Article

General Stochastic Vector Integration - Three Approaches

Moritz Sohns *

Article

General Stochastic Vector Integration: Three Approaches

Moritz Sohns

University of South Africa, University of Oxford, University of Finance and Administration Prague; sohns@mokel.science or sohns@mokel.science

Abstract: The development of stochastic integration theory throughout the 20h century has led to several definitions and approaches of the stochastic integral, particularly for predictable integrands and semimartingale integrators. This survey provides an overview of two prominent approaches in defining the stochastic integral: the classical approach attributed to Itô, Meyer, and Jacod, and the more contemporary functional analytical approach mainly developed by Bichteler and Protter. It also delves into the historical milestones and achievements in this area and analyzes them from a modern perspective. Drawing inspiration from the similarities of existing approaches, this survey introduces a new topology-based approach to the general vector-valued stochastic integral for predictable integrands and semimartingale integrators. This new approach provides a faster, simpler way to define the general integral and offers a self-contained derivation of its key properties without depending on semimartingale decomposition theorems.

Keywords: semimartingales; stochastic integration; semimartingale topology; topological stochastic integral; vector stochastic integration

---

1. Introduction

In the past few decades, the importance of Stochastic Integration has increased immensely in multiple disciplines. This is partially due to the growth of Mathematical Finance, which relies heavily on Stochastic Calculus for pricing and hedging financial derivatives in valuation, risk management and portfolio optimization. Additionally, Stochastic Integration is used in many other fields, such as Engineering, Physics, and Biology. For instance it is used in filtering and control theory, to study how random perturbations affect physical phenomena or to model the impact of stochastic variability on population dynamics. Although basic versions of the stochastic integral are usually sufficient for most of the mentioned applications, some cases need a more advanced version. In such scenarios, the stochastic integral should be defined for multidimensional càdlàg processes as integrators, the integral operator should be linear and continuous, and the set of integrands should include all càglàd processes and as many predictable processes as possible. Two exemplary applications of these more general requirements are the two celebrated Fundamental Theorems of Asset Pricing in continuous time for semimartingale models [33,34,69,70,184].

There are two alternative prominent approaches for defining the stochastic integral and deriving its properties.

The first, known as the classical method, traces the historical development of the stochastic integral. It began with its definition for Brownian Motion integrators [79–84]. Later, with the aid of the Doob-Meyer decomposition [146], it expanded to square integrable martingale integrators [86,114,161] and a broader set of integrands [149–152]. This definition was extended to include locally bounded predictable integrands and general semimartingale integrators [26,89,157]. Jacod [88,89] introduced the general stochastic integral for unbounded predictable integrands that meet specific integrability conditions. J. Jacod's approach centered on the characterization of semimartingale jumps. An alternative method, which provided an equivalent definition of the stochastic integral,

was proposed by Chou et al. [26]. One of the compelling attributes of the general semimartingale integral is that the integrands are closed under a specific topology, namely the semimartingale or Émery topology [37,52]. However, when the integral is extended to the multidimensional case via component-wise summation, the space of integrands loses its closed property. Therefore, the vector integral must be further expanded to regain closedness, which was done by Dellacherie and Meyer [37], Jacod [90], Memin [138], Shiryaev and Cherny [184].

Over time, several simplifications and shortcuts have been introduced to the classical approach. For instance, employing the Fundamental Theorem of Local Martingales has simplified the extension from square integrable martingales to local martingales and, consequently, to semimartingales [42,155]. However, defining the general stochastic vector integral and its properties is still challenging and entails numerous technical proofs. At some point in the theory, any semimartingale needs to be decomposed into a continuous and a purely discontinuous part. This usually requires more advanced tools, so most authors restrict themselves to continuous integrators. Even in textbooks that treat general semimartingales as integrators, the authors usually only present the case of locally bounded predictable integrands. This simplifies the theory significantly. Not only because locally bounded predictable integrands allow for a straightforward application of the monotone class theorem but also because local martingales are stable under stochastic integration for locally bounded predictable integrands which avoids the more complicated theory of sigma martingales.

In contrast to the classical approach, the functional analytic approach presents an alternative. This more recent approach is built on a result by Bichteler [13,14] and Dellacherie [35] (first published in Meyer [156] and inspired by results by Pellaumail [166] and Metivier and Pellaumail [141]), stating that semimartingales are the most general class of integrators for which one can define the stochastic integral for predictable integrands in a way that the integral operator is continuous with respect to specific metrics. This was then further developed by Lenglart [124] and Protter [170,171]. Protter [172] even wrote an entire comprehensive textbook based on the functional analytic approach, covering the results from all the above publications.

The functional analytic approach directly formulates the stochastic integral for one-dimensional semimartingale integrators and càglàd integrands by drawing upon techniques from functional analysis. However, for extending the stochastic integral to general predictable integrands, this approach again relies on the decomposition of semimartingales and utilizes similar, very technical, methods as the classical approach. Also, the functional analytic method does not address the vector-valued scenario. To extend the integral in this context, one must revert to classical methods again.

Owing to these complexities in both approaches, standard textbooks and university lectures rarely cover the general stochastic integral. As a result, topics like the Fundamental Theorems of Asset Pricing and their associated proofs are typically reserved for experts in stochastic analysis.

This survey explores and compares the two approaches, discussing their historical development and highlighting their similarities. The terminology has been slightly adjusted to align with modern terminology, thus avoiding some redundant definitions and notations in an effort to enhance clarity and emphasize common patterns.

Inspired by the commonalities of the two approaches — such as defining the stochastic integral through a feasible extension without using any of their results — this paper introduces a new method for the General Stochastic Integral. It leverages the fact that the integral operator is a linear isometry between topological vector spaces, offering a faster alternative to the two established methods. The caveat of this approach is its high level of abstraction, making the demonstration of its properties challenging. Notably, similar approaches have been previously utilized by Chou et al. [26], Karandikar and Rao [105], and Assefa and Harms [4].

For the classical and functional analytic approaches, we present the definitions and derivations of the stochastic integral without proving their fundamental properties or results. In contrast, we develop the topological stochastic vector integral independently in a self-contained manner, without relying on

previous methods. Furthermore, the survey includes proofs of the basic properties of the stochastic integral.

This survey is organized as follows.

First, we will review some definitions and well-known results from probability theory, topology, and functional analysis that are essential for any definition of the stochastic integral and stochastic analysis in general.

The next chapter will examine the classical approach to stochastic integration. We will, mostly in chronological order, trace the development of the stochastic integral, starting with the classical approach developed by Itô. We will present all definitions and important theorems but adjust the terminology to align with modern approaches to minimize complexity and better highlight the similarities and differences between different methods. Using the Doob decomposition of martingales, we will extend Itô's definition of the stochastic integral to allow for square-integrable martingales as integrators, as Doob, Kunita, Watabe, and Meyer did. We will then follow the various generalizations of the stochastic integral developed by the Strasbourg school around Paul-André Meyer to arrive at the general vector-valued stochastic integral for semimartingale integrators and predictable integrands.

In the following chapter, we review and comment on the functional analytic approach. This approach can be divided into two parts. The definition of the stochastic integral for càglàd integrands fundamentally differs from the classical approach and offers a definition without relying on the decomposition of semimartingales. This will be covered in the first section of that chapter. The second section deals with the extension to predictable integrands, which is similar to classical methods and relies on the decomposition of semimartingales.

The last chapter introduces a third, completely new approach to the general vector-valued stochastic integral for predictable integrands and semimartingale integrators. To that end, we first provide essential preliminaries and define the required metrics and topologies. We then introduce a definition of semimartingales equivalent to the widely used ones and highlight that the set of semimartingales is a topological vector space with respect to the semimartingale topology. After this, we analyze the essential properties of semimartingales and illustrate them with examples.

Next, we examine the general stochastic integral using the new topological approach. Its definition is straightforward and does not require complex proofs. However, studying the basic properties of the stochastic integral relies heavily on the technically involved proof of the Dominated Convergence Theorem, which can be found in the appendix. We will demonstrate all the properties of the stochastic integral that are typically necessary for applications without relying on the decomposition of semimartingales.

Lastly, we explore the stability of local martingales under stochastic integration, as this is highly relevant in finance, and provide a quick overview of the definition of the quadratic variation consistent with this new approach.

In this survey, we focus on predictable integrands. Although several extensions are known in the literature, and the initial versions of the stochastic integral were not restricted to predictable integrands, we concentrate on them because only predictable integrands allow for very general stochastic integral operators that exhibit specific desirable properties, such as the stability of the sigma-martingale property.

2. Preliminaries

2.1. Definitions and Well-Known Results

We always assume a complete filtered probability space $(\Omega, \mathcal{F}, \mathbf{P})$, equipped with a filtration $\mathbb{F} = (\mathcal{F}_t)_{t \geq 0}$ satisfying the *usual conditions*. This means:

- $\mathcal{F}_0$ includes all $\mathbf{P}$-null sets from $\mathcal{F}$.
- The filtration $\mathbb{F} = (\mathcal{F}_t)_{t \geq 0}$ is right-continuous, i.e., $\mathcal{F}_t = \bigcap_{u > t} \mathcal{F}_u$ for all $t \geq 0$.

We use the following definitions and notations:

- For $p \in [1, \infty)$, define the seminorm

$$\|X\|_{L^p} := \left(\mathbf{E}[|X|^p] \right)^{1/p}.$$

The space L^p is then defined as

$$L^p := \{ X : \Omega \to \mathbb{R} \text{ is measurable and } \|X\|_{L^p} < \infty \}.$$

- A *stochastic process* X on $(\Omega, \mathcal{F}, \mathbf{P})$ is a collection of $\mathbb{R}$-valued or $\mathbb{R}^d$-valued random variables $(X_t)_{t \geq 0}$.
- A process X is said to be *adapted* to $\mathbb{F}$ if X_t is $\mathcal{F}_t$-measurable for each t.
- Two stochastic processes X and Y are *modifications* of each other if $X_t = Y_t$ a.s. for all t, and they are *indistinguishable* if for almost every $\omega \in \Omega$, $X_t(\omega) = Y_t(\omega)$ for all t.
- A *stopping time* T is a non-negative random variable such that $\{T \leq t\} \in \mathcal{F}_t$ for each $t \geq 0$.
- For a stopping time T, the *σ-algebra $\mathcal{F}_T$ of events occurring up to time T* is the σ-algebra consisting of those events $A \in \mathcal{F}$ with

$$A \cap \{T \leq t\} \in \mathcal{F}_t \text{ for every } t \geq 0.$$

- A process X is said to be *càdlàg* (respectively *càglàd*) if its paths are right (resp. left) continuous with left (resp. right) limits.
- The *essential supremum* of a random variable X is given by

$$\operatorname*{ess\,sup}_{\omega \in \Omega} X(\omega) = \inf\{c \in \mathbb{R}; \mathbf{P}(X > c) = 0\}.$$

Indicator function
- The function $\mathbb{1}_A$ denotes the indicator function of a subset A:

$$\mathbb{1}_A(x) = \begin{cases} 1 & \text{if } x \in A, \\ 0 & \text{else.} \end{cases}$$

- Let S, T be stopping times. Then, the sets

$$
\begin{aligned}
[\![S, T]\!] &:= \{(\omega, t) \in \Omega \times \mathbb{R}_+ ; S(\omega) \leq t \leq T(\omega)\} \\
[\![S, T[\![&:= \{(\omega, t) \in \Omega \times \mathbb{R}_+ ; S(\omega) \leq t < T(\omega)\} \\
]\!]S, T]\!] &:= \{(\omega, t) \in \Omega \times \mathbb{R}_+ ; S(\omega) < t \leq T(\omega)\} \\
]\!]S, T[\![&:= \{(\omega, t) \in \Omega \times \mathbb{R}_+ ; S(\omega) < t < T(\omega)\} \\
[\![S]\!] &:= [\![S, S]\!].
\end{aligned}
$$

are called *stochastic intervals*.
- For a stopping time T and a process X, $X_t^T := X_{t \wedge T}$ denotes the *stopped process*.
- A family of random variables $(U_\alpha)_{\alpha \in A}$ is *uniformly integrable* if

$$\lim_{n \to \infty} \sup_{\alpha \in A} \int_{\{|U_\alpha| \geq n\}} |U_\alpha| \, d\mathbf{P} = 0.$$

- A real-valued, adapted process $X = (X_t)_{0 \leq t < \infty}$ is called a *martingale* (resp. *supermartingale*, *submartingale*) with respect to the filtration $\mathbb{F}$ if:
  (i) $X_t \in \mathcal{L}^1(\mathbf{P})$;
  (ii) if $s \leq t$, then $\mathbf{E}[X_t | \mathcal{F}_s] = X_s$, a.s. (resp. $\leq X_s$, $\geq X_s$).
- A process X is a *local martingale* if there exists a sequence of stopping times $(T^n)_{n \in \mathbb{N}}$ increasing to ∞ a.s. such that X^{T^n} is a martingale for each n.

- For a càdlàg process Y, the process Y_- is defined to be the process Z that satisfies $Z_t = Y_{t-}$. Furthermore, we set $XY_{0-} := Y_0$.
- For a càdlàg process X the jump process ΔX is defined as $\Delta X_t := X_t - X_{t-}$ with $X_{t-} = \lim_{s \nearrow t} X_s$. Hence, one gets $X_- = X - \Delta X$ and $\Delta X_0 = 0$.
- A property is said to hold *locally* (resp. *prelocally*) for a stochastic process X if there exists a localizing sequence of stopping times $(T^n)_{n \in \mathbb{N}}$ such that the property holds for each X^{T^n} (resp. $X_-^{T^n}$) almost surely.
- The *predictable σ-algebra* is the smallest σ-algebra making all left-continuous, adapted processes measurable. It is denoted by $\mathcal{P}$. A *predictable process* is a stochastic process $X = (X_t)_{t \geq 0}$ such that the mapping

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

  is measurable with respect to the predictable σ-algebra $\mathcal{P}$ on $\Omega \times [0, \infty)$ and the Borel σ-algebra on $\mathbb{R}$. The set of all predictable processes is also denoted by $\mathcal{P}$.
- A càdlàg process A is an *FV-process* if almost all of its paths have finite variation on each compact interval.

Theorem 1 ([153], Section § 2). *The predictable σ-algebra $\mathcal{P}$ on $\Omega \times [0, \infty)$ can be characterized equivalently in the following ways:*

(i) $\mathcal{P}$ *is the smallest σ-algebra such that all left-continuous, adapted processes are measurable.*

(ii) $\mathcal{P}$ *is the σ-algebra generated by all continuous, adapted processes.*

Definition 1. *An adapted process $B = (B_t)_{0 \leq t < \infty}$ taking values in $\mathbb{R}^n$ is called an n-dimensional Brownian motion if*

(i) *for $0 \leq s < t < \infty$, $B_t - B_s$ is independent of $\mathcal{F}_s$ (increments are independent of the past);*

(ii) *for $0 < s < t$, $B_t - B_s$ is a Gaussian random variable with mean zero and variance matrix $(t - s)C$, for a given, non-random matrix C.*

The Brownian motion starts at x if $\mathbf{P}(B_0 = x) = 1$.

Theorem 2 ([172], Theorem I.26). *Let B be a Brownian motion. Then there exists a modification of B which has continuous paths almost surely.*

We assume that all Brownian motions discussed have continuous paths and, unless stated otherwise, the covariance matrix C is the identity matrix, making it a standard Brownian motion.

Theorem 3 ([172], Theorem I.12). *Let X be a right-continuous martingale which is uniformly integrable. Then $Y = \lim_{t \to \infty} X_t$ a.s. exists, $\mathbf{E}[|Y|] < \infty$, and one says Y closes X as a martingale.*

Theorem 4 ([172], Theorem I.13). *Let X be a (right-continuous) martingale. Then $(X_t)_{0 \leq t < \infty}$ is uniformly integrable if and only if $Y = \lim_{t \to \infty} X_t$ exists a.s., $\mathbf{E}[|Y|] < \infty$, and $(X_t)_{0 \leq t \leq \infty}$ is a martingale, where $X_\infty = Y$.*

The following lemma is rarely explicitly mentioned in the literature. It is, however, crucial for our approach.

Lemma 1 ([71], Theorem 7.38). *Let M be a local martingale, T a stopping time and ξ a real $\mathcal{F}_T$-measurable random variable. Then*

$$N = \xi(M - M^T)$$

is a local martingale.

The following theorem goes back to Burkholder [23]. Updated versions with notations aligned to stochastic integration and simplified proofs can be found in [153, Theorem 47], [164, Corollaire VIII-3-14] and [49], also see [175] for similar results.

Theorem 5 ([164], Corollaire VIII-3-14). *Let $(M_t)_{t \in \mathbb{N}}$ be a time-discrete martingale and $(H_t)_{t \in \mathbb{N}}$ a predictable process with $\sup_t |H_t| < 1$ almost surely. For a sequence of time-discrete random times $(T^n)_{n \in \mathbb{N}}$ and $t \in \mathbb{N}$, define*

$$(H * M)_t := \sum_{k=1}^{t} H_{k-1}(M_k - M_{k-1}),$$

where H_k are random variables with respect to $\mathcal{F}_k$. Then, for any $T \in \mathbb{N}$, we have

$$\mathbf{P}\left(\sup_{n \leq T} |(H * M)_n| > \alpha \right) \leq \frac{18}{\alpha} \mathbf{E}|M_T|$$

for all $\alpha > 0$. If M is positive, then 18 can be replaced by 9. If M can be closed by a random variable M_∞, we can replace T by ∞.

2.2. Measurability of Stochastic Processes

For the study of stochastic analysis, stochastic processes and stochastic integration, it is essential to address some fundamental concepts related to the measurability of stochastic processes. Historically, issues with measurability were often treated as an afterthought. Many of the original papers and textbooks on the subject overlook these issues, leaving gaps in their proofs. Here, we give a short overview of the topic and provide the results that close the gaps in the original literature.

In continuous-time stochastic processes, measurability can be more challenging to handle than in discrete-time settings. This is because sets of measure zero, though individually insignificant, can have a positive probability when accumulated over uncountable sets. Before the foundational works of Dynkin [47] and Chung and Doob [27], who introduced the concept of progressive measurability, measurability issues were often neglected. However, the above-mentioned contributions laid the groundwork for handling measurability in more general stochastic processes. In particular, predictable processes – a key subclass of progressively measurable processes – play a crucial role in stochastic integration. The importance of predictable processes later became evident in the work of Courrege [30], Courrège [31] and Meyer [145,146,149,150,151,152].

We start with the basic definitions.

Definition 2. *Let $(X_t)_{t \geq 0}$ be a stochastic process on the filtered probability space $(\Omega, \mathcal{F}, \mathbb{F}, \mathbf{P})$.*

(a) It is called jointly measurable or product measurable if the mapping $X : \Omega \times \mathbb{R}_+ \to \mathbb{R}$ is measurable with respect to the product σ-algebra $\mathcal{F} \otimes \mathcal{B}(\mathbb{R}_+)$, where $\mathcal{B}(\mathbb{R}_+)$ denotes the Borel σ-algebra on $\mathbb{R}_+$.

(b) It is called progressively measurable if the restriction of X to $\Omega \times [0, t]$ is measurable with respect to the product σ-algebra $\mathcal{F}_t \otimes \mathcal{B}([0, t])$, where $\mathcal{F}_t$ is the filtration up to time t.

(c) It is called predictable if it is measurable with respect to the predictable σ-algebra $\mathcal{P}$.

Product measurability ensures that the process can be integrated with respect to both the probability measure on Ω and the Lebesgue measure on $\mathbb{R}_+$. This is a crucial property for defining stochastic integrals, as it guarantees that the integrands are well-behaved in both the probabilistic and temporal dimensions. However, product measurability alone does not guarantee that the process is adapted to the filtration.

Progressive measurability is a stronger condition than product measurability because it requires that the measurability holds for every finite time interval, taking into account the information available up to that time. This condition ensures that the values of the process at each time depend only on the information available up to that time, thereby maintaining the non-anticipative nature of the

process. Progressive measurability is essential for the stochastic integral over any time interval to be a well-defined random variable. Additionally, if a process $(X_t)_{t \geq 0}$ is progressively measurable, then it is necessarily adapted. However, not every adapted process is progressively measurable.

Predictability is an even stronger condition that requires measurability with respect to the predictable σ-algebra, which is generated by left-continuous, adapted processes. Predictable processes are particularly important in stochastic integration because they form the largest class of integrands for which the stochastic integral with respect to semimartingales can be defined in such a way that the integral exhibits certain desirable properties. Every predictable process is progressively measurable, and by extension, product measurable. However, not all progressively measurable processes are predictable, as predictability specifically requires left-continuity or an equivalent structure.

Proofs of these statements can be found in [37, p 120 ff].

Verifying progressive measurability is inherently more challenging than verifying adaptability, although there exist simple and useful sufficient criteria. For instance, the limits of progressively measurable processes are also progressively measurable, and if $(X_t)_{t \geq 0}$ is an adapted process with right- or left-continuous paths, then it has a progressively measurable modification ([186, Proposition 1.13]). Another important property is that for a progressively measurable process X and a finite stopping time T, the random variable X_T is $\mathcal{F}_T$-measurable ([186, Proposition 2.18]).

These definitions are crucial for resolving the measurability issues encountered in the construction of the Itô integral. By ensuring that the process is progressively measurable, we obtain the adaptedness of the process throughout operations such as smoothing. This is because progressive measurability inherently incorporates the adaptedness to the filtration, which is generated by the Brownian motion in the context of the Itô integral.

Theorem 6 (Chung-Doob-Meyer [27,147]). *Let X be an $\mathbb{R}$-valued measurable stochastic process adapted to a filtration $\mathcal{F}$. Then, X has a measurable $\mathcal{F}$-progressive modification.*

The history of this theorem involves significant contributions from several researchers. Initially, [27] proved the existence of a progressively measurable modification under the assumption of separability. Meyer [147] extended this result by removing the separability condition, although his argument was considered complex, demanding and incomplete. The gap in Meyer's proof was later addressed by Dellacherie and Meyer [37], using advanced tools from functional analysis. The theorem plays a pivotal role in the creation of the stochastic integral how it is presented in [186].

We summarize the results:

- Every *predictable* process is *progressively measurable* and *product measurable*.
- Every *progressively measurable* process is *product measurable*, but not necessarily *predictable*.
- A *product measurable* process may not be progressively measurable or predictable. However, for any process that is product measurable and adapted, there exists a modification which is progressively measurable.

Thus, predictability implies progressive measurability, and progressive measurability implies product measurability. Each level of measurability provides increasing levels of structure and guarantees certain desirable properties necessary for defining stochastic integrals.

In this survey, we mostly avoid the aforementioned measurability issues by mainly focusing on predictable processes, which are also necessarily progressively measurable and product measurable. However, for the sake of completeness, these different notions of measurability are included here to provide a comprehensive foundation for the discussion of stochastic integration, especially considering that some of the referenced sources here include the above-mentioned gaps.

3. The Classical Approach

This section outlines the classical approach to stochastic integration and traces its historical development. The goal is to not only review key literature but also to identify the technical challenges

encountered throughout its development. Given that similar obstacles are expected in the new topological approach introduced in this work, a discussion of the classical solutions will provide valuable insights.

For a comprehensive account of the historical progression of stochastic processes and integration theory, see [96,119,158,196].

It should be noted that while defining the stochastic integral is typically straightforward, proving more advanced properties — such as the Dominated Convergence Theorem and change of variable formulas — tends to be the more challenging task.

3.1. The Beginning: Wiener and Doeblin

It is assumed that Norbert Wiener (1894 – 1964), an American mathematician, was the first to define the stochastic integral

$$\int_0^t f(s)\mathrm{d}B_s$$

for a Brownian motion B and smooth deterministic functions f, defined on the positive half-axis, utilizing the concept of "integration by parts" [165,197,198]. Starting with the "natural" formula $\mathrm{d}(fB) = f\mathrm{d}B + B\mathrm{d}f$, one puts

$$\int_0^t f(s)\mathrm{d}B_s = f(t)B_t - \int_0^t f'(s)B_s\mathrm{d}s,$$

where the integral $\int_0^t f'(s)B_s\mathrm{d}s$ is treated as the trajectory-wise (i.e., for each $\omega \in \Omega$) Riemann integral of the continuous functions $f'(s)B_s(\omega)$ defined for $s \geq 0$.

However, since the need for f to be differentiable and deterministic is too stringent and of little practical use, the idea has received little attention.

The German-French mathematician Wolfgang Doeblin (1915 – 1940), later known as Vincent Dobbin, was probably the first to establish an integration theory for non-differentiable integrands, developing a theory quite similar to the one Itô developed independently of Doeblin's findings a few years later. However, Doeblin died shortly after developing his idea during the Second World War while fighting for France, and his results were lost until the early twenty-first century [21].

3.2. Itô's Stochastic Integral

From today's point of view, Kiyosi Itô (1915 – 2008) is mostly considered as the "discoverer" of stochastic integration. He made groundbreaking contributions to the theory in the 1940s and 1950s, independently of earlier work by Wolfgang Doeblin. His work laid the foundation for modern stochastic calculus, culminating in several key publications [79–84], which are extensively summarized in [133]. In these papers, Itô rigorously defined the stochastic integral using Brownian motion as the integrator.

Before Itô's contributions, applied scientists had already been informally working with integrals involving "white noise", the non-existent derivative of Brownian motion. Itô's key innovation was to redefine this product of white noise and time differential as the differential of Brownian motion, which he used as an integrator. This method combined aspects of both the Riemann–Stieltjes integral (regarding the integrator) and the Lebesgue integral (regarding the integrand), which formalizes the construction of the stochastic integral.

The motivation behind Itô's calculus arose from the need to simplify the construction of Markov diffusion processes, which had traditionally been achieved through a more cumbersome multi-step process. This process involved the Hille–Yosida theory, the Riesz representation theorem, and the Kolmogorov extension theorem. Itô's approach eliminated these steps, which allowed for the direct construction of diffusion processes as solutions to stochastic integral equations. The properties of these processes could then be derived through the stochastic integral equations and the now-famous Itô formula.

In this section, we revisit the original approach by Itô towards stochastic integration and only incorporate minor adjustments to align the method with more general and contemporary frameworks; for example, we employ left-continuous simple functions rather than right-continuous ones as was presented by Itô, which subsequently leads to the development of the theory of predictable integrands as opposed to optional integrands. This adjustment allows us to eliminate the assumption that the underlying filtration is quasi-left continuous, thereby harmonizing with the modern setup of stochastic integration. While this alteration has minimal impact on the stochastic integral as defined by Itô, it significantly enhances the framework for further extensions of the stochastic integral, which do not seamlessly integrate with Itô's original setup. The reason is that we want to present a coherent theory in this survey, highlighting the similarities and key differences between the different approaches.

Additionally, we extend the domain of the stochastic integral from the interval $[0, 1]$ to $[0, \infty)$. The original interval can always be recovered by appropriately adjusting the integrand. We avoid defining the stochastic integral over the entire axis $(-\infty, \infty)$, as was done by Doob [44] because it has proven to be irrelevant for most practical applications.

In addition, we've standardized the notation and terminology to align with modern conventions and recent approaches, helping to avoid inconsistent definitions and notations. We've also tackled the issue of measurability right from the start since it has often been a point of confusion and the cause of incomplete proofs in the past (see Section 2.2).

Itô's objective was to develop a definition for the stochastic integral

$$I_t(H) := H \bullet B_t := \int_0^t H(s, \omega) \mathrm{d}B_s$$

that imposed minimal limitations on the process H. Since Brownian motion trajectories (**P**-a.s.) have unbounded variation, an application of the Banach-Steinhaus theorem demonstrates that the integral

$$I_t(H) = H \bullet B_t = \int_0^t H(s, \omega) \mathrm{d}B_s$$

cannot be defined as the trajectory-wise Lebesgue-Stieltjes integral (see, for example, [172, Theorem I.29] and section I.8 in the same book). Itô proposed defining it as the limit (in an appropriate probabilistic sense) of integrals $I_t\left(H^{(n)}\right)$ of suitable functions $H^{(n)}$ for which a natural stochastic integral definition exists. This technique was later used for all extensions of the stochastic integral as well. Itô began by defining the stochastic integral for *elementary predictable processes*.

Definition 3. *A elementary predictable process $H : \Omega \times [0, \infty)$ is defined as*

$$H(t, \omega) = H_0(\omega) \mathbb{1}_{\{0\}}(t) + \sum_{i=1}^n H_i(\omega) \mathbb{1}_{(t_i, t_{i+1}]}(t).$$

where each $H_i(\omega)$ is an $\mathcal{F}_{t_i}$-measurable, bounded, random variable, and $0 = t_1 \leq t_i \leq t_{i+1} < \infty$ for all i. The space of elementary predictable processes is denoted by $\mathcal{E}$.

Remark 1. *Using stopping times T^i instead of deterministic times t_i, yields simple predictable processes. These are used, for example, in the functional analytic approach to define the stochastic integral. However, both approaches yield the same stochastic integral, as both elementary and simple predictable processes span the space of predictable processes (see Theorem 1 or, for more elaborations, [28,176]). Nevertheless, in the literature, the definition for simple predictable processes is often inconsistent. We stick to the most widely-used definition.*

Given a Brownian motion B and an elementary predictable function H adapted to the filtration generated by B, the integral $J^B(H)$ is defined as follows:

Definition 4. *Let $H \in \mathcal{E}$ be represented as*

$$H(t, \omega) = H_0(\omega)\mathbb{1}_{\{0\}}(t) + \sum_{i=1}^{n} H_i(\omega)\mathbb{1}_{(t_i, t_{i+1}]}(t).$$

We define the stochastic integral $J^B(H)_t$ by

$$J_t^B(H) := H \bullet B_t := \int_0^t H \mathrm{d}B_s = \sum_{i=1}^{n} H_i\big(B_{t_{i+1} \wedge t} - B_{t_i \wedge t}\big)$$

and

$$J_\infty^B(H) = \sum_{i=1}^{n} H_i\big(B_{t_{i+1}} - B_{t_i}\big)$$

The stochastic integral, as given above, offers a direct interpretation within financial theory. Consider a process B that represents the price of a financial asset. If an investor rebalances their holdings at deterministic times t_i, maintaining $H_i(\omega)$ shares during the interval $(t_i, t_{i+1}]$, then the value of $J_t^B(H)$ represents the investor's accumulated wealth at time t, which may be positive or negative depending on market movements.[1]

Remark 2. *One of the key innovations in the definition of the stochastic integral, as given above, is that the integral is defined in a "non-anticipating" way. This construction is non-anticipating in two key aspects. Firstly, the integrand is evaluated at the start of each interval, ensuring that the integration process does not depend on future values. Secondly, the evaluated values H_t depend solely on the past observations of the Brownian path up to time t, aligning with the term "adapted" that has become common in the literature. It should be re-emphasized that the Itô integral does not exist, in general, for non-adapted integrands. It does, however, exist for non-predictable or optional integrands (for discontinuous integrators, one usually needs the additional assumption of predictability for the stochastic integral to exist and exhibit the properties one would expect from a stochastic integral).*

Remark 3. *As mentioned above, in the original work, Itô considers step functions that are right-continuous. However, in the context presented here, we use elementary predictable processes, which are by definition left-continuous. For integrators that are continuous processes, such as Brownian motion, the choice between right and left continuity of the step functions used in approximations does not affect the outcome of the integral. Yet, when extending the theory to include integrators with discontinuities, right-continuous step functions do not behave as seamlessly. In particular, in order to ensure that stochastic integrals with respect to martingales do not exhibit any drift, the integrands must be approximable by left-continuous elementary predictable processes. Therefore, even though the distinction is inconsequential for integrals with respect to Brownian motions, we opt for left-continuous functions to maintain consistency with the broader framework. While the use of elementary predictable processes ultimately leads to the concept of the predictable σ-algebra, the use of right-continuous step function would lead to optional processes. Even though there has also been a stochastic integration theory developed for optional processes, it is not considered as important. Itô also adjusted his approach in later publications and switched to left-continuous elementary predictable processes [85].*

For the extension to a larger class of integrands, Itô's central idea was to define appropriate metrics for the set of elementary predictable processes and for a subspace of the space of random variables in such a way that the mapping that maps an elementary predictable process H on its stochastic integral $J^B(H)$ becomes an isometry and then to extend this isometry to the closure of the set of elementary predictable processes via the well-known extension theorems. This technique has been the same for all extensions for the stochastic integral and is, therefore, extremely crucial.

---

1   In practice, the modelling of asset prices typically involves processes that exhibit positivity and may incorporate a drift component, differing from the idealized model of a Brownian motion.

Itô considered a specific metric on the space of adapted functions H that satisfy certain integrability conditions.

Definition 5. *(a)* *For any jointly measurable and adapted process H, we define the seminorm $\|\cdot\|_{\text{Itô}}$ by*

$$\|H\|_{\text{Itô}}^2 := \mathbf{E}\left[\int_0^\infty H^2(s,\omega)\mathrm{d}s\right]. \tag{1}$$

(b) *By $\mathcal{E}_{\text{Itô}}$, we denote the space of elementary predictable processes endowed with the seminorm $\|\cdot\|_{\text{Itô}}$.*

The following result can be demonstrated by leveraging the fact that the integrands are adapted to the filtration of the Brownian motion. As a result, the independence of the Brownian motion's increments allows us to establish the L^2-isometry for elementary predictable processes.

Lemma 2 ([79], Theorem 2.1). *The mapping*

$$J_\infty^{\text{Itô}} : \begin{cases} \mathcal{E}_{\text{Itô}} & \to & L^2(\Omega, \mathcal{F}, \mathbf{P}) \\ H & \mapsto & H \bullet B_\infty \end{cases}$$

is an isometry. Specifically,

$$\|H\|_{\text{Itô}}^2 = \mathbf{E}\left[\int_0^\infty H(s,\omega)^2 \mathrm{d}s\right] = \mathbf{E}\left[(H \bullet B_\infty)^2\right] = \|J_\infty^{\text{Itô}}(H)\|_{L^2}^2$$

In order to be able to extend the stochastic integral from $\mathcal{E}$ to its closure $\overline{\mathcal{E}}$, we need to define a space in which we want to close the subspace $\mathcal{E}$. We follow Itô's original approach and choose

$$\mathcal{E}_0 := \{H \,;\, \text{jointly measurable, adapted and } \|H\|_{\text{Itô}} < \infty\}.$$

Let $\overline{\mathcal{E}}$ denote the closure of $\mathcal{E}$ in $\mathcal{E}_0$ with respect to $\|\cdot\|_{\mathcal{E}}$.

