A New Approach to General Stochastic Integration and Sigma Martingales

by

Moritz Sohns

submitted in accordance with the requirements
for the degree of

Doctor of Philosophy

in the subject

Mathematics

at the

UNIVERSITY OF SOUTH AFRICA

SUPERVISOR: PROF Z I ALI

May 31, 2026

The copyright of this thesis vests in the author. No quotation from it or information derived from it is to be published without full acknowledgment of the source. The thesis is to be used for private study or non-commercial research purposes only.

Published by the University of South Africa (UNISA) in terms of the nonexclusive license granted to UNISA by the author.

Declaration

I, Moritz Sohns, declare that this thesis entitled *A New Approach to General Stochastic Integration and Sigma Martingales*, and the work presented in it, is my own. The work described in this Doctor of Philosophy (PhD) thesis was carried out under the supervision of Prof. Zakaria Idriss Ali in the Department of Mathematical Sciences at the University of South Africa.

This thesis represents original work by the author and has not been submitted in any form for any degree or diploma at any other University. Where use has been made of the work of others, it is duly acknowledged in the text.



Moritz Sohns

May 31, 2026

To my late grandfather Karlheinz, who inspired my love for logic, rational thinking, and mathematics.

Abstract

This thesis develops a new, topology-based construction of the general stochastic vector integral together with a representation theory for sigma martingales, and applies these results to mathematical finance. The construction overcomes technical limitations of the classical and functional analytic approaches to stochastic integration, while the representation theory extends key tools of stochastic analysis to the natural class of sigma martingales and yields, in particular, a new proof of the Second Fundamental Theorem of Asset Pricing in the general semimartingale model.

The first main contribution is a direct construction of the general stochastic vector integral via the semimartingale topology of Émery. Classical approaches and the functional analytic approach require decomposition theorems that usually entail cumbersome and technical proof strategies. We bypass these arguments entirely. The integral is defined as the unique continuous extension of the elementary stochastic integral with respect to the semimartingale topology, and the construction handles bounded and unbounded integrands within a single framework. The vector-valued case follows directly, without a separate componentwise definition. This topological construction simplifies several classical arguments and shortens proofs, while preserving all essential properties of the integral, including associativity, linearity, the standard convergence theorems and even the multi-dimensional general Itô formula.

The thesis further develops a self-contained theory of sigma martingales, a class of processes that generalizes local martingales and arises naturally in financial models with unbounded jumps and discontinuities. The treatment avoids the theory of semimartingale characteristics, which is powerful but technically demanding, and instead develops the required results through stochastic-integral and local-martingale arguments. A new proof of the Ansel–Stricker Lemma is given, together with a novel generalization, which turns out to be essential for transforming financial market models from nominal to discounted form.

The third main contribution is a martingale representation theorem for sigma martingales, which extends the classical Jacod–Yor theorem. Earlier versions of this theorem cover square-integrable martingales or, at most, local martingales. We show that for a

sigma martingale X, the predictable representation property for sigma martingales is equivalent to the absence of non-trivial orthogonal sigma martingales, to the uniqueness of the absolutely continuous sigma martingale measure for X, and to the extremality of the original measure in the convex set of sigma martingale measures. As a direct consequence, we obtain a new proof of the Second Fundamental Theorem of Asset Pricing in the general semimartingale model. To the best of our knowledge, this is the first proof of the Second Fundamental Theorem under the same hypotheses used for the First Fundamental Theorem; previous treatments either imposed additional technical conditions or worked with local martingale measures that need not exist in general markets.

Keywords: Sigma martingale, General stochastic integral, Semimartingale, Asset pricing, Fundamental Theorem of Asset Pricing, Jacod–Yor theorem, Topology-based integration

Acknowledgements

I would like to express my deepest gratitude to my supervisor, Prof. Dr. Zakaria Ali, for his guidance, encouragement, and unwavering support throughout the course of this thesis. His responsiveness, even during weekends, his thoughtful mentorship, and the consistently friendly and constructive exchanges have made a significant impact on both this work and my academic growth.

