with respect to the same filtration) satisfying*
$$\mathbb{E}\left[\int_0^T X_t^2 dA_t\right] < \infty,$$
then there exists a sequence $\{X_t^{(n)}\}_{n\in\mathbb{N}}$ of elementary processes such that
$$\lim_{n\to\infty} \mathbb{E}\left[\int_0^T \left(X_t - X_t^{(n)}\right)^2 dA_t\right] = 0.$$

Proof. We start by assuming that X is bounded. That is, there is some constant $0 < C < \infty$ such that $|X_t(\omega)| \leq C$ for each $t \geq 0$ and $\omega \in \Omega$.

Now, since $t \mapsto A_t(\omega) + t$ is strictly increasing in $t \geq 0$, there is a continuous, strictly increasing inverse function $T_s(\omega)$ defined for $s \geq 0$ such that
$$A_{T_s(\omega)}(\omega) + T_s(\omega) = s, \quad s \geq 0.$$
In particular, this means that $T_s \leq s$ for any $s \geq 0$ and $\{T_s \leq t\} = \{A_t + t \geq s\} \in \mathcal{F}_t$. Thus, for each $s \geq 0$, T_s is an $\{\mathcal{F}_t\}_t$ bounded (by s) stopping time. Taking s as our new time variable, we define a new filtration $\mathcal{G}_s = \mathcal{F}_{T_s}$ and introduce the time changed process $Y_s(\omega) = X_{T_s(\omega)}(\omega)$, which is adapted to $\{\mathcal{G}_s\}_s$ (see, for instance, Proposition 1.2.18, page 9 of [7] for a proof). Now, the approximation result mentioned in Sec.2.2 tells us that, for any $\varepsilon > 0$ and any $R > 0$, there is an elementary process $Y^\varepsilon = \{Y_s^\varepsilon \colon s \geq 0\}$ such that
$$\mathbb{E}\left[\int_0^R (Y_s - Y_s^\varepsilon)^2 ds\right] < \frac{\varepsilon}{2}.$$

Using that X is bounded, we have that

$$\begin{aligned}\mathbb{E}\left[\int_0^\infty Y_s^2 ds\right] &= \mathbb{E}\left[\int_0^\infty \mathbb{I}_{\{T_s \le T\}}(s) X_{T_s}^2 ds\right] \\ &= \mathbb{E}\left[\int_0^{A_T+T} X_{T_s}^2 ds\right] \\ &\le C^2(\mathbb{E}[A_T] + T) < \infty.\end{aligned}$$

Hence, by taking R large enough and setting $Y_s^\varepsilon = 0$ for $s \ge R$ we can obtain

$$\mathbb{E}\left[\int_0^\infty (Y_s - Y_s^\varepsilon)^2 \, ds\right] < \varepsilon.$$

The elementary process Y^ε can be then written as

$$Y_s^\varepsilon(\omega) = \sum_{j=0}^{n-1} e_j(\omega) \mathbb{I}_{[s_j, s_{j+1})}(s)$$

where e_j is $\mathcal{G}_{s_j}$-measurable and bounded (say by some positive constant K) and $0 = s_0 < ... < s_n = R$. From this process we can define a process in the original time variable as follows:

$$X_t^\varepsilon(\omega) := Y_{t+A_t}^\varepsilon = \sum_{j=0}^{n-1} e_j(\omega) \mathbb{I}_{[s_j, s_{j+1})}(t + A_t) = \sum_{j=0}^{n-1} e_j(\omega) \mathbb{I}_{[T_{s_j}, T_{s_{j+1}})}(t).$$

Where in the last step we have used that

$$\begin{aligned}&\{t + A_t \ge s_j\} = \{T_{s_j} \le t\}, \\ &\{t + A_t < s_{j+1}\} = \{t + A_t \ge s_{j+1}\}^c = \{T_{s_{j+1}} \le t\}^c = \{T_{s_{j+1}} > t\}.\end{aligned}$$

Which implies

$$\{t \colon s_j \le A_t + t < s_{j+1}\} = \{t \colon T_{s_j} \le t < T_{s_{j+1}}\}.$$

The process $X^\varepsilon = \{X_t^\varepsilon \colon t \ge 0\}$ is $\mathcal{B} \times \mathcal{F}$-measurable and $\{\mathcal{F}_t\}_t$-adapted by Lemma 1.2.15, page 8, in [7]. However, it is not clear that X^ε is elementary. We will see, however, that it can be approximated by elementary processes. Before doing so, note that

$$
\begin{aligned}
\varepsilon &> \mathbb{E}\left[\int_0^\infty |Y_s - Y_s^\varepsilon|^2\, ds\right] \\
&= \mathbb{E}\left[\int_0^\infty \left|X_{T_s} - Y^\varepsilon_{A_{T_s}+T_s}\right|^2 d(A_{T_s}+T_s)\right] \\
&= \mathbb{E}\left[\int_0^\infty \left|X_t - Y^\varepsilon_{A_t+t}\right|^2 d(A_t+t)\right] \\
&= \mathbb{E}\left[\int_0^\infty |X_t - X_t^\varepsilon|^2\, d(A_t+t)\right] \\
&\geq \mathbb{E}\left[\int_0^T |X_t - X_t^\varepsilon|^2\, dA_t\right].
\end{aligned}
$$

Where we have implicitly defined X_t to be zero whenever $t > T$.

To prove that one can approximate X^ε by elementary processes, we prove that each term $\eta_t := e_j(\omega)\mathbb{I}_{[T_{s_j}, T_{s_{j+1}})}(t)$ can be approximated in such a way.
Now recall that $T_{s_j} \leq T_{s_{j+1}} \leq s_{j+1}$ for each $j \in \{0, 1, ..., n-1\}$ and consider the processes

$$
T_l^{(m)}(\omega) := \sum_{k=0}^{2^{m+1}} \frac{k+1}{2^m}\mathbb{I}_{[k2^{-m},(k+1)2^{-m})}(T_l(\omega)), \quad l \in \{s_j, s_{j+1}\} \tag{2.3.6}
$$

and

$$
\eta_t^{(m)}(\omega) := e_j(\omega)\mathbb{I}_{\left[T^{(m)}_{s_j}(\omega), T^{(m)}_{s_{j+1}}(\omega)\right)}(t).
$$

Then we have that $T_l^{(m)} \searrow T_l$ a.s. for $l \in \{s_j, s_{j+1}\}$ and, since $\{T_{s_j} \leq k2^{-m} < T_{s_{j+1}}\} \in \mathcal{F}_{k2^{-m}}$ and e_j restricted to $\{T_{s_j} \leq k2^{-m} < T_{s_{j+1}}\}$ is $\mathcal{F}_{k2^{-m}}$-measurable (by Lemma 1.2.15 of [7] again), $\eta^{(m)}$ is elementary. Furthermore,

$$
\begin{aligned}
&\mathbb{E}\left[\int_0^T \left(\eta_t - \eta_t^{(m)}\right)^2 dA_t\right] \leq \mathbb{E}\left[\int_0^\infty \left(\eta_t - \eta_t^{(m)}\right)^2 dA_t\right] \\
&= \mathbb{E}\left[\int_0^\infty e_j^2\left(\mathbb{I}_{[T_{s_j}, T_{s_{j+1}})}(t) - \mathbb{I}_{\left[T^{(m)}_{s_j}, T^{(m)}_{s_{j+1}}\right)}(t)\right)^2 dA_t\right] \\
&\leq K^2\mathbb{E}\left[\int_0^\infty \left(\mathbb{I}_{[T_{s_j}, T_{s_{j+1}})}(t) + \mathbb{I}_{\left[T^{(m)}_{s_j}, T^{(m)}_{s_{j+1}}\right)}(t) - 2\mathbb{I}_{[T_{s_j}, T_{s_{j+1}})}(t)\mathbb{I}_{\left[T^{(m)}_{s_j}, T^{(m)}_{s_{j+1}}\right)}(t)\right) dA_t\right] \\
&= K^2\mathbb{E}\left[A_{T_{s_{j+1}}} - A_{T_{s_j}} + A_{T_{s_{j+1}}^{(m)}} - A_{T^{(m)}_{s_{j+1}}} - 2\left(A_{T_{s_{j+1}}} \wedge A_{T^{(m)}_{s_{j+1}}} - A_{T_{s_j}} \vee A_{T^{(m)}_{s_j}}\right)\right] \\
&= K^2\left(\mathbb{E}\left[A_{T^{(m)}_{s_{j+1}}} - A_{T_{s_{j+1}}}\right] + \mathbb{E}\left[A_{T^{(m)}_{s_j}} - A_{T_{s_j}}\right]\right) \xrightarrow{m\to\infty} 0.
\end{aligned}
$$

Where we have used that $T_l^{(m)} \geq T_l$ for each $m \in \mathbb{N}$ and $l \in \{s_j, s_{j+1}\}$, that the process A is increasing and the monotone convergence theorem.

Finally, if X is not necessarily bounded, we define

$$X_t^{(n)}(\omega) := X_t(\omega)\mathbb{I}_{\{|X_t(\omega)| \leq n\}}$$

and thereby obtain a sequence of bounded processes. The dominated convergence theorem implies

$$||X^{(n)} - X||_M^2 = \mathbb{E}\left[\int_0^T X_t^2 \mathbb{I}_{\{|X_t|>n\}} d\langle M\rangle_t\right] \xrightarrow{n\to\infty} 0$$

because $X_t^2\mathbb{I}_{\{|X_t|>n\}} \leq X_t^2$ for each $n \in \mathbb{N}$ and $\mathbb{E}\left[\int_0^T X_t^2 d\langle M\rangle_t\right] < \infty$, which implies that X_t^2 is finite for μ_M-almost all t, ω and hence, that $X_t^2\mathbb{I}_{\{|X_t|>n\}} \to 0$ μ_M-a.e. t, ω as n approaches infinity. □

We shall remark that the progressive measurability condition is not a restriction. Indeed, if $X \in \nu_M$, then X is $\mathcal{B} \times \mathcal{F}$ measurable and $\{\mathcal{F}_t\}_t$-adapted, so it has a progressively measurable modification.

By taking $A_t = \langle M\rangle_t$, we have that the set of elementary processes $\mathcal{E}$ is dens in ν_M with respect to the norm $||\cdot||_M$.

Following the same recipe as in Sec. 2.2, we now define the stochastic integral of an elementary process with respect to M and prove the isometry formula for such processes, among some other properties.

Definition 2.3.2. *The stochastic integral of an elementary process*

$$\phi_t(\omega) = \sum_{j=0}^{n-1} e_j(\omega)\mathbb{I}_{[t_j,t_{j+1})}(t) \in \mathcal{E}$$

with $0 = t_0 < ... < t_n = T$ *is defined as*

$$I_T(\phi)(\omega) = \left(\int_0^T \phi_t dM_t\right)(\omega) := \sum_{j=0}^{n-1} e_j(\omega)\left(M_{t_{j+1}}(\omega) - M_{t_j}(\omega)\right).$$

For $0 \leq s \leq t \leq T$, *we define*

$$I_t(\phi) = \int_0^t \phi_r dM_r := \int_0^T \mathbb{I}_{[0,t)}(r)\phi_r dM_r$$

and

$$\int_s^t \phi_r dM_r := I_t(\phi) - I_s(\phi).$$

Theorem 2.3.5. *Let $X, Y \in \mathcal{E}$ be two elementary processes on $[0,T]$, $\alpha, \beta \in \mathbb{R}$ be any two real numbers and $0 \leq s < t \leq T$. Then the following properties hold $\mathbb{P}$-a.s.:*

(i) $I_0(X) = 0$,

(ii) $I_t(\alpha X + \beta Y) = \alpha I_t(X) + \beta I_t(Y)$,

(iii) $\mathbb{E}\left[I_t(X) \middle| \mathcal{F}_s\right] = I_s(X)$,

(iv) $\mathbb{E}\left[(I_T(X))^2\right] = ||X||_M^2$,

(v) $\langle I(X)\rangle_t = \int_0^t X_r^2 d\langle M\rangle_r$.

Proof. The first two parts can be proved as in the case of the Brownian motion, so we will not treat them here. For the martingale property, part (iii), one only needs to use the fact that M is a martingale and distinguish cases for the possible values of s and t (if they lie on the same interval of the partition of $[0,T]$ considered in the definition of X, if they do not, etc.) and use the tower property for the conditional expectation and the fact that M is a martingale.

For the isometry property, first let us write X as

$$X_t = \sum_{j=0}^{n-1} e_j \mathbb{I}_{[t_j, t_{j+1})}(t)$$

for some $0 = t_0 < ... < t_n = T$ and e_j $\mathcal{F}_{t_j}$-measurable and set $\Delta M_j := M_{t_{j+1}} - M_{t_j}$. Now observe that, since M is a martingale,

$$\begin{aligned}
\mathbb{E}\left[(\Delta M_j)^2 \middle| \mathcal{F}_{t_j}\right] &= \mathbb{E}\left[M_{t_{j+1}}^2 + M_{t_j}^2 - 2M_{t_{j+1}}M_{t_j} \middle| \mathcal{F}_{t_j}\right] \\
&= \mathbb{E}\left[M_{t_{j+1}}^2 \middle| \mathcal{F}_{t_j}\right] + \mathbb{E}\left[M_{t_j}^2 \middle| \mathcal{F}_{t_j}\right] - 2\mathbb{E}\left[M_{t_{j+1}}M_{t_j} \middle| \mathcal{F}_{t_j}\right] \\
&= \mathbb{E}\left[M_{t_{j+1}}^2 \middle| \mathcal{F}_{t_j}\right] + M_{t_j}^2 - 2M_{t_j}\mathbb{E}\left[M_{t_{j+1}} \middle| \mathcal{F}_{t_j}\right] \\
&= \mathbb{E}\left[M_{t_{j+1}}^2 \middle| \mathcal{F}_{t_j}\right] + M_{t_j}^2 - 2M_{t_j}^2 \\
&= \mathbb{E}\left[M_{t_{j+1}}^2 - M_{t_j}^2 \middle| \mathcal{F}_{t_j}\right]. \qquad (2.3.7)
\end{aligned}$$

Thus, for any $j \in \{0, ..., n-1\}$ and using that $\{M_t^2 - \langle M\rangle_t \colon t \geq 0\}$ is a martingale and that e_j is $\mathcal{F}_{t_j}$-measurable as well,

$$\begin{aligned}
\mathbb{E}\left[e_j^2\left(\Delta M_j\right)^2\right] &= \mathbb{E}\left[e_j^2\mathbb{E}\left[\left(\Delta M_j\right)^2\big|\right]\right] \\
&= \mathbb{E}\left[e_j^2\mathbb{E}\left[M_{t_{j+1}}^2 - M_{t_j}^2\big|\mathcal{F}_{t_j}\right]\right] \\
&= \mathbb{E}\Big[e_j^2\Big(\mathbb{E}\left[M_{t_{j+1}}^2 - \langle M\rangle_{t_{j+1}} - \left(M_{t_j}^2 - \langle M\rangle_{t_j}\right)\big|\mathcal{F}_{t_j}\right] \\
&+ \mathbb{E}\left[\langle M\rangle_{t_{j+1}} - \langle M\rangle_{t_j}\big|\mathcal{F}_{t_j}\right]\Big)\Big] \\
&= \mathbb{E}\left[e_j^2\mathbb{E}\left[\langle M\rangle_{t_{j+1}} - \langle M\rangle_{t_j}\big|\mathcal{F}_{t_j}\right]\right] \\
&= \mathbb{E}\left[\mathbb{E}\left[e_j^2\left(\langle M\rangle_{t_{j+1}} - \langle M\rangle_{t_j}\right)\big|\mathcal{F}_{t_j}\right]\right] \\
&= \mathbb{E}\left[e_j^2\left(\langle M\rangle_{t_{j+1}} - \langle M\rangle_{t_j}\right)\right].
\end{aligned}$$

On the other hand, for each $0 \leq i < j < n$, we have, since e_i, e_j and ΔM_i are $\mathcal{F}_{t_j}$-measurable and M is a martingale,

$$\mathbb{E}\left[e_ie_j\Delta M_i\Delta M_j\right] = \mathbb{E}\left[e_ie_j\Delta M_i\mathbb{E}\left[\Delta M_j\big|\mathcal{F}_{t_j}\right]\right] = 0.$$

All in all,

$$\begin{aligned}
\mathbb{E}\left[\left(\int_0^T X_r dM_r\right)^2\right] &= \mathbb{E}\left[\left(\sum_{j=0}^{n-1} e_j\Delta M_j\right)^2\right] \\
&= \sum_{j=0}^{n-1}\mathbb{E}\left[e_j^2\left(\Delta M_j\right)^2\right] + 2\sum_{0\leq i<j<n}\mathbb{E}\left[e_ie_j\Delta M_i\Delta M_j\right] \\
&= \sum_{j=0}^{n-1}\mathbb{E}\left[e_j^2\left(\langle M\rangle_{t_{j+1}} - \langle M\rangle_{t_j}\right)\right] \\
&= \mathbb{E}\left[\sum_{j=0}^{n-1} e_j^2\int_0^T \mathbb{I}^2_{[t_j,t_{j+1})}(t)d\langle M\rangle_t\right] \\
&= \mathbb{E}\left[\int_0^T\sum_{j=0}^{n-1} e_j^2\mathbb{I}^2_{[t_j,t_{j+1})}(t)d\langle M\rangle_t\right] \\
&= \mathbb{E}\left[\int_0^T\left(\sum_{j=0}^{n-1} e_j\mathbb{I}_{[t_j,t_{j+1})}(t)\right)^2 d\langle M\rangle_t\right] \\
&= \mathbb{E}\left[\int_0^T X_t^2 d\langle M\rangle_t\right].
\end{aligned}$$

This establishes the isometry property for elementary processes.

For the quadratic variation, we start by noticing that the process $I(X) = \{I_t(X) \colon 0 \leq t \leq T\}$ is a martingale (by (iii) and the fact that the process is adapted, which can be seen from the definition) with $\mathbb{P}$-a.s. continuous sample paths (because of the continuity of M) and, by the isometry formula, it is square integrable, so the existence of its quadratic variation $\langle I(X)\rangle = \{\langle I(X)\rangle_t \colon 0 \leq t \leq T\}$ is guaranteed.

Now let us assume that, for instance, $0 \leq s < t \leq T$ are such that $t_{m-1} \leq s < t_m$ and $t_{k-1} \leq t < t_k$ with $0 \leq m \leq k \leq n$ and set $\sum_{j=m}^{k-2} e_j \Delta M_j$ to be 0 whenever $m > k-2$.

Then we have that, using similar arguments to the ones seen when proving the isometry formula and writing $\Delta\langle M\rangle_j = \langle M\rangle_{t_{j+1}} - \langle M\rangle_{t_j}$,

$$\begin{aligned}
&\mathbb{E}\left[(I_t(X) - I_s(X))^2 \,\middle|\, \mathcal{F}_s\right] \\
&= \mathbb{E}\left[\left(e_{m-1}\left(M_{t_m} - M_s\right) + \sum_{j=m}^{k-2} e_j \Delta M_j + e_{k-1}\left(M_t - M_{t_{k-1}}\right)\right)^2 \middle|\, \mathcal{F}_s\right] \\
&= \mathbb{E}\left[e_{m-1}^2\left(M_{t_m} - M_s\right)^2 + \sum_{j=m}^{k-2} e_j^2 \Delta M_j^2 + e_{k-1}^2\left(M_t - M_{t_{k-1}}\right)^2 \middle|\, \mathcal{F}_s\right] \\
&= \mathbb{E}\left[e_{m-1}^2\left(\langle M\rangle_{t_m} - \langle M\rangle_s\right) + \sum_{j=m}^{k-2} e_j^2 \Delta\langle M\rangle_j + e_{k-1}^2\left(\langle M\rangle_t - \langle M\rangle_{t_{k-1}}\right)^2 \middle|\, \mathcal{F}_s\right] \\
&= \mathbb{E}\left[\int_s^t X_r^2 d\langle M\rangle_r \middle| \mathcal{F}_s\right].
\end{aligned}$$

Thus, we have shown that

$$\mathbb{E}\left[(I_t(X) - I_s(X))^2 \,\middle|\, \mathcal{F}_s\right] = \mathbb{E}\left[\int_s^t X_r^2 d\langle M\rangle_r \middle| \mathcal{F}_s\right] \tag{2.3.8}$$

for any $0 \leq s \leq t \leq T$. Now, for $0 \leq s \leq t \leq T$, since $I(X)$ is a martingale, we have that, by the same computations to the ones performed in (2.3.7) and using that $\int_0^r X_u^2 d\langle M\rangle_u$ is $\mathcal{F}_r$-measurable for any $0 \leq r \leq T$ (one can see this by just using the definition of X),

$$\begin{aligned}
\mathbb{E}\left[I_t^2(X)\middle|\mathcal{F}_s\right] - I_s^2(X) &= \mathbb{E}\left[I_t^2(X) - I_s^2(X)\middle|\mathcal{F}_s\right] \\
&= \mathbb{E}\left[(I_t(X) - I_s(X))^2\middle|\mathcal{F}_s\right] \\
&= \mathbb{E}\left[\int_s^t X_r^2 d\langle M\rangle_r\middle|\mathcal{F}_s\right] \\
&= \mathbb{E}\left[\int_0^t X_r^2 d\langle M\rangle_r - \int_0^s X_r^2 d\langle M\rangle_r\middle|\mathcal{F}_s\right] \\
&= \mathbb{E}\left[\int_0^t X_r^2 d\langle M\rangle_r\middle|\mathcal{F}_s\right] - \int_0^s X_r^2 d\langle M\rangle_r.
\end{aligned}$$

Which can be rearranged to obtain that

$$\mathbb{E}\left[I_t^2(X) - \int_0^t X_r^2 d\langle M\rangle_r\middle|\mathcal{F}_s\right] = I_s^2(X) - \int_0^s X_r^2 d\langle M\rangle_r.$$

Hence, the process

$$\left\{\int_0^t X_r^2 d\langle M\rangle_r \colon 0 \le t \le T\right\}$$

is a process starting at 0, non-decreasing, adapted and continuous such that

$$\left\{I_t^2(X) - \int_0^t X_r^2 d\langle M\rangle_r \colon 0 \le t \le T\right\}$$

is a martingale. By uniqueness of the quadratic variation, (v) holds. $\square$

As in the case of the Brownian motion, the isometry property tells us that if $\{X^{(n)}\}_n \subset \mathcal{E}$ is convergent sequence with respect to the norm $||\cdot||_M$, then it is a Cauchy sequence of elementary processes, which, in turn implies that the sequence $\left\{I\left(X^{(n)}\right)\right\}_n$ is a Cauchy sequence in $L^2(\Omega)$. Since the latter is a Hilbert space as well, we conclude that there is an $L^2(\Omega)$-limit for the sequence $\left\{I\left(X^{(n)}\right)\right\}_n$, allowing us to make the following definition.

Definition 2.3.3. *Let $X \in \nu_M$ and $\{X^{(n)}\} \subset \mathcal{E}$ be a sequence of elementary processes such that $||X - X^{(n)}||_M \to 0$ as n approaches infinity, then the stochastic integral of X with respect to M is defined as*

$$I_T(X) = \int_0^T X_t dM_t := L^2(\Omega)\text{-}\lim_{n\to\infty} \int_0^T X_t^{(n)} dM_t.$$

For $0 \leq s \leq t \leq T$*, we define*

$$I_t(X) = \int_0^t X_r dM_r := \int_0^T \mathbb{I}_{[0,t)}(r) X_r dM_r$$

and

$$\int_s^t X_r dM_r := I_t(X) - I_s(X).$$

One can easily show that if X is in ν_M, then so is $X\mathbb{I}_{[0,t)(\cdot)}$, so the definition of $I_t(X)$ for $0 \leq t \leq T$ makes sense, and that

$$I_t(X) = \int_0^t X_r dM_r = L^2(\Omega)\text{-}\lim_{n\to\infty} \int_0^T \mathbb{I}_{[0,t)}(r) X_r^{(n)} dM_r.$$

One should observe that the stochastic integral process $\{I_t(X)\colon 0 \leq t \leq T\}$, being an $L^2(\Omega)$-limit of continuous processes, needs not to be continuous as well. However, it turns out that one can find a version whose sample paths are continuous. From now on, we shall consider such version.

Finally, and as it is usual, one shows that the definition of the stochastic integral does not depend on the approximating sequence and that many properties of the stochastic integral of elementary processes are also satisfied by general elements of X.

Theorem 2.3.6. *Let* $X \in \nu_M$ *and* $\{X^{(n)}\}, \{Y^{(n)}\} \subset \mathcal{E}$ *be two approximating sequences for* X *of elementary processes, that is, such that* $||X - X^{(n)}||_M \to 0$ *and* $||X - Y^{(n)}||_M \to 0$ *as* n *approaches infinity, then*

$$L^2(\Omega)\text{-} \int_0^T X_t^{(n)} dM_t = L^2(\Omega)\text{-} \int_0^T Y_t^{(n)} dM_t, \quad \textit{a.s.}$$

Proof. We have that

$$\begin{aligned} \left|\left|I_T\left(X^{(n)}\right) - I_T\left(Y^{(n)}\right)\right|\right|_{L^2(\Omega)} &= \left|\left|I_T\left(X^{(n)} - Y^{(n)}\right)\right|\right|_{L^2(\Omega)} \\ &= \left|\left|X^{(n)} - Y^{(n)}\right|\right|_M \\ &\leq \left|\left|X^{(n)} - X\right|\right|_M + \left|\left|X - Y^{(n)}\right|\right|_M \xrightarrow{n\to\infty} 0, \end{aligned}$$

where the isometry formula has been used. By the a.s. uniqueness of the $L^2(\Omega)$-limit, we conclude that the sequences $\left\{I_T\left(X^{(n)}\right)\right\}_n$ and $\left\{I_T\left(Y^{(n)}\right)\right\}_n$ have the same $L^2(\omega)$-limit a.s. $\square$

Theorem 2.3.7. *Properties (i)-(v) from Theorem 2.3.5 hold for any* $X \in \nu_M$.

Proof. For properties (i), (ii), one only needs to use the isometry formula for simple processes to obtain the desired results. The proofs are the same as in the case of the Brownian motion, so we will not repeat them here.

For property (iii), since the stochastic integral is defined as an $L^2(\Omega)$-limit of martingales (by Theorem 2.3.5), it is, in particular, an $L^1(\Omega)$-limit of martingales and hence, a martingale itself.

We now prove that Eq.(2.3.8) holds for any $X \in \nu_M$. To do so, let $A \in \mathcal{F}_s$, then we have that

$$\begin{aligned}
\mathbb{E}\left[\mathbb{I}_A \left(I_t(X) - I_s(X)\right)^2\right] &= \lim_{n\to\infty} \mathbb{E}\left[\mathbb{I}_A \left(I_t\left(X^{(n)}\right) - I_s\left(X^{(n)}\right)\right)^2\right] \\
&= \lim_{n\to\infty} \mathbb{E}\left[\mathbb{I}_A \int_s^t \left(X_r^{(n)}\right)^2 d\langle M\rangle_r\right] \\
&= \mathbb{E}\left[\mathbb{I}_A \int_s^t X_r^2 d\langle M\rangle_r\right].
\end{aligned}$$

The first equality follows from the fact that $I_t(X)$ and $I_s(X)$ are the $L^2(\Omega)$-limits of the corresponding sequences, whilst the last one follows from

$$\mathbb{E}\left[\int_s^t \left|X_r^{(n)} - X_r\right|^2 d\langle M\rangle_r\right] \le \left|\left|X^{(n)} - X\right|\right|_M^2 \xrightarrow{n\to\infty} 0.$$

Meaning that $\mathbb{I}_{[s,t)}(\cdot)X^{(n)}_{\cdot}$ converges to $\mathbb{I}_{[s,t)}(\cdot)X_{\cdot}$ in norm $||\cdot||_M$ and hence, that their respective norms converge as desired.
Since this holds for any $A \in \mathcal{F}_s$, (2.3.8) is established as well for $X \in \nu_M$.

Setting $s = 0$ and taking expectations in both sides of (2.3.8) proves that

$$\mathbb{E}\left[I_t^2(X)\right] = \int_0^t X_r^2 d\langle M\rangle_r$$

for any $0 \le t \le T$. In particular, for $t = T$, we obtain (iv). The exact same argument for (v) in Theorem 2.3.5 would work to prove this property for $X \in \nu_M$ if it were not for the fact that now there is no guarantee that the integral $\int_0^s X_r^2 d\langle M\rangle_r$ is $\mathcal{F}_s$-measurable. However, since X is $\mathcal{B} \times \mathcal{F}$-measurable and adapted, it has a progressively measurable modification, which gives a representative X' in ν_M which is progressively measurable (that is, $||X - X'||_M = 0$). For such representative, the integral $\int_0^s (X_r')^2\, d\langle M\rangle_r$ is indeed $\mathcal{F}_s$-measurable. □

2.3.3 Further results

In this section we present some results concerning the computation of stochastic integrals and their variations and covariations. In most cases, we will just announce them without proof, as the main purpose of the remaining section is to highlight the basic results focusing on the ideas.

The product rule

Theorem 2.3.8. *Let* $M = \{M_t \colon 0 \le t \le T\}$ *and* $N = \{N_t \colon 0 \le t \le T\}$ *be two continuous square integrable martingales, then, for each* $0 \le t \le T$, *the following holds* $\mathbb{P}$*-a.s.*

$$M_t N_t = M_0 N_0 + \int_0^t N_s dM_s - \int_0^t M_s dN_s + \langle M, N\rangle_t$$

or, in differential form,

$$d\left(M_t N_t\right) = N_t dM_t + M_t dN_t + d\langle M, N\rangle_t.$$

Proof. We prove the result for $M = N$. The general result will follow by polarization; that is,

$$M_t N_t = \frac{1}{4}\left[(M_t + N_t)^2 - (M_t - N_t)^2\right].$$

Now, fix $t \in [0, T]$ and consider a family of partitions of the interval $[0, t]$ of the form $\{0 = t_0 < ... < t_n = t\}$ so that its mesh converges to zero. Then we have that, by setting $\Delta M_j = M_{t_{j+1}} - M_{t_j}$ for $j \in \{0, 1, ..., n-1\}$,

$$M_t^2 - M_0^2 = \sum_{j=0}^{n-1}\left(M_{t_{j+1}}^2 - M_{t_j}^2\right) = 2\sum_{j=0}^{n-1} M_{t_j}\Delta M_j + \sum_{j=0}^{n-1}\left(\Delta M_j\right)^2.$$

Now, since M is a continuous square integrable martingale, we have that $M \in \nu_M$. Moreover, because of the continuity of M (which implies that it is bounded in $[0, t]$) and the fact that it is adapted, the sequence of processes $\phi^{(n)}$ defined by

$$\phi_s^{(n)} = \sum_{j=0}^{n-1} M_{t_j}\mathbb{I}_{[t_j, t_{j+1})}(s)$$

constitutes a sequence of elementary processes that approximate in $||\cdot||_M$ norm the process $\mathbb{I}_{[0,t)}(\cdot)M$. (using the continuity and boundedness, which allows us to use the dominated convergence theorem). Moreover, we also have that

$$\int_0^t \phi_s^{(n)} dM_s = \sum_{j=0}^{n-1} M_{t_j}\Delta M_j$$

so we can conclude that

$$\sum_{j=0}^{n-1} M_{t_j}\Delta M_j \to \int_0^t M_s dM_s$$

as the mesh goes to zero and where the convergence is in $L^2(\Omega)$. On the other hand, Theorem 2.3.2 tells us that $\sum_{j=0}^{n-1}\left(\Delta M_j\right)^2$ converges, in probability, to the quadratic variation $\langle M\rangle_t$. □

Observe that the differential form of this result resembles a bit the usual product rule, however, due to the fact that the sample paths have non-vanishing quadratic variation, higher order terms must be taken into account. In this case it translates into taking into account the covariation of the processes. This result can be seen as a particular case of the much more general result known as the Itô formula, which can be seen as a generalization of the chain rule to functions of unbounded variation (and finite quadratic variation).