Definition 6. *The continuous linear mapping $\overline{J_\infty^{\text{Itô}}}$ from $\overline{\mathcal{E}}$ to $L^2(\Omega, \mathcal{F}, \mathbf{P})$, obtained as the extension of $J_\infty^{\text{Itô}}$ from elementary predictable processes to L^2 by continuity, is called the stochastic integral. That is, for $H \in \overline{\mathcal{E}}$, let $(H^n)_{n \in \mathbb{N}}$ be a sequence of elementary predictable processes in $\mathcal{E}$ that converges to H in $\|\cdot\|_{\mathcal{E}}$. Then the stochastic integral $\overline{J_\infty^{\text{Itô}}}(H)$, denoted by $H \bullet B_\infty$, is defined as the L^2 limit of the stochastic integrals of the approximating sequence, i.e.,*

$$H \bullet B_\infty = \lim_{n \to \infty} H^n \bullet B_\infty,$$

where the convergence is in the $L^2(\Omega, \mathcal{F}, \mathbf{P})$ sense. For each $t \geq 0$, by setting

$$H \bullet B_t = H\mathbb{1}_{[0,t]} \bullet B_\infty,$$

we obtain a well-defined stochastic process, which we call the stochastic integral of H with respect to B.

In his first works, Itô initially defined the stochastic integral for integrands within the topological closure of elementary predictable processes, providing an abstract characterization of possible integrands. With today's knowledge, it is evident that the canonical choice for integrands is the space of predictable processes with a finite L^2-norm. The concept of predictable processes, however, was not formulated at the time of Itô's early work. Subsequent literature typically identifies integrands as the set of adapted processes H satisfying $\|H\|_{\mathcal{E}}^2 < \infty$, with additional considerations necessary to ensure measurability, as discussed in [28,96,101,133].

In the definition above, the set of potential integrands is given as the closure of the elementary predictable processes with respect to the Itô seminorm, which provides a somewhat abstract charac-

terization. The following theorem offers concrete conditions for describing these integrands more explicitly.

Theorem 7 ([85], Theorem 5.2.1). *Let $\mathscr{L}^2$ be the space of jointly measurable, adapted processes with*

$$\mathbf{E}\left[\int_0^\infty H^2(s,\omega)\,\mathrm{d}s\right] < \infty. \tag{2}$$

Then $\overline{\mathcal{E}} = \mathscr{L}^2$. That is, $\mathscr{L}^2$ is a complete metric space, and the space of elementary predictable processes $\mathcal{E}$ is dense within $\mathscr{L}^2$.

For the proof of Theorem 7, which is often combined with the proof of the Itô isometry 8 (see below), formal measurability arguments are important but are often neglected (see Section 2.2) Here, a possible issue is the potential loss of the adaptedness of processes during certain operations, such as smoothing, which involves uncountably many values of the time parameter.

The proof of Theorem 7, and therefore ultimately also of the Itô isometry, relies on approximating the stochastic process X by a sequence of elementary predictable processes X_n that are adapted to the filtration generated by the Brownian motion. These approximations are achieved by truncating X and smoothing the truncated versions. The Itô isometry is then demonstrated for these elementary predictable processes, and by taking limits, it is extended to the original process X. The smoothing operation, essential for making the process pathwise continuous, generally does not preserve the adaptedness of the process due to its reliance on an uncountable set of time values. This problem directly affects the applicability of the Itô isometry and, consequently, the construction of the Itô integral. In the literature, these issues have, for example, been neglected in [133] but also in the more recent [202], see also the discussions in [28,96].

One proposed solution to ensure the measurability is to assume that the process X is separable [56], although this approach is not widely adopted in the literature. The more common approach is to employ the condition of progressive measurability, which is slightly stronger than adaptedness and preserves the condition of adaptedness throughout the smoothing operation, thus ensuring the applicability of the Itô isometry in the construction of the Itô integral [101]. With Theorem 6, one obtains the general result.

By Theorem 7, the Itô Isometry can be naturally extended to $\overline{\mathcal{E}} = \mathscr{L}^2$. For completeness, we mention the result again.

Theorem 8 (Itô Isometry). *Let J^B be the integral operator as previously defined. When $\mathscr{L}^2$ is equipped with the extension of the norm $\|\cdot\|_{\mathcal{E}}$ to $\|\cdot\|_{\mathscr{L}^2}$, then for all $H \in \mathscr{L}^2$, the operator $J^B(H)$ is an isometry from $\mathscr{L}^2$ into $L^2(\Omega, \mathcal{F}, \mathbf{P})$. Specifically, for any $H \in \mathscr{L}^2$, it holds that*

$$\|H\|_{\mathscr{L}^2}^2 = \mathbf{E}\left[\int_0^\infty H(s,\omega)^2\,\mathrm{d}s\right] = \mathbf{E}\left[\left(J_\infty^B(H)\right)^2\right] = \|J_\infty^B(H)\|_{L^2}^2.$$

The set of potential integrands can be extended further via localization (see, for example, Klenke [109, p. 642 ff] or Kuo [118, Chapter 5]). However, at the time of Itô's first articles, the concept of localization was unheard of. We will touch on this topic later.

3.3. The decomposition of martingales and Doob's work

Itô's research was limited to Brownian motions as integrators. The next necessary generalization from today's point of view is to extend the integral to square-integrable martingales as integrators and then via localization to locally square-integrable martingales (which include continuous local martingales). The key objective is to establish metrics analogous to those used for Brownian motion, facilitating an Itô isometry-like result, which would permit the extension of the integral from elementary predictable integrands to a broader class. At the time of Itô's early research, the concept of local

martingales was not yet known, and insights into martingales and their decomposition were still undeveloped. Furthermore, it was necessary to overcome various measurability challenges, particularly for potential classes of integrands. Here's a brief outline of these developments.

In his book [44] and his treatment of Itô's results, Doob highlights three essential properties of Brownian motion that were crucial in Itô's construction of the stochastic integral. These properties are:

1. The increments of Brownian motion are orthogonal. This means that for a stochastic process X and times $s_1 < s_2 \leq t_1 < t_2$, the increments satisfy

$$\mathbf{E}[(X_{t_2} - X_{t_1})(X_{s_2} - X_{s_1})] = 0.$$

This orthogonality simplifies the construction of the stochastic integral for processes with such orthogonal increments.

2. For a Brownian motion B, both B_t and $(B_t^2 - t)_{t \geq 0}$ are martingales. This observation led Doob to propose extending the stochastic integral to a broader class of martingales.

3. Brownian motion has independent increments, a fundamental property used to justify its stochastic process structure.

Using these properties, Doob was able to extend the construction of the stochastic integral to processes with orthogonal increments. By exploiting the orthogonality of the increments, Doob could show that integrals of functions against processes with such increments can be expressed as sums over disjoint time intervals, much like in the Riemann-Stieltjes integral, which greatly simplified the extension.

The second key observation (the martingale properties of $(B_t^2 - t)$ and B_t) motivated Doob to develop a stochastic integral with respect to L^2-martingales. However, Doob's construction hinged on decomposing the square of an L^2-martingale, which was not fully proven at the time.

Doob also examined martingales M for which there exists a deterministic, increasing process F such that $M_t^2 - F(t)$ is again a martingale. He generalized the conditions on integrands from Equation (2) to

$$\int_0^\infty \mathbf{E}\left[H_t^2\right] \mathrm{d}F(t) < \infty, \tag{3}$$

which allowed him to define the stochastic integral for a broader class of integrands.

Doob extended Itô's isometry Theorem 8 to the following relationship:

$$\mathbf{E}\left[\left(\int_0^t H_s \mathrm{d}M_s\right)^2\right] = \mathbf{E}\left[\int_0^t H_s^2 \mathrm{d}F(s)\right],$$

which sufficed to define the stochastic integral similarly to Itô's construction. Moreover, he demonstrated that the martingale property is preserved under this form of stochastic integration [44].

However, a critical challenge was whether, for every martingale M, such a function F existed. To address this, Doob established the now-famous Doob decomposition theorem for submartingales in discrete time: if X_n is a discrete-time submartingale, it can be uniquely decomposed as $X_n = M_n + A_n$, where M is a martingale and A is a process with non-decreasing paths.

Since M^2 is a submartingale for any martingale M, Doob sought to establish a continuous-time analogue of this decomposition to further generalize the stochastic integral to more martingales. Contributions to this topic were made by Gilbert Agnew Hunt's and his papers on Markov processes [74–76]. Hunt's work provided key insights into continuous-time decompositions.

Then, in 1962, the French mathematician P. A. Meyer made significant advancements in the theory of stochastic integrals with two important papers [145,146]. These papers contributed in three key ways:

- Establishing the Doob decomposition for certain subclasses of continuous-time submartingales, under a condition known as "Class (D)" in honor of Doob [145]. He later even proved the uniqueness of this decomposition in [146].
- Introducing a localization procedure that paved the way for further developments and for defining and applying local martingales.
- Investigating the structure of L^2-martingales using Hilbert space methods, which laid the foundation for important future work [86,114,161].

In the closing remarks of his paper, Meyer also suggested further extensions but also challenges that may arise:

Define the norm of a well-adapted step process as this last integral. The preceding relation allows the extension of the stochastic integral by standard Hilbert-space methods to all well-adapted (jointly measurable) processes which have a finite norm and are equal to the limit of a convergent sequence (in the sense of the norm just defined) of well-adapted step processes. Among them are all right continuous well-adapted processes, but it seems hard to show (though it is certainly true) that the full class of well-adapted processes whose norm is finite has been attained by this procedure.

While Meyer's suggestion mentioned the possible extension of the stochastic integral, a mathematically rigorous definition was not provided. Furthermore, subsequent research by Leonard Ornstein and examples presented by Guy Johnson and Lester Helms, especially those involving three-dimensional Brownian motion, showed that not all submartingales belong to Class (D). This observation suggested that Doob's and Meyer's approach might not encompass all martingales as integrators [97].

The formal definition of the "class (D)" is as follows.

Definition 7. *A càdlàg supermartingale Z with $Z_0 = 0$ is of Class (D) if the collection $\{Z_T : T$ a finite valued stopping time$\}$ is uniformly integrable.*

With modern terminology, the key result of Doob was the following theorem.

Theorem 9 (Doob-Meyer Decomposition [172], Theorem III.8). *Let Z be a càdlàg supermartingale of class (D). Then, there exists a unique, increasing, predictable process A with $A_0 = 0$ and a uniformly integrable martingale M, such that*

$$Z = M - A.$$

In the foundational works listed above, Meyer established a decomposition such as the one described, wherein A is a *natural process*. The contemporary version of this formulation was shown by Doléans [39,40], who demonstrated that an increasing process is natural if it is predictable. Alternative proofs of Theorem 9 were presented by Bass [6], Rao [173] and Jakubowski [93].

Rao employed the $\sigma(L^1, L^\infty)$-topology and utilized the Dunford-Pettis compactness criterion to achieve the sought-after continuous time decomposition as a weak-L^1 limit from discrete approximations. To establish the predictability of A, one subsequently uses the theorem by Doléans-Dade.

Bass provided a more fundamental proof, which hinges on the dichotomy between predictable and entirely inaccessible stopping times.

Jakubowski, following a similar approach to Rao, recognized that the predictability of the process A could also be deduced by applying Komlos' Lemma [111].

A concise and modern proof can be found in the recent paper by [10], which combines ideas from Jakubowski [93] and Beiglböck et al. [8] to establish the continuous-time decomposition using a suitable Komlós-type lemma.

Even before a general proof of the Doob decomposition for martingales that are not of class (D) was discovered, Philippe Courrège contributed two critical findings to developing the stochastic

integral [30,31]. He focussed on processes whose paths are right continuous with left limits a.s. and processes whose jumps at stopping times are only allowed to occur at totally inaccessible stopping times; meaning jumps that happen "totally unannounced".

When decomposing a submartingale $M^2 = N + A$ into its Doob-Meyer decomposition, assuming the filtration is quasi-left continuous, the process A is guaranteed to have continuous sample paths. Specifically, in the L^2-isometry,

$$\mathbf{E}\left[\left(\int_0^t H_s \, dM_s\right)^2\right] = \mathbf{E}\left[\int_0^t H_s^2 \, dA_s\right],$$

the process A is guaranteed to have continuous paths under this decomposition. This assumption, as made by Courrège, simplifies the theory significantly by allowing the use of continuous paths for the finite variation process A.

Also, Courrège assumed that the integrands have left-continuous paths and investigated processes that were measurable with respect to the σ-algebra they generate on $\mathbb{R} \times \Omega$, referring to them as "very well adapted" processes. This laid the groundwork for what would later be formalized as the predictable σ-algebra, a concept that would be fully developed and utilized by P. A. Meyer to rigorously define the space of integrands for stochastic integration.

Courrège's assumption of continuous paths in A turned out to be highly beneficial in simplifying the theory and clarifying the structure of martingales.

3.4. The Stochastic Integral for square-integrable martingales

So far, the stochastic integral has been defined primarily for Brownian motion as the integrator. Courrège extended this concept by incorporating additional continuity assumptions on the underlying filtration and utilizing Doob's generalization of the Itô integral, along with the Doob-Meyer decomposition, which provided a "recipe" for defining the stochastic integral with respect to martingales of class (D). The next advancement was to extend the stochastic integral to square-integrable martingales, or L^2-martingales. This development took place in the 1960s, driven mainly by the works of Itô [86], Kunita and Watanabe [114], Motoo and Watanabe [161], who extended the class of possible integrators, and Meyer [149,150,151,152], who explored the potential class of integrands and relaxed some assumptions on the filtration.

The former three works ([86,114,161]) not only marked a significant milestone in the evolution of stochastic integration but also contributed notably to the theory of martingales and stochastic processes: Ito's paper on the transformation of Markov processes by multiplicative functionals laid the foundation for further studies in stochastic integrals and transformations. Motoo and Watanabe investigated the structure of square-integrable additive functionals of Markov processes and introduced a random inner product that facilitated the definition and analysis of stochastic integrals. Finally, Kunita and Watanabe extended the framework to more general settings involving càdlàg adapted processes and general filtrations. This provided a crucial foundation for the construction and characterization of stochastic integrals and greatly influenced modern martingale theory. Combined, their primary contributions to stochastic integration were:

- Establishing the notion of local martingales, providing a localized version of the Doob-Meyer decomposition and a localized version of the stochastic integral.
- Discovering the Hilbert space structure of the space of L^2-martingales.
- Defining a new approach to the stochastic integral, which treats it as an operator on L^2-martingales.

While the notion of local martingales was first introduced by Itô [86], we will first discuss the novel approach before extending the stochastic integral through localization. But we start with the definition of square-integrable martingales and their immediate consequences.

Definition 8. *Let M be a martingale. If* $\mathbf{E}[M_t^2] < \infty$ *for all t, we call M a square-integrable martingale. If* $\sup_t \mathbf{E}[M_t^2] = \mathbf{E}[M_\infty^2] < \infty$, *then we call M an L^2-martingale. The space of L^2-martingales is denoted by $\mathcal{M}^2$.*

Even though sometimes in the literature, the term "square-integrable martingale" is used for what we define as L^2-martingale, the condition for L^2-martingales on the indefinite interval $[0, \infty)$ is stricter than for square-integrable martingales. For instance, an L^2-martingale is uniformly integrable and belongs to class (D). However, this distinction has little impact on the development of the stochastic integral, as every locally square-integrable martingale is also a local L^2-martingale. Nevertheless, L^2-martingales are stable under stochastic integration, a property not generally valid for square-integrable martingales.

A direct consequence of the Doob-Meyer decomposition theorem is the existence of the sharp-bracket process (also known as conditional quadratic variation or predictable quadratic variation). For an L^2-martingale X, the process $-X^2$ is a supermartingale. By the Doob-Meyer theorem, there exist processes M and A such that $-X^2 = -M + A$, where A is an increasing, predictable process and M is a local martingale. The process A is then defined as the sharp-bracket process.

Theorem 10 ([149], Théorème 3). *Let X be an L^2-martingale. There exists a unique, increasing, predictable process $\langle X, X \rangle$ with $\langle X, X \rangle_0 = 0$, such that $X^2 - \langle X, X \rangle$ is a martingale. The process $\langle X, X \rangle$ is called the sharp-bracket process or predictable quadratic variation. If Y is another L^2-martingale, the conditional quadratic covariation of X and Y is defined as*

$$\langle X, Y \rangle = \frac{1}{2}(\langle X + Y, X + Y \rangle - \langle X, X \rangle - \langle Y, Y \rangle).$$

Remark 4. *The predictable covariation process $\langle X, Y \rangle$ can be defined even if the predictable quadratic variation $\langle X, X \rangle$ does not exist. This is because $\langle X, Y \rangle$ can be defined as the compensator of the optional covariation process $[X, Y]$, which will be covered later. Specifically, $\langle X, Y \rangle$ is the unique predictable process such that $[X, Y] - \langle X, Y \rangle$ is a local martingale.*

Remark 5. *From the definition, it is easy to see that for $X, Y \in \mathcal{M}^2$, we have*

$$\mathbf{E}[X_\infty Y_\infty] = \mathbf{E}[\langle X, Y \rangle_\infty].$$

The existence of the sharp-bracket process allowed to define an inner product on the space of stochastic integrals, which was used for the new approach. The approach by Itô, Doob and Courrège is to define an L^2-isometry to extend the stochastic integral from a basic class of integrands to a wider class so the stochastic integral $L = H \bullet M$ is usually defined as a unique continuous extension, often expressed as a limit $L = H \bullet M = \lim_{n \to \infty} H^n \bullet M$ for an appropriate sequence H^n, with respect to suitable metrics. The new approach by Kunita and Watanabe [114], Motoo and Watanabe [161] and later Meyer [149,150,151,152] is to see the stochastic integral as an operator on L^2-martingales, exploiting the Hilbert space structure of L^2, which was given by this newly defined inner product (see Theorem 13). They defined orthogonality and the stochastic integral operator such that orthogonality is preserved under stochastic integration. That is, the stochastic integral $L = H \bullet M$ is the unique square-integrable martingale L that satisfies

$$\mathbf{E}[L_\infty N_\infty] = \mathbf{E}\left[\int_0^\infty H_s \, d\langle M, N \rangle_s\right] \tag{4}$$

for every square-integrable martingale N (see Theorem 13).

This approach was later extended to more general stochastic integrals [41] and even the vector-valued stochastic integral [29]. However, to maintain consistency, we present the definition of the stochastic integral for square-integrable martingales in a way that is aligned with Itô's original approach. The equivalence of these approaches is straightforward to demonstrate.

Remark 6. *Similar to Itô's initial approach, Motoo and Watanabe first defined the stochastic integral for bounded, adapted càdlàg processes as integrands, then extended this class of processes, ending up with optional processes as integrands. This can lead to contradictions when dealing with non-continuous processes with finite variation. Hence, optional processes for integrands became unpopular. We, therefore, consider only predictable processes as integrands, an adjustment later made by Meyer.*

Definition 9. *Let $M \in \mathcal{M}^2$ and $H \in \mathcal{E}$ be represented as*

$$H_t = H_0 \mathbb{1}_{\{0\}} + \sum_{i=1}^{n} H_i \mathbb{1}_{(t_i, t_{i+1}]}(t).$$

We define the stochastic integral $J^M(H)_t$ by

$$J_t^M(H) := H \bullet M_t := \int_0^t H \, dM_s = H_0 M_0 + \sum_{i=1}^{n} H_i \left(M_{t_{i+1} \wedge t} - M_{t_i \wedge t} \right)$$

and

$$J_\infty^M(H) = H_0 M_0 + \sum_{i=1}^{n} H_i \left(M_{t_{i+1}} - M_{t_i} \right).$$

Similar to the other approaches, the definition of the stochastic integral for square-integrable martingales is straight forward. The actual work is finding suitable metrics, demonstrating an Itô-like isometry and showing that the space of elementary predictable processes is dense in a suitable space of integrands.

Definition 10. *Let $M \in \mathcal{M}^2$.*

(a)   For $H \in \mathcal{P}$, we define

$$\|H\|_M^2 := \mathbf{E}\left[\int_0^\infty H_s^2 \mathrm{d}\langle M, M \rangle_s \right]. \tag{5}$$

(b)   By $\mathcal{E}_M$, we denote the space of elementary predictable processes endowed with the seminorm $\|\cdot\|_M$.
(c)   We define

$$L^2(M) := \{ H \in \mathcal{P} \, ; \, \|H\|_M^2 < \infty \}.$$

It is indeed easy to check that $\mathcal{E} \subset L^2(M)$.

Lemma 3 ([29], Lemma 12.1.4). *By endowing $\mathcal{M}^2$ with the seminorm defined by $\|M\|_{\mathcal{M}^2} := \left(\mathbf{E}\left[M_\infty^2 \right] \right)^{\frac{1}{2}} = \left(\mathbf{E}[\langle M, M \rangle_\infty] \right)^{\frac{1}{2}}$, the mapping*

$$J_\infty^M : \begin{cases} \mathcal{E}_M & \to & \mathcal{M}^2 \\ H & \mapsto & H \bullet M \end{cases}$$

becomes an isometry.

It is easy to show that such an integral with respect to an L^2-martingale is again an L^2-martingale, and thus, the integral operator maps elementary predictable functions to the Hilbert space $\mathcal{M}^2$

From the uniqueness of the sharp-bracket, one quickly deduces

$$\langle H \bullet M, H \bullet M \rangle = H^2 \bullet \langle M, M \rangle$$

and thus,

$$\|H \bullet M\|_{\mathcal{M}^2} = \mathbf{E}[\langle H \bullet M, H \bullet M \rangle_\infty] = \mathbf{E}\left[\int_0^\infty H_s^2 \mathrm{d}\langle M, M \rangle_s\right],$$

is the required isometry property. The existence and uniqueness of the extension are then apparent.

Since $L^2(M)$ is the space L^2 of the **P**-measure associated with the integrable increasing process $\langle M, M \rangle$ over the predictable σ-field for the proof, the related Hausdorff space is a Hilbert space. We want to extend the integral to the closure of $\mathcal{E}_M$ in $L^2(M)$. Therefore, we examine the closure $\overline{\mathscr{U}}$ of $\mathscr{U} := L^2(M) \cap \mathcal{E}$. But according to the monotone class theorem $\overline{\mathscr{U}}$ contains all bounded predictable processes and therefore $\overline{\mathscr{U}} = L^2(M)$ and hence, $\mathcal{E}$ is dense in $L^2(X)$:

Lemma 4 ([29], Lemma 12.1.8).

Let $\overline{\mathcal{E}_M}$ denote the closure of $\mathcal{E}_M$ in $L^2(M)$ with respect to $\|\cdot\|_{\mathcal{E},M}$. Then we have

$$\overline{\mathcal{E}_M} = L^2(M).$$

Now, it is easy to define the stochastic integral similar to how it was done for the Itô integral.

Definition 11. *The continuous linear mapping $\overline{J_\infty^M}$ from $L^2(M)$ to $\mathcal{M}^2$, obtained as the extension of J_∞^M from elementary predictable processes to $\mathcal{M}^2$ by continuity, is called the stochastic integral. For each $t \geq 0$, by setting*

$$H \bullet M_t = H \mathbb{1}_{[0,t]} \bullet M_\infty,$$

we obtain a well-defined stochastic process, which we call the stochastic integral of H with respect to M.

The Itô isometry has now the following shape:

Theorem 11 ([29], Lemma 12.1.4). *Let J^M be the integral operator as previously defined. When $\mathscr{L}^2$ is equipped with the extension of the norm $\|\cdot\|_{\mathcal{E}}$ to $\|\cdot\|_{\mathscr{L}^2}$, then for all $H \in \mathscr{L}^2$, the operator $J^B(H)$ constitutes an isometry from $\mathscr{L}^2$ into $\mathcal{M}^2$. Specifically, for any $H \in \mathscr{L}^2$, it holds that*

$$\|H\|_{\mathscr{L}^2}^2 = \mathbf{E}\left[\int_0^\infty H(s,\omega)^2 \mathrm{d}s\right] = \mathbf{E}\left[\left(J_\infty^B(H)\right)^2\right] = \|J_\infty^B(H)\|_{\mathcal{M}^2}^2.$$

Kunita and Watanabe [114], Motoo and Watanabe [161] treated the sharp-bracket process as a pseudo-inner product and, therefore, were able to utilize Hilbert-space methods on the space of L^2-martingales. While the properties of metric spaces or Banach spaces would suffice to define the stochastic integral, some properties of the space of stochastic integrals, such as the martingale representation property, could not be demonstrated without the tools that are exclusive to Hilbert spaces.

By considering the sharp-bracket process as a pseudo inner product, it was only logical to investigate a possible variant of the Cauchy-Schwarz inequality. Such a variant is given by the Kunita-Wantanabe inequality, which is due to Courrège and Meyer [157] but based on a result by Kunita and Watanabe [114].

Theorem 12 (Kunita-Watanabe Inequality, [157] Théorème 21). *If M and N are two square-integrable martingales and H and K are two measurable processes, then we have almost surely:*

$$\int_0^T |H_s| \, |\mathrm{d}\langle M, N \rangle_s| \leq \left(\int_0^T H_s^2 \, \mathrm{d}\langle M, M \rangle_s\right)^{1/2} \left(\int_0^T K_s^2 \, \mathrm{d}\langle N, N \rangle_s\right)^{1/2}.$$

The Kunita-Watanabe inequalities provide crucial insights into the properties of stochastic integrals and are crucial in demonstrating the following theorem, which is essentially the definition of the stochastic integral as it was introduced by Kunita and Watanabe.

Theorem 13 (Kunita and Watanabe, [178] Theorem 28.1). *Let $M \in \mathcal{M}^2$, H be a predictable process such that $H \in L^2(M)$, and N be another square-integrable martingale. Then the stochastic integral $H \bullet M$ is the unique element in $\mathcal{M}^2$ with the property that for every $N \in \mathcal{M}^2$, we have*

$$\mathbf{E}[(H \bullet M)_\infty N_\infty] = \mathbf{E}\left[\int_0^\infty H_t \, \mathrm{d}\langle M, N \rangle_t\right].$$

With the help of Theorem 12, it is straightforward to show that the novel approach by Kunita and Watanabe [114], Motoo and Watanabe [161] is equivalent to the one outlined above. The mapping

$$H \mapsto \mathbf{E}\left[(H \bullet M)_\infty - \int_0^\infty H_s \, \mathrm{d}M_s \, N_s\right]$$

defines a linear functional on $L^2(M)$. By applying the Kunita-Watanabe inequality, it is clear that this function is continuous and vanishes on simple functions.

If the stochastic integral is defined as the unique continuous extension, it follows that the function is zero for all elements in $L^2(M)$.

Furthermore, the right-hand side of the equation in Theorem 13 defines a continuous linear functional in N (given H and M), while the left-hand side represents the inner product of $H \bullet M$ and N in the Hilbert space L^2. By the Riesz-Fréchet representation theorem in Hilbert spaces, for given H and M, there exists a unique element $H \bullet M \in L^2$ that corresponds to this linear functional.

3.5. Localization and Quadratic Variation

Besides the discovery of the sharp-bracket process as a pseudo-inner product, one of the key results of Itô and S. Watanabe was their establishment of local martingales in their studies of multiplicative functionals of Markov processes [86]. With this, it was possible to define the square-bracket process for local L^2-martingales (see [150, page 97 ff] for more details).

Let M be a local L^2-martingale, and let (T^n) be a sequence of stopping times such that M^{T^n} is an L^2-martingale for all n. With this, the process

$$\left(M^{T^{n+1}}\right)^2 - \left\langle M^{T^{n+1}}, M^{T^{n+1}} \right\rangle$$

is by Theorem 10 a uniformly integrable martingale, and we see that the process

$$\left(M_t^{T^n}\right)^2 - \left\langle M^{T^{n+1}}, M^{T^{n+1}} \right\rangle_{t \wedge T^n}$$

is also a martingale. It then follows from the uniqueness of the Doob decomposition theorem that

$$\langle M^{T^n}, M^{T^n} \rangle_t = \langle M^{T^{n+1}}, M^{T^{n+1}} \rangle_{T^n \wedge t}.$$

As T^n tend to infinity for $n \to \infty$, we establish that there exists an increasing process $\langle M, M \rangle$, such that $M^2 - \langle M, M \rangle$ is a local martingale.

The statement can be formalized as follows:

Theorem 14 ([71], Lemma 7.28). *Let M be a locally square-integrable martingale. Then there exists a unique predictable and increasing process $\langle M, M \rangle$ such that $M^2 - \langle M, M \rangle$ is a local martingale with $M_0 = 0$. This process is again called the sharp-bracket process or predictable quadratic variation.*

This statement allows to extend the definition of $\langle M, N \rangle$ to locally square-integrable martingales M, N by setting

$$\langle M, N \rangle = \frac{1}{2}(\langle M + N \rangle - \langle M, M \rangle - \langle N, N \rangle).$$

With similar deductions, it is also possible to prove a version of Doob's decomposition theorem in continuous time without requiring the use of class (D).

Theorem 15 ([38], Theorem VII.12). *Let X be a càdlàg supermartingale. Then X has a unique decomposition:*

$$X = X_0 + M - A$$

where M is a local martingale, A is a predictable, increasing process, and $M_0 = A_0 = 0$.

Inspired by Kunita and Watanabe's work, Paul-André Meyer contributed significantly by writing four seminal papers [149–152]. These works extended the stochastic integral to predictable integrands and locally square-integrable martingales.

In the original publication by Itô [86], the definition of the stochastic integral via Equation (4) also applied to local L^2-martingales. Since the space of local L^2-martingales is neither a Hilbert space nor a Banach space, the classical extension approach does not naturally carry over to local L^2-martingales. Therefore, for the transition from L^2-martingales to locally square-integrable martingales, Meyer [150] applied a similar approach as was done before for the sharp-bracket process via coalescing stopped processes into one process.

Let M be a locally square-integrable martingale, and thus also a local L^2-martingale, and let $L^2(M)$ be defined as on Page 17. By assumption, there exists a sequence $(T^n)_{n \in \mathbb{N}}$ of stopping times such that $T^n \to \infty$ as $n \to \infty$, and M^{T^n} is an L^2-martingale for all n. If $H \in L^2(M^{T^n})$ for all n, we have the relation

$$\left(H \bullet M^{T^{n+1}} \right)^{T^n} = H \bullet M^{T^n}.$$

Thus, the processes $H \bullet M^{T^n}$ can be coalesced into a single process, denoted $H \bullet M$.

This approach was followed by Dellacherie and Meyer [37] and later by Protter [172] in defining the general stochastic integral. Of course, for a formal derivation of the stochastic integral for local square-integrable martingales, it still needs to be shown that the stochastic integral is independent of the choice of stopping times.

These ideas are formalized as follows. Let $L^2_{\text{loc}}(M)$ be the space of processes H such that the increasing process

$$\int_0^t H_s^2 \mathrm{d}\langle M, M \rangle_s$$

is locally integrable. It is easy to see that $H \in L^2_{\text{loc}}(M)$ is equivalent to the existence of an increasing sequence of stopping times T^n, such that $H \in L(M^{T^n})$ for all n.

Theorem 16 (Definition of the Stochastic Integral, [150] p. 98). *Let $M \in \mathcal{M}^2_{\text{loc}}$ and $H \in L^2_{\text{loc}}(M)$. Then, there exists a uniquely defined process L with*

$$L = H \bullet M \text{ on } [\![0, T^n]\!]$$

for each sequence of stopping times T^n such that $M^{T^n} \in \mathcal{M}^2$ and $H \in L^2(M^{T^n})$. We note $H \bullet M := L$ and say that $H \bullet M$ is the stochastic integral of H with respect to M.

In the case where $M \in \mathcal{M}^2$, this definition coincides with that of the stochastic integrals defined in the last section and in the case where M is a Brownian motion, the stochastic integral coincides with the Itô integral.

The approach described here, which uses the sharp-bracket process to define a pseudo-scalar product and, subsequently, a metric, follows Itô's method to extend the integral for square-integrable martingales from elementary predictable processes to general predictable processes. While this process and the localization procedure mentioned earlier are straightforward, the proofs for the properties of the stochastic integral defined by this method are challenging and dense.

An alternative is offered by the approach towards stochastic integration suggested by Motoo and Kunita, as discussed in [114,161] (see Theorems 13 and 17). This method allows the direct definition of the stochastic integral for localized processes, which simplifies the application of Hilbert space methods.

Theorem 17 ([114], Theorem 2.1). *For $M \in \mathcal{M}^2_{\text{loc}}$ and $H \in L^2(M)_{\text{loc}}$, the integral $H \bullet M$ is uniquely characterized by the relation*

$$\langle H \bullet M, N \rangle = H \bullet \langle M, N \rangle \quad \text{for all } N \in \mathcal{M}^2_{\text{loc}}.$$

Besides the definition of the stochastic integral for locally square-integrable martingales, one of the key results of Meyer's work [149–152] was the establishment of the square-bracket process. Since the sharp-bracket process originally existed only for square-integrable martingales and was heavily dependent on the chosen probability measure, it was rather impractical and not applicable in many situations. Meyer extended the scalar product $\langle X, Y \rangle$ by introducing a pseudo-scalar product, denoted by $[X, Y]$, which exists for all martingales, including local martingales. Unlike $\langle X, Y \rangle$, which depends on the probability measure, $[X, Y]$ is invariant under changes of equivalent probability measures. This extension was crucial for the development of semimartingale theory and the further generalization of the stochastic integral.