I am also grateful to the many co-students, supervisors, and lecturers from my previous academic stations, the Technical University of Darmstadt, the University of Oxford, and the University of Finance and Administration in Prague, for the many inspiring conversations and the support I received during my studies.

Contents

Declaration iii

Abstract v

Acknowledgements vii

List of Symbols x

1 Introduction 1

2 Preliminaries 6

2.1 Basic Framework and Stochastic Processes 6

2.2 Martingales and Their Properties 8

2.3 The Predictable σ-Algebra and Related Concepts 12

2.4 Quasimartingales 14

2.5 Measurability, Inequalities, and Supporting Results 15

3 The General Stochastic Vector Integral – Historical Perspectives 20

3.1 The Classical Approach 21

3.2 The Functional Analytic Approach 52

3.3 The Stochastic Vector Integral 60

3.4 Stochastic Integration in Textbooks 65

3.5 Comparison with Recent Developments 67

4 The Topological Stochastic Vector Integral 71

4.1 Semimartingales 71

4.2 Properties of Semimartingales 82

4.3 The General Stochastic Integral 93

4.4 The Dominated Convergence Theorem 100

4.5 Properties of the Stochastic Integral 112

4.6 Stability of Local Martingales Under Stochastic Integration 117

4.7 The Quadratic Variation of a Semimartingale 120

5 The Itô Formula **129**

5.1 Statement and Proof 129

5.2 Applications of the Itô Formula 138

6 Sigma Martingales **148**

6.1 Definition and Properties 150

6.2 The Jacod–Yor Theorem for Sigma Martingales 172

7 Applications to Mathematical Finance **183**

7.1 The General Semimartingale Market Model with Dividends 183

7.2 Admissibility and Equivalent Sigma Martingale Measures 188

7.3 Application: Dividend-Paying Geometric Brownian Motion 191

7.4 Departures from the Classical Assumptions 193

8 Conclusion **195**

Bibliography **198**

Index **215**

List of Symbols

Basic Sets and Numbers

$\mathbb{N}$	Natural numbers
$\mathbb{R}$	Real numbers
$\mathbb{R}_+$	Non-negative real numbers $[0, \infty)$
$\mathbb{Q}$	Rational numbers
$\mathbb{C}$	Complex numbers

Probability and Measure Theory

$(\Omega, \mathcal{F}, \mathbf{P})$	Probability space
$\mathbb{F} = (\mathcal{F}_t)_{t \geq 0}$	Filtration
$\mathbf{P}$	Probability measure
$\mathbf{Q}$	Alternative probability measure
$\mathbf{E}$	Expectation operator
$\mathbf{E}[\cdot \vert \mathcal{F}_t]$	Conditional expectation
$\ll$	Absolute continuity of measures
$\sim$	Equivalence of measures
$\frac{\mathrm{d}\mathbf{Q}}{\mathrm{d}\mathbf{P}}$	Radon-Nikodym derivative
$\mathcal{B}(\mathbb{R})$	Borel σ-algebra on $\mathbb{R}$

Stochastic Processes

$X = (X_t)_{t \geq 0}$	Stochastic process
X_-	Left-continuous version of X
X_{t-}	Left limit at time t
ΔX_t	Jump of X at time t: $X_t - X_{t-}$
X^T	Process stopped at time T
X_t^*	Supremum process: $\sup_{s \in [0,t]} \vert X_s \vert$
X^*	$\sup_{t \geq 0} X_t^*$

Spaces of Processes

$\mathbb{D}$	Space of càdlàg processes (right-continuous with left limits)
$\mathbb{L}$	Space of càglàd processes (left-continuous with right limits)
$\mathcal{S}^d$	d-dimensional semimartingales
$\mathcal{S}_e^d$	Semimartingales with Émery topology
$\mathcal{M}^d$	d-dimensional martingales
$\mathcal{M}_\infty^d$	Bounded d-dimensional martingales
$\mathcal{M}_{\text{loc}}^d$	d-dimensional local martingales
$\mathcal{M}_\sigma^d$	d-dimensional sigma martingales
$\mathcal{M}_0$	Martingales starting at zero
$\mathcal{H}^p$	Hardy space of martingales
$\mathcal{BMO}$	Space of martingales of bounded mean oscillation
$\mathcal{P}$	Set of predictable processes
$\mathcal{L}$	Left-continuous adapted processes
Λ^d	d-dimensional simple predictable processes
Λ_u^d	Simple predictable with uniform topology
$L(X)$	Integrands for semimartingale X
$L^1(X)$	Special integrands for X
$\mathcal{I}(X)$	Stochastic integrals w.r.t. X