The Itô formula

Theorem 2.3.9 (Itô formula). *Let $f\colon [0,T]\times\mathbb{R}^p \to \mathbb{R}$, $(t,x)\mapsto f(t,x)$, be a $\mathcal{C}^{1,2}$ function ($\mathcal{C}^1$ with respect to the first variable and $\mathcal{C}^2$ with respect to the second one) and $M = \{M_t = (M_t^{(1)},...,M_t^{(p)})\colon 0\le t\le T\}$ be a vector of continuous square integrable martingales, then the following holds $\mathbb{P}$-a.s.*

$$f(t,M_t) = f(0,M_0) + \int_0^t \frac{\partial f}{\partial t}(s,M_s)ds + \sum_{i=1}^p \int_0^t \frac{\partial f}{\partial x_i} f(s,M_s)dM_s^{(i)} + \frac{1}{2}\sum_{1\le i,j\le p}\int_0^t \frac{\partial^2 f}{\partial x_i \partial x_j}(s,M_s)d\langle M^{(i)},M^{(j)}\rangle_s \tag{2.3.9}$$

or, in differential form

$$df(t,M_t) = \frac{\partial f}{\partial t}(t,M_t)dt + \sum_{i=1}^p \frac{\partial f}{\partial x_i} f(t,M_t)dM_t^{(i)} + \frac{1}{2}\sum_{1\le i,j\le p} \frac{\partial^2 f}{\partial x_i\partial x_j}(t,M_t)d\langle M^{(i)},M^{(j)}\rangle_t. \tag{2.3.10}$$

Sketch of the proof for the case $p=1$. As in the case of the Brownian motion, fix $t\in[0,T]$ and consider a family of partitions of the form $\Pi = \{0 = t_0 < ... < t_n = t\}$ so that its mesh $|\Pi|$ tends to zero. We have that, by Taylor's formula

$$\begin{aligned} f(t,M_t) - f(0,M_0) &= \sum_{k=0}^{n-1} f(t_{k+1},M_{t_{k+1}}) - f(t_k,M_{t_k}) \\ &= \sum_{k=0}^{n-1}\frac{\partial f}{\partial t}(\xi_k,M_{t_k})\Delta t_k + \sum_{k=0}^{n-1}\frac{\partial f}{\partial x}(t_k,M_{t_k})\Delta M_k + \frac{1}{2}\sum_{k=0}^{n-1}\frac{\partial^2 f}{\partial x^2}(t_k,\eta_k)\left(\Delta M_k\right)^2. \end{aligned}$$

where, $\Delta t_k = t_{k+1}-t_k$, $\Delta M_k = M_{t_{k+1}} - M_{t_k}$ and $\xi_k \in (t_k,t_{k+1})$ and $\eta_k \in \left(M_{t_k}\wedge M_{t_{k+1}}, M_{t_k}\vee M_{t_{k+1}}\right)$ are the intermediate values given by Taylor's theorem. Given that $\partial_t f$ and M are continuous, the map $s\mapsto \partial_t f(s,X_s)$ is also continuous, meaning that

$$\left|\sum_{k=0}^{n-1}\frac{\partial f}{\partial t}(\xi_k, M_{t_k})\Delta t_k - \sum_{k=0}^{n-1}\frac{\partial f}{\partial t}(t_k, M_{t_k})\Delta t_k\right| \longrightarrow 0, \quad \mathbb{P}\text{-a.s.}$$

as the mesh of the partition goes to zero. Since

$$\sum_{k=0}^{n-1}\frac{\partial f}{\partial t}\Delta t_k \xrightarrow{|\Pi|\to 0} \int_0^t \frac{\partial f}{\partial t}(s, M_s)ds,$$

we conclude that

$$\sum_{k=0}^{n-1}\frac{\partial f}{\partial t}(\xi_k, M_{t_k})\Delta t_k \xrightarrow{|\Pi|\to 0} \int_0^t \frac{\partial f}{\partial t}(s, M_s)ds$$

as well.

Now, given that $\partial_x f$ is continuous, we can use the same arguments seen in the previous theorem to see that

$$\sum_{k=0}^{n-1}\frac{\partial f}{\partial x}(t_k, M_k)\Delta M_k \longrightarrow \int_0^t \frac{\partial f}{\partial x}(s, M_s)dM_s$$

as the mesh of the partition goes to zero and where the limit is in $L^2(\Omega)$.

Using the continuity of $\partial_x^2 f$, we see that

$$\left|\sum_{k=0}^{n-1}\frac{\partial^2 f}{\partial x^2}(t_k, \eta_k)\left(\Delta M_k\right)^2 - \sum_{k=0}^{n-1}\frac{\partial^2 f}{\partial x^2}(t_k, M_k)\left(\Delta M_k\right)^2\right| \longrightarrow 0, \quad \mathbb{P}\text{-a.s.}$$

as the mesh goes to zero. The then goal is to show that

$$\left|\sum_{k=0}^{n-1}\frac{\partial^2 f}{\partial x^2}(t_k, M_k)\Delta\langle M\rangle_k - \sum_{k=0}^{n-1}\frac{\partial^2 f}{\partial x^2}(t_k, M_k)\left(\Delta M_k\right)^2\right| \xrightarrow{|\Pi|\to 0} 0 \tag{2.3.11}$$

in some suitable sense. With this and the fact that

$$\sum_{k=0}^{n-1}\frac{\partial^2 f}{\partial x^2}(t_k, M_k)\Delta\langle M\rangle_k \xrightarrow{|\Pi|\to 0} \int_0^t \frac{\partial^2 f}{\partial x^2}(s, M_s)d\langle M\rangle_s$$

one then concludes the proof. The proof of (2.3.11) is based on a couple of technical lemmas regarding the bounds of the quadratic variation and the variation of order four over finite partitions that can be found in [7], Lemmas 1.5.9 and 1.5.10, page 33. □

The covariation formula

In the following chapter, in some occasions we will come across situations where we will have to compute covariations of stochastic with respect to different continuous square integrable martingales. This last result will prove to be useful in such situations.

Moreover, as seen in Proposition 3.2.19, page 144, of [7], this results can be used to characterize the stochastic integral in terms of its quadratic variation, which is a Lebesgue-Stieltjes integral, a much more familiar object.

Theorem 2.3.10. *Let* $M = \{M_t \colon 0 \leq t \leq T\}$ *and* $N = \{N_t \colon 0 \leq t \leq T\}$ *be two continuous square integrable martingales,* $X = \{X_t \colon 0 \leq t \leq T\} \in \nu_M$, $Y = \{Y_t \colon 0 \leq t \leq T\} \in \nu_N$ *be two integrable processes (with respect to the corresponding martingales) and*

$$I_t^M(X) \coloneqq \int_0^t X_s dM_s, \quad I_t^N(Y) \coloneqq \int_0^t Y_s dN_s$$

be the corresponding stochastic integrals. Then the following equality holds $\mathbb{P}$*-a.s.:*

$$\left\langle I^M(X), I^N(Y) \right\rangle_t = \int_0^t X_s Y_s d\langle M, N\rangle_s.$$

Computations like the ones done in Theorem 2.3.5 or Theorem 2.3.7 show that the result does indeed hold for elementary processes X and Y. However, in this case we need to work a bit more to obtain the results for general processes X and Y.

The proof of this result is a consequence of a couple of previously established results: Proposition 3.2.14 and Lemmas 3.2.15 and 3.2.16, pages 142-144 in [7].

Lemma 2.3.2. *Let* M, N *and* X *be as in Theorem 2.3.10, then*

$$\left\langle I^M(X), N \right\rangle_t = \int_0^t X_s d\langle M, N\rangle_s, \quad \mathbb{P}\text{-a.s.}$$

for each $t \in [0, T]$.

Proof of Theorem 2.3.10. Recall that if $X \in \nu_M$, then $I^M(X)$ is a continuous square integrable martingale and similarly for $Y \in \nu_N$. Hence, we have that, by the previous lemma, $d\langle I^M(X), I^N(Y)\rangle_t = X_s d\langle M, I^N(Y)\rangle_t$ and $d\langle M, I^N(Y)\rangle_s = Y_s d\langle M, N\rangle_s$, so

$$\left\langle I^M(X), I^N(Y) \right\rangle_t = \int_0^t X_s d\left\langle M, I^N(Y)\right\rangle_t = \int_0^t X_s Y_s d\langle M, N\rangle_s.$$

Where the equalities hold $\mathbb{P}$-a.s. □

2.3.4 An approximation result

As discussed in the introduction and in Section 2.1, we are interested in giving conditions under which the considered random noises converge (in law, weakly) to the desired noise. In this chapter, the noises considered are represented as stochastic integrals of some processes $f = \{f_t : t \in [0, T]\}$ with respect to some other process, so the problem can be formulated as follows:

Let $W^{(n)} = \left\{W_t^{(n)} : t \in [0, T]\right\}$, $n \in \mathbb{N}$ be a sequence of stochastic processes converging in law to a certain stochastic process $W = \{W_t : t \in [0, T]\}$ as n approaches infinity and assume that the integrals

$$I_t^{(n)} := \int_0^t f_s dW_s^{(n)}, \quad I_t := \int_0^t f_s dW_s$$

are well defined in some sense (maybe as an $L^2(\Omega)$ limit as studied in this chapter or ω by ω by using the Lebesgue-Stieltjes integral). Under which conditions for the process f, the sequence of processes $I^{(n)} = \left\{I_t^{(n)} : t \in [0, T]\right\}$ converge in law (or in some other sense) to the process $I = \{I_t : t \in [0, T]\}$?

Motivated by Donsker's invariance principle, which states that the Brownian motion can be weakly approximated by random walks whose jumps are of finite variance, we will be considering the case where the processes $W^{(n)}$ are given by

$$W_t^{(n)} = \frac{1}{\sqrt{n}} \left[\sum_{j=1}^{[nt]} X_j + (nt - [nt]) X_{[nt]+1} \right], \quad W_0^{(n)} = 0. \tag{2.3.12}$$

Where the sum is taken to be zero whenever whenever $[nt] = 0$ and $\{X_j\}_{j \in \mathbb{N}}$ is a sequence of i.i.d. centered random variables with unitary variance. We will assume, as well, that $\mathbb{E}\left[X_1^4\right] < \infty$. These processes, which are the linear interpolation of a random walk and thus, have continuous sample paths, are known to converge in law to the standard Brownian motion, which we will be denoting by $W = \{W_t : t \in [0, T]\}$, in the space of continuous functions $\mathcal{C}([0, T])$.

Moreover, the sample paths of the process $W^{(n)}$ are of bounded variation, meaning that the integrals $I_t^{(n)}$ can be thought as Lebesgue-Stieltjes integrals for each $\omega \in \Omega$, and this is what we will do. Not only that, but the fact that the sample paths of $W^{(n)}$ are continuous implies, by properties of the Lebesgue-Stieltjes integral, that the sample paths of $I^{(n)}$ are continuous as well whenever f is good enough (for instance, if f has continuous sample paths as well).

A positive result can then be given when, for instance, f is a continuous deterministic process, and this is what we will be proving in this section. To do so, we will use some results regarding the weak convergence of probability measures in metric spaces (mainly, Prohorov's Theorem, Theorems 5.1 and 5.2 in [3], pages 59 and 60, and Billingsley's Criterion, Theorem 12.3 in [4], page 95). For a brief introduction to these results and the techniques that will be used, we refer to [5].

Theorem 2.3.11. *Let $W^{(n)}$ be as in (2.3.12), W be a Brownian motion and f be a continuous deterministic process. Then the integral processes $I^{(n)}$ converge in law to the integral process I, in the space of continuous functions $\mathcal{C}([0,T])$, as n approaches infinity.*

Proof. To prove this result, we will follow the exact same strategy as in Section 3 of [5]. That is, we will be proving the tightness of the sequence $\{I^{(n)}\}_n$ and the converge of the finite dimensional distributions.

Tightness

To prove the tightness, we consider the alternative expression for (2.3.12) which is given by the Donsker kernel:

$$W_t^{(n)} = \frac{1}{\sqrt{n}} \int_0^{ns} \theta(x)dx, \quad \theta(x) = \sum_{j=1}^{\infty} X_j \mathbb{I}_{[j-1,j)}(x).$$

This expression allows us to write $I_t^{(n)}$ as

$$I_t^{(n)} = \sqrt{n} \int_0^t \theta(ns) f_s \, ds,$$

which is a usual Lebesgue integral.

Clearly, the sequence $\{I_0^{(n)}\}_n$ is tight as a sequence of random variables (in $\mathbb{R}$) because $I_0^{(n)} = 0$ for all $n \in \mathbb{N}$. So, by Billingsley's Criterion, to prove the tightness of the sequence $\{I^{(n)}\}_n$ it suffices to show that

$$\mathbb{E}\left[\left| I_t^{(n)} - I_s^{(n)} \right|^\gamma \right] \leq |F(t) - F(s)|^{1+\alpha}$$

for all $s, t \in [0,T]$, for some constants $\gamma \geq 0$, $\alpha > 0$ and for some non-decreasing function F. We will show the latter for $\gamma = 4$, $\alpha = 1$ and $F(t) = Ct$, $t \in [0,T]$ for some positive constant C independent of s, t and n. The result trivially holds for $s = t$, so from now on we shall assume $0 \leq s < t \leq T$.

First observe that, by setting $||f||_\infty = \sup\{|f_t| \colon t \in [0,T]\}$,

$$\begin{aligned}
\mathbb{E}\left[\left|I_t^{(n)} - I_s^{(n)}\right|^4\right] &= n^2 \mathbb{E}\left[\left(\int_s^t \theta(nu) f_u \, du\right)^4\right] \\
&= 24 n^2 \mathbb{E}\left[\int_{[s,t]^4} \theta(nu_1) \cdot ... \cdot \theta(nu_4) f_{u_1} \cdot ... \cdot f_{u_4} \mathbb{I}_{\{u_1 \le ... \le u_4\}}(\boldsymbol{u}) \, d\boldsymbol{u}\right] \\
&= 24 n^2 \int_{[s,t]^4} \mathbb{E}\left[\theta(nu_1) \cdot ... \cdot \theta(nu_4)\right] f_{u_1} \cdot ... \cdot f_{u_4} \mathbb{I}_{\{u_1 \le ... \le u_4\}}(\boldsymbol{u}) \, d\boldsymbol{u} \\
&\le 24 n^2 \int_{[s,t]^4} \left|\mathbb{E}\left[\theta(nu_1) \cdot ... \cdot \theta(nu_4)\right]\right| \left|f_{u_1} \cdot ... \cdot f_{u_4}\right| \mathbb{I}_{\{u_1 \le ... \le u_4\}}(\boldsymbol{u}) \, d\boldsymbol{u} \\
&\le 24 ||f||_\infty^4 n^2 \int_{[s,t]^4} \left|\mathbb{E}\left[\theta(nu_1) \cdot ... \cdot \theta(nu_4)\right]\right| \mathbb{I}_{\{u_1 \le ... \le u_4\}}(\boldsymbol{u}) \, d\boldsymbol{u} \\
&= 24 ||f||_\infty^4 n^2 \int_{[s,t]^4} \left|\mathbb{E}\left[\theta(nu_1) \cdot ... \cdot \theta(nu_4) \mathbb{I}_{\{u_1 \le ... \le u_4\}}(\boldsymbol{u})\right]\right| \, d\boldsymbol{u}. \qquad (2.3.13)
\end{aligned}$$

Where $\boldsymbol{u} = (u_1, ..., u_4)$, $d\boldsymbol{u} = du_1 du_2 du_3 du_4$ and in the second line we have used that the product $\theta(nu_1) \cdot ... \cdot \theta(nu_4) f_{u_1} \cdot ... \cdot f_{u_4}$ remains invariant under permutations of the variables u_1, u_2, u_3, u_4 within the domain of integration considered, $[s,t]^4$, to fix an ordering by making use of the indicator function $\mathbb{I}_{\{u_1 \le ... \le u_4\}}$.

Now observe that

$$\begin{aligned}
&\mathbb{E}\left[\theta(nu_1) \cdot ... \cdot \theta(nu_4) \mathbb{I}_{\{u_1 \le ... \le u_4\}}(\boldsymbol{u})\right] \\
&\quad = \sum_{j_1, ..., j_4} \mathbb{E}\left[X_{j_1} \cdot ... \cdot X_{j_4}\right] \mathbb{I}_{[j_1 - 1, j_1)}(nu_1) \cdot ... \cdot \mathbb{I}_{[j_4 - 1, j_4)}(nu_4) \mathbb{I}_{\{u_1 \le ... \le u_4\}}(\boldsymbol{u}) \\
&\quad = \sum_{i,j} \left[(M_4 - 1)\delta_{ij} + 1\right] \mathbb{I}_{[i-1,i)^2}(nu_1, nu_2) \mathbb{I}_{[j-1,j)^2}(nu_3, nu_4) \mathbb{I}_{\{u_1 \le ... \le u_4\}}(\boldsymbol{u}). \qquad (2.3.14)
\end{aligned}$$

Where $M_4 = \mathbb{E}\left[X_1^4\right]$. Indeed, in the first sum, if some of the j_i is different from the others, then the fact that the random variable X_{j_i} will be independent of the other three and that they are centered will imply that

$$\mathbb{E}\left[X_{j_1} \cdot ... \cdot X_{j_4}\right] = 0.$$

On the other hand, if $j_1 = j_2 \neq j_3 = j_4$, then

$$\mathbb{E}\left[X_{j_1} \cdot ... \cdot X_{j_4}\right] = \mathbb{E}\left[X_{j_1}^2\right] \mathbb{E}\left[X_{j_3}^2\right] = 1.$$

Finally, if $j_1 = ... = j_4$,

$$\mathbb{E}\left[X_{j_1} \cdot ... \cdot X_{j_4}\right] = \mathbb{E}\left[X_{j_1}^4\right] = M_4.$$

All other cases are discarded due to the presence of the indicator $\mathbb{I}_{\{u_1 \leq \ldots \leq u_4\}}$. Observe that, by the Cauchy-Schwarz inequality, $M_4 \geq \left(\mathbb{E}\left[X_1^2\right]\right)^2 = 1$, so the summands in (2.3.14) are non-negative and thus, (2.3.14) will be non-negative, so we can get rid of the absolute value in (2.3.13).

From here, one can follow the exact same procedure described in Theorem 3.1 in [5] to obtain the following bound

$$\mathbb{E}\left[\left|I_t^{(n)} - I_s^{(n)}\right|^4\right] \leq 24||f||_\infty^4 M_4 (t-s)^2,$$

and hence, proving the tightness of $\{I^{(n)}\}_n$.

Convergence of the finite dimensional distributions

Let $0 \leq t_1 < t_2 < \ldots < t_m \leq T$, we will show that the random vectors

$$I_f^{(n)} := \left(I_{t_1}^{(n)}, I_{t_2}^{(n)} - I_{t_1}^{(n)}, \ldots, I_{t_m}^{(n)} - I_{t_{m-1}}^{(n)}\right), \quad n \in \mathbb{N}$$

converge in law to the random vector

$$I_f := \left(I_{t_1}, I_{t_2} - I_{t_1}, \ldots, I_{t_m} - I_{t_{m-1}}\right)$$

as n approaches infinity. From now on, we will assume that $t_1 > 0$ and consider the point $t_0 = 0$ (if $t_1 = 0$ one can proceed similarly without having to consider the point t_0).

Now fix any $i \in \{1, \ldots, m\}$ and $s \in [t_{i-1}, t_i]$. Observe that, if $j \leq [nt_{i-1}]$, then $j \leq nt_{i-1}$ as well (since $[nt_{i-1}] \leq nt_{i-1}$). In particular, since $s \geq t_{i-1}$, we will have that $j \leq ns$, so $\mathbb{I}_{[j-1,j)}(ns) = 0$ for such values of j.
Similarly, if $j > [nt_i] + 1$, we will have that $j - 1 > [nt_i]$. That is, $j - 1$ is an integer strictly greater than the entire part of nt_i and thus, $j - 1 > nt_i$. Since $s \leq t_i$, we will have $j - 1 > ns$, meaning that $\mathbb{I}_{[j-1,j)}(ns) = 0$ for such values of j.
All in all, for these values of s we will have that

$$\theta(ns) = \sum_{j=[nt_{i-1}]+1}^{[nt_i]+1} X_j \mathbb{I}_{[j-1,j)}(ns)$$

and hence,

$$I_{t_i}^{(n)} - I_{t_{i-1}}^{(n)} = \sqrt{n} \int_{t_{i-1}}^{t_i} \theta(ns) f_s \, ds = \sqrt{n} \sum_{j=[nt_{i-1}]+1}^{[nt_i]+1} X_j C_{i,j}^{(n)},$$

where

$$\begin{aligned}
C_{i,j}^{(n)} &:= \int_{t_{i-1}}^{t_i} \mathbb{I}_{[j-1,j)}(ns) f_s ds \\
&= \int_{t_{i-1}\vee\frac{j-1}{n}}^{t_i\wedge\frac{j}{n}} f_s ds \\
&= f_y \left[\left(t_i \wedge \frac{j}{n} \right) - \left(t_{i-1} \vee \frac{j-1}{n} \right) \right] \qquad (2.3.15)
\end{aligned}$$

for some $y \in \left[\left(t_{i-1} \vee \frac{j-1}{n} \right), \left(t_i \wedge \frac{j}{n} \right) \right]$ by the integral mean value theorem (we omit the dependence on i, j and n in y). Observe that

$$\left| C_{i,j}^{(n)} \right| \leq \frac{||f||_\infty}{n} \qquad (2.3.16)$$

which is a uniform bound in both $i \in \{1, ..., m\}$ and $j \in \{[nt_{i-1}] + 1, ..., [nt_i] + 1\}$.

For each $i \in \{1, ..., m\}$, let us consider the random variable

$$\Delta_i^{(n)} = \sqrt{n} \sum_{j=[nt_{i-1}]+1}^{[nt_i]} X_j C_{i,j}^{(n)}$$

where the sum is understood to be zero when $[nt_{i-1}] = [nt_i]$. Then we have that, for any $\varepsilon > 0$, by Chebyshev's inequality,

$$\begin{aligned}
\mathbb{P}\left\{ \left| I_{t_i}^{(n)} - I_{t_{i-1}}^{(n)} - \Delta_i^{(n)} \right| > \varepsilon \right\} &= \mathbb{P}\left\{ \left| C_{i,[nt_i]+1}^{(n)} X_{[nt_i]+1} \sqrt{n} \right| > \varepsilon \right\} \\
&\leq n \frac{\left(C_{i,[nt_i]+1}^{(n)} \right)^2}{\varepsilon^2} \\
&\leq \frac{||f||_\infty^2}{n\varepsilon^2} \xrightarrow{n\to\infty} 0.
\end{aligned}$$

So, if we set $\Delta^{(n)} := \left(\Delta_1^{(n)}, ..., \Delta_m^{(n)} \right)$, we have that

$$\mathbb{P}\left\{ \left|\left| I_f^{(n)} - \Delta^{(n)} \right|\right| > \varepsilon \right\} \leq \sum_{i=1}^m \mathbb{P}\left\{ \left| I_{t_i}^{(n)} - I_{t_{i-1}}^{(n)} - \Delta_i^{(n)} \right| > \frac{\varepsilon}{\sqrt{m}} \right\} \xrightarrow{n\to\infty} 0.$$

Implying that the vectors $I_f^{(n)}$ and $\Delta^{(n)}$ have the same limit in law. However, observe that, for each $n \in \mathbb{N}$, since the random variables $\{X_j\}_j$ are pairwise independent, the components of the vector $\Delta^{(n)}$ are pairwise independent as well. Therefore, the characteristic function of $\Delta^{(n)}$, $\varphi_n(\boldsymbol{u})$, $\boldsymbol{u} = (u_1, ..., u_m) \in \mathbb{R}^m$, will be given by

$$\varphi_n(\boldsymbol{u}) = \prod_{l=1}^{m} \mathbb{E}\left[\exp\left\{iu_l\Delta_l^{(n)}\right\}\right]. \tag{2.3.17}$$

Given that the $\{X_j\}_j$ are pairwise independent, we have that

$$\mathbb{E}\left[\exp\left\{iu\Delta_l^{(n)}\right\}\right] = \prod_{j=[nt_{l-1}]+1}^{[nt_l]} \mathbb{E}\left[\exp\left\{iuY_{l,j}^{(n)}\right\}\right], \tag{2.3.18}$$

where $Y_{l,j}^{(n)} := \sqrt{n}XC_{l,j}^{(n)}$, $u \in \mathbb{R}$, X has the same law as X_j for all j and the product is understood to be one when $[nt_{l-1}] = [nt_l]$. Now let

$$\eta_{l,j}^{(n)}(u) := \mathbb{E}\left[\exp\left\{iuY_{l,j}^{(n)}\right\}\right],$$

since X has unitary variance (in particular, it has finite second order moment), $\eta_{l,j}^{(n)}(u)$ is a twice differentiable function with respect to u. Moreover, we have that

$$\eta_{l,j}^{(n)}(0) = 1, \quad \frac{d\eta_{l,j}^{(n)}}{du}(0) = i\mathbb{E}\left[Y_{l,j}^{(n)}\right] = 0$$

and

$$\frac{d^2\eta_{l,j}^{(n)}}{du^2}(u) = -\mathbb{E}\left[\left(Y_{l,j}^{(n)}\right)^2 \exp\left\{iuY_{l,j}^{(n)}\right\}\right].$$

In particular,

$$\frac{d^2\eta_{l,j}^{(n)}}{du^2}(0) = -n\left(C_{l,j}^{(n)}\right)^2. \tag{2.3.19}$$

Thus, a Taylor expansion yields, for any $u \in \mathbb{R}$,

$$\eta_{l,j}^{(n)}(u) = 1 + \frac{u^2}{2}\frac{d^2\eta_{l,j}^{(n)}}{du^2}(u_0) \tag{2.3.20}$$

for some $u_0 \in [0, u]$ if $u \geq 0$ or some $u_0 \in [u, 0]$ if $u \leq 0$. By the bound found in (2.3.16), we have that

$$\left|\frac{d^2\eta_{l,j}^{(n)}}{du^2}(u_0)\right| \leq \frac{||f||_\infty^2}{n}\mathbb{E}\left[X^2\right] = \frac{||f||_\infty^2}{n} \xrightarrow{n\to\infty} 0.$$

Thus, for n large enough and for fixed $u \in \mathbb{R}$, the natural logarithm of (2.3.20) will be well defined and continuous in a neighbourhood of (2.3.20) and thus, it will make sense to consider the logarithm of (2.3.18) which will be given by

$$L_l^{(n)}(u) := \sum_{j=[nt_{l-1}]+1}^{[nt_l]} \log \left[1 + \frac{u^2}{2} \frac{d^2 \eta_{l,j}^{(n)}}{du^2}(u_0) \right].$$

Using the Taylor expansion for the logarithm (considering, if needed, n large enough so that $\left| \frac{u^2}{2} \frac{d^2 \eta_{l,j}^{(n)}}{du^2}(u_0) \right| < 1$), we will have that

$$\left| L_l^{(n)}(u) - \frac{u^2}{2} \sum_{j=[nt_{l-1}]+1}^{[nt_l]} \frac{d^2 \eta_{l,j}^{(n)}}{du^2}(u_0) \right| \leq \frac{1}{2} \sum_{j=[nt_{l-1}]+1}^{[nt_l]} \left(\xi_{l,j}^{(n)} \right)^2 \tag{2.3.21}$$

where

$$\left| \xi_{l,j}^{(n)} \right| \leq \frac{u^2}{2} \left| \frac{d^2 \eta_{l,j}^{(n)}}{du^2}(u_0) \right| \leq \frac{u^2}{2} \frac{||f||_\infty^2}{n}.$$

Thus (2.3.21) can be bounded by

$$\frac{u^4}{8} \frac{||f||_\infty^4}{n^2} \left([nt_l] - [nt_{l-1}] \right) \xrightarrow{n \to \infty} 0.$$

On the other hand, using that $|e^{ia} - 1| \leq 2|a|$ for any $a \in \mathbb{R}$, we have that

$$\begin{aligned}
\left| \frac{d^2 \eta_{l,j}^{(n)}}{du^2}(u_0) - \frac{d^2 \eta_{l,j}^{(n)}}{du^2}(0) \right| &\leq \mathbb{E} \left[\left(Y_{l,j}^{(n)} \right)^2 \left| \exp\left\{ iu_0 Y_{l,j}^{(n)} \right\} - 1 \right| \right] \\
&\leq 2|u_0| \mathbb{E} \left[\left| Y_{l,j}^{(n)} \right|^3 \right] \\
&\leq 2|u| \, ||f||_\infty^3 n^{-3/2} M_3,
\end{aligned}$$

where $M_3 := \mathbb{E}\left[|X|^3\right]$. With this, we have that

$$\left| \sum_{j=[nt_{l-1}]+1}^{[nt_l]} \left(\frac{d^2 \eta_{l,j}^{(n)}}{du^2}(u_0) - \frac{d^2 \eta_{l,j}^{(n)}}{du^2}(0) \right) \right| \leq 2|u| ||f||_\infty^2 M_3 \, n^{-3/2} \left([nt_l] - [nt_{l-1}] \right) \xrightarrow{n \to \infty} 0.$$

Thus,

$$\lim_{n \to \infty} L_l^{(n)}(u) = \frac{u^2}{2} \lim_{n \to \infty} \sum_{j=[nt_{l-1}]+1}^{[nt_l]} n \frac{d^2 \eta_{l,j}^{(n)}}{du^2}(0).$$

Now observe that, for each $j \in \{[nt_{l-1}]+1, ..., [nt_l]\}$, we have that $j \leq [nt_l] \leq nt_l$, so $\frac{j}{n} \wedge t_l = j$ for these values of j. On the other hand, if $j = [nt_{l-1}]+1$, we have that $\frac{j-1}{n} = \frac{[nt_{l-1}]}{n} \leq t_{l-1}$, so $\frac{j-1}{n} \vee t_{l-1} = t_{l-1}$. Finally, if $j-1 > [nt_{l-1}]$, we will have $j-1 > nt_{l-1}$, leading to $\frac{j-1}{n} \vee t_{l-1} = j-1$. All in all, we obtain that (2.3.15) reduces to

$$C_{l,j}^{(n)} = f_y \frac{[nt_{l-1}]+1-t_{l-1}n}{n}, \quad \text{if} \quad j = [nt_{l-1}]+1$$

or

$$C_{l,j}^{(n)} = \frac{f_y}{n}, \quad \text{if} \quad j > [nt_{l-1}]+1.$$

Bearing this in mind and using (2.3.15) and (2.3.19), one sees that

$$\sum_{j=[nt_{l-1}]+1}^{[nt_l]} n \frac{d^2 \eta_{l,j}^{(n)}}{du^2}(0) = -f_y^2 \frac{([nt_{l-1}]+1-nt_{l-1})^2}{n} - \frac{1}{n} \sum_{j=[nt_{l-1}]}^{[nt_l]} f_y^2$$

where, remember, y depended on l, j and n. Observe that, for each n and l, $[nt_{l-1}] + 1 - nt_{l-1} \in (0,1]$, meaning that

$$\left| f_y^2 \frac{([nt_{l-1}]+1-nt_{l-1})^2}{n} \right| \leq \frac{||f||_\infty^2}{n} \xrightarrow{n\to\infty} 0.$$

On the other hand, $\frac{1}{n}\sum_{j=[nt_{l-1}]}^{[nt_l]} f_y^2$, $n \geq 1$, are nothing else but the Riemann sums of the integral $\int_{t_{l-1}}^{t_l} f^2(s)ds$. All in all, we obtain that

$$\lim_{n\to\infty} L_l^{(n)}(u) = -\frac{u^2}{2}\int_{t_{l-1}}^{t_l} f_s^2 ds.$$

So, for each $l \in \{1, ..., m\}$, (2.3.18) converges to $\exp\left\{-\frac{u^2}{2}\int_{t_{l-1}}^{t_l} f_s^2 ds\right\}$ and (2.3.17) to

$$\prod_{l=1}^{m} \exp\left\{-\frac{u_l^2}{2}\int_{t_{l-1}}^{t_l} f_s^2 ds\right\}$$

as n approaches, which is the characteristic function of the vector I_f and thus, by Lévy's continuity theorem, we obtain the desired result. □

In the proof of this theorem, we have used that for deterministic and continuous functions f, the process I is a process with independent increments such that

$$I_t - I_s \sim \mathcal{N}\left(0, \int_s^t f_u^2 du\right).$$

In particular, the quadratic variation of I, $\langle I\rangle$, is given by

$$\langle I\rangle_t = \int_0^t f_s^2 ds,$$

which is deterministic. Thus, the process I shares many similarities with the Brownian motion. In fact, one can show that the process I can be obtained by considering a deterministic change of time. Bearing this in mind, it is no surprise that the same proof for the convergence towards the Brownian motion (the one seen in [5]) worked for the convergence of the stochastic integrals since, in essence, what we are proving is the convergence of a modified random walk given by

$$I_t^{(n)} = \sqrt{n}\int_0^t \theta(ns) f_s\, ds = \sqrt{n}\sum_{j=1}^{[nt]+1} X_j \int_0^t \mathbb{I}_{[j-1,j)}(ns) f_s\, ds$$

towards a modified Brownian motion, I.