Another major breakthrough was Meyer's realization of the importance of the predictable σ-algebra. While Courrège had touched upon related ideas, Meyer demonstrated that for martingales with finite variation paths (including those with jumps), the stochastic integral should coincide with pathwise Lebesgue-Stieltjes integration. But Meyer showed that this is, in general, only true if the integrand is a predictable process. He further analyzed the jumps of the stochastic integral, noting that for predictable integrands, the stochastic integral has the same jump behavior as the Lebesgue-Stieltjes integral. This finding laid the groundwork for semimartingale theory, which was developed in subsequent years. For this reason, the stochastic integral for optional instead of predictable integrands, which leads to the compensated stochastic integral, received less attention compared to the stochastic integral with predictable integrands. Nevertheless, works such as [20] and the textbooks by Dellacherie and Meyer [37] and He et al. [71] provide a thorough discussion of this concept.

Furthermore, Meyer managed to derive a general change of variables formula without relying on Lévy systems, which was a significant generalization and was later extended to Markov processes in his subsequent papers.

The definition of the square-bracket process relies on the decomposition of martingales into a continuous and a discontinuous part:

Theorem 18 ([150], Proposition 3). *Let M be a local martingale. Then M can be uniquely decomposed into a continuous local martingale part M^c and a purely discontinuous local martingale part M^d, such that*

$$M_t = M^c_t + M^d_t.$$

Here, M^c is a continuous local martingale, and M^d is a purely discontinuous local martingale satisfying $\Delta M^c = 0$ and $\Delta M^d = \Delta M$.

As any continuous local martingale is necessarily locally square-integrable, we conclude with Theorem 14 that $\langle M^c, M^c \rangle$ exists for any local martingale M.

Definition 12. *Let M be a local martingale and let $M = M^c + M^d$ be its canonical decomposition, where M^c and M^d are its continuous and purely discontinuous local martingale components, respectively. The quadratic*

variation, also called the square-bracket process $[M]$, of M is defined as the unique càdlàg, adapted, increasing process starting at zero such that

$$[M]_t = [M^c]_t + \sum_{0 \leq s \leq t} (\Delta M_s)^2 = \langle M^c \rangle_t + \sum_{0 \leq s \leq t} (\Delta M_s)^2.$$

For local martingales M and N, the quadratic covariation is defined as:

$$[M, N]_t = \frac{1}{2}([M + N]_t - [M]_t - [N]_t) = \langle M^c, N^c \rangle_t + \sum_{0 \leq s \leq t} \Delta M_s \Delta N_s.$$

Theorem 19 ([71], Theorem 6.28). *Let $M, N \in \mathcal{M}^2$. Then $[M, N]$ is the unique finite variation process, such that $MN - [M, N]$ is a local martingale and*

$$\Delta[M, N] = \Delta M \Delta N.$$

In [114, Theorem 1.3], it is shown that for any continuous martingale, the sharp-bracket process can be approximated by appropriate sums that remind us of the variance of a stochastic process. This result can be generalized to the quadratic covariation process.

Theorem 20 ([186], Theorem 5.8). *Let M and N be càdlàg local martingales. For any partition $P = \{0 = t_0 < t_1 < \cdots < t_n = t\}$ of the interval $[0, t]$, we have:*

$$[M, N]_t = \lim_{\|P\| \to 0} \sum_{k=0}^{n-1} (M_{t_{k+1}} - M_{t_k})(N_{t_{k+1}} - N_{t_k}),$$

where $\|P\| = \max_{0 \leq k < n}(t_{k+1} - t_k)$ is the mesh of the partition. The limit is taken as the mesh of the partition tends to zero.

The following theorem gives a partial integration formula for stochastic integrals. It allows for an alternative definition of the quadratic variation process $[M, N]$ via stochastic integrals. The alternative definition was, for example, used in the ansatz by Protter [172].

Theorem 21 ([150], Théorème 2). *For local martingales M, N, the product NM can be expressed by the following integration formula:*

$$NM = N_- \bullet M + M_- \bullet N + [M, N],$$

where $N_- \bullet M$ and $M_- \bullet N$ represent the integrals of N with respect to M and M with respect to N using their left-continuous versions N_- and M_- respectively.

Remark 7. *The assumption of square-integrability for the integrator significantly simplifies the definition of the stochastic integral. However, the additional assumption of continuity, as is often made in textbooks and university lectures, while simplifying some aspects of the theory, does not simplify the definition itself. What it does simplify are the proofs of various properties and prerequisites. For example, the existence of quadratic variation is much easier to establish for continuous martingales compared to general martingales. Additionally, the proof of Itô's formula becomes a matter of straightforward calculation when dealing with continuous martingales. This is why many textbooks restrict the definition of the stochastic integral to continuous martingales, even though the theory extends more generally to locally square-integrable martingales.*

Remark 8. *When M is a continuous martingale, the sharp-bracket process $\langle M, M \rangle$ coincides with the square bracket $[M, M]$, and the Itô isometry holds for right-continuous adapted integrands. Since simple right-continuous processes generate the optional σ-algebra, they are dense in the corresponding L^2 space. This allows us to define the stochastic integral for optional integrands in a manner analogous to its definition for predictable integrands. Moreover, simple processes are dense in $L^2(\Sigma_T, \mathrm{d}\langle M, M \rangle \times \mathrm{d}\mathbf{P})$, where Σ_T is the progressive*

σ-algebra. Consequently, the stochastic integral with respect to a continuous martingale is well-defined for any square-integrable progressive integrand, making the restriction to predictable integrands unnecessary. For further details, refer to Shreve and Karatzas [186].

Remark 9. *Even though the predictable quadratic variation and the optional quadratic variation generally coincide only for continuous square-integrable martingales, the seminorms they naturally induce are equivalent (see [29, Theorem 12.2.1]). This equivalence allows the stochastic integral for square-integrable martingales to be defined using the optional quadratic variation, analogous to its construction with the predictable quadratic variation.*

3.6. The Stochastic Integral for Locally Bounded Predictable Integrands

Up to 1970, stochastic integration was primarily associated with Markov processes, and the above-mentioned work by H. Kunita, S. Watanabe, and P. A. Meyer used the crucial assumption that the underlying filtration of σ-algebras was quasi-left continuous. In a paper by Doléans-Dade and Meyer [41], this assumption was eliminated, transforming the theory into a pure martingale theory and severing its connection to Markov processes. Here, it was shown that the assumption of the filtration being left-continuous can be dropped if the integrands are predictable, which leads to the same results. It was also in this publication that the term "semimartingale" was introduced.

Nowadays, there are different definitions for semimartingales. The definition below is the one that was used in the original papers. As we will examine different (equivalent) definitions of semimartingales, we will, for clarity, use the term *classical semimartingale* for the historical and original definition.

Definition 13. *A process X is a classical semimartingale if it has a decomposition of the form*

$$X = M + A,$$

where M is a local martingale, and A is a càdlàg adapted process of almost surely finite variation.

By definition, it is clear that local martingales and processes with finite variations are both classical semimartingales. Additionally, any right-continuous supermartingale or submartingale is also a classical semimartingale, as, by the Doob-Meyer decomposition, a right-continuous supermartingale X can be decomposed as $X = M - A$, where M is a martingale and A is an increasing process and therefore of finite variation.

In their paper, Doléans-Dade and Meyer [41] defined the stochastic integral with respect to continuous local martingales and extended this integral to general classical semimartingales for locally bounded, predictable integrands. The stochastic integral for continuous local martingales had already been defined earlier, as any continuous local martingale is a local L^2-martingale. However, Doléans-Dade and Meyer used a similar but not completely identical approach to this stochastic integral again. They started by defining the stochastic integral for square-integrable martingales again and chose the same approach as Kunita and Watanabe did before. In particular, they used the definition as is implicitly given by Theorem 17 but used the quadratic variation as was done by Meyer [150].

The key idea of the approach by Doléans-Dade and Meyer [41] for the definition of the stochastic integral for local martingales in comparison to earlier approaches was to decompose the local martingale into appropriate processes for which the stochastic integral can be easily defined. This approach subsequently led to significant advancements in the theory of stochastic integration.

The main decomposition result of Doléans-Dade and Meyer [41] is the following:

Proposition 1 ([41], Proposition 4). *Let M be a local martingale. Then there exists a sequence of stopping times $(T^n)_{n \in \mathbb{N}}$ increasing to infinity and two processes N and V, such that:*

- $M = N + V$

- N^{T^n} *is an L^p-martingale for all n and all p with $1 \le p < \infty$.*
- *V is a process of integrable variation.*

Now, the definition of the stochastic integral for locally bounded predictable processes H and classical semimartingales proceeds as follows.

Definition 14 (Stochastic Integral with respect to classical semimartingales for locally bounded integrands)**.** *Consider a classical semimartingale X with the decomposition*

$$X = M + A,$$

where M is a local martingale and A is a process of finite variation. Let H be a predictable and locally bounded process. We can choose a sequence of finite stopping times T^n, increasing to infinity, such that the following holds:

- *For each n, $M^{T^n} = N^{T^n} + V^{T^n}$, where N^{T^n} is an L^p-martingale and V^{T^n} is a process of integrable variation.*
- *For all n, the process H^{T^n} is bounded.*

Then, we define the stochastic integral as well as the two Stieltjes integrals $H \bullet V$ and $H \bullet A$ by setting

$$H \bullet X = H \bullet N^{T^n} + H \bullet V^{T^n} + H \bullet A^{T^n} \quad on \; [\![0, T^n]\!].$$

It remains to show that $H \bullet X$ is well-defined and does not depend on the selection of stopping times $(T^n)_{n \in \mathbb{N}}$ or on the decomposition. For this, it is essential that the stochastic integral with respect to a local martingale with finite variation (which is a trivial or deterministic process) coincides with the Lebesgue-Stieltjes integral. This only holds for predictable integrands H, and the existence of the Stieltjes integral is only guaranteed for bounded processes. Hence, this definition of the stochastic integral cannot be easily extended to unbounded integrands or optional integrands.

Doléans-Dade and Meyer [41] also demonstrated the following crucial result, which allows for the unique characterization of the stochastic integral with local martingale integrators.

Proposition 2 ([41], Proposition 5)**.** *Let M be a local martingale and H a locally bounded, predictable process. Then $H \bullet M$ is the unique local martingale with $M_0 = 0$ such that*

$$[H \bullet M, N] = H \bullet [M, N]$$

holds for all local martingales N.

The proof of this characterization heavily relies on the fact that $H \bullet M$ is a local martingale for a local martingale M. Since this does not generally hold for unbounded H, a different strategy would be necessary to characterize the general stochastic integral.

3.7. The General Real-Valued Stochastic Integral

The stochastic integral was further generalized to accommodate unbounded predictable integrands by Jacod [88,89] and Chou [26], resulting in the comprehensive theory of the stochastic integral in one dimension.

Jacod's approach is based on the characterization of semimartingale jumps, whereas Chou [26] proposed an alternative methodology, yielding an equivalent definition of the stochastic integral. The space of integrands used by Jacod represents the broadest possible space with a meaningful definition for semimartingale integrators; the stochastic integral cannot be extended to a more general class of one-dimensional integrands without sacrificing crucial properties, such as the continuity of the integral operator.

In general, for the definition of the stochastic integral, classical semimartingales have the disadvantage that they may have non-integrable jumps. Therefore, the first attempts to define the stochastic integral for semimartingale integrators were complicated and technical.

One key ingredient is the Fundamental Theorem of Local Martingales, which states that any local martingale can be represented as the sum of a process of finite variation and a local martingale with bounded jumps. It was proven by Jia-an Yan (but appeared in [155]) and Doléans-Dade [42] independently.

Theorem 22 (Fundamental Theorem of Local Martingales [155]). *Let M be a local martingale and $\alpha > 0$. There exist local martingales U, V such that U is an FV-process, the jumps of V are bounded by α, and $M = U + V$.*

Therefore, when defining the stochastic integral for a classical semimartingale $X = M + A$, without loss of generality, M can be assumed to have bounded jumps and, therefore, be locally square-integrable. However, the integral with respect to a locally square-integrable martingale had already been defined. Hence, since the integration operator is linear, for the definition of the stochastic integral with respect to a classical semimartingale, only the stochastic integral with respect to a predictable finite variation process A remains to be defined, but this can easily be defined by pathwise Lebesgue-Stieltjes integration. Nevertheless, it remains to be shown that the stochastic integral for a classical semimartingale $X = M + A$ does not depend on the decomposition $M + A$. This turns out to be more sophisticated than it seems at first glance. In particular, it is not even true for non-predictable integrands. This is one of the reasons why we restrict ourselves to predictable processes as potential integrands.

We start by giving a short overview of the construction of the stochastic integral with only slight modifications in comparison to how it was done by Jacod [88,89].

So far, for some cases, the stochastic integral $H \bullet X$ has already been defined for predictable processes H that are not locally bounded. Firstly, we have the case where X has finite variation, provided that

$$\int_0^t |H_s| \mathrm{d}|X|_s < \infty \text{ for } t < \infty, \tag{6}$$

and secondly, when X is a local martingale, under the assumption that

$$(H \bullet [X, X])_t \text{ is locally integrable.} \tag{7}$$

These two scenarios are referred to as the usual cases of stochastic integrals for predictable processes that are not locally bounded. Therefore, the canonical way to define the general stochastic integral would be to say that H is integrable with respect to X if there exists a decomposition $X = M + A$ such that $H \bullet M$ and $H \bullet A$ exist in the usual sense, and by setting $H \bullet X = H \bullet M + H \bullet A$. This method was adopted by Jacod. His further strategy was to characterize and control the jumps of the classical semimartingale X, which ultimately led to the concept of special semimartingales.

Definition 15. *A classical semimartingale X is called a special semimartingale if there exist a local martingale M and a predictable FV-process A such that*

$$X = M + A.$$

Remark 10. *There are numerous equivalent definitions of special semimartingales in the literature. For example, a classical semimartingale is called special if there exists a decomposition into a local martingale and a process of locally integrable variation. An overview of equivalent criteria can, for example, be found in [29, Theorem 11.6.10].*

Special semimartingales were initially introduced by Meyer [157] and Yoeurp [200]. It is easy to see that the decomposition of a special semimartingale is unique (up to the constant X_0), as any predictable local martingale is constant.

The following theorem is crucial for the definition of the stochastic integral.

Theorem 23 ([88], Théorème 6). *Let X be a classical semimartingale and $H \in \mathcal{P}$. Then*

$$Y_t = \sum_{0 \leq s \leq t} \Delta X_s \mathbb{1}_{\{|\Delta X_s| \geq 1\} \cup \{|H_s \Delta X_s| \geq 1\}} \tag{8}$$

is of finite variation, and $Z = X - Y$ is a special semimartingale.

Now, the stochastic integral can be defined as follows:

Definition 16. *Let X be a classical semimartingale and $X = Y + Z$ be a decomposition as given by Theorem 23, and let the canonical decomposition of Z be given as*

$$Z = M + A$$

with a local martingale M and a predictable FV-process A. If $(H^2 \bullet [M])^{\frac{1}{2}}$ is a process of integrable variation, $\int_0^t |H_s||\mathrm{d}Y_s| < \infty$, and $\int_0^t |H_s||\mathrm{d}A_s| < \infty$, H is said to be X-integrable, denoted $H \in L(X)$, and $H \bullet Y$ and $H \bullet A$ exist in the Stieltjes sense while $H \bullet M$ exists in the local martingale sense. We define the stochastic integral $H \bullet X$ as

$$H \bullet X := H \bullet Y + H \bullet Z = H \bullet Y + H \bullet M + H \bullet A. \tag{9}$$

The stochastic integral is well-defined.

Remark 11. *It is straightforward to see that $L(X)$ forms a vector space and that the corresponding stochastic integral is additive. Let H and K be two elements of $L(X)$; in Equation (9), we take $J = \{s : |\Delta X_s| \vee |H_s||\Delta X_s| \vee |K_s||\Delta X_s| > 1\}$; Theorem 23 then provides a decomposition $X = N + B$ of X, such that $H \bullet N$ and $K \bullet N$ exist in the local martingale sense and $H \bullet B$ and $K \bullet B$ in the Stieltjes sense, ensuring the additivity of the usual stochastic integral.*

Remark 12. *With the definition given above, it is relatively easy to establish a dominated convergence theorem for stochastic integrals. Let H_n be predictable processes that are bounded above in absolute value by a predictable process K and converge pointwise to a process H. If K belongs to $L(X)$, so do the H_n and H, and $H_n \bullet X$ converge to $H \bullet X$ in S. To demonstrate this, we again apply Theorem 23 by taking $J = \{s : |\Delta X_s| \vee |H_s||\Delta X_s| > 1\}$ in Equation (9), reducing it to a dominated convergence theorem for usual stochastic integrals. In the case of the Topological Stochastic Integral, which is not based on the decomposition of classical semimartingales, this is much harder to prove.*

Jacod laid the foundation for general stochastic integration [88,89]. In 1980, Chou, Meyer, and Stricker introduced an alternative approach that leverages the topology of classical semimartingales along with an explicit decomposition result. Jacod later applied a similar method to define vector stochastic integrals [90], and Mémin examined the connections between these integrals and the topology of classical semimartingales [138]. Given the parallels with the topological stochastic integration discussed in Section 9, we briefly outline these approaches here.

The ansatz by Chou, Meyer, and Stricker is based on the work of Jacod [90] and Memin [138], as well as Meyer [155]. One disadvantage of the original work was that it does not simplify certain aspects: for example, it is not immediately clear that integrability and stochastic integrals are invariant under a change of measure, nor even that the integral is a linear operation in H or that the existence of the stochastic integral does not depend on the decomposition of the classical semimartingale X.

Therefore, Chou et al. [26] introduced another definition, which the authors called *the sophisticated definition*.

Let $X_t^* := \sup_{s \le t} X_s$, then the map

$$X \mapsto \|X\|_{\mathrm{ucp}} := \sum_{n=1}^{\infty} 2^{-n} \mathbf{E}[1 \wedge X_n^*]$$

defines the topology of ucp convergence. We then define the *Émery norm* of a classical semimartingale X as

$$\|X\|_{\mathcal{S}} := \sup\{\|H \bullet X\|_{\mathrm{ucp}} : H \in \mathcal{E}, \|H\| \le 1\}.$$

This topology is sometimes also called *semimartingale topology* (or *Émery topology*). We denote $X^n \to X$ in $\mathcal{S}$ when $\|X^n - X\|_{\mathcal{S}} \to 0$.

Here, the supremum is taken over all elementary predictable processes. In the original publication, it was taken over all simple predictable processes (see Definition 20), and this is also how we are going to define it in the section about the Topological Stochastic Integral. However, it is easy to see that these definitions are equivalent. In fact, one could also choose all predictable processes that are bounded by 1.

Now, let $\mathcal{S}$ be the space of all classical semimartingales equipped with the Émery topology. Then $\mathcal{S}$ is a complete metric non-locally convex topological vector space. If X^n converges to X in $\mathcal{S}$, we have for any finite t

$$\lim_n (X^n - X)_t^* = 0 \text{ in probability.}$$

Some immediate results are now the following: If we replace $\mathbf{P}$ with an equivalent probability measure $\mathbf{Q}$, the topology of $\mathcal{S}$ remains the same (theorem of the closed graph). If $\mathbf{Q}$ is only absolutely continuous with respect to $\mathbf{P}$, the identity mapping of $\mathcal{S}_{\mathbf{P}}$ to $\mathcal{S}_{\mathbf{Q}}$ is continuous.

Now, the stochastic integral can be defined as follows.

Definition 17 (Sophisticated Definition of Stochastic Integral)**.** *Let X be a classical semimartingale, H be a predictable process and $H^n = H \mathbb{1}_{\{|H| \le n\}}$. We say that H is X-integrable if $H^n \bullet X$ converges to a classical semimartingale Y in the semimartingale topology. In this case, we say Y is the stochastic integral of H with respect to X and write*

$$Y = H \bullet X.$$

This definition heavily relies on the definition of the stochastic integral for locally bounded predictable integrands. Therefore, the theory as developed by Doléans-Dade and Meyer [41] and presented in Section 3.6 is crucial.

Remark 13. *Since the semimartingale topology does not depend on the underlying probability measure, it is obvious that $L(X)$ is independent of any change of equivalent probability measures.*

Remark 14. *If H belongs to $L(X)$ and $L(Y)$, it belongs to $L(X + Y)$ and $H \bullet (X + Y') = H \bullet X + H \bullet Y$. However, the linearity with respect to the integrands is less obvious than the one for the integrators.*

Remark 15. *The sophisticated definition of the stochastic integral is equivalent to the definition as given in Definition 16, and $H \in L(X)$ is equivalent to H being X-integrable. For details, refer to [26], page 130 d).*

Remark 16. *Although the approach referred to here as the sophisticated definition of the stochastic integral is rarely covered in university lectures or textbooks, certain aspects of it have proven to be extremely valuable. For instance, this approach allows for a simplified proof of the stochastic dominated convergence theorem (see, for example, [29]).*

When Doléans-Dade and Meyer [41] developed the general stochastic integral for bounded integrands, they managed to extend the class of suitable integrands to locally bounded predictable processes via localization. This begs the question of whether the general stochastic integral can be even further generalized by examining integrands that are only locally X-integrable. However, it turns out that local X-integrability already implies general X-integrability and therefore, such an extension is not possible for the general integral. This was demonstrated by [26] via the following theorem:

Theorem 24 ([26], Théorème 4). *Suppose H is predictable. Assume that there is a sequence of increasing stopping times T^k, tending to ∞, processes $J_k \in L(X)$, such that $H = J_k$ on $[\![0, T^k[\![$. Then $H \in L(X)$ and $H \bullet X = J_k \bullet X$ on $[\![0, T^k[\![$.*

The approach towards the general stochastic integral was later further simplified. However, all approaches (including the functional analytic approach) rely on the decomposition of classical semimartingales into a local martingale part and a Stieltjes part. The simplified approach, which is, for example, presented in [29,71] starts with a definition of the stochastic integral for general local martingales and then, via the definition of the stochastic integral for classical semimartingales, extending the space of potential integrands. As it turns out, by doing so, the stochastic integral for local martingales also gets generalized. The definition does not classify the jumps of the classical semimartingale but builds upon the decomposition of a local martingale into a continuous and a discontinuous part (Theorem 18), and on the definition of the stochastic integral for locally square-integrable martingales (and therefore, in particular, for continuous local martingales) as it was outlined in the previous section. We will give a short overview of this simplified procedure.

Definition 18. *Let M be a local martingale and H a predictable process. The stochastic integral $X = H \bullet M$ of H with respect to M is defined as the uniquely defined local martingale possessing the following characteristics:*

- *The continuous part of X, denoted X^c, satisfies $X^c = H \bullet M^c$, where M^c is the continuous martingale part of M.*
- *The processes $\Delta(H \bullet M)$ and $H\Delta M$ are indistinguishable.*

To prove the existence of the stochastic integral, one examines the class of predictable processes H that locally satisfy

$$\mathbf{E}\left[\left(H_0^2 M_0^2 + \int_0^\infty H_s^2 \mathrm{d}[M, M]_s\right)^{\frac{1}{2}}\right] < \infty.$$

By applying the Burkholder-Davis-Gundy inequality (Burkholder et al. [24], Garsia [58]), one can show the equivalence of the norms $M \mapsto \mathbf{E}\left[[M, M]_\infty^{\frac{1}{2}}\right]$ and $M \mapsto \mathbf{E}[|M_\infty^*|]$ and therefore construct an approximating sequence of simple predictable processes. By dominated convergence and by delocalization (pasting together local processes), one can now construct a local martingale that satisfies the condition from Definition 18.

Now, the integral can be extended in the following way:

Definition 19. *Let X be a classical semimartingale and let H be a predictable process such that there exists a decomposition $X = M + A$, where M is a local martingale and A is a finite variation (FV) process, satisfying the following conditions:*

- *$H \in L_{\mathrm{loc}}^2(M)$,*
- *For almost all ω, the paths of $H(\omega)$ are integrable as a Lebesgue-Stieltjes integral with respect to A.*

Then we say H is X-integrable, denoted $H \in L(X)$, and we define the stochastic integral $H \bullet X$ as

$$H \bullet X = H \bullet M + H \bullet A,$$

where $H \bullet M$ is defined according to Definition 18, and $H \bullet A$ as a Lebesgue-Stieltjes integral.

The original publication by Jacod [88,89] introduced the condition that, similar to the definition in Definition 18, the jumps satisfy $H\Delta X = \Delta(H \bullet X)$. Nevertheless, with the help of the Fundamental Theorem of Local Martingales, it is possible to show that the continuous part X^c of a classical semimartingale is indeed uniquely determined, and therefore, the condition on the jumps is a straightforward consequence of the requirements given above.

It is worth noting that for this definition, at least one decomposition of X is required to satisfy the above-mentioned criteria. However, it is possible that some decompositions of X fulfil these requirements while others do not. In fact, it is even the case that if X is a local martingale and $H \in L(X)$, then $H \bullet X$ is not necessarily a local martingale. Nevertheless, there exists at least one decomposition of $X = M + A$ such that $H \bullet M$ is again a local martingale.

As the existence of the stochastic integral may vary across different decompositions, it is not possible to provide simple, necessary and sufficient integral conditions for the space $L(X)$.

The proof of the uniqueness and existence requires more technical details about the spaces of martingales and several properties of the stochastic integral for square-integrable martingales. Usually, one of the key ingredients for the proof of the properties of the stochastic integral (such as the stability of the local martingale property for locally bounded integrands) is the Burkholder-Davis inequality.

4. The Functional Analytic Approach

The approach to the general stochastic integral for classical semimartingales described in Section 3 has some disadvantages. It starts with the definition of the classical semimartingale as a process decomposable into a local martingale and a process of finite variation. Since the local martingale property depends strongly on the choice of the underlying probability measure, it can be assumed that the classical semimartingale property also depends on the selection of an equivalent probability measure. However, this is not the case, as was first shown by Meyer [157].

Furthermore, the meaning of the classical semimartingale for the stochastic integral is not evident from the definition. This means that the fact that classical semimartingales are the largest class of processes for which a stochastic integral can be defined in an "obvious" way seems completely unclear. Also, the derivation of the integral employing the classical method has turned out to be technically very complex and has been criticized many times (see, for example, [171]).

The functional analytic approach offers an alternative to the classical derivation that addresses and largely solves these problems for the stochastic integral for càglàd integrands. It allows to define stochastic integrals with respect to left continuous, adapted processes very quickly and to prove the most properties of the stochastic integral, including a change of variables formula for semimartingales. This integral and these results suffice for a wide variety of applications. However, for the extension of the stochastic integral to general predictable integrands, one needs to resort again to more technical proofs and, in particular, to the decomposition theorems of martingales and semimartingales, including the Bichteler-Dellacherie-Mokobodzki theorem. One of the disadvantages of the functional analytic approach in comparison to the classical approach is that in the case of the functional analytic approach, it is not directly apparent that the stochastic integral with respect to an L^2-martingale yields again an L^2-martingale which is clearly a desirable property for any kind of stochastic integral.

The motivation for the functional analytic approach emerged from the realization that classical semimartingales represent the largest class of possible integrators for which a well-defined integral can be constructed. This was demonstrated by Bichteler [13,14], Dellacherie [35], Meyer [157]. Here, "well-defined" means that if a càdlàg process X is to satisfy a weak form of the bounded convergence theorem and if the stochastic integral is defined canonically for simple predictable processes, then X must be a semimartingale.

This discovery clarified the importance of semimartingales in the theory of stochastic integration. Meyer [156] and Dellacherie [35] were the first to propose and describe the functional analytic technique as it is understood today. Then, Lenglart [124] and Protter [170,171] built upon it. Protter [172] wrote an entire comprehensive textbook based on the functional analytic approach and covers the results from

all publications mentioned above. Furthermore, it provides a set of simplified proofs for well-known results. Letta [125] came up with a concept similar to the functional analytic approach discussed here.

4.1. The Stochastic Integral for Càglàd Integrands

Similar to Itô's approach, in the functional analytic ansatz, one starts by formulating the stochastic integral for a simple type of integrand. Itô chose right continuous (optional) step functions with deterministic step times. As the importance of predictable integrands became apparent, this approach was mostly changed to using elementary predictable processes (as we presented it in this survey). The functional analytic approach usually starts with defining the stochastic integral for simple predictable integrands, which is a slight generalization of elementary predictable processes by allowing the step times to be stopping times instead of deterministic times. However, an ansatz with elementary predictable processes would yield the same integral.

Definition 20. *A simple predictable process $H : \Omega \times [0, \infty)$ is defined as*

$$H(t, \omega) = H_0(\omega) \mathbb{1}_{\{0\}}(t) + \sum_{i=1}^{n} H_i(\omega) \mathbb{1}_{\rrbracket T^i, T^{i+1} \rrbracket}(t). \tag{10}$$

with stopping times $0 \le T^1 \le \ldots \le T^n \le T^{n+1} < \infty$ where each $H_i(\omega)$ is an $\mathcal{F}_{T^i}$-measurable, bounded, random variable. The space of simple predictable processes is denoted by Λ.

Similar to the elementary predictable processes, simple predictable processes are an optimal choice to start with as the space of simple predictable processes Λ generates the predictable σ-algebra, that is, $\mathcal{P} = \sigma(\Lambda)$. And, for any measure μ on $\mathcal{P}$ and any $q \in [1, \infty]$, it is easy to show by dominated convergence that Λ is a dense set of functions in $L^q(\mathcal{P}, \mu)$, where $L^q(\mathcal{P}, \mu)$ consists of all predictable functions f such that the q-th power of the absolute value of f is integrable with respect to the measure μ:

$$L^q(\Sigma_p, \mu) = \left\{ f : \int_\Omega |f|^q \, d\mu < \infty \right\}.$$

For simple predictable processes, there is a canonical way how to define the stochastic integral with respect to a càdlàd process X:

For a stochastic process X and a simple predictable process H with presentation as in Equation (10), let

$$I_Z(X) = H_0 X_0 + \sum_{i=1}^{n} H^i (X_{T^{i+1}} - X_{T^i}) \tag{11}$$

be the stochastic integral with respect to X.

It is easy to see that the stochastic integral does not depend on the choice of the representation of H in Λ.

Any sensible integral operator should be linear and satisfy some kind of continuity. Equation (11) is clearly linear. In order to define or check continuity, one needs to define some kind of topologies first.

Let Λ_u be the space of simple predictable processes endowed with the topology of uniform convergence on $\mathbb{R}_+ \times \Omega$ and let L^0 denote the space of all real-valued random variables on $(\Omega, \mathcal{F}, P)$, endowed with the topology induced by convergence in probability.

To ensure continuity, in the functional analytic approach, one only allows integrators that satisfy this continuity condition:

Definition 21. *A càdlàg, adapted process X is called a good integrator is the mapping $I_{Z(\cdot \wedge t)} : \Lambda_u \to L^0$ is continuous for each $t \ge 0$.*

The continuity required for a suitable integrator is relatively weak, as uniform convergence is a strong condition, while convergence in probability is comparatively weak. Nevertheless, it can be shown that much stronger continuity results hold (see, for example, [106, Theorem 5.89]).

With these definitions, we can state the previously mentioned result that semimartingales represent the largest class of integrators for which a "reasonable" stochastic integral can be defined. This result is now commonly known as the *Bichteler–Dellacherie–Mokobodzki Theorem* and is typically presented in the following form.

Theorem 25 (Bichteler-Dellacherie-Mokobodzki Theorem, [14] Theorem 7.6). *An adapted càdlàg process X is a good integrator if and only if it is a classical semimartingale.*

The theorem was initially shown by Bichteler [13,14], Dellacherie [35], Meyer [157]. The original proofs were later simplified by Beiglböck and Siorpaes [7], which can also be found in the appendix of [29]. While the original proofs heavily relied on the entire machinery of stochastic calculus and stochastic integration, the latter approach requires much fewer prerequisites.

This definition makes it easy to show that good integrators (and hence semimartingales) are closed under a change of equivalent probability measures and closed under localization. The set of good integrators forms a vector space.

The process of defining the stochastic integral for càglàd integrands is similar to Itô's approach:

Definition 22. *Let X be a good integrator and $H \in \Lambda$ represented as*

$$H(t,\omega) = H_0 \mathbb{1}_{\{0\}}(t) + \sum_{i=1}^{n} H_i \mathbb{1}_{\rrbracket T^i, T^{i+1} \rrbracket}(t).$$

We define the stochastic integral $J^X(H)_t$ by

$$J_t^X(H) := H \bullet X_t := \int_0^t H \mathrm{d}X_s = H_0 X_0 + \sum_{i=1}^{n} H_i (X_{T^{i+1} \wedge t} - X_{T^i \wedge t})$$

and

$$J_\infty^X(H) = H_0 B_0 + \sum_{i=1}^{n} H_i \left(X^{T^{i+1}} - X^{T^i} \right)$$

Then $\Lambda \subset \mathbb{L}$ and the stochastic integral process $J_Z(X)$ is in $\mathbb{D}$.

Before extending the definition of $J^X(H)$ to all $H \in \mathbb{L}$, new topologies have to be defined on Λ, $\mathbb{L}$ and $\mathbb{D}$ due to the implicit localization of $J^X(H)(t)$ to the interval $[0,t]$.

Definition 23. *(a)   A series of processes $(X^n)_{n \in \mathbb{N}}$ is said to converge to a process X uniformly on compacts in probability (ucp) if for each $t \geq 0$,*

$$\sup_{0 \leq s \leq t} |X_s^n - X_s| \to 0 \text{ in probability as } n \to \infty.$$

(b)   By Λ_{ucp} we refer to the space Λ of simple predictable processes endowed with the topology induced by the ucp convergence. We define $\mathbb{L}_{\mathrm{ucp}}$, $\mathbb{D}_{\mathrm{ucp}}$ accordingly.

Remark 17. *More details and a metric for the topology induced by the ucp convergence are given in Section 9.*

The following lemmas show why these particular topologies have been chosen.

Lemma 5 ([172], Theorem 10). *The space Λ_{ucp} is dense in $\mathbb{L}_{\mathrm{ucp}}$.*

Lemma 6 ([106], Theorem 2.71). *The space $\mathbb{D}_{\mathrm{ucp}}$ is complete.*

The continuity extends from I_X to J^X.

Lemma 7 ([172], Theorem II.11)**.** *Let X be a good integrator. Then the mapping $J^X : \Lambda_{\text{ucp}} \to \mathbb{D}_{\text{ucp}}$ is continuous.*

Now, Λ_{ucp} is dense in $\mathbb{L}_{\text{ucp}}$, $\mathbb{D}_{\text{ucp}}$ is a complete metric space, and the linear mapping $J^X : \Lambda_{\text{ucp}} \to \mathbb{D}_{\text{ucp}}$ is continuous.

J^X may thus be extended to a continuous linear mapping from $\mathbb{L}$ to $\mathbb{D}$. For any $t \geq 0$, this determines the stochastic integral $\int_{[0,t]} H_s \mathrm{d} X_s = J^X(H)(t)$ of $H \in \mathbb{L}$ with respect to X on $[0,t]$.

Definition 24. *The continuous linear mapping $\overline{J^X}$ from $\mathbb{L}_{\text{ucp}}$ to $\mathbb{D}_{\text{ucp}}$, obtained as the extension of J^X from Λ_{ucp} to $\mathbb{D}_{\text{ucp}}$ by continuity, is called the stochastic integral. That is, for $H \in \mathbb{L}$, let $\{H^n\}$ be a sequence of simple predictable processes that converge to H in ucp. Then the stochastic integral $\overline{J^X}(H)$, denoted by $H \bullet B$, is defined as the ucp limit of the stochastic integrals of the approximating sequence, i.e.,*

$$H \bullet X = \lim_{n \to \infty} H^n \bullet X,$$

where the convergence is uniformly on compacts in probability.

In the classical approach, the quadratic variation of semimartingales is required to define the stochastic integral. The quadratic variation is typically defined via existence theorems (see Definition 12) or through approximating sums (Theorem 20). In the functional analytic approach, however, the quadratic variation is not necessary for defining the stochastic integral. Consequently, we can first define the stochastic integral for càglàd integrands and then define the quadratic variation using the partial integration formula.

Definition 25. *Let X and Y be good integrators. The quadratic variation $[X] := [X, X] := ([X, X]_t)_{t \geq 0}$ is defined by*

$$[X, X] = X^2 - 2X_- \bullet X.$$

The quadratic covariation $[X, Y]$ of X and Y is defined by

$$[X, Y] = XY - X_- \bullet Y - Y_- \bullet X.$$

The stochastic integral, as developed so far, suffices for most applications, such as the Itô Formula, the Doléans-Dade exponential, most applications in derivative pricing and risk management, Lévy's Theorem, or Lévi's Stochastic Area formula. Nevertheless, for more advanced applications, a more general stochastic integral is required.

4.2. The Stochastic Integral for Predictable Integrands

We now discuss the extension from càglàd integrands to general predictable integrands. As with the classical approach, the extension is very technical and depends on carefully chosen metrics, the decomposition of classical semimartingales, and the theory of specific semimartingale spaces.

The complete development of this theory can be found in [172] and was first published in the initial edition of the cited textbook. However, similar techniques were already discussed in [170,171], Jacod [89], and Chou et al. [26].

To extend the stochastic integral from $\mathbb{L}$ integrands to $\mathcal{P}$, it is again necessary to define appropriate topologies on $\mathcal{P}$ and $\mathbb{D}$ in such a way that $\mathbb{L}$ is dense in $\mathcal{P}$ (or to be more precise $\mathbf{b}\mathbb{L}$ in $\mathbf{b}\mathcal{P}$, where $\mathbf{b}$ denotes the restriction to bounded processes) and that the integral operator is continuous. While for the integral for càglàd integrators, it was sufficient to define a topology that was independent of the integrator, it is necessary to adapt the metric to the corresponding integrator process for the extension here.

First, the integral is only defined for a particular class of semimartingales, the $\mathcal{H}^2$-semimartingales, which is a suspace of the special semimartingales.

A classical semimartingale X is a special semimartingale if there exists a decomposition of X into a local martingale and a predictable process of finite variation (see Definition 15). Such a decomposition is always unique.

It is easy to show that a classical semimartingale is a special semimartingale if and only if its supremum process is locally integrable (Meyer [157], Yoeurp [200]).

We follow here closely the presentation in [172].

Definition 26. *Let X be a special semimartingale with canonical decomposition $X = N + A$. The $\mathcal{H}^2$-norm of X is defined as*

$$\|X\|_{\mathcal{H}^2} := \left\| [N,N]_\infty^{\frac{1}{2}} \right\|_{L^2} + \left\| \int_0^\infty |\mathrm{d}A_s| \right\|_{L^2} < \infty.$$

The space of semimartingales $\mathcal{H}^2$ consists of all special semimartingales with finite $\mathcal{H}^2$-norm.

Remark 18. *In the literature, the term $\mathcal{H}^2$-space typically refers specifically to martingales that satisfy similar integrability conditions. Then, for clarity, the corresponding space for semimartingales is often denoted by $\overline{\mathcal{H}}^2$. Since we do not address the martingale space $\mathcal{H}^2$ in this context, we omit this distinction here.*

The first goal is to extend the class of stochastic integrands from $\mathbf{b}\mathbb{L}$ to $\mathbf{b}\mathcal{P}$, with $X \in \mathcal{H}^2$. First, one easily sees that if $H \in \mathbf{b}\mathbb{L}$ and $X \in \mathcal{H}^2$, then the stochastic integral $H \bullet X$ is an $\mathcal{H}^2$-semimartingale. Additionally, if $X = N + A$ is the canonical decomposition of X, then $H \bullet N + H \bullet A$ is the canonical decomposition of $H \bullet X$. Moreover,

$$\|H \bullet X\|_{\mathcal{H}^2} = \left(\int_0^\infty H_s^2 \, \mathrm{d}[N,N]_s \right)^{\frac{1}{2}} \|_{L^2} + \left\| \int_0^\infty |H_s| \, \mathrm{d}|A|_s \right\|_{L^2}.$$

The key idea in extending our integral is to recognize that $[N,N]$ and A are finite variation processes. Therefore, the integrals $\int_0^t H_s \, \mathrm{d}[N]_s$ and $\int_0^t |H_s| \, \mathrm{d}|A|_s$ are well-defined for any $H \in \mathbf{b}\mathcal{P}$ and not just for $H \in \mathbb{L}$.

Definition 27. *For a special semimartingale $X \in \mathcal{H}^2$ with canonical decomposition be $X = N + A$ and $H, J \in \mathcal{P}$, the metric d_X is defined as*

$$d_X(H,J) := \left\| \left(\int_0^\infty (H_s - J_s)^2 \mathrm{d}[N,N]_s \right)^{\frac{1}{2}} \right\|_{L^2} + \left\| \int_0^\infty |H_s - J_s| |\mathrm{d}A_s| \right\|_{L^2}.$$

and a predictable process H is called $(\mathcal{H}^2, X)$-integrable if $d_X(H,0) < \infty$ holds.

By identifying indistinguishable processes, we obtain the following important result:

Theorem 26 ([172], Theorem IV.1). *The space $\mathcal{H}^2$ is a Banach space, and the quotient space $\mathcal{H}^X / \mathcal{N}$, where $\mathcal{H}^X$ is the set of $(\mathcal{H}^2, X)$-integrable processes and $\mathcal{N}$ is the set of constant processes (modulo modification), is a normalized vector space with norm*

$$\|H\|_{d_X} := d_X(H,0).$$

In particular, since the triangle inequality holds, it follows for two $(\mathcal{H}^2, X)$-integrable processes H, J

$$d_X(H,J) \le d_X(H,0) + d_X(J,0) < \infty.$$

Lemma 8 ([172], Theorem IV.2). *The space of bounded càglàd processes $\mathbf{b}\mathbb{L}$ is dense in the space of bounded predictable processes $\mathbf{b}\mathcal{P}$ under $d_X(\cdot, \cdot)$.*

The proof relies on the monotone class theorem, which breaks down for unbounded processes. This is the reason why we restrict ourselves to bounded processes first. This approach is similar to the classical approach, where Doléans-Dade and Meyer [41] defined the stochastic integral first for bounded predictable integrands as the monotone class theorem could be utilized for that (also see [157] for a nice presentation).

Furthermore, one obtains the following important isometry for $X \in \mathcal{H}^2$, $H, J \in \mathbf{bL}$:

$$\|H \bullet X - J \bullet X\|_{\mathcal{H}^2} = d_X(H, J). \tag{12}$$

Theorem 27 ([172], Theorem IV.3). *If $(H^n)_{n \in \mathbb{N}}$ is a Cauchy sequence in $\mathbf{bL}$ with respect to d_X, then $(H^n \bullet X)_{n \in \mathbb{N}}$ is also a Cauchy sequence in $\mathcal{H}^2$.*

With these prerequisites, it is now possible to define the stochastic integral for bounded predictable integrands and $\mathcal{H}^2$-semimartingales.

Definition 28. *Let $X \in \mathcal{H}^2$, $H \in \mathbf{bP}$ and $(H^n)_{n \in \mathbb{N}}$ be a sequence in $\mathbf{bL}$ which converges in d_X to H. The stochastic integral $H \bullet X$ is the unique process $Y \in \mathcal{H}^2$ with*

$$Y = \lim_{n \to \infty} H^n \bullet X,$$

where the convergence is in $\mathcal{H}^2$.

In particular, it is easy to show that the stochastic integral does not depend on the choice of the approximating sequence.

As the next step, the stochastic integral is defined for unbounded predictable integrands. For $X \in \mathcal{H}^2$ and an $(\mathcal{H}^2, X)$-integrable process $H \in \mathcal{P}$, the sequence H^n is defined by $H^n := H\mathbb{1}_{\{|H| \leq n\}}$. Then the stochastic integral as defined above exists for all H^n and, furthermore, the following theorem holds:

Theorem 28 ([172], Theorem IV.14). *Let $X \in \mathcal{H}^2$ and let $H \in \mathcal{P}$ be $(\mathcal{H}^2, X)$-integrable. For $H^n = H\mathbb{1}_{\{|H| \leq n\}}$ the sequence $(H^n \bullet X)_{n \in \mathbb{N}}$ is a Cauchy sequence in $\mathcal{H}^2$.*

With this result, the stochastic integral $H \bullet X$ can be defined for $X \in \mathcal{H}^2$ as

$$H \bullet X := \lim_{n \to \infty} H^n \bullet X = \lim_{n \to \infty} \left(H\mathbb{1}_{|H| \leq n} \right) \bullet X,$$

where the limit is seen in the $\mathcal{H}^2$-sense.

To pass from $\mathcal{H}^2$-semimartingales to general classical semimartingales, some kind of localization procedure (similar to how it was done in Section 3.5) should be applied. However, even locally, a classical semimartingale is, in general, not a special semimartingale (one can even show that every local special semimartingale is already a special semimartingale, cf. [92, Proposition 4.25]). In order to be able to pass from classical semimartingales to special semimartingales, the localization procedure needs to be relaxed further, and we examine prelocal behavior instead, which means, roughly speaking, the process X is not stopped at the random time T, but at a minimally earlier time $T-$. This way, we circumvent the problem that arises with unbounded jumps at the stopping times and obtain additional regularity properties.

The reason for studying prelocal behaviour is the following theorem.

Theorem 29 ([172], Theorem IV.13). *Let X be a classical semimartingale, then X is prelocally in $\mathcal{H}^2$.*

This allows us to define the general (component-wise) stochastic integral.

Definition 29. *Let X be a classical semimartingale and $H \in \mathcal{P}$. The stochastic integral $H \bullet X$ exists if there is a sequence of stopping times increasing towards infinity such that $X^{T^n-} \in \mathcal{H}^2$ and H is $(\mathcal{H}^2, X^{T^n-})$-integrable for all n. In this case, we say H is integrable with respect to X and write $H \in L(X)$. We define the stochastic integral as*

$$H \bullet X = H \bullet (X^{T^n-}), \quad on \; [\![0, T^n[\![,$$

for all n.

It is easy to show here that the integral does not depend on the sequence of stopping times.

It should be noted once again that the definition of the stochastic integral is usually the easy part. The proof for the most important properties of the stochastic integral is, in most cases, much harder.

There is no dedicated approach for extending the component-wise stochastic integral to the integral for vector-valued semimartingales. Hence, the extension uses the same methods as the classical approach.

5. The Stochastic Vector Integral

So far, we have only examined the stochastic integral for one-dimensional integrands and integrators. The obvious way to generalize stochastic integration to allow for multidimensional semimartingales $X = \left(X^1, \ldots, X^d \right)$ and multi-dimensional integrands $H = (H^1, \ldots H^d)$ would be to calculate each component's stochastic integral and build the sum of all components. Thus, the set of suitable integrands would be defined as the class of processes $H = \left(H^1, \ldots, H^d \right)$ such that, for any $i = 1, \ldots, d, H^i \in L(X^i)$ is integrable. In this case, the space $\{H \bullet X ; H \in L(X)\}$ is closed for the one-dimensional processes X and H.

However, contrary to what L. Galtchouk had previously thought, it turns out that the space of stochastic integrals with respect to a given multidimensional semimartingale, when defined as componentwise stochastic integrals, is, in general, not closed in the semimartingale topology for $d > 1$. To get the closedness of the space of stochastic integrals, the concept of the componentwise stochastic integral has to be furhter generalized. Jacod [90] presented this generalization when he developed the stochastic vector integral with regard to a multidimensional semimartingale. Jean Mémin showed in [138] that this construction results in a closed space of stochastic integrals. This concept of the stochastic vector integral is also briefly treated in C. Dellacherie and Paul-André Meyer's book [37].

In [90], the stochastic vector integral is built in implicit form, as opposed to prior publications in which the stochastic integral was first formed for "simple" integrands and then extended to general integrands via extension theorem of limit construction. The explicit approach for stochastic vector integration was first suggested by Shiryaev [185] and then shown in detail in [184]. In [105] and [106], another stochastic integral was developed, which likewise yielded closedness.

A quick summary of the vector stochastic integral's creation is given in the following, which mainly follows the presentation in [29].

The objective is to enable the components to be netted out before constructing the integral. As in the traditional method, the integral is first defined for local martingales and for finite variation processes. Here, it is assumed that all vectors are column vectors; hence $x^\top y$ is the inner product of x and y.

In the one-dimensional case, the construction of the stochastic integral relies on the quadratic variation process $[M, M]$, which represents the cumulative variance of the martingale M. This process provides a foundation for defining integrals and controlling their behavior.

In the multidimensional setting, each pair of components (M^i, M^j) of a vector martingale $M = (M^1, \ldots, M^d)$ has its own pairwise covariation $[M^i, M^j]$. A direct extension of the one-dimensional approach would require a process that can consistently represent these covariations. Here, we introduce

a new finite variation process C to serve as a unifying framework that links all pairwise covariations through an optional process θ^{ij}, so that

$$[M^i, M^j] = \theta^{ij} \bullet C.$$

This construction is essential because it provides a shared "clock" or "driver" for all covariations, ensuring consistency across the components.

In comparison to the one-dimensional case, where $[M, M]$ alone suffices, C generalizes the quadratic variation process to handle the vector case. Instead of relying on multiple individual quadratic variations, we define all covariations in terms of a single process C, modulated by θ^{ij}. This setup is particularly advantageous when ensuring that the space of stochastic integrals is closed, as it captures both self- and cross-dependencies in a coherent way.

The first step is to demonstrate that, given a d-dimensional vector local martingale with components M^i, there exists an adapted process of finite variation C and an optional process θ such that $x^\top \theta x \geq 0$ for any $x \in \mathbb{R}^d$ and $\theta^\top = \theta$ and

$$\left[M^i, M^j \right] = \theta^{ij} \bullet C$$

for all $i, j \in \{1, 2, \ldots, d\}$.

Similar to the procedure for the componentwise integral, the next step is to establish suitable metrics. Here, we choose for a d-dimensional semimartingale M with $M_0 = 0$, a d-dimensional predictable process H, and $p \in [1, \infty)$

$$H \mapsto \mathbf{E}\left[\left(\left(H^\top \theta H \right) \bullet C \right)_\infty^{p/2} \right]^{1/p} =: \|H\|_{L^p(M)}.$$

If $\|H\|_{L^p(M)} < \infty$, we write $H \in L^p(M)$, and $H \in L^p_{\text{loc}}(M)$ if H is locally in $L^p(M)$.

It is straightforward to demonstrate that $\|H\|_{L^p(M)}$ is a norm since it is composed of a Hilbert space norm and an L^p-norm. A dominating convergence argument similarly demonstrates that the simple integrands are dense in the space $L^p(M)$ for each p.

Considering the definitions above, one can see that the set L^p equals the set of processes H for which $\left((H^\top \theta H) \bullet C \right)^{p/2}$ is a process of integrable variation, and as in the case of scalars, the construction for $p = 2$ may be modified by replacing $[M^i, M^j]$ with $\langle M^i, M^j \rangle$. The space $L^2_{\text{loc}}(M)$ perfectly corresponds to the space of predictable processes H, for which $(H^\top \theta H) \bullet C$ is locally integrable.

The following is the definition of the stochastic integral with regards to a local d-dimensional martingale:

Definition 30. *For a local martingale M with $M_0 = 0$ and a predictable process $H \in L^1_{\text{loc}}(M)$, the stochastic integral is defined as the local martingale $X = H \bullet M$ such that, for any local martingale N*

$$[X, N] = \left(H^\top K \right) \bullet C$$

where K is an optional vector process such that $\left[M^i, N \right] = K^i \bullet C$.

It is easy to show that the integral is uniquely defined.

If each component H^i is M^i-integrable, then the relationship $H^\top \bullet M = \sum_i (H^i \bullet M^i)$ holds by the uniqueness and linearity of the integral. This componentwise summation approach ensures consistency in the construction of the vector integral under these conditions.

In cases where the semimartingale M has uncorrelated components, i.e., $\langle M^i, M^j \rangle = 0$ for each $i \neq j$, the vector integral and the componentwise sum are equivalent. Here, the optional symmetric matrix process θ is effectively diagonal, simplifying the isometry so that each component H^i is necessarily M^i-integrable. This correspondence is crucial, as it reaffirms that the vector stochastic

integral not only extends the classical componentwise approach but does so in a way that inherently respects the independence of uncorrelated components.

To complete the construction, we extend the stochastic integral to vector processes with finite variation. For such processes, the vector stochastic integral is defined as follows:

Definition 31. *Let $A = (A^1, \ldots, A^d)$ be an $\mathbb{R}^d$-valued càdlàg process with components of finite variation. Then, there exists a càdlàg increasing process V and a predictable $\mathbb{R}^d$-valued process $v = (v^1, \ldots, v^d)$ such that each component A^i admits the representation*

$$A^i = v^i \bullet V.$$

If A is predictable, then V and v can also be chosen to be predictable.

We define the space $L^{\mathrm{var}}(A)$ as the set of predictable processes $H \in \mathbb{R}^d$ for which the process

$$H^\top \bullet A := (H^\top v) \bullet V$$

is of finite variation.

Remark 19. *A concrete choice for V is given by*

$$V_t := \sum_{i=1}^d \int_{[0,t]} |\mathrm{d}A^i|,$$

which dominates the variation of each component of A, and ensures the existence of such a predictable density v with respect to V.

It is easy to see that $L^{\mathrm{var}}(A) = L^{\mathrm{var}}_{\mathrm{loc}}(A)$.

The notation $H^\top \bullet X$ is not ambiguous; if both integrals are defined for a vector martingale of finite variation X and H, then the integrals agree.

This norm serves as a natural extension of the one-dimensional L^p-norm for the vector case. In the 1-dimensional setting, the norm relies on the quadratic variation $[M, M]$; here, $\|H\|_{L^p(M)}$ incorporates the matrix θ and the process C to handle the multidimensional context.

If $\|H\|_{L^p(M)} < \infty$, we write $H \in L^p(M)$, and if H is locally in $L^p(M)$, we write $H \in L^p_{\mathrm{loc}}(M)$.

The extension to semimartingales, in general, is now evident.

Definition 32. *A vector process H is integrable with regard to a vector semimartingale X (represented by $H \in L(X)$) if there exists a decomposition $X = M + A$ such that $H^\top \bullet M$ and $H^\top \bullet A$ are well defined, as vector local martingale and vector Stieltjes integrals respectively. Consequently, $H^\top \bullet X := H^\top \bullet M + H^\top \bullet A$.*

The integral is independent of the selection of decomposition and vector space $L(X)$ of X-integrable processes.

Remark 20. *If $H = (H^1, \ldots, H^d)$ is a predictable, locally bounded process, then for any decomposition $X = A + M$ where A is a d-dimensional process of finite variation and M is a d-dimensional local martingale, we have $H \in L^{var}(A) \cap L^1_{loc}(M)$. Thus, $L(X)$ includes all locally bounded predictable processes.*

Remark 21. *If $H \in L(X)$, then H may not belong to $L_{var}(A) \cap L^1_{loc}(M)$ for all decompositions $X = A + M$ with A a d-dimensional process of finite variation and M a d-dimensional local martingale, see [184, Example 5.2].*

The following holds with this formulation of the stochastic integral:

Theorem 30 ([138], Théorème II.7). *In the semimartingale topology, the space $\{H \bullet X\}_{H \in L(X)}$ is complete for every semimartingale X.*

The Infinite Dimensional Case

When X is an infinite-dimensional process the theory developed in this survey does not suffice and further generalizations are required.

Grigelionis and Mikulevičius [66], Korezlioglu and Martaias [112], Kunita [115], Métivier and Pellaumail [142], Métivier and Pistone [143], Meyer [154], and many others, introduced and analyzed several important classes of stochastic integrals in depth. Not unexpectedly, the techniques for infinite-dimensional stochastic integration presented in these papers exhibit parallels and differences.

The cylindrical stochastic integration theory proposed by Mikulevicius and Rozovskii [159,160] is one method for approaching this topic. In Donno and Pratelli [43] and De Donno et al. [32], a particular application of this idea was refined and applied to finance. Alternatively, the theory of Random Measures, which is presented in detail in [15], can be extended to cover infinite-dimensional stochastic integration. If X takes values in a Hilbert space, then the construction for infinite-dimensional stochastic integration, as described in [140], is identical to that for finite-dimensional spaces.

6. The Integral with Respect to Random Measures

There is another method, conceptually equivalent, developed by Metivier and Pellaumail [141], Pellaumail [166], Ruiz de Chávez [181] and for special cases by Skorohod [187]. This method involves extending the "stochastic measure" associated with a semimartingale, starting from the integral of elementary predictable, such as the extension of the measure method of Kussmaul [122]. One of the major advantages of this method is its ability to extend effortlessly to martingales with vector values even in infite definition.

The motivation to study the stochastic integral as an integral with respect to random measures stems firstly from the fact that most of the properties of increasing processes are in fact the properties of the associated random measure $dA_t(\omega)$ on $\mathbb{R}_+$: hence the idea of generalizing these properties to random measures on sets larger than $\mathbb{R}_+$ in an appropriate way. Secondly, a certain number of difficulties in the treatment of martingales and semimartingales in the classical approach are due to the fact that the jumps of these processes can be "too large". However, it is possible to represent the jumps greater ε for any semimartingale as a random measure, for which integration is straight-forward. Thus simplifying the theory.

Besides the advantage of the straightforward extension to the vector-valued martingales, this method makes the connection to several important related topics more clear. The most prominent examples are probably the point processes in general and Poisson point measures in particular.

As the key elements of this approach are similar to the classical method and rely on the decomposition of martingales and the study of the jump behavior of semimartingales, we are not going into detail here.

The introduction of random measures as we understand them here (in particular their role in time), and especially random measures with integer values, can be seen in the works of Itô [86], Kunita and Watanabe [117] for Poisson measures, Kunita and Watanabe [117] for the associated measure and a Markov process, Skorokhod [188]. These authors immediately emphasized the connections with martingale theory, implicitly using the notion of dual predictable projection.

Comprehensive treatments of the theory can be found in Chung and Williams [28], Jacod [89], Métivier [139]. While Chung and Williams [28] or also [182] follow a nice didactical and easy-to-grasp approach but restrict themselves to (local) martingales, Métivier [139] defines the integral for semimartingales in such a way that the extension to multi-dimensional (and even infinite-dimensional) stochastic integral is straight-forward.

The definitions of dual predictable projection and Doob-Meyer measure are elaborated, for example, in Jacod [91]. The notion of a Lévy system, which is closely related to random measures, is due to Kunita and Watanabe [117] when X is a Markov process, although the characterization of the

Lévy system as a dual predictable projection came later: see Benveniste and Jacod [11]. The important and related topic of conditional expectation with respect to a positive Doléans measure is discussed in Jacod [87].

The fact that if N is the counting measure of a Poisson point distribution, then $N(A)$ follows a Poisson law, is classic: see for example Kingman [107]. The important characterization theorem (for example Jacod [89, Théorème 3.34]) establishing a connection between Poisson point measures and the dual predictable projection is due to Kunita and Watanabe [117]; generalizations can be found in Meyer [148] and Grigelionis [63].

It was Bremaud [18] who first showed the importance of martingale theory in the study of point processes; see Grigelionis [64], and the review article by Bremaud and Jacod [19] for a more comprehensive bibliography on the subject. Also see Jacod [87], Kabanov et al. [99], Shiryaev [183] for further treatment.

The notion of local characteristics of a semimartingale, which has become extremely important within the last 30 years is based on the study of random measures. It was introduced by Grigelionis [62], then systematized by Jacod [89]. Contributions in the form of characterization theorems were also made by Boy [17], Skorokhod [189], Meyer [144], Jacod [89], and finally Grigelionis [65].

7. Other Types of Stochastic Integrals

Notably, there are several formulations of the stochastic integral, and not all of them generate the same functions under all conditions. There are also the Fisk-Stratonovich integral (Fisk [55] and Stratonovich [190]), the McShane integral (McShane [134,135], Protter [169]); for a comprehensive treatment, see Friz and Hairer [57]) or [174], and the cylindrical stochastic integration (Mikulevicius and Rozovskii [159,160]). In addition, there is a theory for non-predictable integrands (see, for instance, Shreve and Karatzas [186], Gawarecki and Mandrekar [60], or Revuz and Yor [176], which treats the case of a continuous integrator, or Liptser and Shiryaev [129] and Rüdiger [180] for more general cases).

Another alternative to the Itô integral goes back to Lyons [131] and is, for example, presented by Dudley and Norvaiša [45], Lyons and Qian [130], Lyons et al. [132]. Their approach allows to view the stochastic integral as a pathwise Young integral with process of infinite variation (by means of the p-variation norm and without using martingales) and is referred to as processes with rough paths.

Ayed and Kuo [5], Hibino et al. [73], Hwang et al. [77], Jornet [98], Kuo [119], Kuo et al. [120] developed a stochastic integral that permitted non-adapted integrands in their publications.

Toh and Chew [192] presents a different approach to stochastic integration that yields the same integral.

The integral which is today known as Itô-Kurzweil-Henstock integral uses a Riemann approach based on a non-uniform mesh. It was introduced independently by J. Kurzweil and R. Henstock in the 1950s to study the classical deterministic integral [72,121].

In the papers [22,94,113,177], the authors explore various aspects of stochastic integration with respect to cylindrical Lévy processes. These papers collectively contribute to the development of stochastic integration theory, providing significant insights and methodologies for dealing with cylindrical Lévy processes in infinite-dimensional spaces.

Brzeźniak et al. [22] provide a comprehensive framework for stochastic integration with respect to cylindrical Lévy processes in Banach spaces. They establish integrability conditions and apply their findings to stochastic partial differential equations (SPDEs).

Jakubowski and Riedle [94], Riedle [177] introduce a new approach to stochastic integration with respect to cylindrical Lévy processes in Hilbert spaces. They develop a stochastic integral for random integrands valued in the space of Hilbert-Schmidt operators, without requiring moment or boundedness conditions on the integrands or the integrator. This work lays the foundation by establishing an adapted semi-martingale with càdlàg trajectories in a Hilbert space. Kosmala and Riedle [113] extend the theory by introducing a stochastic integral with respect to cylindrical Lévy

processes with finite p-th weak moments for $p \in [1, 2]$. They focus on integrands consisting of p-summing operators between Banach spaces of martingale type p. Their work applies the integration theory to stochastic evolution equations driven by cylindrical Lévy processes, ensuring the existence and uniqueness of solutions under certain conditions.

Another stochastic calculus framework is presented by Aleš Černý and Johannes Ruf in their paper [193]. This framework aims to simplify and unify the treatment of predictable transformations of semimartingales, including changes of variables, stochastic integrals, and their compositions. By introducing the concept of representing a semimartingale Y through another semimartingale X using a predictable function ξ, the authors establish a measure-invariant representation where the increments of Y can be expressed in terms of the increments of X. The framework accommodates both real-valued and complex-valued semimartingales.

8. Stochastic Integration in Textbooks

Itô's approach to stochastic integration primarily focuses on Brownian motions as integrators and càglàd integrands. This method remains the most common approach in most university courses and textbooks, such as those by Ash et al. [3], Bingham and Kiesel [16], Choe et al. [25], Ebenfeld [48], Elliott and Kopp [50], Klenke [109], Musiela and Rutkowski [163], and Øksendal [202]. This stochastic integral provide a foundation for most applications in biology, physics, and finance. Notably, Øksendal [202] and Musiela and Rutkowski [163] introduce a multidimensional variant of the classical Itô integral for Brownian motions.

Expanding beyond Brownian motion, some authors incorporate Poisson processes and Lévy processes into the framework. Tankov [191] and Applebaum [2] define the stochastic integral not only for Brownian motions but also for Poisson random measures, which facilitates the integration of Lévy processes, which are crucial for modelling jumps in financial markets.

Several works extend the stochastic integral to continuous local martingales. Chung and Williams [28] presents the stochastic integral for continuous local martingales using integration with respect to random measures, and furthermore extends the set of potential integrands beyond predictable ones. Additionally, Durrett [46], Émery [54], Irle [78], Kallianpur [103], Kuo [118], Meintrup and S. [137], Revuz and Yor [176], Shreve and Karatzas [186], and Le Gall et al. [123] use the Itô approach to define the integral for (locally) square-integrable martingales or slightly weaker cases like continuous martingales. However, they do not fully extend these results to define the stochastic integral for semimartingales and tend to omit the fundamental theorem of local martingales, which simplifies the extension from locally square-integrable martingales to semimartingales.

Although Émery [54] provides an appendix discussing the general semimartingale integral, the proofs are omitted. Similarly, Dellacherie and Meyer [38] offers a rigorous treatment of semimartingale theory and stochastic integration but focuses more on the probabilistic potential side.

In some texts, like Gīhman and Skorohod [61], Kallenberg [102], the stochastic integral is first defined for continuous local martingales and semimartingales before being extended to discontinuous cases, using the classical approach for one-dimensional integrals. Similarly, Yan [199], Klebaner [108], and Irle [78] follow the classical approach, starting with the Itô integral for Brownian motions and etend it to local martingales and semimartingales. However, they often only sketch or omit proofs.

Some works, such as Shiryaev [185], first define the integral as the Itô integral and later extend it to semimartingales using the classical method. However, they restrict the set of integrands to locally bounded predictable integrands. While the general vector-valued integral is mentioned, no proofs are provided.

Similar to the historical development, Rogers and Williams [178] defines the general stochastic integral by starting with Itô processes and extending to continuous local martingales, martingales, and finally semimartingales. Unlike earlier works, Rogers uses the fundamental theorem of local martingales for semimartingale integration. This book also touches on proof strategies related to the functional analytic approach, though the primary focus remains classical.

Cohen and Elliott [29] provides a detailed discussion of general predictable vector-valued integrands for general semimartingale integrators, an area only a few sources cover. This book extends beyond the traditional scope to present more advanced applications of stochastic integration with semimartingales.

Also, for a general perspective, He et al. [71] define the stochastic integral for predictable processes with respect to semimartingales directly, without introducing the Itô isometry first. However, it does not cover the vector-valued integral. Kannan and Lakshmikantham [104] follow a similar approach. A detailed treatment of general semimartingale theory is covered in Jacod and Shiryaev [92], which extensively discusses limit theorems and the full semimartingale framework.

An alternative approach to the classical treatment is provided by Medvegyev [136], who defines the stochastic integral for general semimartingales by decomposing a local martingale into its continuous and discontinuous components, treating the integral separately for the discontinuous part. For a functional analytic perspective, Bichteler [14] develops a comprehensive L^p theory for semimartingales, covering similar topics in stochastic integration.

Infinite-dimensional stochastic integration is treated by Métivier [140] and Métivier and Pellaumail [142], who introduce a "control process" (unrelated to stochastic control). This approach provides an equivalent definition of the stochastic integral in finite dimensions and extends it to a broader class of integrals in infinite dimensions, a topic also explored by Gyöngy and Krylov [68], Kunita [116], Walsh [194]. The theory of stochastic integration in infinite dimensions, particularly in the context of stochastic partial differential equations, is further discussed in Gawarecki and Mandrekar [59], which extends the classical theory to infinite-dimensional processes.

While most textbooks on stochastic integration follow the classical approach, a few exceptions explore the functional analytic method. Protter [172] and Bichteler [15] take a different path. Though the functional analytic approach is mentioned in several textbooks, such as Rogers and Williams [178], where it is used to justify certain results, detailed development is often limited to appendices, as seen in Émery [54]. Bichteler [14] and Protter [172] provide more in-depth treatments of this approach, particularly with respect to general semimartingales.

Additionally, textbooks such as Grigoriu [67], Koller [110], Mürmann [162], and Prigent [168] closely follow Protter's approach, although they often omit most proofs and detailed explanations.

In Karandikar and Rao [106], the classical integral for Brownian motions is first defined, and then the integral is extended via the functional analytic approach as introduced by Protter [172].

Rüschendorf [182] employs the random measures approach to define the integral for general semimartingales in one dimension. This approach is straightforward and easy to understand, although the book contains numerous typos.