2.4 Final remarks

2.4.1 Extending the class of integrands

As in the case of the Brownian motion, one can extend the class of integrands by weakening condition (iii) in Definition 2.3.1 as follows:

$$\mathbb{P}\left\{\int_0^T X_t^2 d\langle M\rangle_t < \infty\right\} = 1. \tag{2.4.1}$$

As seen in the course of Stochastic Calculus, this can be done by a localization argument, for which one has to prove beforehand that, for any $\{\mathcal{F}_t\}_t$-stopping time τ taking values in $[0,T]$ and any $X \in \nu_M$ (the one from Definition 2.3.1), the following identity holds a.s.

$$\int_0^\tau X_t dM_t = \int_0^T \mathbb{I}_{[0,\tau)}(t) X_t dM_t.$$

Which is usually shown using that any stopping time like τ can be approximated a.s. by a decreasing sequence of stopping times $\{\tau_n\}_n$ similar to the ones seen in Eq.(2.3.6) (the point being that such stopping times only take a finite number of values for each n). We will omit the details and denote by ν_M' the set of processes that satisfy conditions (i), (ii) and the one in (2.4.1). Nevertheless, we shall remark that for general $X \in \nu_M'$, the process $\{I_t(X)\colon 0 \le t \le T\}$ might not be a martingale anymore (in general, it is a local one) and that it is defined as a limit in probability, rather than an $L^2(\Omega)$-limit.

2.4.2 Extending the class of integrators

One can define, using the stochastic integral seen in this chapter, stochastic integrals with respect to more general processes like continuous local martingales or continuous semimartingales. Moreover, they can be extended to local martingales and semimartingales whose sample paths are càdlàg $\mathbb{P}$-a.s.

Definition 2.4.1. *A process $M = \{M_t \colon t \geq 0\}$ is a continuous local martingale if there is a non-decreasing sequence of $\{\mathcal{F}_t\}_t$-stopping times $\{S_n\}_n$ such that $S_n \nearrow \infty$ $\mathbb{P}$-a.s. and such that, for each $n \in \mathbb{N}$, the stopped process $\{M_{t\wedge S_n} \colon t \geq 0\}$ is a continuous martingale.*

Definition 2.4.2. *An adapted continuous process $M = \{M_t \colon t \geq 0\}$ is a semimartingale if it can be decomposed $\mathbb{P}$-a.s. as follows:*

$$M_t = M_0 + N_t + A_t, t \geq 0, \tag{2.4.2}$$

where $N = \{N_t \colon t \geq 0\}$ is a continuous local martingale and $A = \{A_t \colon t \geq 0\}$ is a continuous process of bounded variation (with both processes N and A starting at 0).

For $X \in \nu'_M$, one can construct the stochastic integral of X with respect to the continuous local martingale by constructing a localizing sequence of stopping times $\{T_n\}_n$, $T_n \nearrow \infty$ $\mathbb{P}$-a.s. (these stopping times depend on the stopping times implicit in the definition of local martingale, $\{S_n\}_n$), such the stopped processes $M^{(n)} = \{M_t^{(n)} := M_{t\wedge T_n} \colon t \geq 0\}$, $n \geq 1$, are continuous square integrable martingales and such that $X^{(n)} = \{X_t^{(n)} := X_t \mathbb{I}_{T_n \geq t}\} \in \nu_M$ for each $n \geq 1$. Then one defines the desired integral as one would expect

$$I_t(X) := I_t^{M^{(n)}}\left(X^{(n)}\right), \text{ on } \{0 \leq t \leq T_n\}.$$

This construction is coherent in the sense that if $1 \leq n \leq m$ (which means that $T_n \leq T_m$ $\mathbb{P}$-a.s.), then $I_t(X)$ is well defined on $[0, T_n(\omega)]$ for $\mathbb{P}$-almost all $\omega \in \Omega$, that is,

$$I_t^{M^{(n)}}\left(X^{(n)}\right) = I_t^{M^{(m)}}\left(X^{(m)}\right), \quad 0 \leq t \leq T_n.$$

Moreover, the definition does not depend on the choice of the localizing sequence of stopping times $\{S_n\}_n$.

The stochastic integral of $X \in \nu'_M$ with respect to a continuous semimartingale M with decomposition given by (2.4.2) is then defined as

$$\int_0^t X_s dM_s = X_0 M_0 + \int_0^t X_s dN_s + \int_0^t X_s dA_s.$$

Which is well defined because the decomposition in (2.4.2) is $\mathbb{P}$-a.s. unique.

One can show as well that the Itô and covariation formulas, among many other results seen here, also hold for such integrals.

Chapter 3

Stochastic integrals with respect to random fields

So far, we have seen that we can construct stochastic integrals with respect to stochastic processes that evolved according to a single variable, the time. Such integrals were introduced to define rigorously the concept of stochastic differential equation driven by a random process.

Introducing the notion of stochastic partial differential equations is a bit more delicate task. Indeed, recall that the point of considering partial differential equations is to determine functions that depend on more than one variable and, in most cases, these variables are not symmetric: for instance, the time variable usually takes non-negative values, while the space variable can take any real value (depending on the problem, of course). So we first need to introduce the concept of space-time random perturbation and then see how we integrate with respect to such objects.

We will start by considering the case of the white noise and the isonormal process, which can be thought of the analogous to the Brownian motion in several (possibly infinite) dimensions in some cases, and, at the end of this Chapter (see Section 3.3), we introduce a wider class of random noises (and hence, of integrators), with the drawback that the class of integrable functions will be a bit more limited. This way, we start by considering objects which might be a bit more familiar to the ones already seen (the Brownian motion and the Itô integral with respect to such process), providing a canonical example that will serve us as a model to construct the integral with respect to much more general processes (as done in Section 3.3).

For the construction of the stochastic integral with respect to the space-time white noise, we will follow the first two chapters of [6], while, for the construction with respect to worthy martingale measures, we will follow the second chapter of [11].

3.1 White noise and the isonormal process

3.1.1 White noise

Definition 3.1.1. *A σ-finite measure ν on $\left(\mathbb{R}^k, \mathcal{R}^k\right)$ (being $\mathcal{R}^k$ the Borel σ-field of $\mathbb{R}^k$) is a measure for which there is a sequence of compact sets $\{E_n\}_{n\in\mathbb{N}}$ such that $\nu(E_n) < \infty$ for all $n \in \mathbb{N}$ and such that $E_n \nearrow \mathbb{R}^k$ (that is, the sequence $\{E_n\}_n$ is increasing and $\mathbb{R}^k = \bigcup_n E_n$).*

Definition 3.1.2. *A (Gaussian) white noise on $\mathbb{R}^k$ based on a σ-finite measure ν on $\left(\mathbb{R}^k, \mathcal{R}^k\right)$ is a Gaussian random field*

$$W = \{W(A)\colon A \in \mathcal{R}^k_f\}, \quad \mathcal{R}^k_f \coloneqq \{A \in \mathcal{R}^k \colon \nu(A) < \infty\}$$

defined on some probability space $(\Omega, \mathcal{F}, \mathbb{P})$ with mean function $\mu(A) = 0$ for any $A \in \mathcal{R}^k_f$ and covariance function $C(A, B) \coloneqq \nu(A \cap B)$.

The existence of this process (and all the other Gaussian processes that we will see during this section) can be seen as a consequence of the following result, which is, in turn, a consequence of Kolmogorov's Extension Theorem:

Lemma 3.1.1. *Let $\mathbb{T}$ be an arbitrary set. Given functions $m\colon \mathbb{T} \to \mathbb{R}$ and $C\colon \mathbb{T}^2 \colon \mathbb{R}$ such that $C(t, s) = C(s, t)$ for all $(s, t) \in \mathbb{T}^2$, and C is non-negative definite, there exists a Gaussian random field $G = \{G(t)\colon t \in \mathbb{T}\}$ with mean function m and covariance function C.*

For a proof, see Lemma 1.2.2 in page 4 of [6].

Observe that, from the definition, $W(A)$ is a normal random variable with zero mean and variance $\nu(A)$ (if $\nu(A) = 0$, then $W(A)$ is the constant random variable zero).

It turns out that W, as a set function from $\mathcal{R}^k_f$ to $L^2(\Omega, \mathcal{F}, \mathbb{P})$, can be thought as a vector-valued measure. To see this, we will need the following proposition.

Proposition 3.1.1. *The following assertions hold true:*

(i) If $A, B \in \mathcal{R}^k_f$ are disjoint, then $W(A)$ and $W(B)$ are independent and $W(A \cup B) = W(A) + W(B)$ $\mathbb{P}$-a.s.

(ii) Let $\{A_n\}_n \subset \mathcal{R}^k$ be a decreasing sequence with $\nu(A_1) < \infty$ and let $A = \bigcap_{n\geq 1} A_n$, then

$$L^2(\Omega) - \lim_{n\to\infty} W(A_n) = W(A).$$

(iii) *Let $\{A_n\}_n \subset \mathcal{R}^k$ be an increasing sequence and let $A = \bigcup_{n\geq 1} A_n$ be such that $\nu(A) < \infty$, then*

$$L^2(\Omega) - \lim_{n\to\infty} W(A_n) = W(A).$$

Proof. For the first part, since $A \cap B = \emptyset$ and the mean function μ is null, we have

$$0 = \nu(A \cap B) = C(A, B) = \mathbb{E}[W(A)W(B)].$$

So $W(A)$ and $W(B)$ are uncorrelated Gaussian random variables and thus, independent. To check the finite additivity property, note that

$$\begin{aligned}
\mathbb{E}\big[\big(W(A\cup B) - W(A) - W(B)\big)^2\big] &= \mathbb{E}\left[W^2(A\cup B)\right] + \mathbb{E}\left[W^2(A)\right] \\
&+ \mathbb{E}\left[W^2(B)\right] - 2\mathbb{E}[W(A\cup B)W(A)] - 2\mathbb{E}[W(A\cup B)W(B)] \\
&+ 2\mathbb{E}[W(A)W(B)] \\
&= \nu(A\cup B) + \nu(A) + \nu(B) - 2\nu\,(A) - 2\nu\,(B) + 2\nu(A\cap B) \\
&= 0.
\end{aligned}$$

Where in the last step we have used that ν is a measure and hence, it is additive.

For (ii), first note that the condition $\nu(A_1) < \infty$ implies $A_n \in \mathcal{R}^k_f$ for all $n \geq 1$. On the other hand, for any $B, C \in \mathcal{R}^k_f$ such that $B \subset C$ we have, by (i),

$$\begin{aligned}
W(C) &= W\left((C\cap B) \cup (C \cap B^c)\right) \\
&= W(C\cap B) + W(C\cap B^c) \\
&= W(B) + W(C\backslash B), \quad \mathbb{P}\text{-a.s.}
\end{aligned}$$

So $W(C\backslash B) = W(C) - W(B)$ $\mathbb{P}$-a.s. By sequential continuity of ν, we have that $\cap_{1\leq j\leq n}(A_j\backslash A) = A_n\backslash A$ decreases to the empty set as n approaches infinity and hence

$$\mathbb{E}\left[(W(A_n) - W(A))^2\right] = \mathbb{E}\left[W^2(A_n\backslash A)\right] = \nu(A_n\backslash A) \xrightarrow{n\to\infty} 0.$$

Similarly, for (iii), the fact that $\nu(A) < \infty$ ensures that $A_n \in \mathcal{R}^k_f$ for all $n \geq 1$ and we have that $\bigcap_{1\leq j\leq n}(A\backslash A_j) = A\backslash A_n$ decreases to the empty set as n approaches infinity, so, by sequential continuity,

$$\mathbb{E}\left[(W(A) - W(A_n))^2\right] = \mathbb{E}\left[W^2(A\backslash A_n)\right] = \nu(A\backslash A_n) \xrightarrow{n\to\infty} 0.$$

As was to be shown. □

Hence, the map $W\colon \mathcal{R}^k_f \to L^2(\Omega, \mathcal{F}, \mathbb{P})$, $A \mapsto W(A)$ is a σ-additive vector-valued measure. Indeed, if $\{B_n\}_n \subset \mathcal{R}^k_f$ is a sequence of pairwise disjoint sets such that

$A = \bigcup_{n\geq 1} B_n \in \mathcal{R}_f^k$, then we have that, since $A_n := \bigcup_{1\leq j\leq n} B_j$ verifies $\bigcup_{n\geq 1} A_n = A$ and $A_n \subset A_m$ whenever $n \leq m$,

$$\sum_{j=1}^{n} W(B_j) = W\left(\bigcup_{j=1}^{n} B_j\right) = W(A_n) \xrightarrow{n\to\infty} W(A)$$

in $L^2(\Omega)$. However, this does not mean that for fixed $\omega \in \Omega$ (or for almost any $\omega \in \Omega$), the map $A \mapsto W(A)(\omega)$ is a a real-valued signed measure.

3.1.2 Isonormal process

From now on, H will denote a real, separable Hilbert space with inner product $\langle \cdot, \cdot \rangle_H$ and norm $||\cdot||_H$.

Definition 3.1.3. *An isonormal Gaussian process on H is a Gaussian process $W = \{W(h) \colon h \in H\}$ defined on some complete probability space $(\Omega, \mathcal{F}, \mathbb{P})$ with vanishing mean function and covariance function $C(h, g) := \langle h, g \rangle_H$.*

Thus, for each $h \in H$, $W(h)$ is normally distributed with zero mean and variance $||h||_H^2$. This means that the map $h \mapsto W(h)$ is an isometry from H to $L^2(\Omega, \mathcal{F}, \mathbb{P})$. Moreover, it is $\mathbb{P}$-a.s. linear. Indeed, for any $a, b \in \mathbb{R}$ and any $h, g \in H$,

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

Where we have used that $||ah + bg||_H^2 = a^2||h||_H^2 + b||g||_H^2 + 2ab\langle h, g\rangle_H$ and the bilinearity and symmetry of $\langle \cdot, \cdot \rangle_H$.

All in all, we have that $W \colon H \to L^2(\Omega, \mathcal{F}, \mathbb{P})$ is a linear isometry (in particular, it is continuous).

We observe as well that, if $h, g \in H$ are orthogonal, $\langle h, g\rangle_H = 0$, then the random variables $W(h)$ and $W(g)$ are independent since they are uncorrelated Gaussian random variables. In particular, if $||h||_H = ||g||_H = 1$, $W(h)$ and $W(g)$ are independent standard normal random variables. This fact can be used to give a deeper insight on the structure of isonormal processes.

Proposition 3.1.2. *Let $\{e_n \colon n \geq 1\}$ be a complete orthonormal system (CONS) in H and $\{\xi_n \colon \geq 1\}$ be a sequence of i.i.d. standard Gaussian random variables defined on $(\Omega, \mathcal{F}, \mathbb{P})$. Then, for any $h \in H$, the series*

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

converges in $L^2(\Omega)$ to a random variable, which we will denote by $W(h)$, and the family $\{W(h)\colon h \in H\}$ constitutes an isonormal Gaussian process on H.

On the other hand, given an isonormal Gaussian process $\{W(h)\colon h \in H\}$ and a CONS $\{e_n\colon n \geq 1\}$ on H, the sequence $\{W(e_n)\colon n \geq 1\}$ consists of independent standard Gaussian random variables and

$$W(h) = L^2(\Omega) - \lim_{N\to\infty} \sum_{n=1}^{N} \langle h, e_n\rangle_H W(e_n)$$

for each $h \in H$.

Proof. By Parseval's identity, we have that

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

in particular, the sum is convergent. Thus, for $n > m \geq 0$ and $h \in H$,

$$\begin{aligned}\mathbb{E}\left[\left(\sum_{j=m+1}^{n} \langle h, e_j\rangle_H \xi_j\right)^2\right] &= \sum_{j=m+1}^{n} \langle h, e_j\rangle_H^2 \mathbb{E}\left[\xi_j^2\right] + \sum_{i\neq j} \langle h, e_i\rangle_H \langle h, e_j\rangle_h \mathbb{E}[\xi_i\xi_j] \\ &= \sum_{j=m+1}^{n} \langle h, e_j\rangle_H^2 \xrightarrow{n,m\to\infty} 0.\end{aligned}$$

Where we have used that the random variables $\{\xi_n\}_n$ are pairwise independent and that they have unitary mean. Therefore, the sequence

$$\left\{\sum_{j=1}^{n} \langle h, e_j\rangle_H \xi_j \colon n \geq 1\right\}$$

is Cauchy in $L^2(\Omega)$ and, by completeness, convergent. Let $W(h)$ denote the limit of this sequence, then it is the $L^2(\Omega)$-limit of Gaussian random variables and hence, Gaussian itself. Given that the ξ_j are i.i.d. standard normal random variables, we have that $\sum_{j=1}^{n} \langle h, e_j\rangle_H \xi_j$ is normally distributed with mean 0 and variance $\sum_{j=1}^{n} \langle h, e_j\rangle_H^2$. Hence, the limit is normally distributed with zero mean and variance $||h||_H^2$ (for instance, by Lévy's continuity theorem).

Moreover, by Parseval's identity and the independence of the ξ_j again, for any

$h, g \in H$,

$$\begin{aligned}
\mathbb{E}\left[W(h)W(g)\right] &= \mathbb{E}\left[\sum_{i,j}\langle h, e_i\rangle_H \langle g, e_j\rangle_H \xi_i\xi_j\right] \\
&= \mathbb{E}\left[\sum_{j=1}^{\infty}\langle h, e_j\rangle_H \langle g, e_j\rangle_H \xi_j^2\right] + \mathbb{E}\left[\sum_{i\neq j}\langle h, e_i\rangle_H \langle g, e_j\rangle_H \xi_i\xi_j\right] \\
&= \sum_{j=1}^{\infty}\langle h, e_j\rangle_H \langle g, e_j\rangle_H \mathbb{E}\left[\xi_j^2\right] + \sum_{i\neq j}\langle h, e_i\rangle_H \langle g, e_j\rangle_H \mathbb{E}[\xi_i\xi_j] \\
&= \sum_{j=1}^{\infty}\langle h, e_j\rangle_H \langle g, e_j\rangle_H \\
&= \langle h, g\rangle_H .
\end{aligned}$$

Where the interchange of the expectation operator and the sums is justified due to the fact that the sums converge in $L^2(\Omega)$. Hence, the family $\{W(h)\colon h \in H\}$ constitutes a process on H with vanishing mean function and covariance function $C(h, g) = \langle h, g\rangle_H$.

As for the second part, the observation made before the statement of this proposition tells us that $\{W(e_n)\colon n \geq 1\}$ is a sequence of i.i.d. standard normal random variables. Since $\{e_n\colon n \geq 1\}$ is a CONS, for any $h \in H$ we have

$$h = H - \lim_{N\to\infty} \sum_{n=1}^{N}\langle h, e_n\rangle_H e_n$$

and hence, by linearity and continuity of $h \mapsto W(h)$,

$$\begin{aligned}
W(h) = W\left(H - \lim_{N\to\infty} \sum_{n=1}^{N}\langle h, e_n\rangle_H e_n\right) \\
= L^2(\Omega) - \lim_{N\to\infty} W\left(\sum_{n=1}^{N}\langle h, e_n\rangle_H e_n\right) \\
= L^2(\Omega) - \lim_{N\to\infty} \sum_{n=1}^{N}\langle h, e_n\rangle_H W(e_n).
\end{aligned}$$

Which is what we wanted to show. $\square$

3.1.3 Relation between the white noise and the isonormal process

In the frame of Lebesgue's integration theory, one starts by defining integrals of indicator functions of ν-measurable sets of the form $\mathbb{I}_A$ (being ν the considered measure), which are defined to attain the value $\nu(A)$. This relates measures of measurable sets with simple functions and, from this, one can extend the notion of integral to a higher class of measurable functions. This integral characterizes the measure as well.

A similar procedure can be followed to establish a relation between the white noise and the isonormal process, where the "measure of a set $A \in \mathcal{R}_f^k$ with respect to the white noise" is related to the "integral of a simple function $\mathbb{I}_A$ with respect to the isonormal process" (for this, a particular Hilbert space H must be considered) to then define integrals of more general functions $h \in H$ with respect to the isonormal process and this integral will characterize the corresponding white noise measure as well.

In this section, ν will denote a σ-finite measure on $(\mathbb{R}^k, \mathcal{R}_f^k)$ and $H = L^2(\mathbb{R}^k, \nu)$ with the inner product $\langle h, g\rangle_H = \int_{\mathbb{R}^k} hg\nu(dx)$.

Given an isonormal process on H, one can easily define a Gaussian white noise $\bar{W} = \{\bar{W}(A)\colon A \in \mathcal{R}_f^k\}$ by setting $\bar{W}(A) := W(\mathbb{I}_A)$. In such case, one has that $\mathbb{E}[\bar{W}(A)] = 0$ for any $A \in \mathcal{R}_f^k$ and

$$\begin{aligned}\mathbb{E}\left[\bar{W}(A)\bar{W}(B)\right] &= \mathbb{E}\left[W(\mathbb{I}_A)W(\mathbb{I}_B)\right]\\ &= \langle \mathbb{I}_A, \mathbb{I}_B\rangle_H\\ &= \int_{\mathbb{R}^k} \mathbb{I}_A\mathbb{I}_B\nu(dx)\\ &= \nu(A\cap B)\end{aligned}$$

for any $A, B \in \mathcal{R}_f^k$. Thus, $\bar{W}$ is a white noise based on ν.

Now let $\bar{W}$ denote a white noise based on ν. For each $A \in \mathcal{R}_f^k$ define $W(\mathbb{I}_A) := \bar{W}(A)$ and extend the definition to functions $h = \sum_{j=1}^r c_j\mathbb{I}_{A_j}$ with $r \in \mathbb{N}$, $c_j \in \mathbb{R}$ and $A_j \in \mathcal{R}_f^k$ pairwise disjoint linearly:

$$W(h) := \sum_{j=1}^{r} c_j\bar{W}(A_j).$$

Then one has, for such h,

$$\begin{aligned}
\mathbb{E}\left[W^2(h)\right] &= \sum_{j=1}^{r} c_j^2 \mathbb{E}\left[\bar{W}^2(A_j)\right] + \sum_{i\neq j} c_i c_j \mathbb{E}\left[\bar{W}(A_i)\bar{W}(A_j)\right] \\
&= \sum_{j=1}^{r} c_j^2 \nu(A_j) \\
&= \sum_{j=1}^{r} \int_{\mathbb{R}^k} c_j^2 \mathbb{I}_{A_j}^2 \nu(dx) \\
&= \int_{\mathbb{R}^k} \left(\sum_{j=1}^{r} c_j \mathbb{I}_{A_j}\right)^2 \nu(dx) \\
&= ||h||_H^2.
\end{aligned}$$

Where we have used that $\bar{W}(A_i)$ is independent of $\bar{W}(A_j)$ for $i \neq j$ (because $A_i \cap A_j = \emptyset$ in this case). Moreover, the value $W(h)$ does not depend on the representation of h. That is, if

$$h = \sum_{j=1}^{r} c_j \mathbb{I}_{A_j}, \quad g = \sum_{i=1}^{m} d_i \mathbb{I}_{B_i}$$

with $c_j, d_i \in \mathbb{R}$, $A_i \cap A_j = \emptyset$ and $B_n \cap B_m = \emptyset$ if $i \neq j$ and $n \neq m$ and $h = g$, then $W(h) = W(g)$ $\mathbb{P}$-a.s. Indeed,

$$\begin{aligned}
\mathbb{E}\left[(W(h) - W(g))^2\right] &= ||h||_H^2 + ||g||_H^2 - 2\sum_{i,j} c_j d_i \nu(A_j \cap B_i) \\
&= \int_{\mathbb{R}^k} \left(\sum_{j=1}^{r} c_j^2 \mathbb{I}_{A_j}^2 + \sum_{i=1}^{m} d_i \mathbb{I}_{B_i} - 2\sum_{i,j} c_j d_i \mathbb{I}_{A_j \cap B_i}\right) \nu(dx) \\
&= \int_{\mathbb{R}^k} \left(\sum_{j=1}^{r} c_j \mathbb{I}_{A_j} - \sum_{i=1}^{m} d_i \mathbb{I}_{B_i}\right)^2 \nu(dx) \\
&= 0.
\end{aligned}$$

As in the case of the usual Lebesgue integration, one checks that the map $h \mapsto W(h)$ is linear and hence, by what we have already shown, it defines a linear isometry (as in the case of the Itô integral) on the set of elementary functions that extends to a linear isometry from $H = L^2(\mathbb{R}^k, \nu)$ to $L^2(\Omega, \mathcal{F}, \mathbb{P})$ as an $L^2(\Omega)$-limit by using the fact that simple functions are dense in $L^2(\mathbb{R}^k, \nu)$ with respect to the norm $|| \cdot ||_H$ and a similar procedure to the one seen in the previous chapter.

The only thing left to check is that the resulting process $\{W(h)\colon h \in H\}$ is indeed an isonormal process. For $h \in H$, $W(h)$ is the $L^2(\Omega)$-limit of centered Gaussian random variables, so it is a centered Gaussian random variable itself. Moreover, for $h, g \in H$, using that

$$\begin{aligned} W(h)W(g) &= \frac{1}{4}\left(W(h)+W(g)\right)^2 - \frac{1}{4}\left(W(h)-W(g)\right)^2 \\ &= \frac{1}{4}\left(W(h+g)\right)^2 - \frac{1}{4}\left(W(h-g)\right)^2 \end{aligned}$$

(because $h \mapsto W(h)$ is linear) and the isometry property, one obtains

$$\mathbb{E}\left[W(h)W(g)\right] = \frac{1}{4}||h+g||_H^2 - \frac{1}{4}||h-g||_H^2 = \langle h, g\rangle_H$$

by the polarization identity of the inner product.

For $h \in H$, the random variable $W(h)$ is called the Wiener integral of h with respect to the white noise $\bar{W}$ and it is usually written as

$$\int_{\mathbb{R}^k} h(x)\bar{W}(dx) := W(h).$$

For the sake of simplicity, we shall write

$$W(h) = \int_{\mathbb{R}^k} h(x)W(dx) \tag{3.1.1}$$

and W instead of $\bar{W}$ for the white noise (we will distinguish the latter from the corresponding isonormal process by the context).

As one may expect, the constructed integral can be related to the Itô integral with respect to the Brownian motion:

Lemma 3.1.2. *Let $\{W(h)\colon h \in L^2(\mathbb{R}, \nu)\}$ be the isonormal Gaussian process associated to the white noise on $\mathbb{R}$ based on $\nu(ds) = \mathbb{I}_{[0,T]}(s)ds$, $T > 0$, and let $B = \{B_t\colon t \geq 0\}$ be the continuous version of the standard Brownian motion defined by*

$$B_t = W(\mathbb{I}_{[0,t)}) = W([0,t)), \quad 0 \leq t \leq T. \tag{3.1.2}$$

Then, for all $h \in L^2(\mathbb{R}, \nu)$,

$$W(h) = \int_0^T h(t)dB_t, \quad \mathbb{P}\text{-a.s.}$$

The point being that the integral on the right-hand side is the usual Itô integral with respect to the Brownian motion.

Proof. We first check that (3.1.2) defines a Brownian motion. Indeed, for $0 \leq s_1 < t_1 \leq s_2 < t_2 \leq T$, we have that $[s_1, t_1) \cap [s_2, t_2) = \emptyset$ and hence, the random variables $W([s_1, t_1))$ and $W([s_2, t_2))$ are independent. Moreover, we have, since $[0, s_i) \subset [0, t_i)$ for $i \in \{1, 2\}$

$$B_{t_i} - B_{s_i} = W([s_i, t_i)),$$

meaning that B has independent increments. Finally, given the definition of ν, we have that $B_{t_i} - B_{s_i}$ is normally distributed with zero mean and variance $t_i - s_i$ and, since $[0, 0) = \emptyset$, $B_0 = 0$.

Now, if $h = \mathbb{I}_{[t_1,t_2)}$ with $0 \leq s < t \leq T$, then

$$\int_0^T h(r) dB_r = B_t - B_s = W([s, t)) = W(\mathbb{I}_{[s,t)}) = W(h), \quad \mathbb{P}\text{-a.s.}$$

By linearity, this extends to step functions of the form $h = \sum_{j=0}^{r-1} a_j \mathbb{I}_{[t_j,t_{j+1})}$ with $a_j \in \mathbb{R}$ constants and $0 = t_0 < \ldots < t_r = T$ and, by density and the isometry properties for both the Itô integral and the map $h \mapsto W(h)$, to any $h \in L^2(\mathbb{R}, \nu)$. $\square$

Observe that this construction of the stochastic integral (for deterministic functions) allows as to easily define the stochastic integral with respect to the Brownian motion over $[0, \infty)$ just by considering the measure $\nu(ds) = \mathbb{I}_{[0,\infty)}(s)ds$. Moreover, this procedure provides us a method to construct integrals with respect to other random fields like, for instance, the Brownian sheet.

Definition 3.1.4. *A Brownian sheet on* $\mathbb{R}^k_+ := \{(x_1, \ldots, x_k) \in \mathbb{R}^k \colon x_j \geq 0,\ 1 \leq j \leq k\}$ *is a Gaussian process with vanishing mean and covariance function*

$$C(s, t) = \prod_{j=1}^{k} (s_j \wedge t_j),$$

for all $s = (s_1, \ldots, s_k), t = (t_1, \ldots, t_k) \in \mathbb{R}^k_+$.

Lemma 3.1.3. *Let* $H = L^2(\mathbb{R}^2, \nu)$, $\nu(dx) = \mathbb{I}_{\mathbb{R}^2_+}(x)dx$ *with the usual inner product,* $\{W(A) \colon A \in \mathcal{R}^2_f\}$ *be a white noise based on* ν *and* $\{W(h) \colon h \in H\}$ *be the corresponding isonormal process, then the process* $W = \{W_{t_1,t_2} \colon (t_1, t_2) \in \mathbb{R}^2_+\}$ *defined by*

$$W_{t_1,t_2} := W([0, t_1) \times [0, t_2))$$

is a Brownian sheet.

Proof. It is clear that it is a Gaussian process with vanishing mean. On the other hand, if $(t_1, t_2), (s_1, s_2) \in \mathbb{R}_+^2$,

$$\begin{aligned}
\mathbb{E}\left[W_{t_1,t_2} W_{s_1,s_2}\right] &= \mathbb{E}\left[W([0,t_1) \times [0,t_2)) W([0,s_1) \times [0,s_2))\right] \\
&= \nu\left(([0,t_1) \times [0,t_2)) \cap ([0,t_1) \times [0,t_2))\right) \\
&= \nu\left([0, t_1 \wedge s_1) \times [0, t_2 \wedge s_2)\right) \\
&= (t_1 \wedge s_1) \cdot (t_2 \wedge s_2).
\end{aligned}$$

□

One can show, by Kolmogorov's continuity criterion, that there is a version of the Brownian sheet with continuous sample paths. Considering such version, one can define the stochastic integral with respect to the Brownian sheet of a deterministic function $h \in H$ by

$$\int_{\mathbb{R}_+^2} h(t_1, t_2) dW_{t_1,t_2} := W(h). \tag{3.1.3}$$

3.1.4 Space-time white noise

Now we are ready to define a concept of space-time white noise, which, at the end of the day, it is simply a particular case of the white noise already seen.

Definition 3.1.5. *Let $D \subset \mathbb{R}^k$ be a non-empty open set, then a space-time white noise based on $\nu(dt, dx) = \mathbb{I}_{\mathbb{R}_+}(t)\mathbb{I}_D(x) dt\, dx$ is a centered Gaussian random field $\{W(A), A \in \mathcal{B}^f_{\mathbb{R}_+ \times D}\}$, where $\mathcal{B}^f_{\mathbb{R}_+ \times D}$ are the Borel sets $A \in \mathcal{B}^f_{\mathbb{R}_+ \times D}$ such that $\nu(A) < \infty$, with covariance function*

$$\mathbb{E}\left[W(A)W(B)\right] = \nu(A \cap B), \quad A, B \in \mathcal{B}^f_{\mathbb{R}_+ \times D}.$$

Of course, the results seen so far regarding the properties of the white noise and construction of an isonormal process from it also hold for the space-time white noise.

From now on, we shall assume that $D \subset \mathbb{R}^k$ is also connected (note that it can be bounded or unbounded). As in the previous chapter, we will fix a time horizon $T > 0$ (thus, the considered measure will be $\nu(dt, dx) = \mathbb{I}_{[0,T]}(t)\mathbb{I}_D(x) dt\, dx$) and set $H = L^2([0, T] \times D)$ and $V = L^2(D)$ with the inner products

$$\langle h, g \rangle_H = \int_0^T \int_D h(t, x) g(t, x) dt dx \quad \langle \varphi, \psi \rangle_V = \int_D \varphi(x) \psi(x) dx,$$

for $h, g \in H$ and $\varphi, \psi \in V$. Observe that

$$\langle h, g \rangle_H = \int_0^T \langle h(t, *), g(t, *) \rangle_V dt$$

where "$*$" denotes spatial dependence. From the isonormal process on H associated to W (the space-time white noise), we can define a Gaussian stochastic process $\{W_s(\phi)\colon s \in [0,T], \varphi \in V\}$ by

$$W_s(\phi) := W\left(\mathbb{I}_{[0,s]}(\cdot)\varphi(*)\right)$$

where "$\cdot$" denotes time dependence. Observe that functions of the type $\mathbb{I}_{[0,s]}(\cdot)\varphi(*)$ with $\varphi \in V$ are functions of H.

The goal now is to establish a couple of results so that we can guarantee that the stochastic integral with respect to the space-time white noise is well defined.

Lemma 3.1.4. *The process $\{W_s(\phi)\colon s \in [0,T], \varphi \in V\}$ has the following properties:*

(i) For any $\varphi \in V$, $\{W_s(\varphi)\colon s \in [0,T]\}$ defines a Brownian motion with variance $s||\varphi||_V^2$.

(ii) For all $s,t \in [0,T]$ and $\varphi, \psi \in V$,

$$\mathbb{E}\left[W_s(\varphi)W_t(\psi)\right] = (s \wedge t)\langle\varphi,\psi\rangle_V.$$

Proof. Since $\{W(h)\colon h \in L^2([0,T] \times D, \nu)\}$ is an isonormal process, we have, for all $\varphi, \psi \in V$, $s,t \in [0,T]$,

$$\mathbb{E}\left[W_s(\varphi)\right] = \mathbb{E}\left[W\left(\mathbb{I}_{[0,s]}(\cdot)\varphi(*)\right)\right] = 0,$$

and,

$$\begin{aligned}\mathbb{E}\left[W_s(\varphi)W_t(\psi)\right] &= \langle \mathbb{I}_{[0,s]}(\cdot)\varphi(*), \mathbb{I}_{[0,t]}(\cdot)\psi(*)\rangle_H \\ &= \int_0^T \mathbb{I}_{[0,s]}(r)\mathbb{I}_{[0,t]}(r)dr \int_D \varphi(x)\psi(x)dx \\ &= (s \wedge t)\langle\varphi,\psi\rangle_V.\end{aligned}$$

The desired result is then concluded from the fact that, as noticed before this lemma, the process $\{W_s(\phi)\colon s \in [0,T], \varphi \in V\}$ is Gaussian and hence, so is $\{W_s(\phi)\colon s \in [0,T]\}$ for fixed $\varphi \in V$. □

From now on, we will consider the continuous version of the Brownian motions $\{W_s(\phi)\colon s \in [0,T]\}$, $\varphi \in V$.

We will consider, as well, an underlying right-continuous and complete filtration $\{\mathcal{F}_s\colon s \in [0,T]\}$ consisting of sub-σ-fields of $\mathcal{F}$ such that:

(i) For fixed $s \in [0,T]$ and for any $\varphi \in V$, the random variable $W_s(\varphi)$ is $\mathcal{F}_s$-measurable.

(ii) For any $s \in [0,T]$, the family $\{W_t(\varphi) - W_s(\varphi)\colon t \in [s,T], \varphi \in V\}$ is independent of $\mathcal{F}_s$.

For instance, the sub-σ-fields

$$\mathcal{F}_s \coloneqq \sigma\{W_t(\varphi)\colon 0 \le t \le s, \varphi \in V\}, \quad s \in [0,T]$$

completed so that they contain all $\mathbb{P}$-null sets and the sets contained in such sets verifies the hypotheses (recall that, for fixed $\varphi \in V$, we consider the continuous version of the Brownian motion $\{W_s(\varphi)\colon s \in [0,T]\}$, so the right-continuity condition of the filtration is fulfilled as well).

Lemma 3.1.5. *Let $\{e_j\colon j \ge 1\}$ be a complete orthonormal basis of V, then*

(i) The sequence $\{W_s(e_j)\colon s \in [0,T], j \ge 1\}$ consists of independent standard Brownian motions adapted to the filtration $\{\mathcal{F}_s\}_s$ and, for $s \in [0,T]$, $\{W_{t+s}(e_s) - W_s(e_j)\colon t \in [0, T-s], j \ge 1\}$ is independent of $\mathcal{F}_s$. Moreover, for all $\varphi \in V$ and $s \in [0,T]$,

$$W_s(\varphi) = \sum_{j=1}^{\infty} \langle \varphi, e_j \rangle_V W_s(e_j) \tag{3.1.4}$$

where the series converges $\mathbb{P}$-a.s. and in $L^2(\Omega)$.

(ii) Given a sequence $\{B_s^j\colon s \in [0,T], j \ge 1\}$ of independent standard Brownian motions, for each $\varphi \in V$ and $s \in [0,T]$, the series

$$\tilde{W}_s(\varphi) \coloneqq \sum_{j=1}^{\infty} \langle \varphi, e_j \rangle_V B_s^j$$

converges $\mathbb{P}$-a.s. and in $L^2(\Omega)$. The process $\{\tilde{W}_s(\varphi)\colon s \in [0,T], \varphi \in V\}$ verifies the conclusions of the previous lemma. Further, for $h \in H$, define

$$\tilde{W}(h) \coloneqq \sum_{j=1}^{\infty} \int_0^T \langle h(s,*), e_j \rangle_V \, dB_s^j \tag{3.1.5}$$

where the integrals on the right-hand side are usual Itô integrals and the series converges $\mathbb{P}$-a.s. and in $L^2(\Omega)$. Then the process $\tilde{W} = \{\tilde{W}(\mathbb{I}_A)\colon A \in \mathcal{B}^f_{[0,T]\times D}\}$ is a space-time white noise on $[0,T] \times D$ and $\{\tilde{W}(h)\colon h \in H\}$ is the associated isonormal process.

(iii) Let W be a space-time white noise on $[0,T] \times D$ based on $\nu(dt,dx) = \mathbb{I}_{[0,T]}(t)\mathbb{I}_D(x)dt\,dx$. If in the previous item we take $B_s^j = W_s(e_j)$, then the resulting space-time white noise in the previous item, $\tilde{W}$, coincides with W.

Proof.

Proof of (i):
By the previous lemma, and since $\{e_j : j \geq 1\}$ is an orthonormal basis, we have $\mathbb{E}[W_s(e_j)W_t(e_i)] = (s \wedge t)\delta_{ij}$ for all $s, t \in [0, T]$, $i, j \in \mathbb{N}$. Thus, if $i \neq j$, the processes $\{W_s(e_i) : s \in [0, T]\}$ and $\{W_s(e_j) : s \in [0, T]\}$ are independent (because they are Gaussian uncorrelated processes). The adaptedness and independence properties follow from the conditions that the filtration $\{\mathcal{F}_s\}_s$ must satisfy.

Now, since $\varphi \in V$ and $\{e_j : j \geq 1\}$ is a complete orthonormal basis of V, we have

$$\varphi(x) = \sum_{j=1}^{\infty} \langle \varphi, e_j \rangle_V e_j(x)$$

where the convergence is in V (in norm $|| \cdot ||_V$). By linearity and continuity of the isonormal process, we have

$$W_s(\varphi) = W\left(\mathbb{I}_{[0,s]}(\cdot)\varphi(*)\right) = \sum_{j=1}^{\infty} \langle \varphi, e_j \rangle_V W\left(\mathbb{I}_{[0,s]}(\cdot)e_j(*)\right) = \sum_{j=1}^{\infty} \langle \varphi, e_j \rangle_V W_s(e_j)$$

where the convergence is in $L^2(\Omega)$. By the Khintchine-Kolmogorov convergence theorem, we also have convergence $\mathbb{P}$-a.s. because

$$\sum_{j=1}^{\infty} \mathbb{E}\left[\langle \varphi, e_j \rangle_V^2 W_s^2(e_j)\right] = \sum_{j=1}^{\infty} \langle \varphi, e_j \rangle_V^2 \mathbb{E}\left[W_s^2(e_j)\right] = \sum_{j=1}^{\infty} \langle \varphi, e_j \rangle_V^2 s = s||\varphi||_V^2 < \infty$$

by Parseval's identity and the random variables $W_s(e_i)$ and $W_s(e_j)$ are independent if $i \neq j$.

Proof of (ii):
For fixed $s \in [0, T]$ and $\varphi \in V$, $\{B_s^j \langle \varphi, e_j \rangle_V : j \geq 1\}$ is a sequence of centered independent random variables of finite variance such that, by Parseval's identity

$$\sum_{j=1}^{\infty} \mathbb{E}\left[\left(B_s^j\right)^2 \langle \varphi, e_j \rangle_V^2\right] = s||\varphi||_V^2 < \infty.$$

Again, by the Khintchine-Kolmogorov convergence theorem, the series $\tilde{W}_s(\varphi) := \sum_{j=1}^{\infty} B_s^j \langle \varphi, e_j \rangle_V$ converges $\mathbb{P}$-a.s. and in $L^2(\Omega)$ for fixed $s \in [0, T]$ and $\varphi \in V$. Moreover, the same theorem tells us that $\mathbb{E}[\tilde{W}_s(\varphi)] = 0$, $\mathbb{E}\left[\tilde{W}_s^2(\varphi)\right] = s||\varphi||_V^2$ and, for any $s, t \in [0, T]$, $\varphi, \psi \in V$,

$$\begin{aligned}
\mathbb{E}\left[\tilde{W}_s(\varphi)\tilde{W}_t(\psi)\right] &= \mathbb{E}\left[\left(\sum_{j=1}^{\infty} B_s^j \langle \varphi, e_j \rangle_V\right)\left(\sum_{i=1}^{\infty} B_t^i \langle \psi, e_i \rangle_V\right)\right] \\
&= \sum_{i,j} \langle \varphi, e_j \rangle_V \langle \psi, e_i \rangle_V \mathbb{E}\left[B_s^j B_t^i\right] \\
&= \sum_{i,j} \langle \varphi, e_j \rangle_V \langle \psi, e_i \rangle_V (s \wedge t)\delta_{ij} \\
&= (s \wedge t) \sum_{j=1}^{\infty} \langle \varphi, e_j \rangle_V \langle \psi, e_j \rangle_V \\
&= (s \wedge t)\langle \varphi, \psi \rangle_V.
\end{aligned}$$

As for the second part, fix $h \in H$ and let $X_j = \int_0^T \langle h(s,*), e_j \rangle_V dB_s^j$. The fact that these Itô integrals are well defined (the integrand is a function satisfying the conditions to be Itô integrable) is a consequence of the fact that h and $\langle h(s,*), e_j \rangle_V$ are deterministic (so there is no need to worry about the measurability conditions) and, for the same reason,

$$\begin{aligned}
\mathbb{E}\left[\int_0^T \langle h(s,*), e_j \rangle_V^2 ds\right] &\leq \int_0^T \sum_{j=1}^{\infty} \langle h(s,*), e_j \rangle_V^2 ds \\
&= \int_0^T ||h(s,*)||_V^2 ds \\
&= \int_0^T \int_D h^2(s,x) ds dx \\
&= ||h||_H^2 < \infty.
\end{aligned}$$

Moreover, the fact that the integrand is deterministic implies that, for each $j \in \mathbb{N}$, the random variable X_j is a centered normal random variable with variance $\int_0^T \langle h(s,*), e_j \rangle_V^2 ds$.

Now, if $i \neq j$, by the observation made in (2.3.4) and the covariation formula (Theorem 2.3.10),

$$\mathbb{E}\left[X_i X_j\right] = \mathbb{E}\left[\int_0^T \langle h(s,*), e_i \rangle_V \langle h(s,*), e_j \rangle_V d\langle B^i, B^j \rangle_s\right] = 0.$$

Where, again, we have used (2.3.4) and the independence of the Brownian motions B^i and B^j to conclude that $\langle B^i, B^j \rangle_s = 0$ for all $s \in [0, T]$.

On the other hand, by the isometry formula for the Itô integral and Parseval's identity,

$$\sum_{j=1}^{\infty} \mathbb{E}\left[X_j^2\right] = \sum_{j=1}^{\infty} \int_0^T \langle h(s,*), e_j\rangle_V^2 ds = \int_0^T ||h(s,*)||_V^2 ds = ||h||_H^2 < \infty.$$

Hence, by the Khintchine-Kolmogorov convergence theorem, the series in (3.1.5) converges $\mathbb{P}$-a.s. and in $L^2(\Omega)$. Given that the $L^2(\Omega)$-limit of Gaussian random variables is Gaussian, we obtain that $\tilde{W}(h)$ is Gaussian.

Now, for $h, g \in H$,

$$\begin{aligned}
\mathbb{E}\left[\tilde{W}(h)\tilde{W}(g)\right] &= \mathbb{E}\left[\left(\sum_{j=1}^{\infty} \int_0^T \langle h(s,*), e_j\rangle_V dB_s^j\right)\left(\sum_{i=1}^{\infty} \int_0^T \langle g(s,*), e_i\rangle_V dB_s^i\right)\right] \\
&= \sum_{i,j} \mathbb{E}\left[\int_0^T \langle h(s,*), e_j\rangle_V \langle g(s,*), e_i\rangle_V d\langle B^j, B^i\rangle_s\right] \\
&= \int_0^T \sum_{j=1}^{\infty} \langle h(s,*), e_j\rangle_V \langle g(s,*), e_j\rangle_V ds \\
&= \int_0^T ds \langle h(s,*), g(s,*)\rangle_V \\
&= \langle h, g\rangle_H. \qquad (3.1.6)
\end{aligned}$$

where the covariation formula (along with the independence of the Brownian motions B^i and B^j for $i \neq j$) and Parseval's identity have been used.

So $\{\tilde{W}(h) \colon h \in H\}$ is a centered Gaussian process with covariance function $C(h,g) = \langle h, g\rangle_H$ or, in other words, an isonormal Gaussian process.

It immediately follows from (3.1.6) that $\tilde{W} := \{\tilde{W}(\mathbb{I}_A) \colon A \in \mathcal{B}^f_{[0,T]\times D}\}$ is a space-time white noise. Lastly, we check that $\{\tilde{W}(h) \colon h \in H\}$ is the isonormal process associated to the white noise $\tilde{W}$ we have just defined. Given that $h \mapsto \tilde{W}(h)$ is linear and, by definition,

$$\tilde{W}(\mathbb{I}_A) = \sum_{j=0}^{\infty} \int_0^T \langle \mathbb{I}_A(s,*), e_j\rangle_V dB_s^j, \quad A \in \mathcal{B}^f_{[0,T]\times D}$$

we can extend this last definition to finite linear combinations of indicators of disjoint sets in $\mathcal{B}^f_{[0,T]\times D}$, using the isometry property seen in (3.1.6) for these linear combinations and the fact that they are dense in H, we conclude that $\{\tilde{W}(h) \colon h \in H\}$ is the isonormal process associated to $\tilde{W}$ as desired.

Proof of (iii):

Recall that now $B_s^j = W_s(e_j) = W(\mathbb{I}_{[0,s]}(\cdot)e_j(*))$. Let $A = [0,t] \times F \in \mathcal{B}^f_{[0,T]\times D}$ with $t \in [0,T]$ and F a Borel subset of D, we have, by definition of $\tilde{W}$ and linearity and continuity of W,

$$
\begin{aligned}
\tilde{W}(A) &= \tilde{W}(\mathbb{I}_A) \\
&= \sum_{j=1}^{\infty} \int_0^T \mathbb{I}_{[0,t]}(s)\langle \mathbb{I}_F, e_j\rangle_V dW_s(e_j) \\
&= \sum_{j=1}^{\infty} \langle \mathbb{I}_F, e_j\rangle_V \int_0^t dW_s(e_j) \\
&= \sum_{j=1}^{\infty} \langle \mathbb{I}_F, e_j\rangle_V W_t(e_j) \\
&= \sum_{j=1}^{\infty} \langle \mathbb{I}_F, e_j\rangle_V W(\mathbb{I}_{[0,t]}(\cdot)e_j(*)) \\
&= W\left(\mathbb{I}_{[0,t]}(\cdot)\left[\sum_{j=1}^{\infty}\langle \mathbb{I}_F, e_j\rangle_V e_j(*)\right]\right) \\
&= W(\mathbb{I}_{[0,t]}(\cdot)\mathbb{I}_F(*)) \\
&= W(A).
\end{aligned}
$$

Hence, $\tilde{W}(A) = W(A)$ for A in the generating class of $\mathcal{B}^f_{[0,T]\times D}$, implying that $\tilde{W}$ and W coincide in all $\mathcal{B}^f_{[0,T]\times D}$. □

With this, we are ready to give a rigorous definition of the stochastic integral with respect to space-times white noise.

3.2 Stochastic integral with respect to space-time white noise

3.2.1 The stochastic integral

The results seen in Lemma 3.1.5 might seem a bit futile or unrelated to the topic of constructing a stochastic integral at first sight since they only give a couple of properties regarding the white-noise and the isonormal process.

Recall, however, that the isonormal at some function $h \in H$ can be thought as a Wiener integral (see Eq.(3.1.1), for instance). With this interpretation, equation

(3.1.5) can be interpreted as a representation of the stochastic integral of h with respect to a space-time white noise. This representation, that might seem a bit arbitrary, is nothing but an analogous of the classical Parseval's identity in $V = L^2(D)$ in the context of stochastic analysis. Indeed, for $f, g \in H$, we can write

$$\int_0^T \int_D f(s,y)g(s,y)dsdy = \sum_{j=1}^{\infty} \int_0^T \langle f(s,*), e_j\rangle_V \langle g(s,*), e_j\rangle_V ds.$$

Now, if we take $g(s,y)dsdy = W(ds,dy)$, then we can think that

$$\begin{aligned}
\langle g(s,*), e_j\rangle_V ds &= \int_D g(s,y)e_j(y)dsdy \\
&= d\left(\int_0^s \int_D e_j(y)g(r,y)drdy\right) \\
&= d\left(\int_0^s \int_D e_j(y)W(dr,dy)\right) \\
&= d\left(\int_0^T \int_D \mathbb{I}_{[0,s]}(r)e_j(y)W(dr,dy)\right) \\
&= dW(\mathbb{I}_{[0,s]}(\cdot)e_j(*)) \\
&= dW_s(e_j).
\end{aligned}$$

So

$$\int_0^T \int_D f(s,y)W(ds,dy) = \sum_{j=1}^{\infty} \int_0^T \langle f(s,*), e_j\rangle_V dW_s(e_j).$$

In fact, this last equation (or Eq.(3.1.5)) will motivate the definition of the stochastic integral, where we will extend the validity of such equations to random functions f satisfying some measurability and integrability conditions, as in the case of the usual Itô integral.

In this case, the class of integrable stochastic processes $G = \{G(s,y)\colon s \in [0,T], y \in D\}$ will be such that (in the following, $\mathcal{B}_A$ denotes the Borel σ-field on A)

(i) The map $(s,y,\omega) \mapsto G(s,y,\omega)$ from $[0,T] \times D \times \Omega$ into $\mathbb{R}$ is $\mathcal{B}_{[0,T]} \times B_D \times \mathcal{F}$-measurable (it is jointly measurable).

(ii) For each $s \in [0,T]$, the map $(y,\omega) \mapsto G(s,y,\omega)$ from $D \times \Omega$ into $\mathbb{R}$ is $\mathcal{B}_D \times \mathcal{F}_s$-measurable (it is adapted to the filtration $\{\mathcal{F}_s\}_s$ satisfying the conditions stated before Lemma 3.1.5).

(iii) $\mathbb{E}\left[\int_0^T \int_D G^2(s,y)dyds\right] < \infty$ (the process is square integrable).

From condition (iii), we have that the map $G(s,*,\omega)$ belongs to the Hilbert space $V = L^2(D)$ for $dsd\mathbb{P}$-a.a. ($dsd\mathbb{P}$-almost all) $(s,\omega) \in [0,T] \times \Omega$, so

$$G(s,*,\omega) = \sum_{j=1}^{\infty} \langle G(s,*,\omega), e_j \rangle_V e_j(*), \quad dsd\mathbb{P}\text{-a.a.}$$

Where $\{e_j : j \geq 1\}$ is a complete orthonormal basis of V and the series converges in V.

Definition 3.2.1. *Let $G = \{G(s,y) : s \in [0,T] \times y \in D\}$ be a stochastic process verifying conditions (i)-(iii), then the stochastic integral with respect to the space-time white noise W is the random variable*

$$(G \cdot W)_T := \int_0^T \int_D G(s,y) W(ds,dy) := \sum_{j=1}^{\infty} \int_0^T \langle G(s,*), e_j \rangle_V dW_s(e_j) \tag{3.2.1}$$

where the series converges in $L^2(\Omega)$.

Before we give a couple of properties regarding the integral, we have to make sure that it is indeed well defined.

First we start by checking that the Itô integrals $\int_0^T \langle G(s,*), e_j \rangle_V dW_s(e_j)$, $j \geq 1$ are well defined for such G. According to assumption (i) that G must satisfy, we have

$$(s,\omega) \mapsto \langle G(s,*), e_j \rangle_V = \int_D G(s,y,\omega) e_j(y) dy$$

is $\mathcal{B}_{[0,T]} \times \mathcal{F}$-measurable (this is part of the content of Fubini's theorem). Moreover, by assumption (ii) satisfied by G, for fixed $s \in [0,T]$, the map $\omega \mapsto \langle G(s,*,\omega), e_j \rangle_V$ is $\mathcal{F}_s$-measurable. Finally, by the third assumption and Bessel's inequality,

$$\mathbb{E}\left[\int_0^T \langle G(s,*), e_j \rangle_V^2 ds\right] \leq \mathbb{E} \int_0^T ||G(s,*)||_V^2 ds = \mathbb{E}\left[||G||_H^2\right] < \infty.$$

So the requirements for $\langle G(s,*), e_j \rangle_V$ to be Itô integrable are fulfilled.

We now see that the series does converge in $L^2(\Omega)$. Indeed, given $N > M \geq 1$, by the independence of the Brownian motions $\{W_s(e_j) : s \in [0,T], j \geq 1\}$ and Tonelli's theorem,

$$\begin{aligned}
\mathbb{E}\left[\left(\sum_{j=M+1}^{N} \int_0^T \langle G(s,*), e_j \rangle_V dW_s(e_j)\right)^2\right] &= \sum_{j=M+1}^{N} \mathbb{E}\left[\int_0^T \langle G(s,*), e_j \rangle_V^2 ds\right] \\
&= \mathbb{E}\left[\int_0^T \left(\sum_{j=M+1}^{N} \langle G(s,*), e_j \rangle_V^2\right) ds\right].
\end{aligned}$$

Now observe that, for each N, M, $\sum_{j=M+1}^{N} \langle G(s,*), e_j \rangle_V^2 \leq ||G(s,*)||_V^2$ by Parseval's identity. By the third condition that G must satisfy, we can apply the dominated convergence theorem and the fact that $\sum_{j=M+1}^{N} \langle G(s,*), e_j \rangle_V^2$ is the tail of a convergent series to conclude that

$$\left\{ \sum_{j=1}^{n} \int_0^T \langle G(s,*), e_j \rangle_V dW_s(e_j) \colon n \geq 1 \right\}$$

is a Cauchy sequence in $L^2(\Omega)$ and hence, that it converges in $L^2(\Omega)$.

Observe that from this we cannot conclude, as in Lemma 3.1.5, that the series converges $\mathbb{P}$-a.s. as well, since there is no guarantee that the random variables $\int_0^T \langle G(s,*), e_j \rangle_V dW_s(e_j)$, $j \geq 1$, are mutually independent nor that they are centered (hence, the Khintchine-Kolmogorov convergence theorem need not hold in this setting). However, the same computations seen in Lemma 3.1.5 show that the terms in the series are pairwise orthogonal in $L^2(\Omega)$ due to the independence of the Brownian motions $\{W_s(e_j) \colon s \in [0,T], j \geq 1\}$.

This last observation, with the fact that $L^2(\Omega)$ convergence implies the convergence of the corresponding norms, leads to the following result:

Proposition 3.2.1. *The stochastic integral satisfies the following isometry property:*

$$\mathbb{E}\left[\left(\int_0^T \int_D G(s,y) W(ds,dy) \right)^2 \right] = \mathbb{E}\left[\int_0^T ||G(s,*)||_V^2 ds \right] = \mathbb{E}\left[||G||_H^2 \right].$$

Observe that if $G \in H$ is deterministic, then Lemma 3.1.5 gives that $(G \cdot W)_T = W(G)$, where $\{W(h) \colon h \in H\}$ is the isonormal Gaussian process on H. Thus, the given definition of stochastic integral is compatible with the construction of the isonormal process and extends it to non-deterministic functions.

Lastly, we check that the definition does not depend on the choice of the orthonormal basis $\{e_j \colon j \geq 1\}$ in V.