Rao [174] provides a comprehensive introduction of the general theory of stochastic processes, with a focus on aspects that underpin the study of stochastic integration. This text delves into the theoretical framework of stochastic processes, including semimartingales and martingales.

9. The Topological Stochastic Vector Integral

In this section, we introduce a novel approach to defining the stochastic vector integral through the topology of semimartingales. Traditional approaches to stochastic integration, such as the classical and the extension from càglàd to predictable integrators in the functional analytic approach, rely heavily on the decomposition of semimartingales. However, by utilizing the semimartingale topology, we can bypass this decomposition, allowing for a direct and streamlined definition of the integral. This definition of the stochastic integral is essentially new but builds on ideas and results from the works of Chou et al. [26], [90], [138], Bichteler [14], Bichteler [15], Shiryaev and Cherny [184], Assefa and Harms [4] and Karandikar and Rao [105].

9.1. Semimartingales

Definition 33. *(a)   For a d-dimensional stochastic process*

$$X = (X^1, \ldots, X^d),$$

we put

$$X_t^* = \max_{i \in \{1,\ldots,d\}} \sup_{s \in [0,t]} \left| X_s^i \right| \quad \text{and} \quad X^* = \sup_{t \geq 0} X_t^*.$$

(b)   A d-dimensional stochastic process $H : \Omega \times \mathbb{R}_+ \to \mathbb{R}^d$ is called simple predictable, if it can be represented as

$$H_t(\omega) = H_0(\omega) \mathbb{1}_{\{0\}}(t) + \sum_{i=1}^n H_i(\omega) \mathbb{1}_{\rrbracket T_i, T_{i+1} \rrbracket}(\omega, t), \tag{13}$$

where $n \in \mathbb{N}, (T_i)_{i=1,\ldots,n+1}$ are stopping times with $0 = T_1 \leq T_2 \leq \ldots \leq T_{n+1} < \infty$ and $H_i = (h_i^1, \ldots, h_i^d)^\top$ are $\mathcal{F}_{T_i}$-measurable $\mathbb{R}^d$-valued random variables that are almost surely bounded in all components for all $i = 1, \ldots, n$.
The set of $\mathbb{R}^d$-valued simple predictable processes will be denoted by Λ^d.

(c)   For $H \in \Lambda^d$, we define

$$\|H\|_u := \operatorname{ess\,sup}_{\omega \in \Omega} \max_{j \in \{1,\ldots,d\}} \max_{i \in \{1,\ldots,n\}} \left| h_i^j \right| = \operatorname{ess\,sup}_{\omega \in \Omega}(H^*)$$

and Λ_u^d denotes the set Λ^d equipped with the topology of uniformly convergence on $\mathbb{R} \times L^\infty(\Omega, \mathcal{F}, \mathbf{P}, \mathbb{R}^d)$, that means $H^n \to H$ in Λ_u^d, if $H, H^n \in \Lambda^d$ for all $n \in \mathbb{N}$ and we have $\|H^n - H\|_u \to 0$ for $n \to \infty$.

Remark 22. *A stochastic process is a d-dimensional simple predictable process if and only if each component is a one-dimensional simple predictable process.*

Remark 23. *For each predictable process H, there exists a sequence of simple predictable processes H^n that converges pointwise to a process that is indistinguishable from H. This is a simple consequence of the monotone class theorem (see, for example, [29, Corollary 7.4.3]).*

The definition of the stochastic integral for simple predictable is quite intuitive.

Definition 34. *Let X be an $\mathbb{R}^d$-valued stochastic process with càdlàg paths and $H \in \Lambda^d$ as the one given in (13). We define the stochastic integral $J_X : \Lambda^d \to \mathbb{D}$ by*

$$J_X(H) := H \bullet X := H_0^\top X_0 + \sum_{i=1}^n H_i^\top (X^{T_{i+1}} - X^{T_i}). \tag{14}$$

The processes H and X are called integrand and integrator.

For simplicity, we will treat processes and their equivalence classes interchangeably. Consequently, the subsequent metric spaces should be viewed as quotient spaces.

Definition 35. *(a)   For any left or right continuous process X, we put*

$$\|X\|_{\mathrm{ucp}} := \sum_{n=1}^\infty \frac{1}{2^n} \mathbf{E}[X_n^* \wedge 1]. \tag{15}$$

Then the topology induced by the metric $d_{\mathrm{ucp}}(X, Y) := \|X - Y\|_{\mathrm{ucp}}$ is the topology of ucp convergence.

(b)   The Émery norm of a càdlàg process X is defined as

$$\|X\|_{\mathcal{S}} := \sup_{\substack{K \in \Lambda^d \\ \|K\|_u \leq 1}} \|K \bullet X\|_{\mathrm{ucp}}.$$

The topology induced by the metric $d_{\mathcal{S}}(X, Y) := \|X - Y\|_{\mathcal{S}}$ is called the semimartingale topology or Émery topology.

(c)   For sequences $(H^n)_{n \in \mathbb{N}}$, we write $H^n \xrightarrow{\mathrm{ucp}} H$ if it converges to a process H in the topology of ucp convergence. We use $X^n \xrightarrow{\mathcal{S}} X$ analogously.

Remark 24. *Even though the name suggests otherwise, the Émery norm is not a norm but only a metric (see, for example, [29] page 278 for elaborations).*

Remark 25. *A sequence of processes $(X^n)_{n \in \mathbb{N}}$ converges in the ucp topology to a process X if and only if we have*

$$\mathbf{P}\big((X^n - X)_t^* > K\big) \to 0 \quad \text{as } n \to \infty \text{ for each } t, K > 0.$$

This explains the name ucp (uniformly on compacts in probability) convergence.

Remark 26. *In the definition of X_t^*, the supremum is taken over the uncountable set $[0, t]$. This means that it might not be measurable for arbitrary processes. However, by limiting ourselves to either left-continuous or right-continuous processes, the supremum can be considered over the countable set of rational times, ensuring measurability. Additionally, it has been demonstrated that the supremum is measurable for jointly measurable processes, provided the probability space is complete. For further details and an example of a process that "behaves badly", refer to Pollard et al. [167, page 214].*

In most publications, a semimartingale is defined to be a process which can be written as the sum of a local martingale and an FV-process. In some publications, a semimartingale is defined as a process under which the integral operator is continuous. We give a different definition and will see later that all of these definitions are equivalent. For clarity, we will use the term topological semimartingale to highlight the different definition.

Definition 36. *An $\mathbb{R}^d$-valued stochastic process $X = (X_t)_{t \geq 0}$ is called a topological semimartingale if:*

(a) X is càdlàg and adapted,

(b) $\|X\|_{\mathcal{S}} < \infty$, and

(c)   for every sequence $(\lambda_n) \subset \mathbb{R}$ with $\lambda_n \to 0$, we have $\|\lambda_n X\|_{\mathcal{S}} \to 0$.

The set of all d-dimensional topological semimartingales is denoted by $\mathcal{S}^d$, and we write $\mathcal{S}_e^d$ for the space $\mathcal{S}^d$ endowed with the semimartingale topology.

Remark 27. *The condition that $\|\lambda_n X\|_{\mathcal{S}} \to 0$ for every sequence $\lambda_n \to 0$ expresses the continuity of scalar multiplication at 0 in the semimartingale topology. This ensures that $\mathcal{S}_e^d$ becomes a topological vector space. It also highlights that the semimartingale topology is not induced by a norm — otherwise, this continuity would be automatic. In particular, the mapping $(\lambda, X) \mapsto \lambda X$ is not jointly continuous in the product topology when $\| \cdot \|_{\mathcal{S}}$ is used as a seminorm.*

In the functional analytic approach, as presented by Protter [172], a semimartingale is defined as a "good integrator", which means a process for which the integral operator is continuous with respect to certain topologies. The next theorem shows that this definition is equivalent to ours.

Theorem 31. *Let X be an adapted càdlàg process. Then X is a topological semimartingale if and only if for each sequence of simple predictable processes $(H^n)_{n \in \mathbb{N}}$ with $\|H^n\|_u \to 0$ the random variables $H^n \bullet X_t$ converge for each $t \geq 0$ in probability to 0.*

Proof. The key distinction between the two criteria lies in their convergence requirements: the semimartingale definition demands uniform convergence, whereas the latter criterion requires only pointwise convergence in probability. Therefore, it is sufficient to demonstrate that

$$\sup_{\substack{K \in \Lambda^d \\ \|K\|_u \leq 1}} \mathbf{E}|K \bullet X_t| = \sup_{\substack{K \in \Lambda^d \\ \|K\|_u \leq 1}} \mathbf{E}[K \bullet X]_t^*.$$

The remaining argument then follows from the linearity of both the expectation and the integral operator.

For $t > 0, \lambda > 0$ and $G \in \Lambda^d$ with $\|G\|_u < 1$, define the stopping time T and the indicator function H as

$$T = \inf\{s > 0 \,;\, |G \bullet X|_s > \lambda\} \wedge t, \quad H = \mathbb{1}_{[0,T]}.$$

Then, we have

$$\mathbf{P}(|G \bullet X|_t^* > \lambda) \leq \mathbf{P}[|G \bullet X|_T \geq \lambda] = \mathbf{P}[|GH \bullet X|_t \geq \lambda].$$

By basic probability theory, we have $\mathbf{E}[K \bullet X]_t^* = \int_0^\infty \mathbf{P}(|K \bullet X|_t^* > \lambda)\mathrm{d}\lambda$ and since $GH \in \Lambda^d$ and $\|GH\|_u \leq 1$, we get

$$\sup_{\substack{K \in \Lambda^d \\ \|K\|_u \leq 1}} \mathbf{E}[K \bullet X]_t^* = \sup_{\substack{K \in \Lambda^d \\ \|K\|_u \leq 1}} \int_0^\infty \mathbf{P}(|K \bullet X|_t^* > \lambda)\mathrm{d}\lambda$$

$$\leq \sup_{\substack{K \in \Lambda^d \\ \|K\|_u \leq 1}} \int_0^\infty \mathbf{P}(|K \bullet X|_t \geq \lambda)\mathrm{d}\lambda$$

$$= \sup_{\substack{K \in \Lambda^d \\ \|K\|_u \leq 1}} \mathbf{E}|K \bullet X|_t.$$

The converse inequality is trivial. □

Remark 28. *Theorem 31 states that semimartingales are the càdlàg processes for which the integral operator is continuous. It can be shown that different types of convergence could be chosen in Theorem 31 that would yield an equivalent definition of semimartingales (see [106, Theorem 5.89]).*

9.2. Properties of Semimartingales

Theorem 32. *(a)   An $\mathbb{R}^d$-valued càdlàg process X is a d-dimensional topological semimartingale if and only if all components are one-dimensional topological semimartingales.*

1.   *Local topological semimartingales and processes that are prelocally topological semimartingales are topological semimartingales.*

(b)   Let $\mathbf{Q}$ be a probability measure that is absolutely continuous with respect to $\mathbf{P}$. Then every $\mathbf{P}$-semimartingale is also a $\mathbf{Q}$-semimartingale.

Proof. (a)   If all components are topological semimartingales, the result follows from the triangle inequality. For the converse, we note that since for $K^i \in \Lambda^1$ with $\|K^i\|_u \leq 1$ we have $\tilde{K}^i := (0, \ldots, 0, K^i, 0, \ldots 0) \in \Lambda^d$ with $\|\tilde{K}^i\|_u \leq 1$. Therefore, we get for $\lambda_n > 0$

$$\sup_{\substack{\tilde{K} \in \Lambda^1 \\ \|\tilde{K}\|_u \leq 1}} \left\| \tilde{K} \bullet (\lambda_n X^i) \right\|_{\mathrm{ucp}} \leq \sup_{\substack{K \in \Lambda^d \\ \|K\|_u \leq 1}} \|K \bullet (\lambda_n X)\|_{\mathrm{ucp}}.$$

(b)   We use Theorem 31. Let $(H^n)_{n \in \mathbb{N}}$ be a sequence with $\|H^n\|_u \to 0$. We have to show $H^n \bullet X^t \to 0$ in probability. Let X be a local topological semimartingale with $(T^k)_{k \in \mathbb{N}}$ being a localizing sequence. We define

$$R_n^t = \begin{cases} T^n & T^n \leq t \\ \infty & T^n > t. \end{cases}$$

Now we have

$$\mathbf{P}(|H^n \bullet X_t| \geq \varepsilon) \leq \underbrace{\mathbf{P}\left(\left| H^n \bullet X_t^{T^n} \right| \geq \varepsilon \right)}_{\to 0 \text{ by assumption}} + \underbrace{\mathbf{P}\left(R_n^t < \infty \right)}_{=\mathbf{P}(T^n \leq t) \to 0 \text{ since } T^n \to \infty}.$$

Therefore, X is a topological semimartingale. If X is prelocally a topological semimartingale, then the proof proceeds analogously.

(c)   Since convergence in probability in $\mathbf{P}$ implies convergence in probability in $\mathbf{Q}$, this is obvious.
        □

Historically, semimartingales have been defined as the set of processes that can be written as the sum of a local martingale and an FV-process. In the functional analytical approach, semimartingales have been defined as the processes for which the integral operator is continuous (see Theorem 31). In fact, these definitions are equivalent, which was proved in the late 1970s independently of each other by Klaus Bichteler [13,14] and Claude Dellacherie with contributions by Paul-André Meyer and Gabriel Mokobodzki [36]. In addition to the literature mentioned above, the theorem is also discussed and proven in modern textbooks such as in Cohen and Elliott [29, Theorem 12.3.26], in Karandikar and Rao [106, Theorem 5.59], and in Protter [172, Theorem III.47]. While the latter two employ classical methods, the former adopts an approach by Beiglböck, who has presented alternative proofs in Beiglböck and Siorpaes [7], Beiglböck et al. [8,9].

As the topological stochastic integral does not rely on the decomposition of semimartingales, the theory does not require the Bichteler-Dellacherie-Mokobodzki. However, for the sake of completeness and because it provides many examples of semimartingales, we present it here.

Theorem 33 (Bichteler-Dellacherie-Mokobodzki Theorem). *Let X be a càdlàg process. The following are equivalent:*

(i) X is a topological semimartingale;

(ii) X is a classical semimartingale in each component;

(iii) X is a good integrator in each component.

Before we proceed with the proof of Theorem 33, we will demonstrate two inequalities that are useful when dealing with the topological stochastic integral. The first one, a variant of the Itô inequality, shows that square-integrable martingales are semimartingales. The second, based on the Burkholder inequality, demonstrates the same for "ordinary" martingales. While the latter is sufficient for our purposes here, we will present both for the sake of completeness.

Theorem 34. *Let M be a d-dimensional càdlàg process and $H \in \Lambda^d$.*

(a)   If M is a square-integrable martingale, then we have

$$\mathbf{E}\left[(H \bullet M_t)^2 \right] \leq d \|H\|_u^2 \mathbf{E}\left[\|M_t\|_{\mathbb{R}^d}^2 \right].$$

(b)   If M is a martingale, then the following inequality holds:

$$\alpha \mathbf{P}[(H \bullet X)_t^* \geq \alpha] \leq 18 \|H\|_u d \sum_{j}^{d} \mathbf{E}\left[\left| X_t^j \right| \right].$$

Proof. (a) By Lemma 1, an application of the tower rule and the Cauchy-Schwarz inequality, we get

$$
\mathbf{E}\left[(H \bullet M_t)^2\right] = \mathbf{E}\left[\left(H_0^\top M_0 + \sum_{i=1}^n H_i^\top \left(M_t^{T_{i+1}} - M_t^{T_i}\right)\right)^2\right]
$$

$$
= \mathbf{E}\left[\left(H_0^\top M_0\right)^2 + \sum_{i=1}^n \left(H_i^\top \left(M_t^{T_{i+1}} - M_t^{T_i}\right)\right)^2\right]
$$

$$
\leq \mathbf{E}\left[\|H_0\|_{\mathbb{R}^d}^2 \|M_0\|_{\mathbb{R}^d}^2 + \sum_{i=1}^n \|H_i\|_{\mathbb{R}^d}^2 \left\|M_t^{T_{i+1}} - M_t^{T_i}\right\|_{\mathbb{R}^d}^2\right]
$$

$$
\leq d\|H\|_u^2 \mathbf{E}\left[\|M_0\|_{\mathbb{R}^d}^2 + \sum_{i=1}^n \left\|M_t^{T_{i+1}} - M_t^{T_i}\right\|_{\mathbb{R}^d}^2\right]
$$

$$
= d\|H\|_u^2 \mathbf{E}\left[\|M_t\|_{\mathbb{R}^d}^2\right]
$$

(b) The idea is to transform the time-continuous statement into a time-discrete one and then apply Theorem 5. We first assume $d = 1$ and $\sup_t |H| < 1$. Since $H \in \Lambda$, there is a representation

$$
H = H_0 \mathbb{1}_{\{0\}}(t) + \sum_{i=1}^n H_i \mathbb{1}_{\rrbracket T_i, T_{i+1}\rrbracket},
$$

with stopping times T_i. By the tower rule and Doob's optional stopping theorem, one easily sees that $(M_{T_i}^T)_{i\in\mathbb{N}}$ is a time-discrete martingale for any stopping time T.

By the definition of the stochastic integral for simple predictable integrands, we obtain $H \bullet M_{T_i} = H * M_i$, where $*$ describes the time-discrete stochastic integral as in Theorem 5. An application of Theorem 5 yields the result for $d = 1$. For general d and H, we obtain

$$
\alpha \mathbf{P}[(H \bullet M)_t^* \geq \alpha] = \alpha \mathbf{P}\left[\left|\frac{H}{\|H\|_u} \bullet M\right|_t^* \geq \frac{\alpha}{\|H\|_u}\right]
$$

$$
\leq \alpha \sum_{j=1}^d \mathbf{P}\left[\left|\frac{H^j}{\|H\|_u} \bullet M^j\right|_t^* \geq \frac{\alpha}{d\|H\|_u}\right]
$$

$$
\leq 18\|H\|_u d \sum_{j=1}^d \mathbf{E}\left[|M_t^j|\right].
$$

□

We now outline the proof of the Bichteler-Dellacherie-Mokobodzki Theorem.

Proof of Theorem 33. We first assume $X = M + A$ with a local martingale M and an FV-process A. For the FV-process A with $A_0 = 0$, we have

$$
\mathbf{Var}[H \bullet A] = \mathbf{Var}\left[\sum_{i=1}^n H_i^\top \left(A^{T_{i+1}} - A^{T_i}\right)\right]
$$

$$
\leq \sum_{i=1}^n \sum_{j=1}^d \left|h_i^j\right| \mathbf{Var}\left[A^{j,T_{i+1}} - A^{j,T_i}\right]
$$

$$
\leq \|H\|_u \sum_{j=1}^d \mathbf{Var}\left[A^j\right].
$$

For a series H^n with $\|H^n\|_u \to 0$, we see that $H \bullet A_t$ converges in probability to 0. Thus, A is a semimartingale by Theorem 31. For any martingale M, we can proceed analogously by applying Theorem 34 and see that M is a semimartingale. Via localization, we get that any local martingale is

locally a semimartingale; thus, by Theorem 32, any local martingale is a semimartingale. By Theorem 32 $M+A$ is a semimartingale as well.

The converse is elegantly proven in [7]. The proof can be readily adapted to our context with minimal changes and does not require the definition or properties of the stochastic integral. Therefore, we will not repeat it here; instead, we will refer to the original article. $\square$

Remark 29. *By employing the Fundamental Theorem of Local Martingales, we can assume without loss of generality that any semimartingale can be decomposed into $X = M + A$, where A is a finite variation process, and M is a locally bounded local martingale. Therefore, the first inequality in Theorem 34 would suffice to demonstrate one direction of the Bichteler-Dellacherie-Mokobodzki Theorem.*

Example 1. • *Since a Brownian motion is a martingale, it is also a semimartingale.*
- *A counting process (and hence a Poisson process)[2] is an FV-process and hence a semimartingale. The compensated Poisson Process[3] is a martingale and thus also a semimartingale.*
- *Any Lévy process[4] can be written as a sum of a local martingale and an FV-process[5]. Hence, each Lévy process is a semimartingale.*

Example 2. *Probably the most prominent example of a stochastic process that is not a semimartingale is the fractional Brownian motion B^H with Hurst parameter $H \in (0,1) \setminus \left\{ \frac{1}{2} \right\}$ (see, for example, [12] for a comprehensive treatment). The fact that a fractional Brownian motion cannot be a semimartingale for $H \neq \frac{1}{2}$ has been proved by several authors (see, for example, Liptser and Shiryaev [128, Example 2 of Section 4.9.13], [127, Corollary 2.2] or Rogers [179, section 2]). However, these authors use the classical definition of semimartingales. With our definition, it is possible to present an alternative and short proof for $H < \frac{1}{2}$.*

Let X be a fractional Brownian Motion on the interval $[0,T]$ with $X_0 = 0$. Furthermore, let π be the set of partitions of the interval $[0,T]$.

For a partition $\pi_0 = \{t_0 = 0, \ldots, t_n = T\} \in \pi$ and $N \in \mathbb{N}$, we define

$$H^{\pi_0,N} = \sum_{j=0}^{n-1} -\mathbb{1}_{\{|X_{t_j}| \leq N\}} \frac{X_{t_j}}{N} \mathbb{1}_{(t_j,t_{j+1}]}.$$

Then $H^{\pi_0,N} \in \Lambda^1$ with $\|H^{\pi_0,N}\|_u \leq 1$ and for $\lambda > 0$ on the set

$$A^{\pi_0,N,\lambda} := \{X_T^* \leq N\} \cap \left\{ \sum_{j=0}^{n-1} (X_{t_{j+1}} - X_{t_j})^2 > \frac{2N}{\lambda} + N^2 \right\}$$

we have

$$H^{\pi_0,N} \bullet (\lambda X)_T = -\frac{\lambda}{N} \sum_{j=0}^{n-1} X_{t_j}(X_{t_{j+1}} - X_{t_j})$$

$$= \frac{\lambda}{N} \sum_{j=0}^{n-1} X_{t_j}^2 - X_{t_j} X_{t_{j+1}}$$

$$= \frac{\lambda}{2N} \sum_{j=0}^{n-1} (X_{t_{j+1}} - X_{t_j})^2 - X_T^2$$

$$> \frac{\lambda}{2N} \left(\frac{2N}{\lambda} + N^2 - N^2 \right) = 1.$$

---

[2] see Cohen and Elliott [29] Definition 5.5.16
[3] see Cohen and Elliott [29] Theorem 5.5.18
[4] see Cohen and Elliott [29] Definition 13.5.1
[5] This was initially shown in Lévy [126], see also Itô [85] for an English proof, Kunita and Watanabe [114] for a proof in the classical semimartingale setting or Applebaum [2] Proposition 2.7.1 for a modern textbook treatment.

Therefore, $\mathbf{P}(H^{\pi_0,N} \bullet (\lambda X)_T > 1) \geq \mathbf{P}(A^{\pi_0,N,\lambda})$ *and it remains to show that* $\mathbf{P}(A^{\pi_0,N,\lambda})$ *is greater than some positive constant independent of* λ *for feasible* π_0 *and* N.

It is well known that the quadratic variation of the fractional Brownian Motion with $H < \frac{1}{2}$ is unbounded in probability (see, for example, [51, Lemma 4.1.2]). That means that

$$c := \limsup_{\lambda \to 0} \sup_{\pi} \mathbf{P}\left(\sum_{j=0}^{n-1} \left(X_{t_{j+1}} - X_{t_j} \right) > \frac{1}{\lambda} \right)$$

exists and is greater than 0.

In particular, that means for each $N \in \mathbb{N}$, we can select a partition $\pi_0 = (t_0, \dots, t_n) \in \pi$ such that

$$\mathbf{P}\left(\sum_{j=0}^{n-1} \left(X_{t_{j+1}} - X_{t_j} \right)^2 > \frac{2N}{\lambda} + N^2 \right) > \frac{c}{2}. \tag{16}$$

Since X is continuous, we can define $N \in \mathbb{N}$ in a way such that

$$\mathbf{P}(X_T^* > N) < \frac{c}{4}. \tag{17}$$

With Equation (17) and Equation (16), we conclude

$$\mathbf{P}\left(A^{\pi,N,\lambda} \right) > \frac{c}{4} > 0.$$

Example 3. *Following a similar strategy as above, one can show that for a Brownian motion B,* $|B|^\alpha$ *with* $0 < \alpha < 1$ *is also not semimartingales. A proof of this fact in the classical semimartingale framework can be found in Wang [195] or Yor [201].*

9.3. The General Stochastic Integral

To extend the integral, we are interested in the topological properties of the above-defined spaces. These were initially studied by Émery [52,53] and Memin [138]. We note that the Émery norm was initially defined to take the supremum over all bounded predictable processes and not only simple predictable processes. But it was shown in Memin [138] that these definitions are equivalent.

Definition 37. *For* $X \in \mathcal{S}^d$ *and* $H \in \Lambda^d$, *we put* $\|H\|_X := \|H \bullet X\|_{\mathcal{S}}$. *Then* $d_X(H,K) := \|H - K\|_X$ *is a metric on* Λ^d. *The corresponding topological space is denoted by* Λ_X^d *and for a sequence* $(H^n)_{n \in \mathbb{N}}$ *we will write* $H^n \xrightarrow{X} H$ *if* H^n *converges to a process* H *with respect to* $\| \bullet \|_X$.

Remark 30. *We re-emphasize that we work with equivalence classes of processes. Specifically, the metric* $d_X(H,K) := \|H - K\|_X$ *is defined on the quotient space* $\Lambda^d / \mathcal{N}_X$, *where* $\mathcal{N}_X = \{K \in \Lambda^d ; \|K\|_X = 0\}$.

Theorem 35. *Let* $H^n \in \Lambda^d$ *and* $X^n \in \mathcal{S}^d$ *for all* n.

(a) *If* H^n *converges to* H *in the ucp topology, then it also converges to* H *with respect to the topology induced by the metric* $\|\cdot\|_X$. *This implies that the ucp topology is finer than the topology induced by* $\|\cdot\|_X$.

(b) *If* X^n *converges to* X *in the semimartingale topology, then it also converges to* X *in the ucp topology. This means that the semimartingale topology is finer than the ucp topology.*

Proof. (a) Let $t \in \mathbb{R}_+$ be fixed and $\varepsilon > 0$. Since X is a semimartingale, we can always find a $c > 0$ such that $\|cX\|_{\mathcal{S}} \leq \varepsilon$. Without loss of generality, we assume $H^n \xrightarrow{ucp} 0$ and hence, we can always find a $N \in \mathbb{N}$ such that $\mathbf{P}((H^n)_t^* > c) \leq \varepsilon$ for all $n \geq N$.

For $K \in \Lambda^1$, it is easy to check that we have

$$K \bullet (H^n \bullet X)_t = (KH^n) \bullet X_t.$$

We obtain for all $n \geq N$ and $K \in \Lambda^1$ with $\|K\|_u \leq 1$

$$
\begin{aligned}
\mathbf{P}\big((K \bullet (H^n \bullet X))_t^* > \varepsilon\big) &= \mathbf{P}\big((KH^n \bullet X)_t^* > \varepsilon\big) \\
&= \mathbf{P}\big((KH^n \bullet X)_t^* > \varepsilon \cap (H^n)_t^* > c\big) \\
&\quad + \mathbf{P}\big((KH^n \bullet X)_t^* > \varepsilon \cap (H^n)_t^* \leq c\big) \\
&\leq \mathbf{P}\big((H^n)_t^* > c\big) + \mathbf{P}\big(c(K \bullet X)_t^* > \varepsilon\big) \\
&\leq \varepsilon + \varepsilon
\end{aligned}
$$

Hence we have $\|H^n\|_X \leq 2\varepsilon$ for all $n \geq N$ and $K \bullet (H^n \bullet X)$ converges uniformly on compacts in probability to 0 and thus $\|H^n\|_X \to 0$.

(b)   This follows immediately with

$$
\|X\|_{\mathcal{S}} := \sup_{\substack{K \in \Lambda^d \\ \|K\|_u \leq 1}} \|K \bullet X\|_{\mathrm{ucp}} \geq \|X\|_{\mathrm{ucp}}.
$$

$\square$

Remark 31. *The converse of these statements does not hold (see, Example 4). However, for martingales M^n, M, it can be shown, that $M^n \xrightarrow{\mathcal{H}^1} M$ implies $M^n \xrightarrow{\mathcal{S}} M$ (see, for example, [26]).*

Example 4. *Let $X_t(\omega) = t$ and*

$$
H_t^n(\omega) = \mathbb{1}_{\{0\}}(t) + \sum_{k=0}^n n^2 \left(\frac{k-1}{n^2}\right)^n \mathbb{1}_{\left(\frac{k-1}{n^2}, \frac{k}{n^2}\right]}(t).
$$

Since X is increasing and H^n is positive and zero for all $t > 1$, we have

$$
(H^n \bullet X)_t^* \leq H^n \bullet X_1 = \sum_{k=0}^n n^2 \left(\frac{k-1}{n^2}\right)^n \frac{1}{n^2} < \int_0^1 x^n \mathrm{d}x = \frac{1}{n+1} \to 0.
$$

Hence, we have $H^n \bullet X \xrightarrow{\mathrm{ucp}} 0$ for $n \to \infty$, and one easily sees (again because of the monotony of X and since $H^n \geq 0$) that we also have $H^n \bullet X \xrightarrow{\mathcal{S}} 0$ and thus $H^n \xrightarrow{X} 0$.

But we also have

$$
(H^n)_1^* = \left(\frac{n^2-1}{n^2}\right)^n \to 1.
$$

Hence for no $\varepsilon < 1$ there exists a $N \in \mathbb{N}$ such that $(H^n)_1^* < \varepsilon$ for all $n \geq N$. Thus, we conclude $H^n \xcancel{\xrightarrow{\mathrm{ucp}}} 0$.

The following results were initially shown by Memin [138]. In our setting, we can provide a simplified proof.

Theorem 36. *The space of semimartingales is complete under the semimartingale topology.*

Proof. Let X^n be a Cauchy sequence in $\|\cdot\|_{\mathcal{S}}$. By Theorem 35, X^n converges in ucp to a càdlàg process X. For $H, H^m \in \Lambda^d$ with $\|H^m - H\|_u \to 0$, we get

$$
\begin{aligned}
\|H^m \bullet X - H \bullet X\|_{\mathrm{ucp}} &\leq \|H^m \bullet X^n - H \bullet X^n\|_{\mathrm{ucp}} + \|H \bullet X^n - H \bullet X\|_{\mathrm{ucp}} \\
&\to \|H \bullet X^n - H \bullet X\|_{\mathrm{ucp}} \quad (m \to \infty).
\end{aligned}
$$

Since X^n converges to X with respect to the semimartingale topology, the right-hand side converges to 0 for $n \to \infty$. Therefore, $H^m \bullet X$ converges in ucp, and thus X is a semimartingale. $\square$

Semimartingales, as integrators, are the largest class of processes, such that the integral operator has some desirable properties. The integral, however, has so far only been defined for simple predictable integrands. Now we are going to increase the class of possible integrands.

Let $X \in \mathcal{S}^d$. We examine the operator J_X from Λ_X^d to $\mathcal{S}_e^d$:

$$J_X : \left\{ \begin{array}{ccc} \Lambda_X^d & \rightarrow & \mathcal{S}_e^d \\ H & \mapsto & H \bullet X \end{array} \right.$$

The function is well defined as we work with equivalence classes of functions on a quotient space.

By the definition of the corresponding metrics, it is clear that J_X is a linear isometry between topological vector spaces and hence is uniformly continuous. Therefore, by the completion of maps theorem (see, for example, Jänich [95, page 55]), we can extend the mapping continuously. From classical theory, we know that the predictability of the integrand is an essential property for ensuring desirable features, such as the stability of the local martingale property under integration with respect to locally bounded integrands. Consequently, we choose to close Λ_X^d as a subspace of the set of predictable processes.

This justifies the following definition.

Definition 38. *Let X be a d-dimensional semimartingale, $L(X) := \overline{\Lambda_X^d}$ the closure of Λ_X^d in $\mathcal{P}$ and $\overline{J}_X$ the unique, continuous extension of J_X. Then we define*

$$\overline{J}_X : \left\{ \begin{array}{ccc} L(X) & \rightarrow & \mathcal{S}_e^d \\ H & \mapsto & H \bullet X := \overline{J}_X(H) \end{array} \right. .$$

We will call $H \bullet X$ the stochastic integral of H with respect to X.

Because of the uniqueness of the extension, we conclude that the stochastic integral coincides with the stochastic integral defined via the classical or functional analytic approach.

One might wonder whether we could obtain a broader range of possible integrands by defining the integral for all processes that are only (pre)locally in $L(X)$, similar to how it was done by Meyer (see Section 3.5). However, this is not the case, as the following theorem demonstrates.

Theorem 37. *Let $H \in \mathcal{P}$ be locally in $L(X)$. Then $H \in L(X)$. In other words:*

$$L(X)_{\mathrm{loc}} = L(X).$$

We utilize the following lemma for the proof.

Lemma 9. *Let $(X^n)_{n \in \mathbb{N}}$ be a sequence in $\mathcal{S}^d$, $X \in \mathcal{S}^d$, and $(T^n)_{n \in \mathbb{N}}$ be an increasing sequence of stopping times with $T^n \to \infty$ almost surely. Assume that $(X^n)^{T^m} \to X^{T^m}$ in $\mathcal{S}^d$ for all m. Then $X^n \to X$ in $\mathcal{S}^d$.*

Proof. Assume that $X^n \nrightarrow X$ in $\mathcal{S}^d$. This implies that for any $\varepsilon > 0$ and any $N \in \mathbb{N}$, there exists an $n \geq N$ such that $\|X^n - X\|_{\mathcal{S}^d} > \varepsilon$. That means, that by the definition of the norm $\|\cdot\|_{\mathcal{S}^d}$, there exists a process $K \in \Lambda^d$ with $|K| \leq 1$ such that

$$\|K \bullet (X^n - X)\|_{\mathrm{ucp}} \geq \varepsilon. \tag{18}$$

However, we have

$$\left\| K \bullet (X^n - X)^{T^m} \right\|_{\text{ucp}} \leq \sup_{\substack{H \in \Lambda^d \\ \|H\|_u \leq 1}} \left\{ \left\| H \bullet (X^n - X)^{T^m} \right\|_{\text{ucp}} \right\}$$

$$= \left\| (X^n - X)^{T^m} \right\|_{\mathcal{S}^d}.$$

By assumption, the right-hand side of (18) tends to 0 as $m \to \infty$. This implies that there exists an $\tilde{m} \in \mathbb{N}$ such that

$$\mathbf{P}\left((K \bullet (X^n - X)^{T^m})_t^* > \varepsilon \right) < \frac{\varepsilon}{2}$$

for all $m \geq \tilde{m}$.