Lemma 3.2.1. *The definition of the stochastic integral does not depend on the particular orthonormal basis in V.*

Proof. Consider another orthonormal basis $\{v_i \colon i \geq 1\}$ in V and write $G(s,*) =$

$\sum_{i=1}^{\infty}\langle G(s,*), v_i\rangle_V v_i(*)$. Assuming that the series and integrals can be permuted,

$$\begin{aligned}\sum_{j=1}^{\infty}\int_0^T \langle G(s,*), e_j\rangle_V dW_s(e_j) &= \sum_{j=1}^{\infty}\int_0^T \left\langle \sum_{i=1}^{\infty}\langle G(s,*), v_i\rangle_V v_i, e_j\right\rangle_V dW_s(e_j)\\ &= \sum_{i,j}\int_0^T \langle G(s,*), v_i\rangle_V \langle v_i, e_j\rangle_V dW_s(e_j)\\ &= \sum_{i=1}^{\infty}\int_0^T \langle G(s,*), v_i\rangle_V \left(\sum_{j=1}^{\infty}\langle v_i, e_j\rangle_V dW_s(e_j)\right).\end{aligned}$$

The goal now is to show that, formally speaking,

$$\sum_{j=1}^{\infty}\langle v_i, e_j\rangle_V dW_s(e_j) = dW_s(v_i).$$

Since the map $\varphi \mapsto W_s(\varphi)$ is linear and continuous, we have that $v_i = \sum_{j=1}^{\infty}\langle v_i, e_j\rangle_V e_j$ (where the convergence is in V) implies

$$W_s(v_i) = \sum_{j=1}^{\infty}\langle v_i, e_j\rangle_V W_s(e_j)$$

where this last series converges in $L^2(\Omega)$ and $\mathbb{P}$-a.s. Let $g = \{g_s\colon s \in [0,T]\} \in L^2([0,T]\times\Omega)$ be a jointly measurable, adapted and real-valued process (at the end, $g_s = \langle G(s,*), v_i\rangle_V$), we want to show that, for each $i \geq 1$,

$$\int_0^T g_s dW_s(v_i) = \sum_{j=1}^{\infty} g_s\langle v_i, e_j\rangle_V dW_s(e_j) \tag{3.2.2}$$

where the series converges in $L^2(\Omega)$. If $g_s = X\mathbb{I}_{[s_0,t_0)}(s)$ with X a bounded $\mathcal{F}_{s_0}$-measurable random variable and $0 \leq s_0 < t_0 \leq T$, then

$$\begin{aligned}\int_0^T g_s dW_s(v_i) &= X\left(W_{t_0}(v_i) - W_{s_0}(v_i)\right)\\ &= X\sum_{j=1}^{\infty}\langle v_i, e_j\rangle_V \left(W_{t_0}(e_j) - W_{s_0}(e_j)\right)\\ &= \sum_{j=1}^{\infty}\int_0^T g_s\langle v_i, e_j\rangle_V dW_s(e_j).\end{aligned}$$

For such genera g, we know that there is a sequence of elementary processes $\{g_s^n \colon s \in [0,T], n \geq 1\}$ such that

$$\mathbb{E}\left[\int_0^T (g_s - g_s^n)^2 ds\right] \xrightarrow{n\to\infty} 0.$$

For these g^n, the claim holds by linearity and, by the isometry property, we have

$$\mathbb{E}\left[\left(\int_0^T (g_s - g_s^n) dW_s(v_i)\right)^2\right] = \mathbb{E}\left[\int_0^T (g_s - g_s^n)^2 ds\right] \xrightarrow{n\to\infty} 0.$$

Now, using the independence of the processes $\{W_s(e_j) \colon s \in [0,T], j \geq 1\}$, Tonelli's theorem and Parseval's identity,

$$\begin{aligned}
\mathbb{E}\left[\left(\sum_{j=1}^{\infty}\int_0^T (g_s - g_s^n)\langle v_i, e_k\rangle_V dW_s(e_j)\right)^2\right] &= \mathbb{E}\left[\sum_{j=1}^{\infty}\int_0^T (g_s - g_s^n)^2 \langle v_i, e_j\rangle_V^2 ds\right] \\
&= \sum_{j=1}^{\infty}\langle v_i, e_j\rangle_V^2 \mathbb{E}\left[\int_0^T (g_s - g_s^n)^2 ds\right] \\
&= \mathbb{E}\left[\int_0^T (g_s - g_s^n)^2 ds\right] \xrightarrow{n\to\infty} 0.
\end{aligned}$$

Hence

$$\int_0^T g_s^n dW_s(v_i) = \sum_{j=1}^{\infty}\int_0^T g_s^n \langle v_i, e_j\rangle_V dW_s(e_j)$$

converges, in $L^2(\Omega)$, to

$$\int_0^T g_s dW_s(v_i) \text{ and to } \sum_{j=1}^{\infty}\int_0^T g_s \langle v_i, e_j\rangle_V dW_s(e_j)$$

as n approaches infinity. Thus, by the $\mathbb{P}$-a.s. uniqueness of the $L^2(\Omega)$, (3.2.2) holds.

The only thing left to do is to justify the changes in the order of summation and integration. To do so, fix $M, N \geq 1$, then

$$\sum_{j=1}^{M}\int_0^T \left\langle \sum_{i=1}^{N}\langle G(s,*), v_i\rangle_V v_i, e_j\right\rangle_V dW_s(e_j) = \sum_{i=1}^{N}\sum_{j=1}^{M}\int_0^T \langle G(s,*), v_i\rangle_V \langle v_i, e_j\rangle_V dW_s(e_j). \tag{3.2.3}$$

Fixing N and letting $M \to \infty$ we obtain that the right-hand side of this last expression tends, in $L^2(\Omega)$, to

$$\sum_{i=1}^{N}\int_0^T \langle G(s,*), v_i\rangle_V dW_s(v_i) \tag{3.2.4}$$

by what we have just shown. On the other hand, the left-hand side of (3.2.3) converges, as $M \to \infty$, to

$$\int_0^T \int_D \left(\sum_{i=1}^N \langle G(s,*), v_i \rangle_V v_i(y) \right) W(ds,dy) \tag{3.2.5}$$

in $L^2(\Omega)$ and by definition of the stochastic integral. On the other hand, given that

$$G(s,*) = V - \lim_{N\to\infty} \sum_{i=1}^N \langle G(s,*), v_i \rangle_V v_i(*)$$

for $dsd\mathbb{P}$-a.a. $(s,\omega) \in [0,T] \times \Omega$, we have

$$\lim_{N\to\infty} \mathbb{E} \left[\int_0^T \left\| G(s,*) - \sum_{i=1}^N \langle G(s,*), v_i \rangle_V v_i \right\|_V^2 ds \right] = 0.$$

Hence, by the isometry formula (Proposition 3.2.1), taking limits as N approaches infinity in (3.2.5) and (3.2.4) leads to

$$\int_0^T \int_D G(s,y) W(ds,dy) = \sum_{i=1}^\infty \int_0^T \langle G(s,*), v_i \rangle_V dW_s(v_i),$$

as was to be shown. □

If $0 \le r \le t \le T$ and $A \subset [0,T] \times D$ is a Borel set, then we will write

$$\int_r^t \int_D G(s,y) W(ds,dy) := \int_0^T \mathbb{I}_{[r,t]}(s) G(s,y) W(ds,dy)$$

and

$$\int_A G(s,y) W(ds,dy) := \int_0^T \int_D \mathbb{I}_A(s,y) G(s,y) W(ds,dy),$$

which are well defined whenever G satisfies the hypothesis previously mentioned.

3.2.2 The indefinite integral

As in the case of the Itô integral, one can consider an integral process $\{(G \cdot W)_t \colon t \in [0,T]\}$ that will result in a continuous square integrable martingale. To see this, we will use the fact that the space of continuous square integrable $\{\mathcal{F}_t\}_t$-martingales in $[0,T]$ vanishing at 0, where indistinguishable processes are identified, is a Hilbert space with respect to the inner product

$$\langle M, N\rangle := \mathbb{E}\left[M_T N_T\right] = \mathbb{E}\left[\langle M, N\rangle_T\right].$$

For each $n \geq 1$, consider the continuous square integrable martingale

$$Z^n = \left\{ Z^n_t = \sum_{j=1}^{n} \int_0^t \langle G(s,*), e_j\rangle_V dW_s(e_j) \colon t \in [0,T] \right\}.$$

Which is continuous because we are considering the continuous version of the Itô integrals involved and we are taking finite linear combinations of them. By the independence of the processes $\{W_s(e_j)\colon s \in [0,T], j \geq 1\}$, the quadratic variation of these martingales will be given by

$$\langle Z^n\rangle = \left\{ \langle Z^n_t\rangle = \sum_{j=1}^{n} \int_0^t \langle G(s,*), e_j\rangle_V^2 ds \colon t \in [0,T] \right\}.$$

Since $\{Z^n_T\}_n$ converges in $L^2(\Omega)$ to $(G{\cdot}W)_T$, we deduce that the sequence of continuous square integrable martingales $\{Z^n\}_n$ converges in the corresponding space to a certain limit denoted by $G \cdot W = \{(G \cdot W)_t \colon t \in [0,T]\}$ which will be a continuous square integrable martingale vanishing at $t = 0$. This last process will be called the indefinite integral process of G with respect to W. Since Z^n and $G \cdot W$ are both martingales, this in particular implies that, by Doob's maximal inequality,

$$\begin{aligned}
\mathbb{E}\left[((G \cdot W)_t - Z^n_t)^2\right] &\leq \mathbb{E}\left[\sup_{0 \leq s \leq T} ((G \cdot W)_s - Z^n_s)^2\right] \\
&\leq \mathbb{E}\left[((G \cdot W)_T - Z^n_T)^2\right] \xrightarrow{n \to \infty} 0
\end{aligned}$$

for each $t \in [0,T]$. Hence, we have that Z^n_t converges to $(G \cdot W)_t$ uniformly in $t \in [0,T]$.

Lemma 3.2.2. *For each $t \in [0,T]$,*

$$(G \cdot W)_t = \int_0^T \int_D G(s,y)\mathbb{I}_{[0,t]}(s) W(ds,dy), \quad \mathbb{P}\text{-}a.s.$$

The main point being that the random variable in left-hand side of the equality corresponds to the limit in the space of continuous square integrable martingales vanishing at zero, while the right-hand side is the corresponds to the stochastic integral in Definition 3.2.1.

Proof. Using Definition 3.2.1, we have

$$\begin{aligned}\int_0^T \int_D G(s,y)\mathbb{I}_{[0,t]}(s)W(ds,dy) &= L^2(\Omega) - \lim_{n\to\infty} \sum_{j=1}^n \int_0^T \langle \mathbb{I}_{[0,t]}(s)\langle G(s,*), e_j\rangle_V dW_s(e_j) \\ &= L^2(\Omega) - \lim_{n\to\infty} \sum_{j=1}^n \int_0^t \langle G(s,*), e_j\rangle_V dW_s(e_j) \\ &= L^2(\Omega) - \lim_{n\to\infty} Z_t^n \\ &= (G\cdot W)_t.\end{aligned}$$

The result follows from the fact that the $L^2(\Omega)$ limit of a sequence of random variables is $\mathbb{P}$-a.s. unique. □

Proposition 3.2.2. *The quadratic variation process of $G \cdot W$ is*

$$\left\{ \int_0^t ||G(s,*)||_V^2 ds \colon t \in [0,T] \right\}.$$

Proof. Let us set $Z_t = (G\cdot W)_t$. Observe that, by Cauchy-Schwarz's and the triangle inequalities,

$$\begin{aligned}\mathbb{E}\left[\left|(Z_t^n)^2 - Z_t^2\right|\right] &= \mathbb{E}\left[|Z_t^n - Z_t|\cdot|Z_t^n + Z_t|\right] \\ &\le ||Z_t^n - Z_t||_{L^2(\Omega)} ||Z_t^n + Z_t||_{L^2(\Omega)} \\ &\le ||Z_t^n - Z_t||_{L^2(\Omega)} \left(||Z_t^n||_{L^2(\Omega)} + ||Z_t||_{L^2(\Omega)}\right).\end{aligned}$$

Now, on one hand, by (2.3.4),

$$\mathbb{E}\left[(Z_t^n)^2\right] = \mathbb{E}\left[\sum_{j=1}^n \int_0^t \langle G(s,*), e_j\rangle_V^2 ds\right] \le \mathbb{E}\left[\int_0^t ||G(s,*)||_V^2 ds\right].$$

On the other hand, by the previous lemma and the isometry formula (Proposition 3.2.1),

$$\mathbb{E}\left[Z_t^2\right] = \mathbb{E}\left[\int_0^t ||G(s,*)||_V^2 ds\right].$$

So, all in all,

$$\mathbb{E}\left[\left|(Z_t^n)^2 - Z_t^2\right|\right] \le 2\left(\mathbb{E}\left[\int_0^t ||G(s,*)||_V^2 ds\right]\right)^{1/2} ||Z_t^n - Z_t||_{L^2(\Omega)} \xrightarrow{n\to\infty} 0.$$

So we have that, for each $t\in[0,T]$, $(Z_t^n)^2$ converges to Z_t^2 in $L^1(\Omega)$.

Now observe that the process

$$X = \left\{ X_t := \sum_{j=1}^{\infty} \int_0^t \langle G(s,*), e_j \rangle_V^2 ds = \int_0^t ||G(s,*)||_V^2 ds \colon t \in [0,T] \right\}$$

is adapted, continuous, non-decreasing and vanishing at $t = 0$. Moreover, it satisfies

$$\mathbb{E}\left[|\langle Z^n \rangle_t - X_t| \right] = \mathbb{E}\left[\sum_{j=n+1}^{\infty} \int_0^t \langle G(s,*), e_j \rangle_V^2 ds \right] \xrightarrow{n\to\infty} 0.$$

Where the limit follows from the fact that $\sum_{j=n+1}^{\infty} \int_0^t \langle G(s,*), e_j \rangle_V^2 ds$ is the tail of a convergent series (by Parseval's identity) and that it can be bounded by $\int_0^t ||G(s,*)||_V^2 ds$, which has finite expected value (so we can apply the dominated convergence theorem). Thus, $\langle Z^n \rangle_t$ converges to X_t in $L^1(\Omega)$.

With this in mind, we conclude that

$$\mathbb{E}\left[\left| (Z_t^n)^2 - \langle Z^n \rangle_t - Z_t^2 + X_t \right| \right] \leq \mathbb{E}\left[\left| (Z_t^n)^2 - Z_t^2 \right| \right] + \mathbb{E}\left[|\langle Z^n \rangle_t - X_t| \right] \xrightarrow{n\to\infty} 0.$$

Thus, the process $\{Z_t^2 - X_t \colon t \in [0,T]\}$ is the $L^1(\Omega)$-limit of continuous martingales and hence, a continuous martingale itself (with respect to the same filtration). By uniqueness of the quadratic variation, we conclude that X is the quadratic variation process of $G \cdot W$. □

3.2.3 Some other properties

Given that the stochastic integral is defined via Itô integrals, many of the properties satisfied by the latter are passed onto the former. We shall mention a couple of them in this section without giving the proofs.

Lemma 3.2.3. *Let $G^{(1)}$ and $G^{(2)}$ be two stochastic processes satisfying conditions (i)-(iii) so that their stochastic integrals are well defined and such that, on some $F \in \mathcal{F}$, for $d\mathbb{P}$-a.a. $\omega \in F$, $G^{(1)}(s,y,\omega) = G^{(2)}(s,y,\omega)$ for $dsdy$-a.a. $(s,y) \in [0,T] \times D$ (the sample paths are the same $\mathbb{P}$-a.s. in F), then, for all $t \in [0,T]$,*

$$\mathbb{I}_F \int_0^t \int_D G^{(1)}(s,y) W(ds,dy) = \mathbb{I}_F \int_0^t \int_D G^{(2)}(s,y) W(ds,dy), \quad \mathbb{P}\text{-a.s.}$$

Lemma 3.2.4. *Let $D_1 \subset D$ be a domain (non-empty connected open set) and $\{v_i \colon i \geq 1\}$ an orthonormal basis of $V_1 = L^2(D_1)$. Let $G = \{G(s,y) \colon (s,y) \in [0,T] \times D_1\}$ be a stochastic process satisfying assumptions (i)-(iii) in D_1 instead of D so that the stochastic integral is well defined. We extend G to $[0,T] \times D$ by setting $G(s,y) = 0$*

for all $s \in [0,T]$, $y \in D \backslash D_1$. Then assumptions (i)-(iii) are fulfilled by the extension of G in D and

$$\sum_{i=1}^{\infty} \int_0^T \langle G(s,*), v_i \rangle_{V_1} dW_s(v_i) = \sum_{j=1}^{\infty} \int_0^T \langle G(s,*), e_j \rangle_V dW_s(e_j)$$

where both series converge in $L^2(\Omega)$.

Lemma 3.2.5. *Let τ be an $\{\mathcal{F}_t\}_t$-stopping time (where the filtration satisfies the conditions (i) and (ii) seen before Lemma 3.1.5) with values in $[0,T]$ and let G be a stochastic process satisfying conditions (i)-(iii) so that its stochastic integral is well defined. Then*

$$(G \cdot W)_\tau = \int_0^T \int_D \mathbb{I}_{[0,\tau]}(s) G(s,y) W(ds,dy), \quad \mathbb{P}\text{-}a.s.$$

3.2.4 An approximation result

As in the case of the stochastic integral with respect to the Brownian motion, in this section we will prove an analogous result to the one seen in Section 2.3.4 for the Gaussian white noise. That is, we will be considering a sequence of random noise $\{W_n\}_n$ that will approximate in some sense a Gaussian space-time white noise W and see that, for some suitable family of processes $f = \{f(t,x) \colon (t,x) \in [0,T] \times D\}$, $D \subset \mathbb{R}^k$, the laws of the integrals of f with respect to the noises W_n converge, as n approaches infinity, to the law of the integral of f with respect to the Gaussian white noise.

For this part, we will follow the ideas in [2], but we will assume some extra regularity on the process f in order to simplify the proofs. Thus, from now on, we will assume that $k = 1$ and that $D = [0, L]$ for some $L > 0$ and that the process f is deterministic and continuous. This in particular implies that the function f is in $L^2([0,T] \times [0,L])$ and that, by the observation made right after Proposition 3.2.1 and (3.1.3), the stochastic integral of f with respect to the space-time white noise can be thought as an integral with respect to the Brownian sheet. Hence, we need to provide an example of sequence of stochastic processes that approximates the Brownian sheet. For this last purpose, we recall that the main result of Section 2.3.4, Theorem 2.3.11, was motivated by the fact that the standard Brownian motion could be approximated by the random walk. A similar result also holds in the plane, more particularly,

Theorem 3.2.1. *Let $\{Z_k\}_{k \in \mathbb{N}^2}$ be a sequence of i.i.d. centered random variables with unitary variance. For any $n \in \mathbb{N}$ we define*

$$\theta_n(t,x) = n \sum_{k=(k^1,k^2) \in \mathbb{N}^2} Z_k \mathbb{I}_{[k^1-1,k^1) \times [k^2-1,k^2)}(tn, xn), \quad (t,x) \in [0,T] \times [0,L]$$

and

$$\zeta_n(t,x) = \int_0^t \int_0^x \theta_n(s,y)dyds, \quad (t,x) \in [0,T] \times [0,L].$$

Then $\zeta_n = \{z_n(t,x) \colon (t,x) \in [0,T] \times [0,L]\}$ *converges in law, in the space of continuous functions* $\mathcal{C}([0,T] \times [0,L])$*, to a Brownian sheet as* n *approaches infinity.*

For a proof of this result we refer to [12], Corollary 1 in page 683. Observe that this, in particular, implies that the finite dimensional distributions of ζ_n converge in law to the ones of the Brownian sheet as n approaches infinity.

As in the proof of Theorem 2.3.11, we will require the extra condition that $\mathbb{E}[Z_k^m] < \infty$ for some even integer $m \in \mathbb{N}$ large enough.

With all this, we are now ready to state the main result of this section.

Theorem 3.2.2. *Let* $f = \{f(t,x) \colon (t,x) \in [0,T] \times [0,L]\}$ *be a deterministic, stochastic process and, for each* $n \in \mathbb{N}$*, let*

$$I_n(t,x) = \int_0^t \int_0^x f(s,y)\theta_n(s,y)dyds.$$

Then, the processes $I_n = \{I_n(t,x) \colon (t,x) \in [0,T] \times [0,L]\}$ *converge in law to the process* $I = \{I(t,x) \colon (t,x) \in [0,T] \times [0,L]\}$ *defined by*

$$I(t,x) = \int_0^t \int_0^x f(s,y)W(ds,dy),$$

in the space of continuous functions $\mathcal{C}([0,T] \times [0,L])$*, as* n *approaches infinity.*

To prove this result, we will again prove that the sequence $\{I_n\}_n$ is tight and that the corresponding finite dimensional distributions converge to the ones of I. In the meantime, we will be proving as well that, for each $n \in \mathbb{N}$, the process I_n has a continuous modification, so that we can think of such processes as random functions in $\mathcal{C}([0,T] \times [0,L])$.

To prove the tightness of the sequence, we will be using the following result, which generalizes the already used Billingsley's Criterion (see [13], Proposition 2.3 in page 95).

Theorem 3.2.3. *Let* $\{X_n\}_{n\in\mathbb{N}}$ *be a family of random functions in* $\mathcal{C}([0,T] \times [0,L])$*. The family is tight if there exist* $q, p > 0$*,* $\delta > 2$ *and a positive constant* C *such that*

$$\sup_{n\geq 1} \mathbb{E}[|X_n(0,0)|^q] < \infty,$$

and, for every $t, s \in [0,T]$ *and* $x, y \in [0,L]$*,*

$$\sup_{n\geq 1} \mathbb{E}[|X_n(t,x) - X_n(s,y)|^p] \leq C\left(|x-y| + |t-s|\right)^{\delta}. \tag{3.2.6}$$

In our setting, the first condition is clearly satisfied, so we will only have to worry about the second one.

Now recall that in the proof of Theorem 2.3.11, it sufficed to prove that, for any $0 \leq t_1 < ... < t_m \leq T$, the sequence of random vectors

$$I_f^{(n)} := \left(I_{t_1}^{(n)}, I_{t_2}^{(n)} - I_{t_1}^{(n)}, ..., I_{t_m}^{(n)} - I_{t_{m-1}}^{(n)} \right)$$

converged in law to the random vector

$$I_f := \left(I_{t_1}, I_{t_2} - I_{t_1}, ..., I_{t_m} - I_{t_{m-1}} \right)$$

as n approaches infinity. This was done so that we could consider a slight modification of the vectors $I_f^{(n)}$, the vectors $\Delta^{(n)}$, which had the property that its components were mutually independent. To do this, the fact that we were working in a one dimensional space, $[0, T]$, was strongly used. Unfortunately, there is no natural way to extend these arguments in the context of several dimensions, so we will have to resort to other types of arguments to prove the convergence of the finite dimensional distributions. This is the purpose of the lemma that we state below.

Lemma 3.2.6. *Let $(F, ||\cdot||)$ be a normed vector space and $\{J^n\}_{n\in\mathbb{N}}$ and J linear maps from F to $L^1(\Omega)$ such that*

$$\sup_{n\geq 1} \mathbb{E}\left[|J^n(f)|\right] \leq C||f||$$

and

$$\mathbb{E}\left[|J(f)|\right] \leq C||f||$$

for any $f \in F$ and for some positive constant C. Moreover, assume that there is a dense subset D of F with respect to the norm $||\cdot||$ such that $J^n(f)$ converges in law to $J(f)$ as n approaches infinity for any $f \in D$. Then, for any $f \in F$, $J^n(f)$ converges in law to $J(f)$ as n approaches infinity.

Proof. Recall that a sequence of random variables $\{X_n\}_{n\in\mathbb{N}}$ converges in law to a random variable X if, and only if, for any Lipschitz function $g \colon \mathbb{R} \to \mathbb{R}$,

$$\mathbb{E}\left[g(X_n)\right] \xrightarrow{n\to\infty} \mathbb{E}\left[g(X)\right].$$

Thus, we shall see that, for any $\varepsilon > 0$, there is $n \in \mathbb{N}$ large enough such that

$$|\mathbb{E}\left[g\left(J^n(f)\right)\right] - \mathbb{E}\left[g\left(J(f)\right)\right]| < \varepsilon, \tag{3.2.7}$$

where g is any Lipschitz function as before. Consider any $h \in D$ such that $||f-h|| < \frac{\varepsilon}{3L_g C}$, where $L_g > 0$ is the Lipschitz constant of g, and apply the triangle inequality to obtain

$$\begin{aligned}|\mathbb{E}[g(J^n(f))] - \mathbb{E}[g(J(f))]| \leq & |\mathbb{E}[g(J^n(f))] - \mathbb{E}[g(J^n(h))]| \\ & + |\mathbb{E}[g(J^n(h))] - \mathbb{E}[g(J(h))]| \\ & + |\mathbb{E}[g(J(h))] - \mathbb{E}[g(J(f))]| \,.\end{aligned}$$

Now observe that

$$\begin{aligned}|\mathbb{E}[g(J^n(f))] - \mathbb{E}[g(J^n(h))]| &\leq \mathbb{E}[|g(J^n(f)) - g(J^n(h))|] \\ &\leq L_g \mathbb{E}[|J^n(f-h)|] \\ &\leq L_g C ||f-h|| < \frac{\varepsilon}{3}\end{aligned}$$

Similarly,

$$|\mathbb{E}[g(J(f))] - \mathbb{E}[g(J(h))]| \leq L_g C ||f-h|| < \frac{\varepsilon}{3}.$$

Finally, given that $J^n(h)$ converges in law to $J(h)$ for $h \in D$, we have that, for n large enough,

$$|\mathbb{E}[g(J^n(h))] - \mathbb{E}[g(J(h))]| < \frac{\varepsilon}{3}.$$

Thus, for n large enough, we obtain (3.2.7) as desired. □

The following, and last, preliminary result we introduce allows us to bound the moments involved in (3.2.6) and to apply the previous lemma in the case where $(F, ||\cdot||) = (L^2([0,T]\times[0,L]), ||\cdot||_2)$ where $||\cdot||_2$ is the usual norm in $L^2([0,T]\times[0,L])$.

Lemma 3.2.7. *There is a positive constant C_m such that, for any function $f \in L^2([0,T]\times[0,L])$, we have*

$$\mathbb{E}\left[\left(\int_0^T \int_0^L f(t,x)\theta_n(t,x)dxdt\right)^m\right] \leq C_m \left(\int_0^T \int_0^L f^2(t,x)dxdt\right)^{\frac{m}{2}}$$

for any $n \in \mathbb{N}$.

Proof. We have that

$$\begin{aligned}&\mathbb{E}\left[\left(\int_0^T \int_0^L f(t,x)\theta_n(t,x)dxdt\right)^m\right] \qquad (3.2.8)\\ &= \int_{[0,T]^m\times[0,L]^m} f(t_1,x_1)\cdot\ldots\cdot f(t_m,x_m)\mathbb{E}\left[\prod_{j=1}^m \theta_n(t_j,x_j)\right] dt_1\ldots dt_m dx_1\ldots dx_m,\end{aligned}$$

with

$$\mathbb{E}\left[\prod_{j=1}^{m}\theta_n(t_j,x_j)\right] = n^m \sum_{k_1,\ldots,k_m\in\mathbb{N}^2} \mathbb{E}\left[Z_{k_1}\cdot\ldots\cdot Z_{k_m}\right]\prod_{j=1}^{m}\mathbb{I}_{[k_j^1-1,k_j^1)}(t_jn)\mathbb{I}_{[k_j^2-1,k_j^2)}(x_jn).$$

Now observe that $\mathbb{E}\left[Z_{k_1}\cdot\ldots\cdot Z_{k_m}\right]=0$ when there is some $j\in\{1,...,m\}$ such that $k_j\neq k_i$ for all $i\in\{1,...,m\}\backslash\{j\}$. Thus, we have that

$$\mathbb{E}\left[\prod_{j=1}^{m}\theta_n(t_j,x_j)\right] = n^m \sum_{(k_1,\ldots,k_m)\in A^m} \mathbb{E}\left[Z_{k_1}\cdot\ldots\cdot Z_{k_m}\right]\prod_{j=1}^{m}\mathbb{I}_{[k_j^1-1,k_j^1)}(t_jn)\mathbb{I}_{[k_j^2-1,k_j^2)}(x_jn)$$

where A^m is the set of $(k_1,...,k_m)\in(\mathbb{N}^2)^m$ such that for all $l\in\{1,...,m\}$, there is some $j\in\{1,...,m\}\backslash\{l\}$ such that $k_j=k_l$. In this set A^m, we have that

$$\left|\mathbb{E}\left[Z_{k_1}\cdot\ldots\cdot Z_{k_m}\right]\right| \leq \mathbb{E}\left[\left|Z_{k_1}\cdot\ldots\cdot Z_{k_m}\right|\right] = \mathbb{E}\left[|Z|^{\alpha_1}\right]\cdot\ldots\mathbb{E}\left[|Z|^{\alpha_r}\right]$$

where $\alpha_i\in\{0,1,...,m\}$ for each $i\in\{1,...,r\}$, $\alpha_1+...+\alpha_r=m$ and Z is a random variable with the same law as Z_k for each $k\in\mathbb{N}^2$. Given that $\mathbb{E}\left[Z^m\right]<\infty$, we have that $\left|\mathbb{E}\left[Z_{k_1}\cdot\ldots\cdot Z_{k_m}\right]\right|\leq C_m$ for some positive constant C_m which does not depend on $\alpha_1,...,\alpha_r$ and for all $(k_1,...,k_m)\in A^m$. Thus, we have

$$\left|\mathbb{E}\left[\prod_{j=1}^{m}\theta_n(t_j,x_j)\right]\right| \leq n^m C_m \sum_{(k_1,\ldots,k_m)\in A^m} \prod_{j=1}^{m}\mathbb{I}_{[k_j^1-1,k_j^1)}(t_jn)\mathbb{I}_{[k_j^2-1,k_j^2)}(x_jn).$$

Now observe that

$$\sum_{(k_1,\ldots,k_m)\in A^m} \prod_{j=1}^{m}\mathbb{I}_{[k_j^1-1,k_j^1)}(t_jn)\mathbb{I}_{[k_j^2-1,k_j^2)}(x_jn) \leq \mathbb{I}_{D^m}(t_1,...,t_m;x_1,...,x_m)$$

where D^m denotes the set of points $(t_1,...,t_m;x_1,...,x_m)\in[0,T]^m\times[0,L]^m$ such that, for all $l\in\{1,...,m\}$, there is some $j\in\{1,...,m\}\backslash\{l\}$ such that $|t_j-t_l|<n^{-1}$ and $|x_j-x_l|<n^{-1}$ and that, if there is some $r\neq j,l$ such that $|t_l-t_r|<n^{-1}$ and $|x_l-x_r|<n^{-1}$, then $|t_j-t_r|<2n^{-1}$ and $|x_j-x_r|<2n^{-1}$. Indeed, suppose there is some $(k_1,...,k_m)\in A^m$ such that

$$\prod_{j=1}^{m}\mathbb{I}_{[k_j^1-1,k_j^1)}(t_jn)\mathbb{I}_{[k_j^2-1,k_j^2)}(x_jn)\neq 0.$$

This in particular means that $\mathbb{I}_{[k_j^1-1,k_j^1)}(t_jn)\mathbb{I}_{[k_j^2-1,k_j^2)}(x_jn)\neq 0$ for all $j\in\{1,...,m\}$ and thus, that

$$(t_j,x_j)\in\left[\frac{k_j^1-1}{n},\frac{k_j^1}{n}\right)\times\left[\frac{k_j^2-1}{n},\frac{k_j^2}{n}\right).$$

Now recall that, in A^m, for each $j \in \{1, ..., m\}$, there is some $l \in \{1, ..., m\} \backslash \{j\}$ such that $k_l = k_j$. For such pair (j, l), we will have that, letting $k_l = k_j = k = (k^1, k^2)$,

$$(t_j, x_j), (t_l, x_l) \in \left[\frac{k^1 - 1}{n}, \frac{k^1}{n}\right) \times \left[\frac{k^2 - 1}{n}, \frac{k^2}{n}\right).$$

Hence, $|t_j - t_l| < n^{-1}$ and $|x_j - x_l| < n^{-1}$. As for the second part, suppose there is some $r \neq j, l$ verifying the previously mentioned property. Then, the claim follows directly from the triangle inequality.