Since $T^m \to \infty$ almost surely, for any $\delta > 0$, there exists an $\overline{m}$ such that $\mathbf{P}(T^m < t) < \frac{\varepsilon}{2}$ for all $m \geq \overline{m}$.

Thus, for all $m \geq \max\{\tilde{m}, \overline{m}\}$, we obtain

$$\mathbf{P}((K \bullet (X^n - X))_t^* > \varepsilon) \leq \mathbf{P}\left((K \bullet (X^n - X)^{T^m})_t^* > \varepsilon \right) + \mathbf{P}(T^m < t) < \varepsilon,$$

which is a contradiction to Equation (18). Therefore, we conclude that $X^n \to X$ in $\mathcal{S}^d$. $\square$

Proof of Theorem 37. Since $H \in L(X)_{\text{loc}}$, there exists a sequence of stopping times T^n tending to infinity such that $H^{T^n} \in L(X)$.

As $H^{T^n} \in L(X)$, there exists a sequence $(H^{n,m})_{m \in \mathbb{N}}$ with $H^{n,m} \in \Lambda^d$ and $H^{n,m} \xrightarrow{X} H^{T^n}$ as $m \to \infty$, which is equivalent to $H^{n,m} \bullet X \xrightarrow{\mathcal{S}} H^{T^n} \bullet X$.

Without loss of generality, we can assume there exists a sequence $(H^{\tilde{n}})_{\tilde{n} \in \mathbb{N}}$ such that for all m

$$(H^{\tilde{n}} \bullet X)^{T^m} \xrightarrow{\mathcal{S}} (H \bullet X)^{T^m}.$$

By Lemma 9, we have $H^{\tilde{n}} \bullet X \xrightarrow{\mathcal{S}} H \bullet X$, and thus $H \in L(X)$. $\square$

Remark 32. *For any semimartingale X, the space $\{H \bullet X\}$ with $H \in L(X)$ is complete with respect to the semimartingale topology. This result was first established by Memin [138]. An alternative proof is presented in [29, Theorem A.6.17]. For the topological stochastic integral, this follows directly from the definition of the stochastic integral.*

So far, we have defined the stochastic integral in great generality, but we hardly have any information about the set $L(X)$. The following theorem provides more information.

Theorem 38. *Let $X \in \mathcal{S}^d$, then*

$$\mathbb{L} \subset \mathbf{b}\mathcal{P} \subset \mathbf{b}\mathcal{P}_{\text{loc}} \subset L(X) = L(X)_{\text{loc}} \subset \mathcal{P}.$$

Proof. By definition, $L(X)$ consists of predictable processes. A simple application of the monotone class theorem yields $\mathbf{b}\mathcal{P} \subset L(X)$ (a fact that is also shown in the proof of Theorem A1). With Theorem 37, we get $\mathbf{b}\mathcal{P}_{\text{loc}} \subset L(X)$. $\square$

Remark 33. *Usually, we even have $\mathbb{L} \subsetneq L(X) \subsetneq \mathcal{P}$. In Example 5, we provide examples of predictable processes that are not in $L(X)$ and demonstrate that there are processes in some $L(X)$ that are not càglàd processes.*

Remark 34. *Corollary 1 and Theorem A4 will provide a more detailed understanding of $L(X)$.*

Remark 35. *Even though, for a given semimartingale X, there are typically predictable processes that do not belong to $L(X)$, every predictable process can be approximated pointwise almost surely by simple predictable*

processes (see Remark 23), and therefore by elements of $L(X)$. However, such pointwise approximation does not, in general, imply that the limit process lies in $L(X)$, as integrability with respect to X is not preserved under pointwise convergence.

The following theorem will be instrumental in the proof of the properties of the stochastic integral. As, at this stage, we do not have the basic properties at hand, we need to resort to a different proof than the one that is used in the standard literature. Since it is a long and technical proof, we have included it in the appendix.

Theorem 39 (Dominated Convergence Theorem). *Let $X \in \mathcal{S}^d$ be a semimartingale, and let $H^m = (\tilde{H}^{1,m}, \dots \tilde{H}^{d,m}) \in \mathcal{P}$ be processes converging a.s. to a limit H. If there exists a process $G = (\tilde{G}^1, \dots, \tilde{G}^d)$ with $\tilde{G}^i \in L(X)$ for all $i \in \{1, \dots, d\}$ such that $\left|\tilde{H}^{i,m}\right| \le G^i$ for all $i \in \{1, \dots, d\}$ and all $m \in \mathbb{N}$, then H^m, H are in $L(X)$ and $H^m \bullet X$ converges to $H \bullet X$ in the semimartingale topology (and hence also in ucp).*

The proof can be found in Section A.

Remark 36. *The condition requiring G to be componentwise integrable is crucial, as demonstrated in the example in Example 9. This is because one cannot simply reduce the values of a component without potentially compromising integrability: higher values in one component might be necessary to offset high values in other components, as will be illustrated in Example 8*

The dominated convergence theorem has a useful corollary that can help comprehend the space of integrands more thoroughly.

Corollary 1. *Let X be a semimartingale and $(H^n)_{n \in \mathbb{N}}$ be a sequence of uniformly locally bounded predictable processes, that is, for a given sequence $(T^n)_{n \in \mathbb{N}}$ we have $\sup_m \left|(H^m)^{T^n}\right| < K_n$, for some constants K_n. Suppose that $H^n_t \to H_t$ almost surely for every t. Then $H \in L(X)$ and $H^n \xrightarrow{X} H$.*

As there is no version of the Dominated Convergence Theorem, which guarantees the integrability of a process that isn't componentwise integrable, the following result is very useful. The proof can be found in the appendix.

Theorem 40. *Let $X \in \mathcal{S}^d$ and H a d-dimensional predictable process. Then $H \in L(X)$ if and only if $H^n \bullet X$ with $H^n = H\mathbb{1}_{\{\|H\|_u \le n\}}$ converges in the semimartingale topology.*

9.4. Properties of the Stochastic Integral

In this section, we are going to examine the most important properties of the stochastic integral.

The following theorem lists most of the important properties of the stochastic integral. All these properties are well known. However, because of our different approach, our proofs differ from the ones in the mentioned literature.

Theorem 41. *Let X be a d-dimensional semimartingale and $H \in L(X)$.*

(a) *For $\alpha, \beta \in \mathbb{R}$ and $J \in L(X)$, we have $\alpha H + \beta J \in L(X)$ and*

$$(\alpha H + \beta J) \bullet X = \alpha H \bullet X + \beta J \bullet X. \tag{19}$$

Thus, $L(X)$ is a vector space.

(b) *For $Y \in \mathcal{S}^d$ and $H \in L(X) \cap L(Y)$, we have $H \in L(X + Y)$ and*

$$H \bullet (X + Y) = H \bullet X + H \bullet Y.$$

(c) *The process $(\Delta(H \bullet X)_s)_{s \ge 0}$ is indistinguishable from $(H_s^\top (\Delta X_s))_{s \ge 0}$.*

(d) We have
$$(H \bullet X)^T = H \mathbb{1}_{[\![0,T]\!]} \bullet X = H \bullet (X^T).$$

(e) Assume $X = (X^1, \ldots, X^d)^\top \in \mathcal{S}^d$ and $H = (H^1, \ldots, H^d)^\top$ with $H^i \in L(X^i)$ for all i. Furthermore, let $K = (K^1, \ldots, K^d)^\top$ be a predictable d-dimensional process and we set $Y^i := H^i \bullet X^i$, $Y := (Y^1, \ldots, Y^d)^\top$, $G^i := K^i H^i$ and $G = (G^1, \ldots, G^d)^\top$. Then $K^i \in L(Y^i)$ for all i if and only if $G^i \in L(X^i)$ for all i, and $K \in L(Y)$ if and only if $G \in L(X)$. In both cases, we have

$$K \bullet Y = G \bullet X.$$

(f) If X is an FV-process, then $H \bullet X$ is indistinguishable from the Lebesgue-Stieltjes integral, computed path-by-path.

(g) For a probability measure $\mathbf{Q}$ with $\mathbf{Q} \ll \mathbf{P}$, $H \in L(X)$ holds under $\mathbf{Q}$ as well, and $H_{\mathbf{Q}} \bullet X = H_{\mathbf{P}} \bullet X$, $\mathbf{Q}$-almost surely.

Proof. (a) Let $\mathbb{A}$ be the set for which 41 holds. It is clear that $\Lambda^d \subset \mathbb{A}$. With the help of the bounded convergence theorem A1, we can show $b\mathcal{P} \subset \mathbb{A}$ and an application of Theorem 40 yields the result.

(b) Analogous to 41.

(c) Let $\mathbb{A}$ be the set of processes for which 41 holds. First, we show that $\Lambda^d \subset \mathbb{A}$. We assume $H \in \Lambda^d$ and the stochastic integral with respect to X is given by

$$H \bullet X = H_0^\top X_0 + \sum_{i=1}^n H_i^\top \left(X^{T_{i+1}} - X^{T_i} \right),$$

with H_i as in Definition 33.
Hence we have

$$\Delta(H \bullet X)_t = (H \bullet X)_t - (H \bullet X)_{t-}$$
$$= \sum_{i=1}^n H_i^\top \left(X_t^{T_{i+1}} - X_{t-}^{T_{i+1}} - \left(X_t^{T_i} - X_{t-}^{T_i} \right) \right)$$
$$= \sum_{i=1}^n H_i^\top \left(\Delta X_t^{T_{i+1}} - \Delta X_t^{T_i} \right),$$

which implies $\Delta(H \bullet X)_s = (H \bullet \Delta X)_s$.
Furthermore, for the stopped process, we have

$$\Delta X_t^{T_i} = X_t^{T_i} - X_{t-}^{T_i} = \begin{cases} X_{T_i} - X_{T_i} = 0 & \text{for } t > T_i \\ \Delta X_t & \text{for } t \leq T_i. \end{cases}$$

That implies

$$\Delta X_t^{T_{i+1}} - \Delta X_t^{T_i} = \mathbb{1}_{]\!]T_i, T_{i+1}]\!]} \Delta X_t,$$

and thus

$$\Delta(H \bullet X)_t = \sum_{i=1}^n \mathbb{1}_{]\!]T_i, T_{i+1}]\!]} H_i^\top \Delta X_t$$
$$= \left(\sum_{i=1}^n \mathbb{1}_{]\!]T_i, T_{i+1}]\!]} H_i^\top \right) \Delta X_t = H_t^\top \Delta X_t.$$

We conclude $\Lambda^d \subset \mathbb{A}$.

Now let $(H^n)_{n \in \mathbb{N}}$ be a uniformly bounded, increasing sequence converging pointwise to a process H. By the bounded convergence theorem A1, we have

$$\Delta(H \bullet X)_t = \lim_{n \to \infty} \Delta(H^n \bullet X)_t = \lim_{n \to \infty} (H_t^n)^\top (\Delta X_t) = H^\top (\Delta X_t). \tag{20}$$

Since $\mathbb{Q} \cap [0, \infty)$ is countable, (20) holds almost everywhere for all rational $t \geq 0$. Since $H \bullet X$ is càdlàg, (20) holds almost surely for all $t \geq 0$. Thus, we have $H \in \mathbb{A}$ and can apply the monotone class theorem, which yields $\mathbf{b}\mathcal{P} \in \mathbb{A}$.

For $H \in L(X)$, we define $H^n := H\mathbb{1}_{\{\|H\| \leq n\}}$ and the result follows with Theorem 40.

1. The result is obvious for $H \in \Lambda^d$ and a bounded stopping time T taking only finitely many values.

    For a bounded stopping time T, we can approximate T from above by stopping time taking finitely many values. For a potentially unbounded stopping time, we can define $T^n = T \wedge n$, and the result still holds.

    For general integrands, we precede as before with the monotone class theorem and then with Theorem 40.

2. See Theorem A5.

3. The proof is again analogous to the ones above, with the only difference that the dominated convergence theorem for the Lebesgue-Stieltjes integral needs to be applied.

4. For a probability measure $\mathbf{Q}$ that is absolutely continuous with respect to $\mathbf{P}$, convergence in probability under $\mathbf{P}$ implies convergence in probability under $\mathbf{Q}$. Hence, the result is immediate $\square$

Example 5. *Let $H_t := \frac{1}{t}$ and $X_t := t$. Clearly, X is a semimartingale and since H is a deterministic process, it is also predictable.*

Assume $H \in L(X)$, then $H \bullet X$ exists and coincides with the Lebesgue-Stieltjes Integral $\int_{[0,t]} \frac{1}{s} ds$. Since this integral does not exist, we conclude $H \notin L(X)$.

Example 6. *Let $H_t := \mathbb{1}_{[1,\infty)}$ and $X_t := t$. Then, it is easy to see that we have $H \in L(X)$, but $H \notin \mathbb{L}$.*

In many publications on stochastic integration, the multidimensional integral is often defined as the sum of integrals for each component. However, Example 8 demonstrates that an integrand being integrable doesn't necessarily imply its componentwise integrability. Nevertheless, when an integrand is componentwise integrable, the subsequent theorem confirms that the stochastic integral is indeed the sum of the integrals of its components

Theorem 42. *Let X be a d-dimensional semimartingale and*

$$H = (H^1, \ldots, H^n)^\top$$

be a stochastic process with $H^i \in L(X^i)$ for all i. Then we have $H \in L(X)$ and

$$H \bullet X = \sum_{i=1}^d H^i \bullet X^i.$$

Proof. This follows directly from the linearity. $\square$

We provide an example illustrating that $L(X)$ encompasses more than just the componentwise integrable processes with respect to X. Additionally, this example demonstrates that the space $H \bullet X$, where H represents componentwise integrable processes, is not closed under the semimartingale topology. Further details and a comprehensive proof are available in [29, Example 12.5.1] or [184, Example 6.4].

Example 7. *Let W^1, W^2 be two independent Brownian motions. We put $H_t = t$ and define the two-dimensional process X by*

$$X = \begin{pmatrix} W^1 \\ (1-H) \bullet W^1 + H \bullet W^2. \end{pmatrix}$$

Then the space

$$L_C(X) = \left\{ \sum_{i=1}^{2} K^i \bullet X^i \, ; \, K^i \in L(X^i) \right\}$$

is not closed in the semimartingale topology.

9.5. Stability of Local Martingales Under Stochastic Integration

To the best of our knowledge, the subsequent criterion for the convergence of a sequence of local martingales to a local martingale has not been documented in existing literature. Nevertheless, this criterion streamlines the proofs of several established theorems. For instance, it offers a more straightforward approach to the general stochastic integral as presented in [172] and facilitates our proof in Theorem 43.

Lemma 10. *Let $(X^n)_{n \in \mathbb{N}}$ be a sequence of local martingales, which converges to a process X in ucp. If $(\sup_n (X^n)_t^*)_{t \in \mathbb{R}_+}$ is locally integrable, then X is a local martingale.*

Proof. Without loss of generality, we assume all processes to be one-dimensional. Because of the ucp-convergence, we can conclude that X is also càdlàg and adapted. By assuming a suitable subsequence, we can also assume that the convergence is almost surely on compact subsets, and thus

$$M_t := \sup_n (X^n)_t^*$$

is also càdlàg and adapted. Furthermore, X^* is increasing, and we have

$$|\Delta M_t| \leq 2 \sup_n (X^n)_t^*.$$

By assumption, the right-hand side is locally integrable and thus M is also locally integrable.

Now, we can find a sequence $(T^k)_{k \in \mathbb{N}}$ such that $(X^n)^{T^k}$ is a martingale for all n and k and M^{T^k} is integrable for all k. Furthermore, because of $X^{T^k} \leq M$, we can apply the dominated convergence theorem and obtain for every bounded stopping time τ

$$\mathbf{E}\left[X_\tau^{T^k} \right]$$
$$= \mathbf{E}\left[\lim_{n \to \infty} (X^n)_\tau^{T^k} \right] = \lim_{n \to \infty} \mathbf{E}\left[(X^n)_\tau^{T^k} \right] = \lim_{n \to \infty} \mathbf{E}\left[X_0^{T^k} \right] = \mathbf{E}\left[X_0^{T^k} \right].$$

Hence, X^{T^k} is a martingale for all k, and thus, we conclude that X is a local martingale. □

Remark 37. *We have found a criterion for the limit of a sequence of local martingales to be a local martingale again. In fact, there are numerous examples of criteria for that kind of problem in the literature. A detailed description with examples and counterexamples is given in the chapter 'Some Particular Problems of Martingale Theory' in Kabanov et al. [100].*

Theorem 43. *Let X be a local martingale (locally square-integrable martingale) and let $H \in \mathcal{P}$ be locally bounded. Then, the stochastic integral $H \bullet X$ is also a local martingale (locally square-integrable martingale).*

Proof. Let X be a local martingale. Without loss of generality, we assume there exists a K with $\|H\|_u \leq K$. Now let $(H^n)_{n \in \mathbb{N}}$ be a sequence with $H^n \in \Lambda^d$ and $H^n \xrightarrow{X} H$.

By putting $\tilde{H}^n := H^n \wedge K$ (where the minimum is taken in each component), we obtain a uniformly bounded sequence $\tilde{H}^n$ with $\tilde{H}^n \xrightarrow{X} H$.

By Lemma 1, $\tilde{H}^n \bullet X$ is a local martingale for all n. Furthermore, by Lemma 10 and since a process is locally uniformly bounded if and only if its jump process is locally uniformly bounded, it suffices to show that $M_t := \sup_n (\Delta \tilde{H} \bullet X^n)_t^*$ is locally integrable.

By Theorem 41 from Theorem 41, we obtain

$$\sup_n \Delta(\tilde{H}^n \bullet X) = \sup_n (H^n)^\top \Delta X \leq K \Delta X,$$

which is, by assumption, locally integrable.

For X being a locally square-integrable martingale or L^2-martingale, we again consider a sequence of simple predictable processes converging to H, then we apply Theorem 34 and proceed analogously. $\square$

Remark 38. *Since each $H \in \mathbb{L}$ is locally bounded, $H \bullet M$ is always a local martingale for a local martingale M and $H \in \mathbb{L}$.*

Remark 39. *For a semimartingale $X = M + A$ with $M \in \mathcal{M}_{\mathrm{loc}}$ and an FV-process A and $H \in L(M)$, one cannot assume that $H \bullet M \in \mathcal{M}_{\mathrm{loc}}$. However, there is always a decomposition $X = \overline{M} + \overline{A}$ with $\overline{M} \in \mathcal{M}_{\mathrm{loc}}$ and an FV-process A such that $H \bullet \overline{M} \in \mathcal{M}_{\mathrm{loc}}$. This is a direct consequence of the Fundamental Theorem of Local Martingales and Theorem 43. For details, we refer to Shiryaev and Cherny [184, Remarks after Definition 3.9].*

9.6. The Quadratic Variation of a Semimartingale

The quadratic variation plays a central role in the development of the general stochastic integral in both the classical approach and the extension from left-continuous integrands to general predictable integrands in the functional analytic approach.

As the quadratic variation of semimartingales can be defined without the more complicated concepts of integration with respect to general predictable integrands (as opposed to stochastic integration with respect to càglàd integrands) and integration of non-componentwise integrable multidimensional stochastic processes, the concept of the quadratic variation is usually not extended to multidimensional processes and is often omitted in the literature on multidimensional stochastic integration (see [90,184]). We also defined the stochastic integral without resorting to the quadratic variation. However, for the sake of completeness, we will define a multidimensional version of the quadratic variation that generalizes the concept of one-dimensional quadratic variation and mention some interesting facts and examples for further treatment and development of the theory. Our approach is consistent with the so-called *tensor quadratic variation* as it was presented by Métivier [139].

Definition 39. *Let $X, Y \in \mathcal{S}^d$ and*

$$XY^\top = \begin{pmatrix} X^1 Y^1 & \cdots & X^1 Y^d \\ \vdots & \ddots & \vdots \\ X^d Y^1 & \cdots & X^d Y^d \end{pmatrix}.$$

Then we define quadratic covariation of X and Y as

$$[X, Y] := \begin{pmatrix} X^1 Y^1 - X^1_- \bullet Y^1 - Y^1_- \bullet X^1 & \cdots & X^1 Y^d - X^1_- \bullet Y^d - Y^d_- \bullet X^1 \\ \vdots & \ddots & \vdots \\ X^d Y^1 - X^d_- \bullet Y^1 - Y^1_- \bullet X^d & \cdots & X^d Y^d - X^d_- \bullet Y^d - Y^d_- \bullet X^d \end{pmatrix}.$$

The quadratic variation of X is defined as $[X, X]$ and denoted by $[X]$ or $[X, X]$.

In the literature, the quadratic variation is often defined to be the limit of particular processes (see, for example, [108, Section 8.5]), which is an equivalent definition. There are even more ways to define the quadratic variation, especially for martingales (see, for example, [29, Chapter 11], [172, Chapter II Corollary 2 of Theorem 27] or [1, Definition 1.15]).

One easily sees that the $(X, Y) \mapsto [X, Y]$ is bilinear and symmetric. Therefore, we get the *polarization identity*

$$[X, Y] = \frac{1}{2}([X + Y, X + Y] - [X, X] - [Y, Y]) \tag{21}$$

In the following, we are going to present the main properties of the quadratic (co-)variation. In order to justify the name quadratic variation by the approximating property, we need the following definition.

Definition 40. *A finite set σ of stopping times, which satisfies $0 = T_1 \leq T_2 \leq T_k < \infty$ is called a random partition. If $(\sigma_n)_{n \in \mathbb{N}}$ is a sequence of random partitions $\sigma_n := (T_1^n, \ldots, T_{k_n}^n)$, $k_n \in \mathbb{N}$, then we say that $(\sigma_n)_{n \in \mathbb{N}}$ converges to the identity if $\lim_{n \to \infty} \sup_{i=1,\ldots,k_n} T_i^n = \infty$ almost surely and*

$$\|\sigma_n\| := \sup_{i=1,\ldots,k_n} |T_{i+1}^n - T_i^n| \xrightarrow{n \to \infty} 0 \quad \mathbf{P} \text{ - a.s.}$$

holds.

Theorem 44. *Let $X, Y \in \mathcal{S}^d$. The process $[X, Y]$ is, in each component, an FV-process, a semimartingale and has the following properties.*

(a) $[X, Y]_0 = X_0 Y_0$ *and* $\Delta [X, Y]_{i,j} = \Delta X^i \Delta Y^j$.

(b)   *For a sequence $(\sigma_n)_{n \in \mathbb{N}}$ of random partitions that tends to the identity, we have*

$$X_0 Y_0 + \sum_{i=1}^{k_n} (X^{T_{i+1}^n} - X^{T_i^n})(Y^{T_{i+1}^n} - Y^{T_i^n}) \to [X, Y], \quad n \to \infty, \tag{22}$$

where σ_n is given by $0 = T_1^n \leq T_2^n \leq \ldots T_{k_n}^n = T$.

(c)   *Let T be a stopping time. Then we have*

$$[X^T, Y] = [X, Y^T] = [X^T, Y^T] = [X, Y]^T.$$

(d)   *The quadratic variation $[X, X]$ is a positive, increasing process.*

(e)   *If X is an FV-process, we have*

$$[X, Y]_t = X_0 Y_0^\top + \sum_{0 < s \leq t} \Delta X_s \Delta Y_s^\top.$$

Proof. As the set of semimartingales form an algebra, it is easy to see that each component is a semimartingale, and by Theorem 44, it is an FV-process.

We can, for all properties, without loss of generality, assume $d = 1$.

(a)   By definition, $[X, Y]_0 = X_0 Y_0 - X_- \bullet Y_0 - Y_- \bullet X_0 = X_0 Y_0$, and the first equation is clear. For the second equation, using Theorem 41,

$$\begin{aligned}
\Delta X \Delta Y &= (X - X_-)(Y - Y_-) = XY - X_- Y - Y_- X + X_- Y_- \\
&= (XY - X_- Y_-) - X_- Y - Y_- X + 2 X_- Y_- \\
&= \Delta(XY) - X_-(Y - Y_-) - Y_-(X - X_-) \\
&= \Delta(XY) - X_- \Delta Y - Y_- \Delta X \\
&= \Delta(XY) - \Delta(X_- \bullet Y) - \Delta(Y_- \bullet X),
\end{aligned}$$

where the equality holds almost surely. This proves the result.

(b) By the polarization identity, it suffices to consider the case where $X = Y$, and without loss of generality, we can assume $X_0 = 0$.

Let $(\sigma_n)_{n \in \mathbb{N}}$ be a sequence of partitions that converges to the identity, with $\sigma_n = \{T_1^n, T_2^n, \ldots, T_{k_n}^n\}$. We define X^{σ_n} as follows:

$$X_t^{\sigma_n} = X_{T_i^n} \text{ for } t \in]\!] T_i^n, T_{i+1}^n]\!] \text{ for each } i.$$

For a fixed $N \in \mathbb{N}$, we set $R = \inf\{t \geq 0 : |X_t| \geq N\}$. Since X is càdlàg, we have $|X^{\sigma_n} \mathbb{1}_{[\![0,R]\!]}| \leq N$ for all n, and $X_t^{\sigma_n} \mathbb{1}_{[\![0,R]\!]} \to X_{t-} \mathbb{1}_{[\![0,R]\!]}$ almost surely as $n \to \infty$.

Now, by expressing $(X^R)_t^2$ as a telescoping sum, we obtain:

$$
\begin{aligned}
&(X^R)_t^2 \\
&= (X_0^R)^2 + 2 \sum_i X_{T^i}^R (X_{T^{i+1} \wedge t}^R - X_{T^i \wedge t}^R) + \sum_i (X_{T^{i+1} \wedge t}^R - X_{T^i \wedge t}^R)^2 \\
&= (X_0^R)^2 + 2\Big((X_{\sigma_n} \mathbb{1}_{[\![0,R]\!]}) \bullet X\Big)_t + \sum_i (X_{T^{i+1} \wedge t \wedge R} - X_{T^i \wedge t \wedge R})^2.
\end{aligned}
\tag{23}
$$

By the assumptions on σ_n, we note that only finitely many terms of this sum are nonzero. Define $Z(\sigma_n, t) := \sum_i (X_{T_{i+1}^n \wedge t} - X_{T_i^n \wedge t})^2$. According to Proposition 4.32 from Potter's book, we have:

$$(X^{\sigma_n} \mathbb{1}_{[\![0,R]\!]}) \bullet X \to X_- \bullet X_t^R$$

as $n \to \infty$ in the semimartingale topology. Furthermore,

$$\{Z(\sigma_n, R \wedge t)\}_{t \geq 0} \to \{Z(R)_t\}_{t \geq 0}$$

in ucp for some càdlàg process $Z(R)$.

Since N was arbitrary, we now have a family of processes $\{Z(R_N)\}_{N \in \mathbb{N}}$, where $R_N = \inf\{t \geq 0 : |X_t| \geq N\}$. It is easy to verify that if $M > N$ (and thus $R_M \geq R_N$), then $Z(R_N) = Z(R_M)$ on the interval $[\![0, R_N]\!]$. Hence, by pasting these processes together, we can define a single process Z such that $Z = Z(R_N)$ on $[\![0, R_N]\!]$ for all N, and $Z(\sigma_n) \to Z$ in ucp as $n \to \infty$.

Thus, on each interval $[\![0, R]\!]$, equation (23) implies $Z = [X, X]$, and therefore $Z = [X, X]$ in general.

1. We assume $d = 1$. By Theorem 41, we get

$$
\begin{aligned}
[X, Y]^T &= (XY)^T - (X_- \bullet Y)^T - (Y_- \bullet X)^T \\
&= X^T Y^T - X_-^T \bullet Y^T - Y_-^T \bullet X^T = [X^T, Y^T].
\end{aligned}
$$

Furthermore,

$$
\begin{aligned}
&[X^T, Y] \\
&= X^T Y - (X_-^T) \bullet Y - Y_- \bullet (X^T) \\
&= X^T (Y^T + Y - Y^T) - (\mathbb{1}_{\{t > T\}} X_-^T + \mathbb{1}_{[0,T]} X_-) \bullet Y - Y_- \bullet (X^T) \\
&= (XY)^T + \underbrace{X^T (Y - Y^T) - (\mathbb{1}_{\{t > T\}} X_-^T) \bullet Y}_{=0} \\
&\quad - (X_- \bullet Y)^T - (Y_- \bullet X)^T \\
&= [X, Y]^T.
\end{aligned}
$$

Thus,

$$[X^T, Y] = [X, Y]^T = [Y, X]^T = [Y^T, X] = [X, Y^T]$$

and everything is shown.

2. This follows with $X = Y$ directly from the summation representation in Theorem 44.

3. X is a semimartingale, and a stochastic integral with respect to X coincides pathwise with the Lebesgue-Stieltjes integral. The formula for partial integration in the Lebesgue-Stieltjes integral yields

$$X_t^2 = \int_0^t X_{s-} \, dX_s + \int_0^t X_s \, dX_s.$$

The formula for partial integration for semimartingales, in turn, yields

$$X_t^2 = 2 \int_0^t X_{s-} \, dX_s + [X, X].$$

Equating the two equations for X^2, we obtain

$$\int_0^t X_s \, dX_s = \int_0^t X_{s-} \, dX_s + [X, X]_t,$$

thus

$$[X, X]_t = \int_0^t (X_s - X_{s-}) \, dX_s = \int_0^t \Delta X_s \, dX_s = \sum_{s \le t} (\Delta X_s)^2.$$

□

Theorem 45. *For $X, Y \in \mathcal{S}^1, H \in L(X)$ and $K \in L(Y)$, we have*

$$[H \bullet X, K \bullet Y]_t = \int_0^t H_s K_s \mathrm{d}[X, Y]_s \quad (t \ge 0). \tag{24}$$

Proof. Without loss of generality, assume $X_0 = Y_0 = 0$. We first consider that H is of the form $H = \mathbb{1}_{[\![0,T]\!]}$ for a stopping time T. Now, by Theorem 44, we have

$$[H \bullet X, Y] = [\mathbb{1}_{[\![0,T]\!]} \bullet X, Y] = [X^T, Y] = [X, Y]^T = \mathbb{1}_{[\![0,T]\!]} \bullet [X, Y] = H \bullet [X, Y].$$

If H is of the form $H = U\mathbb{1}_{]\!]S,T]\!]}$ with stopping times $S \le T$ and an $\mathcal{F}_S$-measurable random variable U, we get

$$\begin{aligned} [H \bullet X, Y] &= [U(\mathbb{1}_{]\!]S,T]\!]} \bullet X), Y] = [U(X^T - X^S), Y] \\ &= U([X^T, Y] - [X^S, Y]) = U([X, Y]^T - [X, Y]^S) \\ &= U(\mathbb{1}_{]\!]S,T]\!]} \bullet [X, Y]) = H \bullet [X, Y]. \end{aligned}$$

With the usual monotone-class argument, this can be extended to any locally bounded H. And finally, Theorem 40 provides the result for $H \in L(X)$. □

Remark 40. *Theorem 45 has a multi-dimensional version as well ([184, Theorem 4.19]). However, the multidimensional variant requires some more technical definitions, and we, therefore, stick to the much simpler version in one dimension.*

Remark 41. *It would have been possible to define the quadratic covariation of two local martingales, M and N, as the unique FV-process $[M, N]$ such that $MN^\top - [M, N]$ is a local martingale in each component, and $\Delta[M, N]_{i,j} = \Delta M^i \Delta N^j$.*

By the integration by parts formula and Theorem 43, it is established that $MN^\top - [M, N]$ behaves as a local martingale in each component. Furthermore, according to Theorem 44, the relation $\Delta[M, N]_{i,j} = \Delta M^i \Delta N^j$ is clear. The uniqueness of $[M, N]$ is derived from the fact that any local martingale of finite variation must be a pure jump process. This condition regarding the jumps then secures the uniqueness.

Now, we have the tools at hand to give some more examples.

Example 8. *Let W be a standard Brownian motion. We define*

$$H_t = \begin{pmatrix} -\frac{1}{\sqrt{t}} \\ \frac{1}{\sqrt{t}} \end{pmatrix} \quad \text{and} \quad X_t = \begin{pmatrix} W_{t \wedge 1} \\ W_{t \wedge 1} + \mathbb{1}_{[1,\infty)} \bullet W_t \end{pmatrix}.$$

We assume H is componentwise integrable with respect to X. Then by Theorem 44, $[H^1 \bullet X^1, H^1 \bullet X^1]$ is a semimartingale. But since

$$[H^1 \bullet X^1, H^1 \bullet X^1]_t = \int_0^{t \wedge 1} \frac{1}{s} \mathrm{d}s = \infty,$$

we obtain a contradiction, and hence H is not componentwise integrable with respect to X.

However, we have $H \in L(X)$. To show that, we define

$$H_t^n = \begin{pmatrix} -\frac{1}{\sqrt{t + \frac{1}{n}}} \\ \frac{1}{\sqrt{t + \frac{1}{n}}} \end{pmatrix}.$$

Then, one easily sees that H^n is Cauchy in the X-topology and converges to H.

Remark 42. *There are several criteria for determining when an integrand is componentwise integrable. One such criterion is discussed in Remark 43. Moreover, all locally bounded predictable integrands are necessarily integrable (see Theorem 38). Additionally, if $H \in L(X)$ and $[X^i, X^j] = 0$ for all i, j with $i \neq j$, then $H^i \in L(X^i)$ for all i.*

The next example shows that the condition of componentwise integrability in the dominant convergence theorem cannot be dropped.