Next, we have that $\mathbb{I}_{D^m}(t_1, ..., t_m; x_1, ..., x_m)$ can be bounded by a finite sum (whose number of summands depends only on m) of products of indicators of the form

$$\mathbb{I}_{[0,n^{-1})^2}(|t_j - t_l|, |x_j - x_l|), \quad \mathbb{I}_{[0,2n^{-1})^2}(|t_j - t_r|, |x_j - x_r|),$$

or

$$\mathbb{I}_{[0,n^{-1})^2}(|t_j - t_l|, |x_j - x_l|)\mathbb{I}_{[0,n^{-1})^2}(|t_l - t_r|, |x_l - x_r|)\mathbb{I}_{[0,2n^{-1})^2}(|t_j - t_r|, |x_j - x_r|)$$

Moreover, in each product, all of the variables $(t_1, ..., t_m; x_1, ..., x_m)$ appear and they do it in only one of the previously specified indicators.
Indeed, in D^m, it can happen that, for $l = 1$, the corresponding j is 2 and that there is no $r \neq j, l$ such that $|t_l - t_r| < n^{-1}$ and $|x_l - x_r| < n^{-1}$, that for $l = 3$, the corresponding j is 5 and that for $r = 4$ and $r = 6$, we have $|t_l - t_r| < n^{-1}$ and $|x_l - x_r| < n^{-1}$, that for $l = 7$, the corresponding j is 8 and that $r = 9$ is the only one that satisfies $|t_l - t_r| < n^{-1}$ and $|x_l - x_r| < n^{-1}$, etc. In this case, the corresponding product of indicators is

$$\mathbb{I}_{[0,n^{-1})^2}(|t_2 - t_1|, |x_2 - x_1|) \cdot \mathbb{I}_{[0,n^{-1})^2}(|t_5 - t_6|, |x_5 - x_6|) \cdot \mathbb{I}_{[0,2n^{-1})^2}(|t_3 - t_4|, |x_3 - x_4|)$$
$$\times \mathbb{I}_{[0,n^{-1})^2}(|t_8 - t_7|, |x_8 - x_7|)\mathbb{I}_{[0,n^{-1})^2}(|t_7 - t_9|, |x_7 - x_9|)\mathbb{I}_{[0,2n^{-1})^2}(|t_8 - t_9|, |x_8 - x_9|)...$$

The point of doing this being that we can group the pairs (t_j, x_j) in groups of two or three. Taking into account all the possibilities (which are finite and depend only on the number m) leads to the bound for $\mathbb{I}_{D^m}(t_1, ..., t_m; x_1, ..., x_m)$.

Thus, (3.2.8), or its absolute value, can be bounded by a finite sum of products of terms of the form (up to some positive multiplicative factor C_m)

$$n^2 \int_{[0,T]^2 \times [0,L]^2} |f(t_j, x_j)||f(t_l, x_l)|\mathbb{I}_{[0,n^{-1})}(|t_j - t_l|)\mathbb{I}_{[0,n^{-1})}(|x_j - x_l|)dt_j dt_l dx_j dx_l, \quad (3.2.9)$$

$$n^2 \int_{[0,T]^2 \times [0,L]^2} |f(t_j, x_j)||f(t_r, x_r)|\mathbb{I}_{[0,2n^{-1})}(|t_j - t_r|)\mathbb{I}_{[0,2n^{-1})}(|x_j - x_r|)dt_j dt_r dx_j dx_r, \quad (3.2.10)$$

or

$$n^3 \int_{[0,T]^3\times[0,L]^3} |f(t_j,x_j)||f(t_l,x_l)||f(t_r,x_r)|$$
$$\times \mathbb{I}_{[0,n^{-1})}(|t_j-t_l|)\mathbb{I}_{[0,n^{-1})}(|t_l-t_r|)\mathbb{I}_{[0,2n^{-1})}(|t_j-t_r|)$$
$$\times \mathbb{I}_{[0,n^{-1})}(|x_j-x_l|)\mathbb{I}_{[0,n^{-1})}(|x_l-x_r|)\mathbb{I}_{[0,2n^{-1})}(|x_j-x_r|)dt_jdt_ldt_rdx_jdx_ldx_r. \quad (3.2.11)$$

Observe that, in each summand, the total number of terms that constitutes the product is such that the factors n^α, $\alpha \in \{1,2\}$, in the previous expressions, when multiplied, give n^m (that is, $n^{\alpha_1}\cdot ...n^{\alpha_s} = n^m$, where s is the number of elements that constitutes each product). Thus, it suffices to show that the first two types of terms, (3.2.9) and (3.2.10), can be bounded, up to some factor independent of n, by $||f||_2^2$, while the terms of the form (3.2.11) can be bounded, again, up to some factor independent of n, by $||f||_2^3$.

Using that, for any $a,b\geq 0$, $ab \leq 2ab \leq a^2+b^2$, we obtain that the integrals in terms like (3.2.9) can be bounded as follows

$$\int_{[0,T]^2\times[0,L]^2} |f(t_j,x_j)||f(t_l,x_l)|\mathbb{I}_{[0,n^{-1})}(|t_j-t_l|)\mathbb{I}_{[0,n^{-1})}(|x_j-x_l|)dt_jdt_ldx_jdx_l$$
$$\leq \int_{[0,T]^2\times[0,L]^2} f^2(t_j,x_j)\mathbb{I}_{[0,n^{-1})}(|t_j-t_l|)\mathbb{I}_{[0,n^{-1})}(|x_j-x_l|)dt_jdt_ldx_jdx_l$$
$$+\int_{[0,T]^2\times[0,L]^2} f^2(t_l,x_l)\mathbb{I}_{[0,n^{-1})}(|t_j-t_l|)\mathbb{I}_{[0,n^{-1})}(|x_j-x_l|)dt_jdt_ldx_jdx_l$$
$$=2\int_{[0,T]^2\times[0,L]^2} f^2(t_l,x_l)\mathbb{I}_{[0,n^{-1})}(|t_j-t_l|)\mathbb{I}_{[0,n^{-1})}(|x_j-x_l|)dt_jdt_ldx_jdx_l$$
$$=2\int_0^T\int_0^L f^2(t_l,x_l)\left(\int_0^T \mathbb{I}_{[0,n^{-1})}(|t_j-t_l|)dt_j\right)\left(\int_0^L \mathbb{I}_{[0,n^{-1})}(|x_j-x_l|)dx_j\right)dx_ldt_l$$
$$\leq \frac{2}{n^2}\int_0^T\int_0^L f^2(t_l,x_l)dx_ldt_l.$$

The second type of terms, (3.2.10), can be treated in a similar way (an extra factor 2 might appear, but this is no problem). Finally, for the integrals in (3.2.11), we have, using that $abc \leq ab^2+ac^2$ for any $a,b,c\geq 0$,

$$\int_{[0,T]^3\times[0,L]^3} |f(t_j,x_j)||f(t_l,x_l)||f(t_r,x_r)|$$
$$\times \mathbb{I}_{[0,n^{-1})}(|t_j-t_l|)\mathbb{I}_{[0,n^{-1})}(|t_l-t_r|)\mathbb{I}_{[0,2n^{-1})}(|t_j-t_r|)$$
$$\times \mathbb{I}_{[0,n^{-1})}(|x_j-x_l|)\mathbb{I}_{[0,n^{-1})}(|x_l-x_r|)\mathbb{I}_{[0,2n^{-1})}(|x_j-x_r|)dt_jdt_ldt_rdx_jdx_ldx_r$$

$$\leq 2\int_{[0,T]^3\times[0,L]^3}|f(t_l,x_l)|f^2(t_j,x_j)\mathbb{I}_{[0,n^{-1})}(|t_j-t_l|)\mathbb{I}_{[0,n^{-1})}(|t_l-t_r|)\mathbb{I}_{[0,2n^{-1})}(|t_j-t_r|)$$
$$\times\, \mathbb{I}_{[0,n^{-1})}(|x_j-x_l|)\mathbb{I}_{[0,n^{-1})}(|x_l-x_r|)\mathbb{I}_{[0,2n^{-1})}(|x_j-x_r|)dt_jdt_ldt_rdx_jdx_ldx_r$$
$$\leq 2\int_{[0,T]^3\times[0,L]^3}|f(t_l,x_l)|f^2(t_j,x_j)\mathbb{I}_{[0,n^{-1})}(|t_j-t_l|)\mathbb{I}_{[0,n^{-1})}(|t_l-t_r|)$$
$$\times\, \mathbb{I}_{[0,n^{-1})}(|x_j-x_l|)\mathbb{I}_{[0,n^{-1})}(|x_l-x_r|)dt_jdt_ldt_rdx_jdx_ldx_r$$
$$= 2\int_{[0,T]^2\times[0,L]^2}|f(t_l,x_l)|f^2(t_j,x_j)\mathbb{I}_{[0,n^{-1})}(|t_j-t_l|)\mathbb{I}_{[0,n^{-1})}(|x_j-x_l|)$$
$$\times\left(\int_0^T\mathbb{I}_{[0,n^{-1})}(|t_l-t_r|)dt_r\right)\left(\int_0^L\mathbb{I}_{[0,n^{-1})}(|x_l-x_r|)dx_r\right)dt_jdt_ldx_jdx_l$$
$$\leq \frac{2}{n^2}\int_{[0,T]^2\times[0,L]^2}|f(t_l,x_l)|f^2(t_j,x_j)\mathbb{I}_{[0,n^{-1})}(|t_j-t_l|)\mathbb{I}_{[0,n^{-1})}(|x_j-x_l|)dt_jdt_ldx_jdx_l$$
$$= \frac{2}{n^2}\int_0^T\int_0^L|f(t_l,x_l)|$$
$$\times\left(\int_0^T\int_0^L f^2(t_j,x_j)\mathbb{I}_{[0,n^{-1})}(|t_j-t_l|)\mathbb{I}_{[0,n^{-1})}(|x_j-x_l|)dx_jdt_j\right)dx_ldt_l.$$

By the Cauchy-Schwartz inequality, this last expression can be bounded by

$$\frac{2||f||_2}{n^2}\left[\int_0^T\int_0^L\left(\int_0^T\int_0^L f^2(t_j,x_j)\mathbb{I}_{[0,n^{-1})}(|t_j-t_l|)\mathbb{I}_{[0,n^{-1})}(|x_j-x_l|)dx_jdt_j\right)^2dx_ldt_l\right]^{\frac{1}{2}}$$

But

$$\int_0^T\int_0^L\left(\int_0^T\int_0^L f^2(t_j,x_j)\mathbb{I}_{[0,n^{-1})}(|t_j-t_l|)\mathbb{I}_{[0,n^{-1})}(|x_j-x_l|)dx_jdt_j\right)^2dx_ldt_l$$
$$= \int_{[0,T]^3\times[0,L]^3}f^2(t_j,x_j)f^2(t_p,x_p)\mathbb{I}_{[0,n^{-1})}(|t_j-t_l|)\mathbb{I}_{[0,n^{-1})}(|x_j-x_l|)$$
$$\times\, \mathbb{I}_{[0,n^{-1})}(|t_p-t_l|)\mathbb{I}_{[0,n^{-1})}(|x_p-x_l|)dt_jdt_pdt_ldx_jdx_pdx_l$$
$$\leq \int_{[0,T]^3\times[0,L]^3}f^2(t_j,x_j)f^2(t_p,x_p)\mathbb{I}_{[0,n^{-1})}(|t_j-t_l|)\mathbb{I}_{[0,n^{-1})}(|x_j-x_l|)dt_jdt_pdt_ldx_jdx_pdx_l$$
$$= \int_{[0,T]^2\times[0,L]^2}f^2(t_j,x_j)f^2(t_p,x_p)$$
$$\times\left(\int_0^T\mathbb{I}_{[0,n^{-1})}(|t_j-t_l|)dt_l\right)\left(\int_0^L\mathbb{I}_{[0,n^{-1})}(|x_j-x_l|)dx_l\right)dt_jdt_pdx_jdx_p$$
$$\leq \frac{1}{n^2}\int_{[0,T]^2\times[0,L]^2}f^2(t_j,x_j)f^2(t_p,x_p)dt_jdt_pdx_jdx_p = \frac{||f||_2^4}{n^2}.$$

Proving that terms like (3.2.11) can be bounded, up to some factor, by $||f||_2^3$ and thus, finishing the proof of this lemma. $\square$

With all this, we are now ready to prove Theorem 3.2.2.

Proof of Theorem 3.2.2.

Tightness and existence of a continuous version
For fixed $(t,x) \in [0,T]\times[0,L]$, consider the function $H_{t,x}(s,y) = \mathbb{I}_{[0,t]}(s)\mathbb{I}_{[0,x]}(y)f(s,y)$, $(s,y) \in [0,T]\times[0,L]$. Then, for any $(t,x),(t',x') \in [0,T]\times[0,L]$, we have that

$$I_n(t',x') - I_n(t,x) = \int_0^T \int_0^L \left(H_{t',x'}(s,y) - H_{t,x}(s,y)\right)\theta_n(s,y)dyds.$$

Since f is continuous, we have that $H_{t,x} \in L^2([0,T]\times[0,L])$ and similarly for $H_{t',x'}$, so their difference will be in $L^2([0,T]\times[0,L])$. Then, by Lemma 3.2.7, we have that

$$\mathbb{E}\left[|I_n(t',x') - I_n(t,x)|^m\right] \leq C_m \left(\int_0^T \int_0^L \left(H_{t',x'}(s,y) - H_{t,x}(s,y)\right)^2 dyds\right)^{\frac{m}{2}} \tag{3.2.12}$$

for some positive constant C_m. Now observe that

$$\left(\mathbb{I}_{[0,t']}(s)\mathbb{I}_{[0,x']}(y) - \mathbb{I}_{[0,t]}(s)\mathbb{I}_{[0,x]}(y)\right)^2 \leq \mathbb{I}_{[0,t\wedge t']}(s)\mathbb{I}_{[x\wedge x',x\vee x']}(y) + \mathbb{I}_{[t\wedge t',t\vee t']}(s)\mathbb{I}_{[0,x\wedge x']}(y).$$

One can draw a picture to convince himself about this last inequality. With this in mind, we have that

$$\left(H_{t',x'}(s,y) - H_{t,x}(s,y)\right)^2 \leq ||f||_\infty^2 \left(\mathbb{I}_{[0,t\wedge t']}(s)\mathbb{I}_{[x\wedge x',x\vee x']}(y) + \mathbb{I}_{[t\wedge t',t\vee t']}(s)\mathbb{I}_{[0,x\wedge x']}(y)\right),$$

which implies that (3.2.12) can be bounded by

$$\begin{aligned} C_m||f||_\infty^m &\left(\int_0^T \int_0^L \left(\mathbb{I}_{[0,t\wedge t']}(s)\mathbb{I}_{[x\wedge x',x\vee x']}(y) + \mathbb{I}_{[t\wedge t',t\vee t']}(s)\mathbb{I}_{[0,x\wedge x']}(y)\right) dyds\right)^{\frac{m}{2}} \\ &\leq C_m||f||_\infty^m \left((t\wedge t')|x-x'| + (x\wedge x')|t-t'|\right)^{\frac{m}{2}} \\ &\leq (T\vee L)^{\frac{m}{2}} C_m||f||_\infty^m \left(|x-x'| + |t-t'|\right)^{\frac{m}{2}}. \end{aligned}$$

Thus, for $m \in \mathbb{N}$ even and large enough, we have that the sequence is tight and, by Kolmogorov's Continuity Criterion, that there is a continuous version of the processes I_n for each $n \in \mathbb{N}$.

Convergence of the finite dimensional distributions
First observe that, by Hölder's inequality and Lemma 3.2.7,

$$\mathbb{E}\left[\left(\int_0^T\int_0^L g(t,x)\theta_n(t,x)dxdt\right)^2\right] \leq \left(\mathbb{E}\left[\left(\int_0^T\int_0^L g(t,x)\theta_n(t,x)dxdt\right)^m\right]\right)^{\frac{2}{m}}$$
$$\leq C_m\int_0^T\int_0^L g^2(t,x)dxdt = C_m||g||_2^2$$

for some positive constant C_m and for any $g \in L^2([0,T]\times[0,L])$. Now consider the normed space $(F, ||\cdot||) = (L^2([0,T]\times[0,L]), ||\cdot||_2)$ and the linear operators

$$J^n(g) = \int_0^T\int_0^L g(s,y)\theta_n(s,y)dyds, \quad J(g) = \int_0^T\int_0^L g(s,y)W(ds,dy).$$

Which are well defined for any $g \in F$. By the observation we have just made and Hölder's (or Cauchy-Schwarz's) inequality, one has that

$$\sup_{n\geq 1}\mathbb{E}\left[|J^n(g)|\right] \leq C||g||_2$$

for some positive constant C and for all $g \in F$. Moreover, Hölder's inequality again and the isometry formula, Proposition 3.2.1, imply that

$$\mathbb{E}\left[|J(g)|\right] \leq C||g||_2$$

as well. Now recall that simple functions of the form

$$g(t,x) = \sum_{i=0}^{k-1} g_i\mathbb{I}_{(t_i,t_{i+1}]}(t)\mathbb{I}_{(x_i,x_{i+1}]}(x) \tag{3.2.13}$$

with $k \geq 1$, $g_0,...,g_{k-1} \in \mathbb{R}$, $0 = t_0 < ... < t_k = T$ and $0 = x_0 < ... < x_k = L$ are dense in F with respect to the norm $||\cdot||_2$.

Now recall that the finite dimensional distributions of a sequence of random functions $\{X_n\}_n$ in $\mathcal{C}([0,T]\times[0,L])$ converge in law to the ones of another random function X in the same space if, and only if, for any $m \geq 1$, $a_1,...,a_m \in \mathbb{R}$ and $(s_1,y_1),...,(s_m,y_m) \in [0,T]\times[0,L]$, the random variables

$$\sum_{j=1}^m a_jX_n(s_j,y_j)$$

converge, in law, to the random variable

$$\sum_{j=1}^m a_jX(s_j,y_j).$$

In our case, $X_n = I_n$, so

$$\begin{aligned}\sum_{j=1}^{m} a_j X_n(s_j, y_j) &= \sum_{j=1}^{m} a_j \int_0^{s_j} \int_0^{y_j} f(t,x)\theta_n(t,x)dxdt \\ &= \int_0^T \int_0^L \left(\sum_{j=1}^{m} a_j \mathbb{I}_{[0,s_j]}(t)\mathbb{I}_{[0,y_j]}(x) f(t,x) \right) \theta_n(t,x)dxdt \\ &= J^n(K),\end{aligned}$$

where

$$K(t,x) = f(t,x) \sum_{j=1}^{m} a_j \mathbb{I}_{[0,s_j]}(t)\mathbb{I}_{[0,y_j]}(x),$$

which is an element of $L^2([0,T] \times [0,L])$ because f is so. Similarly, we have that, since $X = I$,

$$\sum_{j=1}^{m} a_j X(s_j, y_j) = J(K)$$

for the same function K.

So, by Lemma 3.2.6, it suffices to show that the finite dimensional distributions of $J^n(g)$ converge to the ones of $J(g)$ for g of the form (3.2.13). For such functions g, one has

$$\begin{aligned}J^n(g) &= \int_0^T \int_0^L \left(\sum_{i=0}^{k-1} g_i \mathbb{I}_{(t_i,t_{i+1}]}(t)\mathbb{I}_{(x_i,x_{i+1}]}(x) \right) \theta_n(t,x)dxdt \\ &= \sum_{i=0}^{k-1} g_i \int_{t_i}^{t_{i+1}} \int_{x_i}^{x_{i+1}} \theta_n(t,x)dxdt. \qquad (3.2.14)\end{aligned}$$

But notice that

$$\sum_{i=0}^{k-1} g_i \int_{t_i}^{t_{i+1}} \int_{x_i}^{x_{i+1}} \theta_n(t,x)dxdt = \sum_{i=0}^{k-1} g_i h(Y_1^n(i), Y_2^n(i), Y_3^n(i), Y_4^n(i)) \qquad (3.2.15)$$

where $h\colon \mathbb{R}^4 \to \mathbb{R}$ is the continuous function defined by $h(x,y,z,w) = x - y - z + w$

and

$$
\begin{aligned}
Y_1^n(i) &= \int_0^{t_{i+1}} \int_0^{x_{i+1}} \theta_n(t,x)dxdt, \\
Y_2^n(i) &= \int_0^{t_i} \int_0^{x_{i+1}} \theta_n(t,x)dxdt, \\
Y_3^n(i) &= \int_0^{t_{i+1}} \int_0^{x_i} \theta_n(t,x)dxdt, \\
Y_4^n(i) &= \int_0^{t_i} \int_0^{x_i} \theta_n(t,x)dxdt.
\end{aligned}
$$

By Theorem 3.2.1, the random vector

$$(Y_1^n(0), Y_2^n(0), Y_3^n(0), Y_4^n(0), ..., Y_1^n(k-1), Y_2^n(k-1), Y_3^n(k-1), Y_4^n(k-1))$$

converges in law to the random vector

$$(Y_1(0), Y_2(0), Y_3(0), Y_4(0), ..., Y_1(k-1), Y_2(k-1), Y_3(k-1), Y_4(k-1)),$$

with

$$
\begin{aligned}
Y_1(i) &= \int_0^{t_{i+1}} \int_0^{x_{i+1}} W(dt,dx), \\
Y_2(i) &= \int_0^{t_i} \int_0^{x_{i+1}} W(dt,dx), \\
Y_3(i) &= \int_0^{t_{i+1}} \int_0^{x_i} W(dt,dx), \\
Y_4(i) &= \int_0^{t_i} \int_0^{x_i} W(dt,dx).
\end{aligned}
$$

Since convergence in law is preserved under continuous transformations, we conclude that (3.2.14) converges in law to

$$
\begin{aligned}
\sum_{i=0}^{k-1} g_i h(Y_1(i), Y_2(i), Y_3(i), Y_4(i)) &= \sum_{i=0}^{k-1} g_i \int_{t_i}^{t_{i+1}} \int_{x_i}^{x_{i+1}} W(dt,dx) \\
&= \int_0^T \int_0^L g(t,x) W(dt,dx) = J(g).
\end{aligned}
$$

Finishing the proof. □

3.3 More classes of integrators and generalizations

As in the case of the stochastic integral with respect to the Brownian motion (or any other continuous square integrable martingale), the class integrands for which the integral with respect to the space-time white noise can be extended by considering processes $G = \{G(s,y) \colon s \in [0,T], y \in D\}$ satisfying the corresponding measurability conditions and the condition

$$\mathbb{P}\left\{\int_0^T \int_D G^2(s,y)dyds < \infty\right\} = 1.$$

However, the resulting process $(G \cdot W)$ is no longer a martingale, but a local one and the resulting integral is defined as a limit in probability, rather than an $L^2(\Omega)$-limit.

As mentioned in the previous chapter, this is done by a localisation argument and, again, we will omit the details.

Nevertheless, the purpose of this section will be to briefly see how we extend the class of integrators. As mentioned at the beginning of the chapter, for this part we will follow the construction provided by John B. Walsh in [11].

The main object of this section will be what we will call martingale measures, which might be though as a generalization of the white noise in $\mathbb{R}_+ \times D$ (being $D \subset \mathbb{R}^k$ a non-empty open connected set) seen in previous sections.

3.3.1 Martingale measures and worthy measures

As mentioned after Prop.3.1.1, one of the issues that forces us to develop the already seen theory of stochastic integration with respect to the white noise is that the latter cannot be thought (ω by ω) as a real-valued signed measure. However, it could be thought as a vector-valued measure. More precisely, given that the white noise was defined with respect to a σ-finite measure, the white noise was a particular case of what is called a σ-finite L^2-valued measure, which we define in a more general setting below.

From now on, we will work on $(E, \mathcal{E})$ where E is a Polish space and $\mathcal{E}$ the Borel σ-field on E and $\mathcal{A} \subset \mathcal{E}$ will be an algebra of sets (family of sets containing the empty set, closed by complementation and by finite unions).

Definition 3.3.1. *A set function $U \colon \mathcal{A} \times \Omega \to \mathbb{R}$ is (finitely) additive if $U(A \cup B, \omega) = U(A,\omega) + U(B,\omega)$ $\mathbb{P}$-a.s. $\omega \in \Omega$ whenever $A \cap B = \emptyset$ for $A, B \in \mathcal{A}$.*

In the previous definition and in the following ones, and as it is customary, we will omit the dependence on $\omega \in \Omega$. From now on, we will assume that $\mu(A) \coloneqq \mathbb{E}\left[U^2(A)\right] < \infty$ for any $A \in \mathcal{A}$, making U an L^2-valued set function.

Observe that in the case where U is a white noise, μ corresponds to the underlying σ-finite measure under which it is defined. This will be much clearer with the following two definitions, where we will define the σ-finiteness and σ-additivity of such functions in terms of μ, which we shall think of as a usual deterministic measure.

Definition 3.3.2. *A set function $U\colon \mathcal{A}\times\Omega \to \mathbb{R}$ is σ-finite if there exists an increasing sequence $\{E_n\}_{n\in\mathbb{N}} \subset \mathcal{E}$ such that*

(i) $\bigcup_n E_n = E$;

(ii) for all $n \in \mathbb{N}$, $\mathcal{E}_n \subset \mathcal{A}$, where $\mathcal{E}_n := \{E_n \cap S\colon S \in \mathcal{E}\}$ is a sub-σ-field of $\mathcal{E}$;

(iii) for all $n \in \mathbb{N}$, $\sup\{\mu(A)\colon A \in \mathcal{E}_n\} < \infty$.

Definition 3.3.3. *A σ-finite additive set function U is countably additive on $\mathcal{E}_n$ as an L^2-valued set function if for any decreasing sequence $\{A_j\}_j \subset \mathcal{E}_n$ with $\bigcap_j A_j = \emptyset$ we have $\lim_{j\to\infty} \mu(A_j) = 0$.*

If U is countably additive on each $\mathcal{E}_n$ and $A \in \mathcal{E}\backslash\mathcal{A}$, then we define $U(A)$ as the L^2-limit as n approaches infinity of $U(A \cap E_n)$ whenever this limit exists. If $A \in \mathcal{A}$, then the value of $U(A)$ obtained via this new definition coincides with the original value $U(A)$, providing a possible extension of U to sets $A \in \mathcal{E}\backslash\mathcal{A}$.

From now on, we will assume that all σ-finite countably additive (on each $\mathcal{E}_n$) set functions U have been extended in such way.

Definition 3.3.4. *A σ-finite L^2-valued measure is a σ-finite countably additive (on each $\mathcal{E}_n$) set function.*

Definition 3.3.5. *Let $\{\mathcal{F}_t\}_t$ be a right-continuous filtration. A process $M = \{M_t(A)\colon t \geq 0, A \in \mathcal{A}\}$ is a martingale measure if*

(i) $M_0(A) = 0$ $\mathbb{P}$-a.s. for each $A \in \mathcal{A}$;

(ii) for $t > 0$, M_t is a σ-finite L^2-valued measure;

(iii) $M(A) = \{M_t(A)\colon t \geq 0\}$ is an $\{\mathcal{F}_t\}_t$-martingale for each $A \in \mathcal{A}$.

Observe that, since M_t is a σ-finite L^2-valued measure for each $t > 0$, the martingales $\{M_t(A)\colon t \geq 0\}$, $A \in \mathcal{A}$, are square integrable when we restrict ourselves to one of the $(E_n, \mathcal{E}_n)$, that is, if we only consider $A \in \mathcal{E}_n \subset A$. Moreover, by doing this, the σ-finite L^2-valued measures M_t, $t > 0$, will be finite in the sense that the corresponding measure $\mu_t(A) := \mathbb{E}\left[M_t^2(A)\right]$ will be finite (because of (iii) in Definition 3.3.2).

In the case of the space-time white noise, the corresponding martingale measure would be constructed from the processes

$$M_t(A) = W([0,t) \times A), \quad t \geq 0, A \in \mathcal{A}. \tag{3.3.1}$$

Indeed, by recalling that the underlying measure was given by $\nu(ds, dx) = \mathbb{I}_{\mathbb{R}_+}(s)\mathbb{I}_D(x)dsdx$, we have that the process M defined above is Gaussian with mean and covariance functions

$$\mathbb{E}\left[M_t(A)\right] = 0, \quad \mathbb{E}\left[M_s(A)M_t(B)\right] = (s \wedge t)|A \cap B|, \tag{3.3.2}$$

where $|\cdot|$ denotes the Lebesgue measure in D. In particular, for each $t > 0$, M_t is a σ-finite L^2-valued measure and, for each $A \in \mathcal{A}$, the process $\{M_t(A) \colon t \geq 0\}$ is a Brownian motion with variance $t|A|$.

From now on, we shall assume that M is finite (M_t is finite for each $t > 0$) by restricting ourselves to one of the $(E_n, \mathcal{E}_n)$ if necessary. We will restrict our study to a finite time interval $[0,T]$, $T > 0$, as well.

Definition 3.3.6. *The covariance functional of a martingale measure M is*

$$Q_t(A, B) = \langle M(A), M(B)\rangle_t, \quad A, B \in \mathcal{A}$$

From the definition, it is easy to check that the covariance functional is symmetric with respect to A, B: $Q_t(A,B) = Q_t(B,A)$ $\mathbb{P}$-a.s. for each $t \geq 0$ (which is a consequence of the corresponding property for the covariation process). One can also check that it is biadditive: for fixed A and $t \geq 0$, $Q_t(A,\cdot)$ and $Q_t(\cdot, A)$ are (finitely) additive set functions, which can be proved by using the additivity of the martingale measure M and the bilinearity of the covariation process.

We also have the following Cauchy-Schwarz-type inequality

$$|Q_t(A,B)|^2 \leq Q_t(A,A)Q_t(B,B).$$

Which is a consequence of the corresponding property for the covariation process. In the above inequality it is implicitly stated that $Q_t(A,A) \geq 0$ for any $A \in \mathcal{A}$, which is a consequence of the fact that the quadratic variation process is non-negative.

From the covariance functional one can then define a set function defined on rectangles of the form $\Lambda = A \times B \times (s,t] \subset \mathcal{E} \times \mathcal{E} \times [0,T]$ as follows[1]:

$$Q(\Lambda) = Q_t(A,B) - Q_s(A,B) \tag{3.3.3}$$

[1]Recall that all σ-finite L^2-valued measures on $\mathcal{A}$ can be extended to $\mathcal{E}$.

and extend this last definition by additivity to finite unions of pairwise disjoint rectangles $\Lambda_j = A_j \times B_j \times (s_j, t_j]$, $j = 1, ..., n$:

$$Q\left(\bigcup_{j=1}^{n} \Lambda_j\right) := \sum_{j=1}^{n} Q(\Lambda_j).$$

As it is done in the case where Q is a product measure, one can check that this extension is well defined. That is, if

$$\Lambda = \bigcup_{j=1}^{n} R_j, \quad R_j = A_j \times B_j \times (s_j, t_j];$$
$$\Lambda' = \bigcup_{i=1}^{m} R'_i, \quad R'_i = A'_i \times B'_i \times (s'_i, t'_i].$$

With $R_i \cap R_j = \emptyset$ and $R'_i \cap R'_j = \emptyset$ for $i \neq j$ and $\Lambda = \Lambda'$, then $Q(\Lambda) = Q(\Lambda')$ $\mathbb{P}$-a.s. (this can be proved by finding a common refinement for Λ and Λ' and using the biadditivity of the covariance functional).