Example 9. *Let H, X be the processes from the previous example and $K = (0, H^2)$ with H^2 being the second component of H. Clearly, we have $|K^i| \leq |H^i|$ for $i = 1, 2$ and by Example 8, we also have $H \in L(X)$ but we have $K \notin L(X)$. This demonstrates that the condition of componentwise integrability cannot be dropped in the dominated convergence theorem.*

Remark 43. *One might criticize that the topological approach does not directly yield a description of the class of integrands via integrability conditions. However, such criteria are straightforward to derive, particularly for componentwise integrability.*

Let $X \in \mathcal{S}^1$ and $H \in \mathcal{P}$, and consider a decomposition of X as $X = N + A$. Assume that, locally,

$$\mathbf{E}\left[\int_0^\infty H_s^2 \, \mathrm{d}[N, N]_s\right] + \mathbf{E}\left[\int_0^\infty |H_s| \, |\mathrm{d}A_s|\right]^2 < \infty$$

holds. Then $H \in L(X)$.

This can be demonstrated by setting $H^n = \mathbb{1}_{\{|H| \leq n\}} H$. Then $H^n \in L(X)$, and it remains to show that $(H^n)_{n \in \mathbb{N}}$ is a Cauchy sequence with respect to d_X. To this end, we first establish that $(H^n \bullet X)_{n \in \mathbb{N}}$ is Cauchy in $\mathcal{H}^2$ (see for example [172, Theorem IV.14]). Next, we demonstrate that $H^n \bullet X$ converges in the semimartingale topology ([170, Theorem VI.4.19]). The rest follows with the equivalence of the $\mathcal{H}^2$ and X norms ([172, Corollary of Theorem IV.24 and Theorem V.14]).

Appendix A. The Technical Proofs

In this section, we prove the Dominated Convergence Theorem and some limit theorems providing information about $L(X)$, an important convergence theorem for multi-dimensional stochastic integration and, as it is required for the proofs of the former theorems, the associativity of the integral. The strategy for the proof of the dominated convergence theorem is similar to the popular proof

of Caratheodory's extension theorem and is as follows: First, we define two metrics which are less abstract than the ones we used to define the stochastic integral. These metrics will enable us to show that the Dominated Convergence Theorem holds for bounded predictable processes, a statement which is also known as Bounded Convergence Theorem. Then, we define the class of integrands for which the dominated convergence theorem holds and show that this class coincides with $L(X)$ (in one dimension).

As the dominated convergence theorem requires componentwise integrability, we can, without loss of generality, assume for its proof that $d = 1$.

Definition A1. *Let X be a one-dimensional semimartingale.*

(a) *For $H \in \Lambda^1$, we define*

$$\|H\|_{\Lambda,X} = \sup\{\|G \bullet X\|_{\text{ucp}} ; G \in \Lambda^1, |G| \le |H|\},$$

where $|G| \le |H|$ means $|G_t| \le |H_t|$ for all $t \ge 0$.

(b) *For $H \in \mathbf{b}\mathcal{P}$, we define*

$$\|H\|_{\mathbf{b}\mathcal{P},X} := \inf\left\{\sum_{n=1}^{\infty}\|G^n\|_{\Lambda,X} ; G^n \in \Lambda^1, |H| \le \sum_{n=1}^{\infty}|G^n|\right\}.$$

(c) *Furthermore, we define $\mathbb{A} \subset \mathbf{b}\mathcal{P}$ to be the closure of Λ^1 with respect to $\|\cdot\|_{\mathbf{b}\mathcal{P},X}$.*

Remark A1. *It is easy to check that by identifying indistinguishable processes, $\|\cdot\|_{\Lambda,X}$ becomes a norm and $\|\cdot\|_{\mathbf{b}\mathcal{P},X}$ a metric.*

Remark A2. *We will see later in the proof of the bounded convergence theorem that $\mathbb{A}$ coincides with $\mathbf{b}\mathcal{P}$.*

Lemma A1. *Let $(H^n)_{n\in\mathbb{N}}$ be a sequence of simple predictable processes such that $|H^n| \le |H^{n+1}| \le K$ for all n, where K is a constant. For each $\varepsilon > 0$ there exists an n_0 such that we have*

$$\|H^m - H^n\|_{\Lambda,X} \le \varepsilon \text{ for all } m, n \ge n_0. \tag{A1}$$

Proof. Let $\varepsilon > 0$ and assume that equation (A1) does not hold. Then we can extract subsequences $(H^{n_k})_{k\in\mathbb{N}}, (H^{m_k})_{k\in\mathbb{N}}$ such that $(G^k)_{k\in\mathbb{N}} := (H^{n_k} - H^{m_k})_{k\in\mathbb{N}}$ is a bounded, monotone sequence with $\left\|H^{n_k} - H^{m_k}\right\|_{\Lambda,X} > \varepsilon$ for all $k \in \mathbb{N}$. Since $(H^k)_{k\in\mathbb{N}}$ is monotone and bounded, $(G^k)_{k\in\mathbb{N}}$ converges pointwise almost surely to 0. By proceeding analogously to the proof of Theorem 31, we obtain $\left\|H^{n_k} - H^{m_k}\right\|_{\Lambda,X} \to 0$, which is a contradiction. $\square$

Lemma A2. *Assume positive processes $H, (H^n)_{n\in\mathbb{N}} \in \mathbf{b}\mathcal{P}$ such that $H = \sum_{j=1}^{\infty} H^j$. Then we have*

$$\left\|\sum_{j=1}^{\infty} H^j\right\|_{\mathbf{b}\mathcal{P},X} \le \sum_{j=1}^{j} \infty \left\|H^j\right\|_{\mathbf{b}\mathcal{P},X}.$$

Proof. We have

$$\left\|\sum_{j=1}^{n} H^j\right\|_{\mathbf{b}\mathcal{P},X} \le \sum_{j=1}^{n}\left\|H^j\right\|_{\mathbf{b}\mathcal{P},X} \le \sum_{j=1}^{\infty}\left\|H^j\right\|_{\mathbf{b}\mathcal{P},X}.$$

The result follows from the continuity of the metric. $\square$

Lemma A3. *Suppose $H \in \mathbf{b}\mathcal{P}$ and let $(H^n)_{n\in\mathbb{N}}$ be a sequence in $\mathbf{b}\mathcal{P}$ such that*

$$\|H^n - H\|_{\mathbf{b}\mathcal{P},X} \to 0.$$

Then we have $H \in \mathbb{A}$.

Proof. For each H^n, there exists a sequence $(H^{n^m})_{m \in \mathbb{N}} \in \Lambda^1$ such that

$$\left\| H^{n^m} - H^n \right\|_{\mathbf{b}\mathcal{P},X} \to 0 \quad \text{for} \quad m \to \infty.$$

Choosing a diagonal sequence yields a sequence $(H^{n^{m^k}})_{k \in \mathbb{N}}$ with

$$\left\| H - H^{n^{m^k}} \right\|_{\mathbf{b}\mathcal{P},X} \to 0 \quad \text{for} \quad k \to \infty.$$

Hence, $H \in \mathbb{A}$. $\square$

Lemma A4. *Let $H \in \mathbf{b}\mathcal{P}$ and $(H^n)_{n \in \mathbb{N}} \in \mathbb{A}$ with $|H^n| \le |H^{n+1}| \le |H|$ for all $n \in \mathbb{N}$ and $H \in \mathbf{b}\mathcal{P}$ such that $H^n \to H$ pointwise almost surely. Then we have $H \in \mathbb{A}$ and*

$$\|H - H^n\|_{\mathbf{b}\mathcal{P},X} \to 0.$$

Proof. By the definition of $\mathbb{A}$, for each $H^n \in \mathbb{A}$, there exists a sequence

$$(G^{n,m})_{m \in \mathbb{N}} \in \Lambda^1,$$

such that $\|G^{n,m} - H^n\|_{\mathbf{b}\mathcal{P},X} \to 0$ for $m \to \infty$. Without loss of generality, we assume all sequences $(G^{n,m})_{m \in \mathbb{N}}$ to be increasing. By assumption, for each $\varepsilon > 0$ and each n, there exists an $m_n(\varepsilon)$, such that $\|G^m - H^n\|_{\mathbf{b}\mathcal{P},X} < \varepsilon$ for all $m \ge m_n(\varepsilon)$.

We put

$$\overline{m}_n(\varepsilon) := \max\left\{ \overline{m}_{n-1}\left(\frac{\varepsilon}{3}\right) + 1, m_n\left(\frac{\varepsilon}{3}\right) \right\} \quad \text{and} \quad \tilde{G}^n := G^{\overline{m}_n(\varepsilon)}.$$

As $\tilde{G}^m - \tilde{G}^n \in \Lambda^1$, we have $\|\tilde{G}^m - \tilde{G}^n\|_{\mathbf{b}\mathcal{P},X} = \|\tilde{G}^m - \tilde{G}^n\|_{\Lambda,X}$ and by Lemma A1, there exists an $\tilde{n}(\varepsilon) \in \mathbb{N}$, such that $\|\tilde{G}^m - \tilde{G}^n\|_{\Lambda,X} < \frac{\varepsilon}{3}$ for all $n, m \ge \tilde{n}(\varepsilon)$.

We obtain for $m, n \ge \max\{\tilde{n}(\varepsilon), \overline{m}_n(\varepsilon)\}$

$$\|H^m - H^n\|_{\mathbf{b}\mathcal{P},X}$$
$$= \left\| H^m - \tilde{G}^m + (\tilde{G}^n - H^n) + (\tilde{G}^m - \tilde{G}^n) \right\|_{\mathbf{b}\mathcal{P},X}$$
$$\le \underbrace{\left\| H^m - \tilde{G}^m + \right\|_{\mathbf{b}\mathcal{P},X}}_{\le \frac{\varepsilon}{3}} + \underbrace{\left\| \tilde{G}^n - H^n \right\|_{\mathbf{b}\mathcal{P},X}}_{\le \frac{\varepsilon}{3}} + \underbrace{\left\| \tilde{G}^m - \tilde{G}^n \right\|_{\mathbf{b}\mathcal{P},X}}_{\le \frac{\varepsilon}{3}}.$$

Hence, for each ε, there exists an $n(\varepsilon) \in \mathbb{N}$, such that $\|H^m - H^n\|_{\mathbf{b}\mathcal{P},X} \le \varepsilon$ for all $m, n \ge n(\varepsilon)$.

Now we put $n^k = n\left(2^{-k}\right)$ and obtain a sequence $(H^{n^k})_{k \in \mathbb{N}}$ such that

$$\left\| H^{n^{k+1}} - H^{n^k} \right\|_{\mathbf{b}\mathcal{P},X} \le 2^{-k}.$$

Because of

$$H - H^{n^k} = \lim_{n \to \infty} H^n - H^{n^k} = \sum_{m=k}^{\infty} H^{n^{m+1}} - H^{n^m},$$

we obtain with Lemma A2

$$\left\| H - H^{n^k} \right\|_{\mathbf{b}\mathcal{P},X} \le \sum_{m=k}^{\infty} \left\| H^{n^{m+1}} - H^{n^m} \right\|_{\mathbf{b}\mathcal{P},X} \le 2^{-k+1}.$$

Hence, $\left\| H - H^{n^k} \right\|_{\mathbf{b}\mathcal{P},X} \to 0$ and thus, every subsequence of $(H^n)_{n\in\mathbb{N}}$ has a further subsequence that converges to H under $\|\cdot\|_{\mathbf{b}\mathcal{P},X}$. By Lemma A3, this proves the result and implies $H \in \mathbb{A}$. $\quad\square$

Theorem A1 (Bounded Convergence Theorem). *Let $H \in \mathbf{b}\mathcal{P}$ and $(H^n)_{n\in\mathbb{N}}$ be a sequence in $\mathbf{b}\mathcal{P}$ such that $|H^n| \le K$ for a constant K and $H^n \to H$ pointwise almost surely. Then we have*

$$H^n \bullet X \xrightarrow{\;\mathcal{S}\;} H \bullet X.$$

Proof. Let $(G^n)_{n\in\mathbb{N}}$ be a sequence in $\mathbb{A}$ converging uniformly to a process $G \in \mathbf{b}\mathcal{P}$. That means, there exists a sequence $(a_n)_{n\in\mathbb{N}}$ in $\mathbb{R}$ such that $(G^n - G)^* \le a_n$ and $a_n \to 0$. Now we have

$$\|G^n - G\|_{\mathbf{b}\mathcal{P},X} \le \|(G^n - G)^*\|_{\mathbf{b}\mathcal{P},X} \le \|a_n\|_{\mathbf{b}\mathcal{P},X} = |a_n| \|X\|_{\mathrm{ucp}} \to 0.$$

By Lemma A3, we have $G \in \mathbb{A}$ and hence $\mathbb{A}$ is closed under uniform convergence. Clearly, $\mathbb{A}$ also contains all constant functions. Together with Lemma A4, we can apply the monotone class theorem and obtain $\mathbb{A} = \mathbf{b}\mathcal{P}$. Thus, we have $H^n \in \mathbb{A}$ for all n. We put $\tilde{G}^n := \sup_{m\ge n}|H^m - H|$ and $\hat{G}^n := \tilde{G}^1 - \tilde{G}^n$ and have $\left|\hat{G}^n\right| \le \left|\tilde{G}^1\right|$ and $\hat{G}^n \to \tilde{G}^1$ pointwise almost surely. By Lemma A4, we obtain

$$\|H^n - H\|_{\mathbf{b}\mathcal{P},X} \le \left\|\tilde{G}^n\right\|_{\mathbf{b}\mathcal{P},X} = \left\|\hat{G}^n - \tilde{G}^1\right\|_{\mathbf{b}\mathcal{P},X} \to 0.$$

Hence, we deduce $\|H^n - H\|_{\mathbf{b}\mathcal{P},X} \to 0$.

Furthermore, it is easy to check that we have

$$\|H^n - H\|_X \le \|H^n - H\|_{\mathbf{b}\mathcal{P},X}.$$

Thus, $\|H^n - H\|_X \to 0$, which implies $H^n \xrightarrow{\;X\;} H$. $\quad\square$

Definition A2. *Let $X \in \mathcal{S}^1$. We define $L^1(X) \subset L(X)$ to be the set of predictable processes G such that, for each sequence of processes $H^n \in \mathbf{b}\mathcal{P}$ with $|H^n| \le |G|$ and converging pointwise almost surely to 0, we have $\|H^n\|_X \to 0$ for $n \to \infty$.*

For the proof of the dominated convergence theorem for $L^1(X)$, we need a simple property of the stochastic integral. We will examine this property in greater generality in Theorem A5 and Theorem 41.

Theorem A2. *Let X be a one-dimensional semimartingale, $H \in \Lambda^1$ and $K \in L(X)$. Then we have*

$$H \bullet (K \bullet X) = (HK) \bullet X.$$

Proof. This is a simple calculation for $K \in \Lambda^1$ (see, Theorem 41). Taking limits yields the result:

$$H \bullet (K \bullet X) = \lim_{n\to\infty} H \bullet (K^n \bullet X) = \lim_{n\to\infty} (HK^n) \bullet X = (HK) \bullet X$$

with $K^n \in \Lambda^1$ such that $K^n \xrightarrow{\;X\;} K$. $\quad\square$

Theorem A3 (Dominated Convergence Theorem for $L^1(X)$). *Let $X \in \mathcal{S}^1$ and $(H^n)_{n\in\mathbb{N}}$ be a sequence of predictable processes converging pointwise almost surely to a process H such that $|H^n| \le |G|$ with $G \in L^1(X)$. Then $\|H^n - H\|_X \to 0$.*

Proof. We put $Y^n := (H^n - H) \bullet X$. Now let $(K^n)_{n\in\mathbb{N}}$ be a sequence in Λ^1 with $|K^n| \le 1$. Then $|K^n(H^n - H)| \le 2|G|$ and by the definition of $L^1(X)$, we obtain

$$K^n \bullet Y^n = K^n(H^n - H) \bullet X \xrightarrow{\;\mathrm{ucp}\;} 0.$$

Hence, Y^n converges in the semimartingale topology and thus $\|H^n - H\|_X \to 0$. $\square$

Lemma A5. *Let $X \in \mathcal{S}^1$, $(G^n)_{n \in \mathbb{N}}$ be a sequence in $\mathbf{b}\mathcal{P}$ with*

$$\sum_{n=1}^{\infty} \|G^n\|_X < \infty \quad \text{and} \quad G := \sum_{n=1}^{\infty} |G^n|.$$

Furthermore, let $(H^n)_{n \in \mathbb{N}}$ be a sequence in $\mathbf{b}\mathcal{P}$ with $|H^n| \le |G|$.

(a)   *Let $(\lambda^n)_{n \in \mathbb{N}}$ be a sequence of positive numbers such that $|\lambda^n H^n| \le |G|$. Then*

$$\sup_{n \in \mathbb{N}} \|\lambda^n H^n\|_X < \infty.$$

(b)   *If $H^n \to 0$ pointwise almost surely, then $\|H^n\|_X \to 0$ and thus $G \in L^1(X)$ holds.*

(c)   *The set $A = \{\omega \in \Omega \, ; \, G = \infty\}$ is predictable and we have*

$$\|\mathbb{1}_A H\|_X = 0$$

   for all $H \in \mathbf{b}\mathcal{P}$.

Proof. We put $K^n := \sum_{k=1}^{n} \left| G^k \right|$. For any $\lambda > 0$ and $H \in \mathbf{b}\mathcal{P}$ with $\lambda |H| \le |G|$, $\lambda(H \wedge K^n) \vee (-K^n)$ converges pointwise almost surely to λH and we have $\lambda(H \wedge K^n) \vee (-K^n) \le |\lambda H|$.

Hence, we can apply bounded convergence and obtain

$$\|\lambda H\|_X = \lim_{n \to \infty} \|\lambda H \wedge K^n \vee (-K^n)\|_X.$$

Furthermore, we have

$$\lambda(H \wedge K^n) \vee (-K^n) \le \sum_{k=1}^{n} |\lambda H| \wedge \left| G^k \right|$$

and thus

$$\begin{aligned}
\|\lambda H\|_X &= \lim_{n \to \infty} \|\lambda(H \wedge K^n) \vee (-K^n)\|_X \\
&\le \lim_{n \to \infty} \sum_{k=1}^{n} \left\| |\lambda H| \wedge \left| G^k \right| \right\|_X \\
&= \sum_{n=1}^{\infty} \| |\lambda H| \wedge |G^n| \|_X,
\end{aligned}$$

where we used the triangle inequality of the metric $\|\cdot\|_X$.

Hence, we obtain

$$\sup_{m \in \mathbb{N}} \|\lambda^m H^m\|_X \le \sup_{m \in \mathbb{N}} \sum_{n=1}^{\infty} \| |\lambda^m H^m| \wedge G^n \|_X \le \sum_{n=1}^{\infty} \|G^n\|_X < \infty.$$

That proves (a). For (b), we proceed analogously and obtain, by putting $\lambda^m = 1$,

$$\lim_{m \to \infty} \|H^m\|_X \le \lim_{m \to \infty} \sum_{n=1}^{\infty} \| |H^m| \wedge G^n \|_X,$$

and as each summand of the right-hand side is bounded by $\|G^n\|_X$ and, furthermore, $\sum_{n=1}^{\infty} \|G^n\|_X < \infty$, we can commute the sum and the limes and obtain

$$\lim_{m \to \infty} \|H^m\|_X \le \sum_{n=1}^{\infty} \lim_{m \to \infty} \| |H^m| \wedge G^n \|_X.$$

Since $|H^m| \wedge G^n \leq G^n$ for all m and H^m converges pointwise almost surely to 0, we can apply bounded convergence and obtain

$$\||H^m| \wedge G^n\|_X \xrightarrow{X} 0 \quad \text{for} \quad m \to \infty.$$

Hence,

$$\lim_{m \to \infty} \|H^m\|_X \leq \sum_{n=1}^{\infty} \lim_{m \to \infty} \||H^m| \wedge G^n\|_X = 0.$$

For (c) we use (a) with a sequence $(\lambda^n)_{n \in \mathbb{N}}$ such that $\lim_{n \to \infty} \lambda^n \to \infty$. As $|\lambda^n \mathbb{1}_A H| \leq |G|$, we obtain $\sup_{n \in \mathbb{N}} \|\lambda^n \mathbb{1}_A H\|_X < \infty$ and thus $\|\mathbb{1}_A H\|_X = 0$. $\quad \square$

Theorem A4. *We have $L^1(X) = L(X)$. That is, the stochastic dominated convergence theorem holds for all $H \in L(X)$.*

Proof. Let $H \in L(X)$. By definition, there exists a sequence $(H^n)_{n \in \mathbb{N}} \in \Lambda^1$ with $H^n \xrightarrow{X} H$. That means, for each $\varepsilon > 0$, there exists an $n_0 \in \mathbb{N}$ such that $\|H^m - H^n\|_X \leq \varepsilon$ for all $m, n \geq n_0$. We put $H^1 := 0$, and by passing to subsequences, we can assume that

$$\left\|H^{n+1} - H^n\right\|_X \leq 2^{-n} \quad \text{for all } n \in \mathbb{N}.$$

We put

$$G^n := H^{n+1} - H^n \quad \text{and} \quad G := \sum_{n=1}^{\infty} |G^n|.$$

We have $\sum_{n=1}^{\infty} \|G^n\|_X \leq \sum_{n=1}^{\infty} 2^{-n} = 1$ and we put $A := \{\omega \in \Omega \, ; \, G = \infty\}$. By Lemma A5, $\mathbb{1}_{A^c} G \in L^1(X)$. Since $\sum_{n=1}^{\infty} G^n$ is absolutely convergent outside of A, we can define

$$\tilde{H} := \lim_{n \to \infty} \mathbb{1}_{A^c} H^n = \sum_{n=1}^{\infty} \mathbb{1}_{A^c} G^n.$$

Since $|\mathbb{1}_{A^c} H^n| \leq |\mathbb{1}_{A^c} G|$, we have $H \in L^1(X)$ and by dominated convergence, we obtain $\left\|\mathbb{1}_{A^c} H^n - \tilde{H}\right\|_X \to 0$. By Lemma A5, we have $\|\mathbb{1}_A H^n\|_X = 0$, hence,

$$\lim_{n \to \infty} H^n \bullet X = \lim_{n \to \infty} \mathbb{1}_{A^c} H^n \bullet X = \tilde{H} \bullet X$$

and thus $\left\|H - \tilde{H}\right\|_X = 0$. We conclude $H = \tilde{H}$ and hence $H \in L^1(X)$. $\quad \square$

Lemma A6. *Let $X \in \mathcal{S}^d$ and $H \in L(X)$. If $H_n = H \mathbb{1}_{\{\|H\|_u \leq n\}}$ then $H_n \bullet X \to H \bullet X$ in the semimartingale topology.*

Proof. The proof is analogous to Theorem 31. $\quad \square$

For the proof of the important Theorem A6, we need to show the associativity of the stochastic integral first.

Theorem A5. *Let $X = (X^1, \ldots, X^d)^\top \in \mathcal{S}^d$ and $H = (H^1, \ldots, H^d)^\top$ with $H^i \in L(X^i)$ for all i. Furthermore, let $K = (K^1, \ldots, K^d)^\top$ be a predictable d-dimensional process and we set $Y^i := H^i \bullet X^i$, $Y := (Y^1, \ldots, Y^d)^\top$, $G^i := K^i H^i$ and $G = (G^1, \ldots, G^d)^\top$. Then $K^i \in L(Y^i)$ for all i if and only if $G^i \in L(X^i)$ for all i, and $K \in L(Y)$ if and only if $G \in L(X)$. In both cases, we have*

$$K \bullet Y = G \bullet X.$$

Proof. For $H, K \in \Lambda^d$, this is a basic calculation.

Now, we suppose $G^i = K^i H^i \in L(X^i)$ with $i \in \{1, \ldots, d\}$. By Remark 23, there exists a $\tilde{K}^n = (\tilde{K}^{1,n}, \ldots, \tilde{K}^{d,n})^\top \in \Lambda^d$, such that $|\tilde{K}^{i,n}| \leq |K^i|$ and $\tilde{K}_t^{i,n} \to K_t^i$ for all t almost surely. Now, by the first part and by the dominated convergence theorem 39, we have

$$\tilde{K}^{i,n} \bullet Y^i = \left(\tilde{K}^{i,n} H^i\right) \bullet X^i \to \left(K^i H^i\right) \bullet X^i = G^i \bullet X^i.$$

Hence, $\tilde{K}^{i,n} \bullet Y^i$ is Cauchy and we have $K^i \in L(Y^i)$. Since this holds for all $i \in \{1, \ldots, d\}$, we obtain $K \bullet Y = G \bullet X$.

Conversely, assume $K^i \in L(Y^i)$ and let $\tilde{G}^{i,n} \in \Lambda^1$ be a sequence such that $\tilde{G}^{i,n} \to G^i = K^i H^i$ pointwise almost surely. We can write

$$\tilde{G}^{i,n} = \left(\frac{\tilde{G}^{i,n}}{H^i} \mathbb{1}_{\{H^i \neq 0\}}\right) H^i$$

and by proceeding analogously, we obtain $\frac{\tilde{G}^{i,n}}{H^i}\mathbb{1}_{\{H^i \neq 0\}} \in L(Y^i)$ and

$$\frac{\tilde{G}^{i,n}}{H^i}\mathbb{1}_{\{H^i \neq 0\}} \bullet Y^i = \tilde{G}^{i,n} \bullet X.$$

By dominated convergence, this tends to $K^i \bullet Y^i$ and we have $K^i H^i = G^i \in L(X^i)$.

For $G \in L(X)$, we define $\tilde{K}^n = K\mathbb{1}_{\|K\| \leq n}$, $\tilde{G}$ accordingly. As all $\tilde{K}^n$ and $\tilde{G}^n$ are bounded and hence componentwise integrable, we get

$$\|K \bullet Y - G \bullet Y\|_{\mathcal{S}}$$
$$\leq \|K \bullet Y - \tilde{K}^n \bullet Y\| + \|(\tilde{K}^n) \bullet Y - \tilde{G}^n \bullet X\|_{\mathcal{S}} + \|\tilde{G}^n \bullet X - G \bullet Y\|_{\mathcal{S}}$$

For the middle term, we can apply bounded convergence and see that it vanishes. The first and last tend to 0 (Lemma A6). □

Theorem A6. *Let $X \in \mathcal{S}^d$ and H a d-dimensional predictable process. Then $H \in L(X)$ if and only if $H^n \bullet X$ with $H^n = H\mathbb{1}_{\{\|H\|_u \leq n\}}$ converges in the semimartingale topology.*

Proof. The one direction has already been shown with Theorem A6. For the other direction, follow [184] and note that Remark 32 ensures the existence of a $K \in L(X)$ with $Z = K \bullet X$. Let

$$G = \sum_{n=1}^{\infty} \frac{1}{n}\mathbb{1}_{\{n-1 \leq \|H\| < n\}}$$

where G is predictable with $0 < G \leq 1$. By Theorem A5 and the definition of the semimartingale topology, it is easy to see that

$$G \bullet (H_n \bullet X) \xrightarrow{\mathcal{S}} G \bullet Z = (GK) \bullet X.$$

Since $\|GH\| \leq 1$, we can apply dominated convergence together with Theorem A5 and obtain

$$G \bullet (H_n \bullet X) = (GH_n) \bullet X \xrightarrow{\mathcal{S}} (GH) \bullet X.$$

Hence, $(GH) \bullet X$ coincides with $(GK) \bullet X$.

Since $K = G^{-1}GK$ is an element of $L(X)$, Theorem A5 states that G^{-1} also belongs to $L((GK) \bullet X)$, which is identical to $L((GH) \bullet X)$. A subsequent application Theorem A5 shows that H (represented as $G^{-1}GH$) is within $L(X)$, leading to

$$H \bullet X = G^{-1} \bullet ((GH) \bullet X) = G^{-1}((GK) \bullet X) = K \bullet X = Z,$$

which concludes the proof. □

References

1. A. Aksamit and M. Jeanblanc. *Enlargement of Filtration with Finance in View*. SpringerBriefs in Quantitative Finance. Springer International Publishing, 2017. ISBN 9783319412559. doi:10.1007/978-3-319-41255-9.

2. D. Applebaum. *Lévy Processes and Stochastic Calculus (Cambridge Studies in Advanced Mathematics)*. Cambridge University Press, 2nd edition, 2009. ISBN 9780521738651.

3. Robert B Ash, B Robert, Catherine A Doleans-Dade, and A Catherine. *Probability and measure theory*. Academic press, 2000.

4. J. Assefa and P. Harms. Cylindrical stochastic integration and applications to financial term structure modeling, 2023.

5. Wided Ayed and Hui-Hsiung Kuo. An extension of the itô integral. *Communications on Stochastic Analysis*, 2 (3):5, 2008.

6. R. F. Bass. The doob-meyer decomposition revisited. *Canad. Math. Bull.*, 39(2):150, 1996.

7. M. Beiglböck and P. Siorpaes. Riemann-integration and a new proof of the Bichteler–Dellacherie theorem. *Stoch. Process. Appl.*, 124(3):1226–1235, 2014. ISSN 0304-4149. http://dx.doi.org/10.1016/j.spa.2013.10.001.

8. M. Beiglböck, W. Schachermayer, B. Veliyev, et al. A direct proof of the Bichteler–Dellacherie theorem and connections to arbitrage. *The Annals of Probability*, 39(6):2424–2440, 2011. ISSN 0091–1798.

9. M. Beiglböck, W. Schachermayer, and B. Veliyev. A short proof of the Doob-Meyer theorem. *Stochastic Processes and their Applications*, 122(4):1204–1209, 2012. ISSN 0304-4149. http://dx.doi.org/10.1016/j.spa.2011.12.001.

10. Mathias Beiglboeck, Walter Schachermayer, and Bezirgen Veliyev. A short proof of the doob–meyer theorem. *Stochastic Processes and their applications*, 122(4):1204–1209, 2012.

11. A. Benveniste and J. Jacod. Systèmes de lévy des processus de markov. *Invent. Math.*, 21(3):183–198, 1973.

12. F. Biagini, Y. Hu, B. Øksendal, and T. Zhang. *Stochastic Calculus for Fractional Brownian Motion and Applications*. Probability and Its Applications. Springer London, 2008. ISBN 9781846287978.

13. K. Bichteler. Stochastic integrators. *Bulletin of the American Mathematical Society*, 1(5):761–765, 1979. 10.1090/s0273-0979-1979-14655-x.

14. K. Bichteler. Stochastic integration and L^p-theory of semimartingales. *The Annals of Probability*, pages 49–89, 1981. ISSN 0091–1798.

15. K. Bichteler. *Stochastic Integration with Jumps*. Number Bd. 89 in Encyclopedia of Mathematics and its Applications. Cambridge University Press, 2002. ISBN 9780521811293. 10.1017/cbo9780511549878.

16. N.H. Bingham and R. Kiesel. *Risk-Neutral Valuation: Pricing and Hedging of Financial Derivatives*. Springer Finance. Springer London, 2013. ISBN 9781447138563.

17. L. Boy. Stochastic integration and martingales in banach spaces. *Sém. Prob. Rennes*, 9:49–72, 1977.

18. P. Bremaud. Poisson processes in systems with dependent components. *IEEE Trans. Inf. Theory*, 18:676–686, 1972.

19. P. Bremaud and J. Jacod. Processus ponctuels et martingales: résultats récents sur la modélisation et le filtrage. *Adv. Appl. Probab.*, 2:362–416, 1977.

20. James K Brooks and David Neal. The optional stochastic integral. In *Seminar on Stochastic Processes, 1988*, pages 45–54. Springer, 1989. 10.1007/978-1-4612-3698-6_2.

21. B. Bru and M. Yor. Comments on the life and mathematical legacy of wolfgang doeblin. *Finance and Stochastics*, 6(1):3–47, 2002. ISSN 0949–2984. 10.1007/s780-002-8399-0.

22. Zdzisław Brzeźniak, Erika Hausenblas, and Jianliang Zhu. 2d stochastic navier-stokes equations driven by jump noise. *Stochastic Processes and their Applications*, 118:2066–2092, 2008. URL https://doi.org/10.1016/j.spa.2008.09.009.

23. D. L. Burkholder. Martingale Transforms. *The Annals of Mathematical Statistics*, 37(6):1494 – 1504, 1966. 10.1214/aoms/1177699141. URL https://doi.org/10.1214/aoms/1177699141.

24. Donald L Burkholder, Burgess J Davis, and Richard F Gundy. Integral inequalities for convex functions of operators on martingales. In *Proceedings of the Sixth Berkeley Symposium on Mathematical Statistics and Probability*, volume 2, pages 223–240. Univ. California Press Berkeley, Calif., 1972.

25. Geon Ho Choe et al. *Stochastic analysis for finance with simulations*. Springer, 2016. 10.1007/978-3-319-25589-7.

26. C. S. Chou, P.-A. Meyer, and C. Stricker. Sur les integrales stochastiques de processus previsibles non bornes. In *Séminaire de Probabilités XIV 1978/79*, pages 128–139. Springer, 1980. 10.1007/bfb0089478.

27. K. L. Chung and J. L. Doob. Fields, optionality and measurability. *Amer. J. Math.*, 87:397–424, 1965. ISSN 0002-9327. 10.2307/2373011. URL https://doi.org/10.2307/2373011.

28. K. L. Chung and R. J. Williams. *Introduction to stochastic integration*. Probability and its Applications. Birkhäuser Boston, Inc., Boston, MA, second edition, 1990. ISBN 0-8176-3386-3. 10.1007/978-1-4612-4480-6. URL https://doi.org/10.1007/978-1-4612-4480-6.

29. S. N. Cohen and R. J. Elliott. *Stochastic Calculus and Applications*. Probability and Its Applications. Springer New York, 2015. ISBN 9781493928675. 10.1007/978-1-4939-2867-5.

30. Philippe Courrege. Intégrales stochastiques et martingales de carré intégrable. *Séminaire Brelot-Choquet-Deny. Théorie du potentiel*, 7:1–20, 1963.

31. Philippe Courrège. Décomposition des martingales de carré intégrable. pages No. 6, 14, 1964.

32. Marzia De Donno, Paolo Guasoni, and Maurizio Pratelli. Super-replication and utility maximization in large financial markets. *Stochastic processes and their applications*, 115(12):2006–2022, 2005. 10.1016/j.spa.2005.06.010.

33. F. Delbaen and W. Schachermayer. A general version of the fundamental theorem of asset pricing. *Mathematische Annalen*, 300(1):463–520, 1994. ISSN 0025-5831. 10.1007/BF01450498.

34. F. Delbaen and W. Schachermayer. The fundamental theorem of asset pricing for unbounded stochastic processes. *Mathematische Annalen*, 312(2):215–250, 1998. ISSN 0025-5831. 10.1007/s002080050220.

35. C. Dellacherie. Un survol de la theorie de l'integrale stochastique. *Stochastic Processes and their Applications*, 10(2):115–144, 1980. ISSN 0304-4149. http://dx.doi.org/10.1016/0304-4149(80)90017-4.

36. C. Dellacherie. Mesurabilite des debuts et theoreme de section: Le lot a la portee de toutes les boures. In *Séminaire de Probabilités XV 1979/80*, pages 351–370. Springer, 1981.

37. C. Dellacherie and P.-A. Meyer. *Probabilities and Potential*. North-Holland Mathematics Studies. Hermann, Amsterdam, 1978. ISBN 9780720407013.

38. C. Dellacherie and P.-A. Meyer. *Probabilities and Potential, B: Theory of Martingales*. North-Holland Mathematics Studies. Elsevier Science, 1982. ISBN 9780444865267.

39. Catherine Doléans. Processus croissants naturels et processus croissants très-bien-mesurables. *C. R. Acad. Sci. Paris Sér. A-B*, 264:A874–A876, 1967. ISSN 0151-0509.

40. Catherine Doléans. Existence du processus croissant natural associé à un potentiel de la classe (D). *Z. Wahrscheinlichkeitstheorie und Verw. Gebiete*, 9:309–314, 1968. 10.1007/BF00531754. URL https://doi.org/10.1007/BF00531754.

41. C. Doléans-Dade and P.-A. Meyer. Intégrales stochastiques par rapport aux martingales locales. In *Séminaire de Probabilités IV Université de Strasbourg*, pages 77–107. Springer, 1970.

42. Catherine Doléans-Dade. On the existence and unicity of solutions of stochastic integral equations. *Z. Wahrscheinlichkeitstheorie und Verw. Gebiete*, 36(2):93–101, 1976. 10.1007/BF00533992. URL https://doi.org/10.1007/BF00533992.

43. Marzia De Donno and Maurizio Pratelli. Stochastic integration with respect to a sequence of semimartingales. In *In Memoriam Paul-André Meyer*, pages 119–135. Springer, 2006. 10.1007/978-3-540-35513-7_10.

44. Joseph L Doob. *Stochastic processes*, volume 7. Wiley New York, 1953.

45. R. M. Dudley and R. Norvaiša. *Concrete functional calculus*. Springer Monographs in Mathematics. Springer, New York, 2011. ISBN 978-1-4419-6949-1. 10.1007/978-1-4419-6950-7. URL https://doi.org/10.1007/978-1-4419-6950-7.

46. Richard Durrett. *Stochastic calculus*. Probability and Stochastics Series. CRC Press, Boca Raton, FL, 1996. ISBN 0-8493-8071-5. A practical introduction.

47. E. B. Dynkin. *Markov processes. Vols. I, II*, volume 122 of *Die Grundlehren der mathematischen Wissenschaften, Band 121*. Springer-Verlag, Berlin-Göttingen-Heidelberg; Academic Press, Inc., Publishers, New York, 1965. Translated with the authorization and assistance of the author by J. Fabius, V. Greenberg, A. Maitra, G. Majone.

48. S. Ebenfeld. *Grundlagen der Finanzmathematik: mathematische Methoden, Modellierung von Finanzmärkten und Finanzprodukten*. Schäffer-Poeschel, 2007. ISBN 9783791026343.

49. D. A. Edwards. A note on stochastic integrators. In *Mathematical Proceedings of the Cambridge Philosophical Society*, volume 107, pages 395–400. Cambridge University Press, 1990. 10.1017/s0305004100068663.

50. Robert J. Elliott and P. Ekkehard Kopp. *Mathematics of financial markets*. Springer Finance. Springer-Verlag, New York, second edition, 2005. ISBN 0-387-21292-2.

51. Paul Embrechts and Makoto Maejima. *Selfsimilar processes*. Princeton Series in Applied Mathematics. Princeton University Press, Princeton, NJ, 2002. ISBN 0-691-09627-9.

52. M. Émery. Une topologie sur l'espace des semimartingales. In *Séminaire de Probabilités XIII*, pages 260–280. Springer, 1979.

53. M. Émery. Métrisabilité de quelques espaces de processus aléatoires. *Séminaire de probabilités de Strasbourg*, 14:140–147, 1980.

54. Michel Émery. *Stochastic calculus in manifolds*. Universitext. Springer-Verlag, Berlin, 1989. ISBN 3-540-51664-6. 10.1007/978-3-642-75051-9. URL https://doi.org/10.1007/978-3-642-75051-9. With an appendix by P.-A. Meyer.

55. Donald L. Fisk. Quasi-martingales. *Trans. Amer. Math. Soc.*, 120:369–389, 1965. ISSN 0002-9947. 10.2307/1994531. URL https://doi.org/10.2307/1994531.

56. Avner Friedman. *Stochastic differential equations and applications. Vol. 1*. Probability and Mathematical Statistics, Vol. 28. Academic Press [Harcourt Brace Jovanovich, Publishers], New York-London, 1975.

57. Peter Friz and Martin Hairer. A course on rough paths. *Preprint*, 2014. 10.1007/978-3-319-08332-2.

58. Adriano M. Garsia. *Martingale inequalities: Seminar notes on recent progress*. Mathematics Lecture Note Series. W. A. Benjamin, Inc., Reading, Mass.-London-Amsterdam, 1973.

59. L. Gawarecki and V. Mandrekar. *Stochastic Differential Equations in Infinite Dimensions: with Applications to Stochastic Partial Differential Equations*. Probability and Its Applications. Springer Berlin Heidelberg, 2010. ISBN 9783642161940. URL https://books.google.de/books?id=yqlcgl3KumQC.

60. Leszek Gawarecki and Vidyadhar Mandrekar. Partial differential equations as equations in infinite dimensions. In *Stochastic Differential Equations in Infinite Dimensions*, pages 3–16. Springer, 2011.

61. Ĭ. Ĭ. Gīhman and A. V. Skorohod. *Stochastic differential equations*. Ergebnisse der Mathematik und ihrer Grenzgebiete [Results in Mathematics and Related Areas], Band 72. Springer-Verlag, New York-Heidelberg, 1972. Translated from the Russian by Kenneth Wickwire.

62. B. Grigelionis. On the representation of integer-valued random measures by means of stochastic integrals. *Lit. Math. J.*, 9:93–108, 1969.

63. B. Grigelionis. On the representation of integer-valued random measures by means of stochastic integrals with respect to the poisson measure. *Lit. Math. J.*, 11:93–108, 1971.

64. B. Grigelionis. Stochastic point processes and martingales. *Lit. Math. J.*, 15:101–114, 1975.

65. B. Grigelionis. A generalization of the representation theorem for martingales. *Lit. Math. J.*, 17:75–85, 1977.

66. B Grigelionis and R Mikulevičius. Stochastic evolution equations and densities of the conditional distributions. In *Theory and Application of Random Fields*, pages 49–88. Springer, 1983.

67. Mircea Grigoriu. *Stochastic calculus*. Birkhäuser Boston, Inc., Boston, MA, 2002. ISBN 0-8176-4242-0. 10.1007/978-0-8176-8228-6. URL https://doi.org/10.1007/978-0-8176-8228-6. Applications in science and engineering.

68. István Gyöngy and Nicolai V Krylov. On stochastics equations with respect to semimartingales ii. itô formula in banach spaces. *Stochastics*, 6(3-4):153–173, 1982.

69. J. M. Harrison and S. R. Pliska. Martingales and stochastic integrals in the theory of continuous trading. *Stochastic processes and their applications*, 11(3):215–260, 1981. ISSN 0304–4149. 10.1016/0304-4149(81)90026-0.

70. J. M. Harrison and S. R. Pliska. A stochastic calculus model of continuous trading: Complete markets. *Stochastic Processes and their Applications*, 15(3):313–316, 1983. ISSN 0304–4149. 10.1016/0304-4149(83)90038-8.

71. Sheng Wu He, Jia Gang Wang, and Jia An Yan. *Semimartingale theory and stochastic calculus*. Kexue Chubanshe (Science Press), Beijing; CRC Press, Boca Raton, FL, 1992. ISBN 7-03-003066-4.

72. R. Henstock. The efficiency of convergence factors for functions of a continuous real variable. *J. London Math. Soc.*, 30:273–286, 1955. ISSN 0024-6107. 10.1112/jlms/s1-30.3.273. URL https://doi.org/10.1112/jlms/s1-30.3.273.

73. Shinya Hibino, Hui Hsiung Kuo, and Kimiaki Saitô. A stochastic integral by a near-martingale. *Communications on Stochastic Analysis*, 12(2):197, 2018. 10.31390/cosa.12.2.07.

74. GA Hunt. Markoff processes and potentials i. *Illinois Journal of Mathematics*, 1(1):44–93, 1957.

75. GA Hunt. Markoff processes and potentials ii. *Illinois Journal of Mathematics*, 1(3):316–369, 1957. 10.1215/ijm/1255380388.

76. GA Hunt. Markoff processes and potentials iii. *Illinois Journal of Mathematics*, 2(2):151–213, 1958. 10.1215/ijm/1255381342.

77. Chii-Ruey Hwang, Hui-Hsiung Kuo, Kimiaki Saitô, and Jiayu Zhai. A general itô formula for adapted and instantly independent stochastic processes. *Communications on Stochastic Analysis*, 10(3):5, 2016.

78. Albrecht Irle. *Finanzmathematik*. Studienbücher Wirtschaftsmathematik. Springer Spektrum, Wiesbaden, third edition, 2012. ISBN 978-3-8348-1574-3; 978-3-8348-8314-8. 10.1007/978-3-8348-8314-8. URL https://doi.org/10.1007/978-3-8348-8314-8. Die Bewertung von Derivaten. [The evaluation of derivatives].

79. K. Itô. Stochastic integral. *Proc. Imp. Acad.*, 20(8):519–524, 1944. 10.3792/pia/1195572786.

80. K. Itô. On a stochastic integral equation. *Proc. Japan Acad.*, 22(2):32–35, 1946. 10.3792/pja/1195572371.

81. K. Itô. Stochastic differential equations in a differentiable manifold. *Nagoya Math. J.*, 1:35–47, 1950.

82. K. Itô. On a formula concerning stochastic differentials. *Nagoya Math. J.*, 3:55–65, 1951.

83. K. Itô. Multiple Wiener integral. *Journal of the Mathematical Society of Japan*, 3(1):157–169, 1951.

84. K. Itô. *On Stochastic Differential Equations*, volume 4. American Mathematical Soc., 1951.

85. K. Itô. *Lectures on Stochastic Processes*, volume 24. Tata Institute of Fundamental Research, Bombay, 1961.

86. K. Itô. Transformation of markov processes by multiplicative functionals. *Ann. Inst. Fourier*, 15(1):15–30, 1965.

87. J. Jacod. Calculus of semimartingales and applications to the theory of continuous martingales. *Z. Wahrscheinlichkeitstheorie verw. Gebiete*, 25:37–53, 1972.

88. J. Jacod. Sur la construction des intégrales stochastiques et les sous-espaces stables de martingales. In *Séminaire de Probabilités XI*, pages 390–410. Springer, 1977.

89. J. Jacod. *Calcul stochastique et problèmes de martingales*. Number Nr. 714 in Lecture Notes in Mathematics. Springer Verlag, Berlin-Heidelberg-New York, 1979. ISBN 9780387092539.

90. J. Jacod. Intégrales stochastiques par rapport à une semimartingale vectorielle et changements de filtration. In *Séminaire de Probabilités XIV 1978/79*, pages 161–172. Springer, 1980.

91. Jean Jacod. Multivariate point processes: predictable projection, radon-nikodym derivatives, representation of martingales. *Zeitschrift für Wahrscheinlichkeitstheorie und verwandte Gebiete*, 31(3):235–253, 1975.

92. Jean Jacod and Albert N. Shiryaev. *Limit theorems for stochastic processes*, volume 288 of *Grundlehren der mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Springer-Verlag, Berlin, second edition, 2003. ISBN 3-540-43932-3. 10.1007/978-3-662-05265-5. URL https://doi.org/10.1007/978-3-662-05265-5.

93. Adam Jakubowski. An almost sure approximation for the predictable process in the Doob-Meyer decomposition theorem. In *Séminaire de Probabilités XXXVIII*, volume 1857 of *Lecture Notes in Math.*, pages 158–164. Springer, Berlin, 2005. 10.1007/978-3-540-31449-3_11. URL https://doi.org/10.1007/978-3-540-31449-3_11.

94. Adam Jakubowski and Markus Riedle. Stochastic integration with respect to cylindrical Lévy processes. *Ann. Probab.*, 45(6B):4273–4306, 2017. ISSN 0091-1798. 10.1214/16-AOP1164. URL https://doi.org/10.1214/16-AOP1164.

95. K. Jänich. *Topology*. Undergraduate Texts in Mathematics. Springer-Verlag New York Inc., 1984. ISBN 9781461270188.

96. Robert Jarrow and Philip Protter. A short history of stochastic integration and mathematical finance: the early years, 1880–1970. 45:75–91, 2004. 10.1214/lnms/1196285381. URL https://doi.org/10.1214/lnms/1196285381.

97. Guy Johnson and LL Helms. Class *d* supermartingales. *Bulletin of the American Mathematical Society*, 69(1):59–62, 1963.

98. Marc Jornet. On the ayed-kuo stochastic integration for anticipating integrands. *Stochastic Analysis and Applications*, pages 1–25, 2022. 10.1080/07362994.2022.2104730.

99. Y. Kabanov, R. Liptzer, and A. N. Shiryaev. Absolute continuity and singularity of probability measures in terms of stochastic integrals. *Math. Sb.*, 104:227–247, 1977.

100. Y. M. Kabanov, A. N. Shiryaev, J. M. Stojanov, and R. S. Liptser. *From Stochastic Calculus to Mathematical Finance: The Shiryaev Festschrift*. Springer, 2006. ISBN 9783540307822. 10.1007/978-3-540-30788-4.

101. Svenja Kaden and Jürgen Potthoff. Progressive stochastic processes and an application to the Itô integral. *Stochastic Anal. Appl.*, 22(4):843–865, 2004. ISSN 0736-2994. 10.1081/SAP-120037621. URL https://doi.org/10.1081/SAP-120037621.

102. O. Kallenberg. *Foundations of Modern Probability*. Springer, 2. edition, 2002. ISBN 9780387953137. 10.1007/978-1-4757-4015-8.

103. Gopinath Kallianpur. *Stochastic filtering theory*, volume 13 of *Applications of Mathematics*. Springer-Verlag, New York-Berlin, 1980. ISBN 0-387-90445-X.

104. Dhandapani Kannan and Vangipuram Lakshmikantham. *Handbook of stochastic analysis and applications*. CRC Press, 2001.

105. Rajeeva L. Karandikar and B. V. Rao. On the second fundamental theorem of asset pricing, 2016. URL https://doi.org/10.31390/cosa.10.1.05.

106. Rajeeva L. Karandikar and B. V. Rao. *Introduction to stochastic calculus*. Indian Statistical Institute Series. Springer, Singapore, 2018. ISBN 978-981-10-8317-4; 978-981-10-8318-1. 10.1007/978-981-10-8318-1. URL https://doi.org/10.1007/978-981-10-8318-1.

107. J. F. C. Kingman. *Completely random measures*, volume 21. 1967.

108. F. C. Klebaner. *Introduction to Stochastic Calculus With Applications (2nd Edition)*. ICP, 2. edition, 2005.

109. A. Klenke. *Probability Theory: A Comprehensive Course*. Universitext. Springer London, 2008. ISBN 9781848000483.

110. M. Koller. *Life Insurance Risk Management Essentials: Life Insurance Risk Management Essentials*. EAA Series. Springer, 2011. ISBN 9783642207211. 10.1007/978-3-642-20721-1.

111. J. Komlós. A generalization of a problem of Steinhaus. *Acta Math. Acad. Sci. Hungar.*, 18:217–229, 1967. ISSN 0001-5954. 10.1007/BF02020976. URL https://doi.org/10.1007/BF02020976.

112. Hayri Korezlioglu and Claude Martaias. Stochastic integration for operator valued processes on hilbert spaces and on nuclear spaces. *Stochastics: An International Journal of Probability and Stochastic Processes*, 24(3): 171–219, 1988. 10.1080/17442508808833515.

113. Tomasz Kosmala and Markus Riedle. Stochastic integration with respect to cylindrical Lévy processes by p-summing operators. *J. Theoret. Probab.*, 34(1):477–497, 2021. ISSN 0894-9840. 10.1007/s10959-019-00978-x. URL https://doi.org/10.1007/s10959-019-00978-x.

114. H. Kunita and S. Watanabe. On square integrable martingales. *Nagoya Mathematical Journal*, 30:209–245, 1967. 10.1017/s0027763000012484.

115. Hiroshi Kunita. Stochastic integrals based on martingales taking values in hilbert space. *Nagoya Mathematical Journal*, 38:41–52, 1970. 10.1017/s0027763000013507.

116. Hiroshi Kunita. *Stochastic flows and stochastic differential equations*, volume 24. Cambridge university press, 1997.

117. Hiroshi Kunita and Takesi Watanabe. Markov processes and Martin boundaries. I. *Illinois J. Math.*, 9:485–526, 1965. ISSN 0019-2082. URL http://projecteuclid.org/euclid.ijm/1256068151.

118. H.-H. Kuo. *Introduction to Stochastic Integration (Universitext)*. Springer, 1. edition, 2005. ISBN 9780387287201.

119. Hui-Hsiung Kuo. The itô calculus and white noise theory: a brief survey toward general stochastic integration. *Communications on Stochastic Analysis*, 8(1):8, 2014.

120. Hui-Hsiung Kuo, Anuwat Sae-Tang, and Benedykt Szozda. The itô formula for a new stochastic integral. *Communications on Stochastic Analysis*, 6(4):7, 2012.

121. Jaroslav Kurzweil. Generalized ordinary differential equations and continuous dependence on a parameter. *Czechoslovak Mathematical Journal*, 7(3):418–449, 1957. 10.21136/cmj.1957.100258.

122. A. U. Kussmaul. *Stochastic integration and generalized martingales*. Research Notes in Mathematics, No. 11. Pitman Publishing, London-San Francisco, Calif.-Melbourne, 1977.

123. Jean-François Le Gall et al. *Brownian motion, martingales, and stochastic calculus*, volume 274. Springer, 2016.

124. E. Lenglart. Semi-martingales et intégrales stochastiques en temps continu. *Revue du CETHEDEC*, 20(75): 91–160, 1983.

125. Giorgio Letta. *Martingales et intégration stochastique*. Springer, 1984.

126. P. Lévy. *Théorie de l'addition des variables aléatoires*. Paris : Gauthier-Villars, 1937.

127. S. J. Lin. Stochastic analysis of fractional Brownian motions. *Stochastics: An International Journal of Probability and Stochastic Processes*, 55(1-2):121–140, 1995.

128. R. Liptser and A. Shiryaev. *Theory of Martingales*, volume 49 of *Mathematics and its Applications*. Kluwer Academic Publishers, Dordrecht, 1989. ISBN 978-0-7923-0395-4.

129. Robert S Liptser and Albert N Shiryaev. *Statistics of random processes: I. General theory*, volume 5. Springer Science & Business Media, 2013.

130. Terry Lyons and Zhongmin Qian. *System control and rough paths*. Oxford Mathematical Monographs. Oxford University Press, Oxford, 2002. ISBN 0-19-850648-1. 10.1093/acprof:oso/9780198506485.001.0001. URL https://doi.org/10.1093/acprof:oso/9780198506485.001.0001. Oxford Science Publications.

131. Terry J. Lyons. Differential equations driven by rough signals. *Rev. Mat. Iberoamericana*, 14(2):215–310, 1998. ISSN 0213-2230. 10.4171/RMI/240. URL https://doi.org/10.4171/RMI/240.

132. Terry J. Lyons, Michael Caruana, and Thierry Lévy. *Differential equations driven by rough paths*, volume 1908 of *Lecture Notes in Mathematics*. Springer, Berlin, 2007. ISBN 978-3-540-71284-8; 3-540-71284-4. Lectures from

the 34th Summer School on Probability Theory held in Saint-Flour, July 6–24, 2004, With an introduction concerning the Summer School by Jean Picard.

133. HP McKean. Stochastic integrals academic press. *New York*, 1969.

134. Edward James McShane. *Stochastic Calculus and Stochastic Models*. Academic Press, New York, 1974.

135. EJ McShane. Stochastic differential equations. *Journal of Multivariate Analysis*, 5(2):121–177, 1975.

136. P. Medvegyev. *Stochastic Integration Theory*. Oxford Graduate Texts in Mathematics. OUP Oxford, 2007. ISBN 9780199215256. 10.1093/oso/9780199215256.001.0001.

137. D. Meintrup and Schäffler S. *Stochastik: Theorie und Anwendungen (Statistik und ihre Anwendungen) (German Edition)*. Springer, 1. edition, 2005. ISBN 9783540216766. 10.1007/b137972.

138. J. Memin. Espaces de semi martingales et changement de probabilité. *Probability Theory and Related Fields*, 52 (1):9–39, 01 1980. ISSN 1432-2064. 10.1007/BF00534184.

139. M. Métivier. *Semimartingales: A Course on Stochastic Processes*, volume 2 of *De Gruyter studies in mathematics*. XI, Berlin, New York, 1982. ISBN 9783110086744.

140. M. Métivier. Semimartingales. In *Semimartingales*. de Gruyter, 2011.

141. M. Metivier and J. Pellaumail. Mesures stochastiques à valeurs dans des espaces L_0. *Zeitschrift für Wahrscheinlichkeitstheorie und verwandte Gebiete*, 40:101–114, 1977.

142. M. Métivier and J. Pellaumail. *Stochastic integration*. Academic Press, 2014.

143. M. Métivier and G. Pistone. Une formule d'isométrie pour l'intégrale stochastique hilbertienne et équations d'évolution linéaires stochastiques. *Zeitschrift für Wahrscheinlichkeitstheorie und Verwandte Gebiete*, 33(1):1–18, 1975.

144. P. A. Meyer. Intégrales stochastiques iv–v. *Sém. de Prob.*, X–XI.

145. P.-A. Meyer. A decomposition theorem for supermartingales. *Illinois Journal of Mathematics*, 6(2):193–205, 1962. 10.1215/ijm/1255632318.

146. P.-A. Meyer. Decomposition of supermartingales: The uniqueness theorem. *Illinois Journal of Mathematics*, 7 (1):1–17, 1963. 10.1215/ijm/1255637477.

147. P.-A. Meyer. *Probabilités et Potentiel*. Hermann, Paris, 1966.

148. P.-A. Meyer. On the multiplicative decomposition of positive supermartingales. pages 103–116, 1967.

149. P.-A. Meyer. Intégrales stochastiques I. *Séminaire de Probabilités de Strasbourg*, 39:72–94, 1967.

150. P.-A. Meyer. Intégrales stochastiques II. *Séminaire de probabilités de Strasbourg*, 1:95–117, 1967.

151. P.-A. Meyer. Intégrales stochastiques III. *Séminaire de probabilités de Strasbourg*, 1:118–141, 1967.

152. P.-A. Meyer. Intégrales stochastiques IV. *Séminaire de probabilités de Strasbourg*, 1:142–162, 1967.

153. P.-A. Meyer. *Martingales and stochastic integrals I.*, volume 284 of *Lecture notes in mathematics*. Springer, 1972. 10.1007/bfb0069773.

154. P. A. Meyer. Notes sur les intégrales stochastiques. IV. Caractérisation de BMO par un opérateur maximal. pages 446–481, 1977. 10.1007/bf01932733. URL https://doi.org/10.1007/bf01932733.

155. P.-A. Meyer. Le théorème fondamental sur les martingales locales. *Séminaire de Probabilités XI*, 581:463–464, 1977.

156. P.-A. Meyer. Caracterisation des semimartingales, d'apres dellacherie. In C. Dellacherie, P.-A. Meyer, and M. Weil, editors, *Séminaire de Probabilités XIII*, volume 721 of *Lecture Notes in Mathematics*, pages 620–623. Springer Berlin Heidelberg, 1979. ISBN 978-3-540-09505-7. 10.1007/BFb0070899.

157. P.-A. Meyer. Un cours sur les intégrales stochastiques. In M. Émery and M. Yor, editors, *Séminaire de probabilités 1967 - 1980*, volume 1771 of *Lecture Notes in Mathematics*, pages 174–329. Springer Berlin Heidelberg, 2002. ISBN 978-3-540-42813-8. 10.1007/978-3-540-45530-1_11.

158. Paul-André Meyer. *Les processus stochastiques de 1950 à nos jours*, pages 813–848. Birkhäuser, Basel, 2000.

159. R Mikulevicius and BL Rozovskii. Martingale problems for stochastic pdes. *Stochastic partial differential equations: six perspectives*, 64:243–326, 1999.

160. Remigijus Mikulevicius and Boris L Rozovskii. Normalized stochastic integrals in topological vector spaces. In *Séminaire de probabilités XXXII*, pages 137–165. Springer, 1998. 10.1007/bfb0101756.

161. M. Motoo and S. Watanabe. On a class of additive functionals of Markov processes. *Journal of Mathematics of Kyoto University*, 4(3):429–469, 1965.

162. Michael Mürmann. *Wahrscheinlichkeitstheorie und stochastische Prozesse*. Springer, 2014.

163. M. Musiela and M. Rutkowski. *Martingale Methods in Financial Modelling (Stochastic Modelling and Applied Probability)*. Springer, 2nd edition, 2011. ISBN 9783540209669.

164. Jacques Neveu. *Martingales à temps discret*. Masson et Cie, Éditeurs, Paris, 1972.

165. Raymond Edward Alan Christopher Paley, Norbert Wiener, and Antoni Zygmund. Notes on random functions. *Mathematische Zeitschrift*, 37(1):647–668, 1933. 10.1007/bf01474606.

166. J. Pellaumail. *Sur l'intégrale stochastique et la décomposition de Doob-Meyer*. Société mathématique de France, 1973.

167. D. Pollard, R. Gill, and B. D. Ripley. *A User's Guide to Measure Theoretic Probability*. Cambridge Series in Statistica. Cambridge University Press, 2002. ISBN 9780521002899.

168. Jean-Luc Prigent. *Weak convergence of financial markets*. Springer Finance. Springer-Verlag, Berlin, 2003. ISBN 3-540-42333-8. 10.1007/978-3-540-24831-6. URL https://doi.org/10.1007/978-3-540-24831-6.

169. P. Protter. A comparison of stochastic integrals. *The Annals of Probability*, 7(2):276–289, 1979. ISSN 0091–1798. 10.1214/aop/1176995088.

170. P. Protter. Semimartingales and stochastic differential equations. *Lecture Notes 3rd Chilean Winter School in Probab. and Statist., Technical Report*, pages 85–25, 1985.

171. P. Protter. Stochastic integration without tears. *Stochastics: An International Journal of Probability and Stochastic Processes*, 16(3-4):295–325, 1986. 10.1080/17442508608833378.

172. P. Protter. *Stochastic Integration and Differential Equations: Version 2.1 (Stochastic Modelling and Applied Probability)*. Springer, 2010. ISBN 9783642055607.

173. K. Murali Rao. On decomposition theorems of Meyer. *Math. Scand.*, 24:66–78, 1969. ISSN 0025-5521. 10.7146/math.scand.a-10920. URL https://doi.org/10.7146/math.scand.a-10920.

174. Malempati M Rao. *Stochastic processes: general theory*, volume 342. Springer Science & Business Media, 2013.

175. Murali Rao. Doob's decomposition and Burkholder's inequalities. In *Séminaire de Probabilités, VI (Univ. Strasbourg, année universitaire 1970–1971; Journées Probabilistes de Strasbourg, 1971)*, Lecture Notes in Math., Vol. 258, pages 198–201. Springer, Berlin-New York, 1972.

176. D. Revuz and M. Yor. *Continuous Martingales and Brownian Motion*, volume 293 of *Grundlehren der mathematischen Wissenschaften A series of comprehensive studies in mathematics*. Springer, Berlin, Heidelberg, 3rd edition, 1999. ISBN 9783540643258.

177. Markus Riedle. Stochastic integration with respect to cylindrical Lévy processes in Hilbert spaces: an L^2 approach. *Infin. Dimens. Anal. Quantum Probab. Relat. Top.*, 17(1):1450008, 19, 2014. ISSN 0219-0257. 10.1142/S0219025714500088. URL https://doi.org/10.1142/S0219025714500088.

178. L. Rogers and D. Williams. *Diffusions, Markov Processes and Martingales: Volume 2, Itô Calculus*. Cambridge Mathematical Library. Cambridge University Press, 2000. ISBN 9780521775939.

179. L. C. G. Rogers. Arbitrage with fractional Brownian motion. *Mathematical Finance*, 7(1):95–105, 1997. ISSN 0960–1627.

180. Barbara Rüdiger. Stochastic integration with respect to compensated Poisson random measures on separable Banach spaces. *Stochastics and Stochastic Reports*, 76(3):213–242, 2004.

181. J. Ruiz de Chávez. Le théorème de Paul Lévy pour des mesures signées. In *Seminar on probability, XVIII*, volume 1059 of *Lecture Notes in Math.*, pages 245–255. Springer, Berlin, 1984. 10.1007/BFb0100048. URL https://doi.org/10.1007/BFb0100048.

182. Ludger Rüschendorf. *Stochastic processes and financial mathematics*. Springer, 2023.

183. A. N. Shiryaev. Absolute continuity and singularity of probability measures in functional spaces. pages 209–225, 1980.

184. A. N. Shiryaev and A. Cherny. Vector stochastic integrals and the fundamental theorems of asset pricing. *Proceedings of the Steklov Institute of Mathematics-Interperiodica Translation*, 237:6–49, 2002.

185. Albert N. Shiryaev. *Essentials of stochastic finance*, volume 3 of *Advanced Series on Statistical Science & Applied Probability*. World Scientific Publishing Co., Inc., River Edge, NJ, 1999. ISBN 981-02-3605-0. 10.1142/9789812385192. URL https://doi.org/10.1142/9789812385192. Facts, models, theory, Translated from the Russian manuscript by N. Kruzhilin.

186. S. E. Shreve and I. Karatzas. *Brownian Motion and Stochastic Calculus (Graduate Texts in Mathematics)*. Springer, korrigierter 8. druck, 2. edition, 2008. ISBN 9780387976556.

187. A. V. Skorohod. A remark on square integrable martingales. *Teor. Verojatnost. i Primenen.*, 20:199–202, 1975. ISSN 0040-361x.

188. A. V. Skorokhod. Studies in the theory of random processes. pages viii+199, 1965. Translated from the Russian by Scripta Technica, Inc.

189. A. V. Skorokhod. Integral representation of martingales. *Z. Wahrscheinlichkeitstheorie verw. Gebiete*, 11: 367–388, 1969.

190. RL Stratonovich. A new representation for stochastic integrals and equations. *SIAM Journal on Control*, 4(2): 362–371, 1966. 10.1137/0304028.

191. P. Tankov. *Financial Modelling with Jump Processes, Second Edition*. Chapman & Hall/CRC Financial Mathematics Series. Taylor & Francis, 2003. ISBN 9780203485217. 10.1201/9780203485217.

192. Tin-Lam Toh and Tuan-Seng Chew. The riemann approach to stochastic integration using non-uniform meshes. *Journal of mathematical analysis and applications*, 280(1):133–147, 2003. ISSN 0022–247X. 10.1016/s0022-247x(03)00059-3.

193. Aleš Černý and Johannes Ruf. Simplified stochastic calculus via semimartingale representations. *Electron. J. Probab.*, 27:Paper No. 3, 32, 2022. 10.1214/21-ejp729. URL https://doi.org/10.1214/21-ejp729.

194. John B Walsh. An introduction to stochastic partial differential equations. In *École d'Été de Probabilités de Saint Flour XIV-1984*, pages 265–439. Springer, 1986.

195. A. T. Wang. Quadratic variation of functionals of Brownian motion. *The Annals of Probability*, pages 756–769, 1977. ISSN 0091–1798.

196. Shinzo Watanabe. The Japanese contributions to martingales. *J. Électron. Hist. Probab. Stat.*, 5(1):13, 2009. ISSN 1773-0074.

197. Norbert Wiener. Differential-space. *Journal of Mathematics and Physics*, 2(1-4):131–174, 1923.

198. Norbert Wiener. The Dirichlet problem. *Journal of Mathematics and Physics*, 3(3):127–146, 1924. 10.1002/sapm192433127.

199. Jia-An Yan. *Introduction to Stochastic Finance*. Universitext. Springer Singapore, 1 edition, 2018. ISBN 978-981-13-1657-9. 10.1007/978-981-13-1657-9.

200. Ch Yoeurp. Décompositions des martingales locales et formules exponentielles. In *Séminaire de Probabilités X Université de Strasbourg*, pages 432–480. Springer, 1976.

201. M. Yor. Un exemple de processus qui n'est pas une semi-martingale. *Temps Locaux*, 52:219–222, 1978.

202. B. Øksendal. *Stochastic Differential Equations: An Introduction with Applications (Universitext)*. Springer, 6. edition, 2010. ISBN 9783540047582.