This last set function Q is positive definite in the following sense: if $a_1, ..., a_n \in R$, $A_1, ..., A_n \in \mathcal{E}$ are pairwise disjoint and $s \leq t$, then

$$\begin{aligned}\sum_{i,j} a_i a_j Q(A_i \times A_j \times (s,t]) &= \sum_{i,j} a_i a_j \left[\langle M(A_i), M(A_j)\rangle_t - \langle M(A_i), M(A_j)\rangle_s\right] \\ &= \left\langle \sum_i a_i M(A_i), \sum_j a_j M(A_j)\right\rangle_t - \left\langle \sum_i a_i M(A_i), \sum_j a_j M(A_j)\right\rangle_s \geq 0 \qquad (3.3.4)\end{aligned}$$

where we have used the bilinearity of the covariation process and the fact that the quadratic variation is a non-decreasing process.

In the case of the Gaussian white noise, we obtain that

$$Q_t(A, B) = t|A \cap B|.$$

To see this, we first observe that if $|A| = 0$ (or $|B| = 0$), then $M_t(A)$ (or $M(B)$), as defined in (3.3.1), vanishes for all $t \geq 0$ with probability 1, so the covariance functional (Definition 3.3.6) vanishes as well. On the other hand, since $A \cap B \subset A$ (or $A \cap B \subset B$), we have that $|A \cap B| \leq |A| = 0$ (or $|A \cap B| \leq |B| = 0$), obtaining the desired result when A or B are null sets with respect to the Lebesgue measure.

On the other hand, if $|A \cap B| = \sqrt{|A| \cdot |B|}$, we will have that

$$\int_D \mathbb{I}_A(x)\mathbb{I}_B(x)dx = \sqrt{\int_D \mathbb{I}_A(x)dx \int_D \mathbb{I}_B(x)dx}.$$

That is, the equality in the Cauchy-Schwarz inequality is attained, meaning that $\mathbb{I}_A(x) = \lambda \mathbb{I}_B(x)$ for some constant $\lambda \in \mathbb{R}$. But $\mathbb{I}_A(x)$ only takes the values 1 and 0, meaning that λ can only be one of those values.
If $\lambda = 1$, then we have that $\mathbb{I}_A(x) = \mathbb{I}_B(x)$ for all $x \in D$, so $A = B$ and thus, $M(A)$ and $M(B)$ define the same process, meaning that $Q_t(A, B)$ will simply be the quadratic variation of $M(A)$, which is $t|A| = t|A \cap B|$.
If $\lambda = 0$, then this implies that $\mathbb{I}_A(x) = 0$ for all $x \in D$, implying that A is a null set, which is what we have studied in the previous scenario.

Now let us consider the case where

$$\rho := \frac{|A \cap B|}{\sqrt{|A| \cdot |B|}} \in (0, 1)$$

and define the processes

$$W_t^{(1)} = \frac{M_t(A)}{\sqrt{|A|}}, \quad W_t^{(2)} = \frac{M_t(B)}{\sqrt{|B|}},$$

which are standard Brownian motions with correlation

$$\mathbb{E}\left[W_s^{(1)} W_t^{(2)}\right] = \rho(s \wedge t).$$

This can be seen by using Eq.(3.3.2).

Observe that these are well defined because $\rho > 0$ implies that, by the Cauchy-Schwarz inequality, $0 < |A \cap B| \leq \sqrt{|A| \cdot |B|}$. The very same inequality and the fact that $\rho < 1$ implies that $W^{(1)}$ and $W^{(2)}$ are two different process. Finally, define

$$Z_t := \frac{W_t^{(2)} - \rho W_t^{(1)}}{\sqrt{1 - \rho^2}}.$$

We first see that $Z = \{Z_t \colon t \in [0, T]\}$ is a Gaussian process, more particularly, it defines a standard Brownian motion independent of $W^{(1)}$. To do so, consider any vector $(Z_{t_1}, ..., Z_{t_n})$ with $t_1, ..., t_n \in [0, T]$ and any real scalars $a_1, ..., a_n \in \mathbb{R}$. Then observe that

$$a_1 Z_{t_1} + ... + a_n Z_{t_n} = b_1 W(A_1) + ... + b_m W(A_m)$$

for some real scalars $b_1, ..., b_m$, some sets $A_1, ..., A_m \in \mathcal{A}$ and where W denotes the Gaussian white noise. Given that W is Gaussian by definition, the right-hand side of this last equation is a Gaussian random variable, implying that the left-hand side is so and thus, proving that Z is a Gaussian process. Its mean and covariance functions are given by

$$\mathbb{E}\left[Z_t\right] = 0$$

and

$$
\begin{aligned}
\mathbb{E}\left[Z_s Z_t\right] &= \frac{1}{1-\rho^2}\mathbb{E}\left[W_s^{(2)}W_t^{(2)} - \rho W_s^{(2)}W_t^{(1)} - \rho W_s^{(1)}W_t^{(2)} + \rho^2 W_s^{(1)}W_t^{(1)}\right] \\
&= \frac{s\wedge t - \rho^2(s\wedge t) - \rho^2(s\wedge t) + \rho^2(s\wedge t)}{1-\rho^2} \\
&= s\wedge t.
\end{aligned}
$$

To see that it is independent of $W^{(1)}$, it suffices to show that they are uncorrelated.

$$
\begin{aligned}
\mathbb{E}\left[W_s^{(1)}Z_t\right] &= \frac{1}{\sqrt{1-\rho^2}}\mathbb{E}\left[W_t^{(2)}W_s^{(1)}\right] - \frac{\rho}{\sqrt{1-\rho^2}}\mathbb{E}\left[W_s^{(1)}W_t^{(1)}\right] \\
&= \frac{\rho}{\sqrt{1-\rho^2}} s\wedge t - \frac{\rho}{\sqrt{1-\rho^2}} s\wedge t \\
&= 0.
\end{aligned}
$$

Finally, using all this, we have that

$$
\langle W^{(1)}, W^{(2)}\rangle_t = \rho\rangle W^{(1)}\rangle_t + \rho\sqrt{1-\rho^2}\langle W^{(1)}, Z\rangle_t = \rho t.
$$

Meaning that

$$
\langle M(A), M(B)\rangle_t = \sqrt{|A|\cdot|B|}\rho t = t|A\cap B|,
$$

as was to be shown. Observe that, in the case of the Gaussian white noise, the covariance functional is deterministic (which can be though of an analogous property of the Brownian motion, whose quadratic variations is deterministic). Moreover, the covariance functional coincides with the underlying measure considered for the white noise which, in particular, is a measure.

What we have shown for the case of the Gaussian white noise is not true in general, that is, the set function Q cannot always be extended to a signed measure on $\mathcal{E}\times\mathcal{E}\times\mathcal{B}$ (being $\mathcal{B}$ the Borel σ-field on $[0,T]$). For instance, there is no guarantee that the σ-additivity holds. To solve this problem, we will dominate Q by some measure, which will allow us to prove, for example, the σ-additivity.

Definition 3.3.7. *A signed measure K on $\mathcal{E}\times\mathcal{E}\times\mathcal{B}$ is positive definite if, for each bounded measurable function $f\colon E\times[0,T]\to\mathbb{R}$,*

$$
\int_{E\times E\times[0,T]} f(x,s)f(y,s)K(dx\,dy\,ds) \geq 0 \tag{3.3.5}
$$

whenever the above integral makes sense.

When K is symmetric in x and y ($K(dx\,dy\,ds) = K(dy\,dx\,ds)$), one has the following Cauchy-Schwartz' and Minkowski's inequalities (the proofs are essentially the same as the standard ones, so we will omit them)

$$|(f,g)_K|^2 \leq (f,f)_K(g,g)_K,$$
$$(f+g,f+g)_K{}^{1/2} \leq (f,f)_K{}^{1/2} + (g,g)_K{}^{1/2}.$$

Where

$$(f,g)_K := \int_{E\times E\times[0,T]} f(x,s)g(y,s)K(dx\,dy\,ds).$$

Definition 3.3.8. *A martingale measure M is worthy if there is a random σ-finite measure $K(\Lambda,\omega)$, $\Lambda \in \mathcal{E}\times\mathcal{E}\times\mathcal{B}$, $\omega\in\Omega$, such that*

(i) K is positive definite and symmetric with respect to x and y ($K(dx\,dy\,ds) = K(dy\,dx\,ds)$);

(ii) for fixed $A,B\in\mathcal{E}$, $\{K(A\times B\times[0,t])\colon t\geq 0\}$ is predictable (it is measurable with respect to the σ-field generated by all $\{\mathcal{F}_t\}_t$-adapted processes with $\mathbb{P}$-a.s. left-continuous sample paths);

(iii) for all n, $\mathbb{E}\left[K(E_n\times E_n\times[0,T])\right] < \infty$;

(iv) for any rectangle Λ, $|Q(\Lambda)| \leq K(\Lambda)$.

We call K the dominating measure of M.

Proposition 3.3.1. *If M is a worthy martingale measure as in Definition 3.3.8, then the corresponding set function Q can be extended to a (random) positive definite signed measure on $\mathcal{E}\times\mathcal{E}\times\mathcal{B}$ (as in Definition 3.3.7) whose total variation, $|Q|$, will be bounded by the dominating measure K.*

Proof. Since $\mathcal{E}$ is separable (can be generated by a countable collection of sets, say $\{A_n\}_n$), we shall first restrict ourselves to a countable sub-algebra $\mathcal{G}$ of $\mathcal{E}\times\mathcal{E}\times\mathcal{B}$ formed by sets of the form $A\times B\times(s,t]$ with $A,B\in\{A_n\}_n$ and $s,t\in\mathbb{Q}\cap[0,T]$ satisfying $\sigma(\mathcal{G}) = \mathcal{E}\times\mathcal{E}\times\mathcal{B}$ and upon which $Q(\cdot,\omega)$ is finitely additive (this is ensured by the fact that the sub-algebra is constituted of rectangles).

In this sub-algebra, we have that $\lambda := Q + K$ is a non-negative finitely additive set function dominated by the measure $2K$, so it is a countably additive in the sub-algebra. That is, if $\{R_n\}_n$ is a sequence of pairwise disjoint sets in $\mathcal{G}$ whose union is contained in the same sub-algebra, then $\lambda\left(\bigcup_j R_j\right) = \sum_j \lambda(A_j)$. Indeed, one can easily check that, by finite additivity, if $A \subset B$ with $A, B \in \mathcal{G}$, then $\lambda(B \backslash A) = \lambda(B) - \lambda(A)$. So, by taking $B = \bigcup_j R_j$ and $A = \bigcup_{j=1}^N R_j$ for fixed N, we have

$$
\begin{aligned}
0 &\leq \lambda\left(\bigcup_{j\geq 1} R_j\right) - \sum_{j=1}^N \lambda(R_j) \\
&= \lambda\left(\bigcup_{j\geq 1} R_j\right) - \lambda\left(\bigcup_{j=1}^N R_j\right) \\
&= \lambda\left(\bigcup_{j\geq 1} R_j \Big\backslash \bigcup_{j=1}^N R_j\right) \\
&= \lambda\left(\bigcup_{j=N+1}^{\infty} R_j\right) \\
&\leq 2K\left(\bigcup_{j=N+1}^{\infty} R_j\right) \xrightarrow{N\to\infty} 0.
\end{aligned}
$$

However, the previous computation holds $\mathbb{P}$-a.s., where the implicit null set might depend on the family of rectangles $\{R_j\}_j$, for the following argument, we must see that there is a $\mathbb{P}$-null set, M, independent of the choice of $\{R_j\}_j$ for which the previous computations hold for $\omega \in \Omega \backslash M$.
For each R_i, R_j, $i \neq j$, the additivity property

$$\lambda(R_i \cup R_j) = \lambda(R_i) + \lambda(R_j)$$

holds for every $\omega \in \Omega \backslash M_{i,j}$ with $\mathbb{P}(M_{i,j}) = 0$. However, we have that $R_i = A_{n_i}$, $R_j = A_{n_j}$ for some $n_i, n_j \in \mathbb{N}$. In other words, the null sets involved in the previous computations depend only on the choice of the sub-algebra $\mathcal{G}$, but not on the choice of the particular family of disjoint sets $\{R_j\}_j$. To make this last statement a bit more clear, let us rephrase it in another way. Given the sub-algebra $\mathcal{G} = \{A_n\}_n$, we have that, given any two disjoint sets A_{n_i} and A_{n_j} in $\mathcal{G}$, there is a $\mathbb{P}$-null set $M_{i,j}$ so that

$$\lambda(A_{n_i} \cup A_{n_j}) = \lambda(A_{n_i}) + \lambda(A_{n_j})$$

holds for all $\omega \in \Omega \backslash M_{i,j}$. Now this will imply that, given any sequence of disjoint rectangles $\{R_j\}_j \subset \mathcal{G}$ whose union is also in $\mathcal{G}$,

$$\sum_{j=1}^N \lambda(R_j) = \sum_{j=1}^N \lambda(A_{n_j}) = \lambda\left(\bigcup_{j=1}^N A_{n_j}\right)$$

holds for some pairwise disjoint sets $\{A_{n_j} : j = 1, ..., N\}$ and for all $\omega \in \Omega \backslash M$, where

$$M := \bigcup_{i,j} M_{i,j}.$$

Which is a countable (because $\mathcal{G}$ is countable) union of $\mathbb{P}$-null sets and thus, a $\mathbb{P}$-null set. Indeed, say, for instance and to convince ourselves, that $N = 3$. Then we have that

$$\lambda(A_{n_1} \cup A_{n_2}) = \lambda(A_{n_1}) + \lambda(A_{n_2})$$

outside $M_{1,2}$. Now, since $\mathcal{G}$ is an algebra, we will have that $A_{n_1} \cup A_{n_2} = A_{n_m}$ for some $n_m \in \mathbb{N}$ and $A_{n_m} \cap A_{n_3} = \emptyset$ (because $A_{n_1} \cap A_{n_3} = \emptyset$ and $A_{n_2} \cap A_{n_3} = \emptyset$), so we will have

$$\lambda(A_{n_1} \cup A_{n_2} \cup A_{n_3}) = \lambda(A_{n_m} \cup A_{n_3}) = \lambda(A_{n_m}) + \lambda(A_{n_3})$$

outside $M_{m,3}$. Finally, since we have $\lambda(A_{n_m}) = \lambda(A_{n_1}) + \lambda(A_{n_2})$ outside $M_{1,2}$, we will have

$$\lambda(A_{n_1} \cup A_{n_2} \cup A_{n_3}) = \lambda(A_{n_1}) + \lambda(A_{n_2}) + \lambda(A_{n_3})$$

outside $M_{1,2} \cup M_{m,3}$ for some $m \in \mathbb{N}$, leading to the definition of our set M. Outside this set M, the previous computations regarding the σ-additivity of λ hold for any ω independently of $\{R_j\}_j$.

So we obtain that λ is $\mathbb{P}$-a.s. countably additive in $\mathcal{G}$ and thus, it is a pre-measure on $\mathcal{G}$. Then, $\mathbb{P}$-a.s., the desired result is a direct consequence of Carathéodory's extension theorem, which will give us an extension for λ and thus, an extension for Q as a signed measure on $\sigma(\mathcal{G}) = \mathcal{E} \times \mathcal{E} \times \mathcal{B}$.

Moreover, its total variation, $|Q|$, will be bounded by K ($|Q|(\Lambda) \leq K(\Lambda)$ for any $\Lambda \in \mathcal{E} \times \mathcal{E} \times \mathcal{B}$) by (iv) in Definition 3.3.8 and will be positive definite (in the sense of Definition 3.3.7). Indeed, (3.3.4) allows us to prove the property for indicator functions and, by a usual approximating procedure, the result extends to any bounded measurable function f for which the integral in (3.3.5) (with Q instead of K) is well defined. □

In the following, we shall refer to this extension as Q, without distinguishing it from the original one. Observe that, for such extension, one has, since Q is positive definite,

$$\begin{aligned}(f, f)_Q &= |(f, f)_Q| \\ &\leq \int_{E \times E \times [0,T]} |f(x, s)| \cdot |f(y, s)||Q|(dx\, dy\, ds) \\ &\leq \int_{E \times E \times [0,T]} |f(x, s)| \cdot |f(y, s)|K(dx\, dy\, ds) \\ &= (|f|, |f|)_K. \end{aligned} \tag{3.3.6}$$

As previously mentioned, for the case of the Gaussian white noise there is no need to consider any dominating measure K, since Q itself already defines a measure.

With this, we are ready to start talking about integrals with respect to worthy measures.

3.3.2 Integration with respect to worthy measures

The program for this part will be quite similar to the one already seen for the Itô integral with respect to continuous square integrable martingales. We will first define the class of functions for which the integral will be defined and then define the stochastic integral for a rather simple class of functions that will be dense in the previous class with respect to some topology. Finally, the stochastic integral for any integrable function will be defined as the limit of stochastic integrals of the simpler class.

Let us fix a worthy martingale measure with covariance functional Q and dominating measure K.

Definition 3.3.9. *A function $f\colon E\times[0,T]\times\Omega\to\mathbb{R}$ is elementary if it is of the form*

$$f(x,s,\omega)=X(\omega)\mathbb{I}_{(a,b]}(s)\mathbb{I}_A(x) \tag{3.3.7}$$

for some $0\leq a<b\leq T$, $A\in\mathcal{E}$ and X a bounded $\mathcal{F}_a$-measurable random variable. A function f is simple if it is a finite linear combination of elementary functions. The class of simple functions will be denoted by $\mathcal{S}$.

Definition 3.3.10. *The predictable σ-field on $\Omega\times E\times[0,T]$ is the σ-field generated by $\mathcal{S}$, $\sigma(\mathcal{S})$. A function will be predictable if it is $\sigma(\mathcal{S})$-measurable.*

The following theorem, which we state without proof, summarizes the approximating procedure needed to construct the stochastic integral for the predictable functions f such that $||f||_M<\infty$ with

$$||f||_M^2\coloneqq\mathbb{E}\left[(|f|,|f|)_K\right]=\mathbb{E}\left[\int_{E\times E\times[0,T]}|f(x,s)||f(y,s)|K(dx\,dy\,ds)\right].$$

Such class of functions will be denoted by $\mathcal{P}_M$.

Theorem 3.3.1. *The following statements hold true:*

(i) $||\cdot||_M$ is a norm (where we identify indistinguishable functions) and $(\mathcal{P}_M,||\cdot||_M)$ is a Banach space.

(ii) The set of simple functions $\mathcal{S}$ is dense in $\mathcal{P}_M$ with respect to the norm $||\cdot||_M$.

Definition 3.3.11. *The stochastic integral of an elementary function f given by Eq.(3.3.7) with respect to the martingale measure M, which we shall denote by $f \cdot M$, is the process given by*

$$(f \cdot M)_t(B) := X\left[M_{t\wedge b}(A\cap B) - M_{t\wedge a}(A\cap B)\right], \quad B \in \mathcal{E}, t\in[0,T]. \tag{3.3.8}$$

For $f \in \mathcal{S}$, $f \cdot M$ is defined by linearity.

Observe that, for fixed $B \in \mathcal{E}$, the integral process $t \mapsto (f \cdot M)_t(B)$ is simply the integral of a predictable process with respect to a square integrable martingale.

As in the case of the stochastic integral with respect to the white noise and with respect to continuous square integrable martingales, we have an isometry formula for the integral of simple processes that will be crucial when proving that that the stochastic integral of a simple function is well defined (Lemma 3.3.2) and the existence of a process that will be called the stochastic integral for a function $f \in \mathcal{P}_M$.

Lemma 3.3.1. *Let $f, g \in \mathcal{S}$ be any two simple functions, then*

$$\mathbb{E}\left[(f\cdot M)_t(B)(g\cdot M)_t(B)\right] = \mathbb{E}\left[\int_{B\times B\times[0,t]} f(x,s)g(y,s)Q(dx\,dy\,ds)\right]$$

for any $t \in [0,T]$ and $B \in \mathcal{E}$. In particular, for $f = g$,

$$\mathbb{E}\left[(f\cdot M)_t^2(B)\right] = \mathbb{E}\left[\int_{B\times B\times[0,t]} f(x,s)f(y,s)Q(dx\,dy\,ds)\right].$$

Proof. We will first prove the result for $f = g$. The general result will follow from polarization and the fact that Q is symmetric. Let

$$f(x,s,\omega) = \sum_{j=1}^n f_j(x,s,\omega), \quad f_j(x,s,\omega) = X_j(\omega)\mathbb{I}_{I_j}(s)\mathbb{I}_{A_j}(x). \tag{3.3.9}$$

Where X_j is bounded and $\mathcal{F}_{a_j}$-measurable, $I_j = (a_j, b_j]$ with $0 \le a_j < b_j \le T$ and $A_j \in \mathcal{E}$ for each $j \in \{1, ..., n\}$. Then we have that

$$\begin{aligned}(f\cdot M)_t^2(B) &= (f_i\cdot M)_t(B)(f_j\cdot M)_t(B)\\ &= \sum_{1\le i,j\le n} X_iX_j\Delta_iM(B)\,\Delta_jM(B)\end{aligned} \tag{3.3.10}$$

for any $t \in [0,T]$, $B \in \mathcal{E}$ and where $\Delta_jM(B) = M_{t\wedge b_j}(A_j\cap B) - M_{t\wedge a_j}(A_j\cap B)$. Expanding the product $\Delta_iM(B)\,\Delta_jM(B)$ we obtain

$$\Delta_iM(B)\,\Delta_jM(B) = M_{t\wedge b_j}M_{t\wedge b_i} - M_{t\wedge b_j}M_{t\wedge a_i} - M_{t\wedge a_j}M_{t\wedge b_i} + M_{t\wedge a_j}M_{t\wedge a_i}$$

where $M_{t\wedge b_j} = M_{t\wedge b_j}(A_j \cap B)$, $M_{t\wedge b_i} = M_{t\wedge b_i}(A_i \cap B)$, etc. On the other hand

$$\begin{aligned}
&\int_{B\times B\times[0,t]} f(x,s)f(y,s)Q(dx\,dy\,ds)\\
&= \sum_{1\le i,j\le n} X_iX_j \int_{B\times B\times[0,t]} \mathbb{I}_{(a_i,b_i]\cap(a_j,b_j]}(s)\mathbb{I}_{A_i}(x)\mathbb{I}_{A_j}(y)Q(dx\,dy\,ds)\\
&= \sum_{1\le i,j\le n} X_iX_jQ\left(A_i\cap B\times A_j\cap B\times[0,t]\cap(a_i,b_i]\cap(a_j,b_j]\right). \qquad (3.3.11)
\end{aligned}$$

We will see that the expectation of the summand $(i,j)\in\{1,...n\}^2$ in (3.3.10) coincides with the expectation of the same summand in (3.3.11).

Before doing so, we see that, for instance, from (2.3.3), one can show that, if $0\le a<b\le T$ are deterministic and using the same notation as in (2.3.3),

$$\langle M,N\rangle_t^{(a)} = \langle M^{(a)},N\rangle_t = \langle M,N^{(a)}\rangle_t = \langle M^{(a)},N^{(a)}\rangle_t,$$

where $Z^{(a)} := \{Z_t^{(a)} := Z_{t\wedge a} \colon t\in[0,T]\}$ for any process Z (and analogously with b), and that

$$\langle M,N\rangle_t^{(a)} = \langle M^{(b)},N^{(a)}\rangle_t = \langle M^{(a)},N^{(b)}\rangle_t.$$

With this in mind, observe that, for fixed $(i,j)\in\{1,...,n\}^2$, if $a_i<b_i\le a_j<b_j$,

$$\begin{aligned}
\langle M_{\cdot\wedge b_j}, M_{\cdot\wedge b_i}\rangle_t &= \langle M_{\cdot\wedge a_j}, M_{\cdot\wedge b_i}\rangle_t\\
\langle M_{\cdot\wedge b_j}, M_{\cdot\wedge a_i}\rangle_t &= \langle M_{\cdot\wedge a_j}, M_{\cdot\wedge a_i}\rangle_t
\end{aligned}$$

and thus,

$$(f_i\cdot M)_t(B)(f_j\cdot M)_t(B) = X_iX_j\left[M_1 - M_3 + M_2 - M_4\right],$$

where

$$\begin{aligned}
M_1 &= M_{t\wedge b_j}M_{t\wedge b_i} - \langle M_{\cdot\wedge b_j}, M_{\cdot\wedge b_i}\rangle_t,\\
M_2 &= M_{t\wedge a_j}M_{t\wedge a_i} - \langle M_{\cdot\wedge a_j}, M_{\cdot\wedge a_i}\rangle_t,\\
M_3 &= M_{t\wedge a_j}M_{t\wedge b_i} - \langle M_{\cdot\wedge a_j}, M_{\cdot\wedge b_i}\rangle_t,\\
M_4 &= M_{t\wedge b_j}M_{t\wedge a_i} - \langle M_{\cdot\wedge b_j}, M_{\cdot\wedge a_i}\rangle_t.
\end{aligned}$$

We also observe that the corresponding summand in (3.3.11) vanishes for any value of $t\in[0,T]$ since $(a_i,b_i]\cap(a_j,b_j]=\emptyset$.
Now, if $t\le a_i$, then $M_1=M_2=M_3=M_4$, so the summands (i,j) in (3.3.10) and (3.3.11) vanish and thus, their expectations coincide. For $a_i<t\le a_j$, we have that

$M_1 = M_3$ and $M_2 = M_4$, so the corresponding summands coincide again.
If $t > a_j$, the martingale property for M_1 and the law of total expectation tell us that

$$\begin{aligned}\mathbb{E}\left[X_i X_j M_1\right] &= \mathbb{E}\left[X_i X_j \mathbb{E}\left[M_1 \middle| \mathcal{F}_{a_j}\right]\right] \\ &= \mathbb{E}\left[X_i X_j \left(M_{a_j \wedge b_j} M_{a_j \wedge b_i} - \langle M_{\cdot \wedge b_j}, M_{\cdot \wedge b_i} \rangle_{a_j}\right)\right].\end{aligned}$$

Where we have used that X_i is $\mathcal{F}_{a_i}$-measurable and that $\mathcal{F}_{a_i} \subset \mathcal{F}_{a_j}$ since $a_i < a_j$, so $X_i X_j$ is $\mathcal{F}_{a_j}$-measurable. Similar computations for M_2, M_3 and M_4 lead to the fact that

$$\mathbb{E}\left[X_i X_j M_1\right] = \mathbb{E}\left[X_i X_j M_3\right], \quad \mathbb{E}\left[X_i X_j M_2\right] = \mathbb{E}\left[X_i X_j M_4\right].$$

So the expected value of the corresponding summand in (3.3.10) vanishes, coinciding with the expectation of the corresponding summand in (3.3.11).

If $a_i \leq a_j < b_i \leq b_j$, we will have that

$$\begin{aligned}&(f_i \cdot M)_t(B)(f_j \cdot M)_t(B) \\ &= X_i X_j \left[M_1 - M_3 + M_2 - M_4\right] + X_i X_j \left(\langle M_{\cdot \wedge b_j}, M_{\cdot \wedge b_i} \rangle_t - \langle M_{\cdot \wedge a_j}, M_{\cdot \wedge b_i} \rangle_t\right) \\ &= X_i X_j \left[M_1 - M_3 + M_2 - M_4\right] \\ &+ X_i X_j \left(\langle M(A_j \cap B), M(A_i \cap B) \rangle_{t \wedge b_i} - \langle M(A_j \cap B), M(A_i \cap B) \rangle_{t \wedge a_j}\right) \\ &= X_i X_j \left[M_1 - M_3 + M_2 - M_4\right] + X_i X_j Q\left(A_i \cap B \times A_j \cap B \times [0,t] \cap (a_j, b_i]\right).\end{aligned}$$

For $t \leq a_j$, one can follow the same procedure as before to see that both summands vanish, so the corresponding summands in (3.3.10) and (3.3.11) vanish. For $t > a_j$, the expected value of $X_i X_j \left[M_1 - M_3 + M_2 - M_4\right]$ will vanish (by the martingale property), while the expected value of the second term in this last equation will be the same as the expected value of the summand (i, j) in (3.3.11). So they coincide again.

Finally, if $a_i \leq a_j < b_j \leq b_i$, we will have that

$$\begin{aligned}&(f_i \cdot M)_t(B)(f_j \cdot M)_t(B) \\ &= X_i X_j \left[M_1 - M_3 + M_2 - M_4\right] + X_i X_j \left(\langle M_{\cdot \wedge b_j}, M_{\cdot \wedge b_i} \rangle_t - \langle M_{\cdot \wedge a_j}, M_{\cdot \wedge b_i} \rangle_t\right) \\ &= X_i X_j \left[M_1 - M_3 + M_2 - M_4\right] \\ &+ X_i X_j \left(\langle M(A_j \cap B), M(A_i \cap B) \rangle_{t \wedge b_j} - \langle M(A_j \cap B), M(A_i \cap B) \rangle_{t \wedge a_j}\right) \\ &= X_i X_j \left[M_1 - M_3 + M_2 - M_4\right] + X_i X_j Q\left(A_i \cap B \times A_j \cap B \times [0,t] \cap (a_j, b_j]\right).\end{aligned}$$

And the same computations will lead to the desired result. □

Lemma 3.3.2. *The stochastic integral $f \cdot M$ does not depend on the particular choice of the representation of $f \in \mathcal{S}$. Moreover, for fixed $B \in \mathcal{E}$ and $t \in [0, T]$, the map $f \mapsto (f \cdot M)_t(B)$ is linear for $f \in \mathcal{S}$.*

Proof. We will only prove the first part, the linearity part follows immediately from the definition.

Let $f, g \in \mathcal{S}$ be any two simple functions such that $f = g$. By the previous lemma we have that, for any $t \in [0, T]$ and any $B \in \mathcal{E}$, setting $A = B \times B \times [0, t]$,

$$
\begin{aligned}
&\mathbb{E}\left[((f \cdot M)_t(B) - (g \cdot M)_t(B))^2\right] \\
&= \mathbb{E}\left[\int_A f(x,s)f(y,s)dQ + \int_A g(x,s)g(y,s)dQ - 2\int_A f(x,s)g(y,s)dQ\right] \\
&= \mathbb{E}\left[\int_A (f(x,s) - g(x,s))\,(f(y,s) - g(y,s))\,dQ\right] \\
&= 0.
\end{aligned}
$$

Where and $dQ = Q(dx\,dy\,ds)$ and where we have used that Q is symmetric with respect to its two first arguments. So $(f \cdot M)_t(B) = (g \cdot M)_t(B)$ $\mathbb{P}$-a.s. $\square$

Lemma 3.3.3. *For $f \in \mathcal{S}$, $f \cdot M$ is a worthy martingale measure. Its covariance functional and dominating measure, Q_f and K_f respectively, are given by*

$$Q_f(dx\,dy\,ds) = f(x,s)f(y,s)\,Q(dx\,dy\,ds), \tag{3.3.12}$$

$$K_f(dx\,dy\,ds) = |f(x,s)f(y,s)|K(dx\,dy\,ds). \tag{3.3.13}$$

Moreover, for any $B \in \mathcal{E}$ and $t \in [0, T]$,

$$\mathbb{E}\left[(f \cdot M)_t^2(B)\right] \leq ||f||_M^2. \tag{3.3.14}$$

Proof. Let $f \in \mathcal{S}$ be any simple function like in (3.3.9), then we have that $(f \cdot M)_0(B) = 0$ $\mathbb{P}$-a.s. for any $B \in \mathcal{E}$ since the X_j are bounded and the differences in (3.3.8) vanish for each $j \in \{1, ..., n\}$. On the other hand, the fact that the X_j are bounded and independent of $B \in \mathcal{E}$ tells us that $B \mapsto X_j\, M_{t \wedge b_j}(A_j \cap B)$ and $B \mapsto X_j\, M_{t \wedge a_j}(A_j \cap B)$ are σ-finite L^2-valued measures. In particular, their difference will also be a σ-finite L^2-valued measure and so it will be their sum.

Further, given that stopped martingales are martingales and finite linear combinations of martingales are martingales, we have that $M_{t \wedge b_j}(A_j \cap B) - M_{t \wedge a_j}(A_j \cap B)$ is a martingale and, since X_j is $\mathcal{F}_{a_j}$-measurable, the process $(f_j \cdot M)(B) := \{(f_j \cdot M)_t(B) \colon t \in [0, T]\}$ for fixed $B \in \mathcal{E}$ is a martingale as well. Indeed, if $t \leq a_j$, then $(f_j \cdot M)_t(B) = 0$, so it is adapted and the martingale property $\mathbb{E}\left[(f_j \cdot M)_t(B)|\mathcal{F}_s\right] = (f_j \cdot M)_s(B)$ is satisfied for $0 \leq s \leq t \leq a_j$. On the other hand, if $t > a_j$, then the random variable X_j is $\mathcal{F}_t$ measurable (because $\mathcal{F}_{a_j} \subset \mathcal{F}_t$), so $(f_j \cdot M)_t(B)$ is $\mathcal{F}_t$-measurable. Moreover, if $0 \leq s \leq a_j < t$, then, by the tower property,

$$
\begin{aligned}
\mathbb{E}\left[(f_j \cdot M)_t(B)\middle|\mathcal{F}_s\right] &= \mathbb{E}\left[X_j \mathbb{E}\left[M_{t \wedge b_j}(A_j \cap B) - M_{t \wedge a_j}(A_j \cap B)\middle|\mathcal{F}_{a_j}\right]\middle|\mathcal{F}_s\right] \\
&= 0 \\
&= (f_j \cdot M)_s(B).
\end{aligned}
$$

Similarly, if $a_j < s \leq t$,

$$\begin{aligned}
\mathbb{E}\left[(f_j \cdot M)_t(B)\middle|\mathcal{F}_s\right] &= X_j \mathbb{E}\left[M_{t\wedge b_j}(A_j \cap B) - M_{t\wedge a_j}(A_j \cap B)\middle|\mathcal{F}_s\right] \\
&= X_j \left(M_{s\wedge b_j}(A_j \cap B) - M_{s\wedge a_j}(A_j \cap B)\right) \\
&= (f_j \cdot M)_s(B).
\end{aligned}$$

So $(f_j \cdot M)(B)$ is a martingale and, all in all, $f \cdot M$ is a martingale measure.

To see that it is worthy, we only need to provide a dominating measure satisfying the properties in Definition 3.3.8. To do so, we first compute the covariation of $f \cdot M$. For this, we will check that, for any $B, C \in \mathcal{E}$, the process

$$(f \cdot M)_t(B)\,(f \cdot M)_t(C) - \int_{B\times C\times[0,t]} f(x,s) f(y,s) Q(dx\,dy\,ds) \tag{3.3.15}$$

is a martingale. Using the same notation as in Lemma 3.3.1 and defining $M^B_{t\wedge b_i} = M_{t\wedge b_i}(A_i \cap B)$, $M^C_{t\wedge b_j} = M_{t\wedge b_j}(A_j \cap C)$, etc., we have that

$$\begin{aligned}
(f \cdot M)_t(B)\,(f \cdot M)_t(C) &= \sum_{1\leq i,j\leq n} X_i X_j \Delta_i M(B) \Delta_j M(C) \\
&= \sum_{1\leq i,j\leq n} (f_i \cdot M)_t(B)(f_j \cdot M)_t(C) \\
&= \sum_{1\leq i,j\leq n} X_i X_j \left[M^B_{t\wedge b_i} M^C_{t\wedge b_j} - M^B_{t\wedge b_i} M^C_{t\wedge a_j} - M^B_{t\wedge a_i} M^C_{t\wedge b_j} + M^B_{t\wedge a_i} M^C_{t\wedge a_j}\right].
\end{aligned}$$

While

$$\begin{aligned}
&\int_{B\times C\times[0,t]} f(x,s) f(y,s) Q(dx\,dy\,ds) \\
&\qquad = \sum_{1\leq i,j\leq n} X_i X_j Q\left(A_i \cap B \times A_j \cap C \times [0,t] \cap I_i \cap I_j\right).
\end{aligned}$$

As in Lemma 3.3.1, we will see that, for each $(i,j) \in \{1,...,n\}^2$,

$$(f_i \cdot M)_t(B)(f_j \cdot M)_t(C) - X_i X_j Q\left(A_i \cap B \times A_j \cap C \times [0,t] \cap I_i \cap I_j\right) \tag{3.3.16}$$

is a martingale.

Let us start by assuming that $a_j < b_j \leq a_i < b_i$. Then

$$Q\left(A_i \cap B \times A_j \cap C \times [0,t] \cap I_i \cap I_j\right) = 0$$

while

$$(f_i \cdot M)_t(B)(f_j \cdot M)_t(C) = X_i X_j \left[M_1 - M_3 + M_2 - M_4\right], \tag{3.3.17}$$

where

$$
\begin{aligned}
M_1 &= M^B_{t\wedge b_i} M^C_{t\wedge b_j} - \langle M^B_{\cdot\wedge b_i}, M^C_{\cdot\wedge b_j}\rangle_t, \\
M_2 &= M^B_{t\wedge a_i} M^C_{t\wedge a_j} - \langle M^B_{\cdot\wedge a_i}, M^C_{\cdot\wedge a_j}\rangle_t, \\
M_3 &= M^B_{t\wedge a_i} M^C_{t\wedge b_j} - \langle M^B_{\cdot\wedge a_i}, M^C_{\cdot\wedge b_j}\rangle_t, \\
M_4 &= M^B_{t\wedge b_i} M^C_{t\wedge a_j} - \langle M^B_{\cdot\wedge b_i}, M^C_{\cdot\wedge a_j}\rangle_t.
\end{aligned}
$$

If $t \le a_i$, one can easily show that (3.3.17) vanishes, so it is $\mathcal{F}_t$-measurable and the martingale property is satisfied for $0 \le s \le t$. On the other hand, for $t > a_i$, the product $X_i X_j$ is $\mathcal{F}_t$-measurable, so (3.3.17) is adapted. Moreover, and as previously argued (distinguishing the cases cases $0 \le s \le a_i < t$ and $a_i < s \le t$), one shows that

$$
\mathbb{E}\left[(f_i \cdot M)_t(B)(f_j \cdot M)_t(C)\middle|\mathcal{F}_s\right] = (f_i \cdot M)_t(B)(f_j \cdot M)_s(C).
$$

If $a_j \le a_i < b_j \le b_i$, then

$$
Q\left(A_i \cap B \times A_j \cap C \times [0,t] \cap I_i \cap I_j\right) = Q\left(A_i \cap B \times A_j \cap C \times [0,t] \cap (a_i, b_j]\right)
$$

and

$$
\begin{aligned}
&(f_i \cdot M)_t(B)(f_j \cdot M)_t(C) = X_i X_j \left[M_1 - M_3 + M_2 - M_4\right] \\
&+ X_i X_j \left(\langle M(A_i \cap B), M(A_j \cap C)\rangle_{t\wedge b_j} - \langle M(A_i \cap B), M(A_j \cap C)\rangle_{t\wedge a_i}\right) \\
&= X_i X_j \left[M_1 - M_3 + M_2 - M_4\right] + X_i X_j Q\left(A_i \cap B \times A_j \cap C \times [0,t] \cap (a_i, b_j]\right).
\end{aligned}
$$

So (3.3.16) becomes

$$
X_i X_j \left[M_1 - M_3 + M_2 - M_4\right].
$$

Following similar procedures to the ones already seen, one then shows that this is a martingale. The case $a_j \le a_i < b_i \le b_j$ is done similarly. So we have established (3.3.12).

Since Q is dominated by K, then it is immediate to see that Q_f is dominated by K_f, which is clearly a symmetric and positive definite measure such that $|Q_f(\Lambda)| \le K_f(\Lambda)$ for any rectangle Λ. Moreover, for each $n \in \mathbb{N}$,

$$
\begin{aligned}
&\mathbb{E}\left[K_f(E_n \times E_n \times [0,T])\right] = \mathbb{E}\left[\int_{E_n \times E_n \times [0,T]} |f(x,s) f(y,s)| K(dx\,dy\,ds)\right] \\
&\le \sum_{1\le i,j\le n} \mathbb{E}\left[|X_i||X_j| \int_{A_i \cap E_n \times A_j \cap E \times I_i \cap I_j} K(dx\,dy\,ds)\right] \\
&= \sum_{1\le i,j\le n} \mathbb{E}\left[|X_i||X_j| K(A_i \cap E_n \times A_j \cap E \times I_i \cap I_j)\right] \\
&\le R^2 \mathbb{E}\left[K(E_n \times E_n \times [0,T]])\right] < \infty
\end{aligned}
$$

where $R > 0$ is a positive constant such that $|X_i(\omega)| \leq R$ for almost every $\omega \in \Omega$ and for any $i \in \{1, ..., n\}$. Thus, $f \cdot M$ is a worthy martingale measure with covariance functional Q_f and dominating measure K_f. Moreover, by using (2.3.4) and inequality (3.3.6), we obtain (3.3.14) for elementary functions f. □

Now, for any $f \in \mathcal{P}$, by Theorem 3.3.1 there is a sequence of simple functions $\{f_n\}_n \subset \mathcal{S}$ such that $||f - f_n||_M \to 0$ as n approaches infinity. This will mean that, by (3.3.14), for any $t \in [0, T]$ and any $B \in \mathcal{E}$,

$$\begin{aligned}\mathbb{E}\left[((f_n \cdot M)_t(B) - (f_m \cdot M)_t(B))^2\right] &= \mathbb{E}\left[\ ((f_n - f_m) \cdot M)_t^2(B)\right] \\ &\leq ||f_n - f_m||_M^2 \xrightarrow{n,m\to\infty} 0.\end{aligned}$$

So the sequence of random variables $\{(f_n \cdot M)_t(B)\}_n$ is Cauchy in $L^2(\Omega)$ and thus, convergent to a random variables denoted by $(f \cdot M)_t(B)$. In particular, for fixed $B \in \mathcal{E}$, the latter converges for any $t \in [0, T]$. By completeness of the space of square integrable martingales, the limiting process $\{(f \cdot M)_t(B) \colon t \in [0, T]\}$ will be a square integrable martingale starting at 0. As usual, one checks that the limit does not depend on the choice of the approximating sequence. Of course, the obtained integral acts linearly on $\mathcal{P}_M$.

As usual, the stochastic integral $(f \cdot M)_t(A)$ is written as

$$\int_0^t \int_A f(x, s) M(dx\, ds)$$

as well.

Lastly, we check that the limiting process $f \cdot M \coloneqq \{(f \cdot M)_t(B) \colon t \in [0, T], B \in \mathcal{E}\}$ is a worthy martingale measure whenever $f \in \mathcal{P}_M$.

Theorem 3.3.2. *If $f \in \mathcal{P}_M$, then $f \cdot M$ is a worthy martingale measure with covariance functional Q_f and dominating measure K_f given by*

$$Q_f(dx\, dy\, ds) = f(x, s) f(y, s)\, Q(dx\, dy\, ds), \tag{3.3.18}$$

$$K_f(dx\, dy\, ds) = |f(x, s) f(y, s)| K(dx\, dy\, ds). \tag{3.3.19}$$

Moreover, if $f, g \in \mathcal{P}_M$ and $A, B \in \mathcal{E}$, then

$$\langle (f \cdot M)(A), (g \cdot M)(B) \rangle_t = \int_{A\times B\times[0,t]} f(x, s) g(y, s) Q(dx\, dy\, ds) \tag{3.3.20}$$

and

$$\mathbb{E}\left[(f \cdot M)_t^2(A)\right] \leq ||f||_M^2.$$

Proof. Let $\{f_n\}_n \subset \mathcal{S}$ be a sequence of simple functions such that $||f - f_n||_M \to 0$ as n approaches infinity. For each $n \in \mathbb{N}$, and $A, B \in \mathcal{E}$, the process

$$(f_n \cdot M)_t(A)(f_n \cdot M)_t(B) - \int_{A\times B\times[0,t]} f_n(x,s)f_n(y,s)Q(dx\,dy\,ds) \tag{3.3.21}$$

is a martingale by Lemma 3.3.3. Since both $(f_n \cdot M)_t(A)$ and $(f_n \cdot M)_t(B)$ converge in $L^2(\Omega)$ to, respectively, $(f \cdot M)_t(A)$ and $(f \cdot M)_t(B)$, we have that their product converges in $L^1(\Omega)$ to $(f \cdot M)_t(A)(f \cdot M)_t(B)$. On the other hand,

$$\begin{aligned}
&\mathbb{E}\left[\left|\int_{A\times B\times[0,t]} (f_n(x,s)f_n(y,s) - f(x,s)f(y,s))\,Q(dx\,dy\,ds)\right|\right] \\
&\le \mathbb{E}\left[\int_{E\times E\times[0,T]} |f_n(x,s)||f_n(y,s) - f(y,s)|K(dx\,dy\,ds)\right] \\
&+ \mathbb{E}\left[\int_{E\times E\times[0,T]} |f(y,s)||f_n(x,s) - f(x,s)|K(dx\,dy\,ds)\right] \\
&= \mathbb{E}\left[(|f_n|, |f_n - f|)_K\right] + \mathbb{E}\left[(|f|, |f_n - f|)_K\right] \\
&\le ||f_n||_M||f_n - f||_M + ||f||_M||f_n - f||_M \\
&= ||f_n - f||_M\left(||f_n||_M + ||f||_M\right) \xrightarrow{n\to\infty} 0.
\end{aligned}$$

Where we have used the Cauchy-Schwarz inequality for the norm $||\cdot||_M$. So the integral in (3.3.21) converges in $L^1(\Omega)$ to

$$\int_{A\times B\times[0,t]} f(x,s)f(y,s)Q(dx\,dy\,ds)$$

and thus, (3.3.21) converges in $L^1(\Omega)$ to

$$(f \cdot M)_t(A)(f \cdot M)_t(B) - \int_{A\times B\times[0,t]} f(x,s)f(y,s)Q(dx\,dy\,ds).$$

Given that the $L^1(\Omega)$-limit of martingales is a martingale, we conclude that the latter is a martingale, proving that (3.3.18) holds. It immediately follows that (3.3.19) and (3.3.2) hold as well and (3.3.20) follows from polarization (we have proved the case $f = g$ by proving (3.3.18)).

Finally, we check that $f \cdot M$ is a martingale measure. Let $\{A_n\}_n \subset \mathcal{E}$ be a decreasing sequence such that $\cap_n A_n = \emptyset$, then, by (3.3.2) and the monotone convergence theorem,

$$\mathbb{E}\left[(f\cdot M)_t^2(A_n)\right] \le \mathbb{E}\left[\int_{A_n\times A_n\times[0,t]} |f(x,s)f(y,s)|K(dx\,dy\,ds)\right] \xrightarrow{n\to\infty} 0,$$

as desired. $\square$

3.4 An observation and application to SPDEs

3.4.1 The two definitions coincide

Observe that when the martingale measure corresponds to the one given by the space-time white noise, in principle, the definition of stochastic integral studied in the previous section, given in [11], and the one by [6] seen in Section 3.2.1 need not be the same when the latter is restricted to predictable functions.

Fortunately, this is not the case. Indeed, consider an elementary process f as in (3.3.7) with $A \subset D$ (being D the set where the space-time white noise takes place). Walsh's definition tells us that

$$(f \cdot W)_t(B) = X\left[W([0, t \wedge b] \times A \cap B) - W([0, t \wedge a] \times A \cap B)\right],$$

while Definition 3.2.1 leads to

$$\begin{aligned}
\int_0^t \int_D f(x,s) W(dx\, ds) &= \sum_{j=1}^{\infty} \int_0^t \langle f(s,*), e_j \rangle_V dW_s(e_j) \\
&= X \sum_{j=1}^{\infty} \int_{t\wedge a}^{t \wedge b} \langle \mathbb{I}_A, e_j \rangle_V dW_s(e_j) \\
&= X \sum_{j=1}^{\infty} \langle \mathbb{I}_A, e_j \rangle_V \left[W_{t\wedge b}(e_j) - W_{t \wedge a}(e_j)\right] \\
&= X \left[W_{t\wedge b}\left(\sum_{j=1}^{\infty} \langle \mathbb{I}_A, e_j \rangle_V e_j \right) - W_{t\wedge a}\left(\sum_{j=1}^{\infty} \langle \mathbb{I}_A, e_j \rangle_V e_j \right) \right] \\
&= X \left[W_{t\wedge b}\left(\mathbb{I}_A\right) - W_{t\wedge a}\left(\mathbb{I}_A\right)\right] \\
&= X \left[W\left([0, t\wedge b] \times A\right) - W\left([0, t \wedge a] \times A\right)\right].
\end{aligned}$$

Thus, both definitions coincide for elementary functions and, by linearity, for simple functions. The result for general $f \in \mathcal{P}_M$ then follows by density.

3.4.2 Stochastic Partial Differential Equations

Recall that one of the purposes of constructing the stochastic integral with respect to the space-time white noise or, in general, any worthy martingale measure, was to rigorously define the concept of stochastic partial differential equation driven by such noises. In this section we shall briefly discuss how this is done in the case where the partial differential operator is linear.

In the case of ordinary differential equations, to define the stochastic analog, see Eq.(1.0.2), we made use of its integral representation, Eq.(1.0.1), to then exploit the already constructed stochastic integral with respect to a stochastic process (Brownian motion or, in general, càdlàg semi-martingales) to give a meaning to such expressions.

As one might suspect, this approach cannot be used to define the concept of SPDE, since, in general, there is no integral representation of a partial differential equation. However, in some cases (for instance, when the PDE is linear), the solution of a PDE has an integral representation, which we shall use to define the solution of a randomly perturbed PDE.

In the following, and to focus on the main ideas, we will omit the details regarding the conditions for which the assertions that we will make hold true. Moreover, we will be considering SPDEs defined by linear partial differential operators of the form

$$\mathcal{L} = \mathcal{L}_x,$$

where $\mathcal{L}_x$ is a linear partial differential operator in the spatial variable, like in the case of the Laplace equation,

$$\mathcal{L} = \frac{\partial}{\partial t} + \mathcal{L}_x,$$

like in the case of the heat equation, or

$$\mathcal{L} = \frac{\partial^2}{\partial t^2} + \mathcal{L}_x,$$

like in the case of the wave equation.

It is well-known that, usually, to find a solution to the problem

$$\mathcal{L}u(t,x) = f(t,x), \quad (t,x) \in (0,T] \times D,$$

where $D \subset \mathbb{R}^k$ is a bounded or unbounded domain and with some initial condition at $t = 0$ and, possibly, some boundary conditions on ∂D, one first finds a fundamental solution or a Green's function associated to the operator $\mathcal{L}$ and the corresponding boundary conditions, say $\Gamma = \Gamma(t,x;s,y)$ (this function might not depend on t and s, as in the case of the Laplace operator). Then the solution to the considered problem is given by

$$u(t,x) = I_0(t,x) + \int_0^t \int_D \Gamma(t,x;s,y) f(s,y) dy ds, \quad (t,x) \in (0,T] \times D.$$

where $I_0(t,x)$ is the solution to the homogeneous PDE, $\mathcal{L}u(t,x) = 0$, with the same initial and boundary conditions.

Replacing $f(t,x)$ by the Gaussian white noise, $\dot{W}(t,x)$, or some other suitable noise, will lead to what is called a stochastic partial differential equation. In this situation, it is natural to make the following reasoning

$$\begin{aligned}\int_0^t \int_D \Gamma(t,x;s,y) f(s,y) dyds &= \int_0^t \int_D \Gamma(t,x;s,y) \dot{W}(s,y) dyds \\ &= \int_0^t \int_D \Gamma(t,x;s,y) W(ds,dy).\end{aligned}$$

Of course, some regularity condition on Γ must be satisfied so that the integral $\int_0^t \int_D \Gamma(t,x;s,y)W(ds,dy)$ can be defined as in Def.3.2.1 or as seen in Section 3.3.2. When W is a space-time white noise, a sufficient condition is that, for all $(t,x) \in [0,T] \times D$, the functions

$$[0,T] \times D \ni (s,y) \mapsto \Gamma(t,x;s,y)\mathbb{I}_{[0,t)}(s),$$

or

$$D \ni y \mapsto \Gamma(x;y),$$

when $f = f(x)$ and $\Gamma = \Gamma(x;y)$ does not depend on the time variables, are Borel measurable functions and belong to $L^2([0,T] \times D)$ and $L^2(D)$, respectively. This formulation will lead to what are called *random field* solutions:

Definition 3.4.1. *Let W be a Gaussian space-times white noise on $[0,T] \times D$, or, in general, a worthy martingale measure. The random field solution to the SPDE $\mathcal{L}u = \dot{W}$ on $[0,T] \times D$, with the specified initial and boundary conditions, is the random field*

$$u(t,x) = I_0(t,x) + \int_0^t \int_D \Gamma(t,x;s,y)W(ds,dy),$$

or

$$u(x) = I_0(x) + \int_D \Gamma(x;y)W(dy),$$

where $I_0(t,x)$ and $I_0(x)$ are the solutions to the corresponding homogeneous PDEs, $\mathcal{L}u = 0$, with the same initial and boundary conditions.

Observe that these definitions have been made for the case where the problem is linear and the noise considered is additive. Nevertheless, these ideas can be used to define concept of solution of a SPDE of the form

$$\mathcal{L}u(t,x) = \sigma\left(t,x,u(t,x)\right)\dot{W}(t,x) + b(t,x,u(t,x)), \quad t,x \in (0,T] \times D$$

with some given initial and boundary conditions and where $\mathcal{L}$ is a linear partial differential operator as the ones previously considered. In such cases, a random field solution is defined as a jointly measurable, adapted and real valued process $u = \{u(t,x) \colon (t,x) \in [0,T] \times D\}$ satisfying

$$\begin{aligned} u(t,x) = I_0(t,x) &+ \int_0^t \int_D \Gamma(t,x;s,y)\sigma(s,y,u(s,y))W(ds,dy) \\ &+ \int_0^t \int_D \Gamma(t,x;s,y)b(s,y,u(s,y))dyds \end{aligned}$$

where $(t,x) \in [0,T] \times D$. It is implicitly assumed that the functions Γ, σ and b satisfy some regularity conditions so that the integrals on the right-hand side of this last equation can be defined as studied in previous sections.

It is worth mentioning that when the partial differential operator $\mathcal{L}$ is not linear or spatial dimensions considered are high enough, the PDEs considered might not have function-valued solutions, thus, making sense of expressions of the form $\mathcal{L}u = \sigma\dot{W} + b$ might be a difficult task. For instance, the function $\Gamma = \Gamma(t,x;s,y)$ for the heat equation is not square integrable when the spatial dimension is equal or greater than two.

Chapter 4

Good integrators

In this work, we have focused all our efforts on defining stochastic integrals with respect to martingales and martingale measures as limits of Riemann-Stieltjes sums. However, why do we restrict ourselves to such processes? Can we go a step further and reproduce the same procedure to construct new integrals with respect to some other processes?

To answer these questions, we shall ask ourselves what properties must possess the constructed integral in order to be a “good integral”. More specifically, consider a stochastic process $M = \{M_t \colon t \in [0,T]\}$ which induces a stochastic integral I_M, which we will think as an operator acting on some space of stochastic processes that returns a random variable. What properties should the operator I_M have? For instance, it should be a linear operator so that we can say it extends the already known notions of integral.

Some authors (see, for instance, [10], p.52) assert that, apart from the linearity property, the integral operator should satisfy some version of the Bounded Convergence Theorem, which is quite reasonable since this is satisfied for the classical Lebesgue and Lebesgue-Stieltjes integrals. This property is not only chosen for this purpose: as seen in the previous chapters, the stochastic integrals considered arise as limits of integrals of much simpler processes. These limits were totally justified by the Bounded Convergence Theorem, so it is natural to ask for this property to be fulfilled as well for general processes, not only for the simple ones.
As mentioned a couple times, even though the stochastic integral seen here has been constructed as an $L^2(\Omega)$-limit, when one extends the class of integrable functions with respect to a given process M, the resulting integral is a limit in probability. Hence, a version of this theorem to be satisfied is that the uniform convergence of a sequence of processes $X^{(n)}$ towards a certain process X implies the convergence in probability. One can think that we ask for uniform convergence of the sequences $X^{(n)}$ because the considered integral is constructed via Riemann-Stieltjes sums.

Now let $\mathcal{S}_u$ denote the space of simple functions of the form

$$X_t(\omega) = X_0 \mathbb{I}_{\{0\}}(t) + \sum_{j=1}^{n} e_j(\omega) \mathbb{I}_{(t_j, t_{j+1}]}(t)$$

for some $0 = t_1 < ... < t_{n+1} = T$, $e_j \in L^\infty \left(\Omega, \mathcal{F}_{t_j}, \mathbb{P}\right)$ endowed with the topology of the uniform convergence and let $L^0(\mathbb{P})$ be the space of random variables endowed with the topology of the convergence in probability. For given process M and $X \in \mathcal{S}_u$, we define the linear mapping $I_M \colon \mathcal{S}_u \to L^0(\mathbb{P})$ as follows

$$I_M(X) = X_0 M_0 + \sum_{j=1}^{n} e_j (M_{t_{j+1}} - M_{t_j}).$$

A "good integrator" $M = \{M_t \colon t \in [0, T]\}$ is then defined as a process such that its integral operator, $I_M \colon \mathcal{S} \to L^0(\mathbb{P})$, is a continuous linear mapping with the considered topologies (the continuity condition is equivalent to the Bounded Convergence Theorem property).

It turns out that if M is a good integrator, then M is a semimartingale with càdlàg sample paths. More specifically,

Theorem 4.0.1 (Bichteler-Dellacherie)**.** *An adapted, càdlàg process M is a good integrator if, and only if, it is a semimartingale. That is, M is a good integrator if, and only if, it can be written as $M = N + A$ where N is a local martingale and A is a process whose sample paths are of bounded variation.*

For a proof of this result we refer to [10], Theorem 47, page 146.

With this, we can give an answer to the previously made questions: semimartingales are the most general processes that induce an integral operator satisfying desirable properties like linearity and the Bounded Convergence Theorem.

Of course, this does not give an answer to why we limit our study of stochastic integration with respect to random fields to integrals with respect to martingale measures, but, at least, gives us an idea of why it is the case. In addition, martingale measures are good enough objects to model a large part of the scenarios that one can witness.

Bibliography

[1] Xavier Bardina and Maria Jolis. "Weak convergence to the multiple Stratonovich integral". In: *Stochastic Processes and their Applications* 90.2 (2000), pp. 277–300. ISSN: 0304-4149. DOI: https://doi.org/10.1016/S0304-4149(00)00045-4. URL: https://www.sciencedirect.com/science/article/pii/S0304414900000454.

[2] Xavier Bardina, Maria Jolis, and Lluís Quer-Sardanyons. "Weak Convergence for the Stochastic Heat Equation Driven by Gaussian White Noise". In: *Electronic Journal of Probability* 15.none (2010), pp. 1267–1295. DOI: 10.1214/EJP.v15-792. URL: https://doi.org/10.1214/EJP.v15-792.

[3] Patrick Billingsley. *Convergence of probability measures*. Second. Wiley Series in Probability and Statistics: Probability and Statistics. A Wiley-Interscience Publication. New York: John Wiley & Sons Inc., 1999. ISBN: 0-471-19745-9.

[4] Patrick. Billingsley. *Convergence of probability measures*. eng. Wiley series in probability and mathematical statistics. New York: Wiley, 1968. ISBN: 0471072427.

[5] Salim Boukfal Lazaar. "Weak convergence of the Lazy Random Walk to the Brownian motion". In: *Reports@SCM* 8.1 (Jan. 2024), pp. 11–19. DOI: 10.2436/20.2002.02.34. URL: https://revistes.iec.cat/index.php/reports/article/view/151016.

[6] Robert C. Dalang and Marta Sanz-Solé. *Stochastic Partial Differential Equations, Space-time White Noise and Random Fields*. 2024. arXiv: 2402.02119.

[7] Ioannis Karatzas and Steven E. Shreve. *Brownian motion and stochastic calculus*. eng. 2nd ed. Graduate texts in mathematics ; 113. New York: Springer, 1991. ISBN: 0387976558.

[8] Davar Khoshnevisan. *Multiparameter Processes: An Introduction to Random Fields*. Monographs in Mathematics. Springer, 2002. ISBN: 9780387954592. URL: https://books.google.es/books?id=ZRpdJFM3I8EC.

[9] Bernt Øksendal. *Stochastic differential equations : an introduction with applications*. eng. 6th ed., corr. 5th print. Universitext. Berlin: Springer-Verlag, 2007. ISBN: 9783540047582.

[10] Philip E. Protter. *Stochastic Integration and Differential Equations*. eng. Second edition, version 2.1. Vol. 21. Stochastic modelling and applied probability. Berlin, Heidelberg: Springer Berlin / Heidelberg, 2004. ISBN: 3540003134.

[11] John B. Walsh. "An introduction to stochastic partial differential equations". In: *École d'Été de Probabilités de Saint Flour XIV - 1984*. Springer Berlin Heidelberg, 1986, pp. 265–439.

[12] Michael J. Wichura. "Inequalities with Applications to the Weak Convergence of Random Processes with Multi-Dimensional Time Parameters". In: *The Annals of Mathematical Statistics* 40.2 (1969), pp. 681–687. DOI: 10.1214/aoms/1177697741. URL: https://doi.org/10.1214/aoms/1177697741.

[13] Marc Yor. "Le drap brownien comme limite en loi de temps locaux lineaires". fre. In: *Séminaire de Probabilités XVII 1981/82*. Lecture Notes in Mathematics. Berlin, Heidelberg: Springer Berlin Heidelberg, 2006, pp. 89–105. ISBN: 3540122893.