Integration and Operations

$H \bullet X$	Stochastic integral
$\int_0^t H_s \, \mathrm{d}X_s$	Alternative notation for stochastic integral
$[X, Y]$	Quadratic covariation
$[X]$ or $[X, X]$	Quadratic variation
$[X, Y]^c$	Continuous part of quadratic covariation
$\langle X, Y \rangle$	Predictable quadratic covariation
$\mathrm{Var}[A]$	Total variation of process A

Norms and Metrics

$\lVert \cdot \rVert_p$	L^p norm
$\lVert \cdot \rVert_{\mathcal{H}^p}$	Hardy space norm
$\lVert \cdot \rVert_u$	Uniform norm
$\lVert \cdot \rVert_{\text{ucp}}$	Norm for ucp topology
$\lVert \cdot \rVert_{\mathcal{S}}$	Émery norm
$\lVert \cdot \rVert_X$	Norm induced by semimartingale X
d_{ucp}	Metric for ucp topology

$d_{\mathcal{S}}$	Semimartingale metric
d_X	Metric induced by X

Convergence

$\xrightarrow{\text{ucp}}$	Uniformly on compacts in probability
$\xrightarrow{\mathcal{S}}$	In semimartingale topology
$\xrightarrow{X}$	In X-topology
$\xrightarrow{\mathbf{P}}$	In probability
$\xrightarrow{\text{a.s.}}$	Almost surely

Special Functions and Operators

$\mathcal{E}(X)$	Stochastic exponential (Doléans-Dade)
$\mathbf{1}_A$, charf_A	Indicator/characteristic function
$a \wedge b$	Minimum of a and b
$a \vee b$	Maximum of a and b
$x^+ = x \vee 0$	Positive part
$x^- = (-x) \vee 0$	Negative part
$\text{sgn}(x)$	Sign function

Intervals and Sets

$[a, b]$	Closed interval
(a, b)	Open interval
$[a, b)$, $(a, b]$	Half-open intervals
$[\![a, b]\!]$	Stochastic interval
$]\!]a, b]\!]$	Open stochastic interval

Relations

$\perp$	Orthogonality of processes
$X^\perp$	Processes orthogonal to X
esssup	Essential supremum
liminf	Limit inferior
limsup	Limit superior

Function Spaces

$\mathcal{C}^k$	k-times continuously differentiable functions
$\mathcal{C}_c$	Continuous functions with compact support
$\mathcal{C}_b$	Bounded continuous functions

L^p	Lebesgue space
L^∞	Essentially bounded functions

Financial Mathematics

$S = (S^0, S^1, \ldots, S^d)$	Price processes
φ	Trading strategy
V_t	Wealth process
$\mathcal{P}_\sigma(X)$	sigma martingale measures for X
NFLVR	No Free Lunch with Vanishing Risk
PRP	Predictable Representation Property

Chapter 1

Introduction

Stochastic integration is a cornerstone of modern applied mathematics and plays a central role in fields such as mathematical finance, physics, engineering, biology, and control theory. Over the past century, the theory has developed into a broad and flexible framework. It is now taught at universities worldwide, not only in mathematics but also in disciplines such as physics, biology, economics, computer science, statistics, and quantitative finance. Despite its wide adoption, there is no single unified approach to how the subject is introduced. The scope of the theory, both in terms of admissible integrands and integrators, varies between treatments. Multiple steps and generalizations are possible, and different texts and curricula stop at different stages in the development. Two prominent frameworks have emerged: the classical approach and the functional analytic approach. Each has its own formulation, methodology, and limitations. A range of generalizations and special cases has also been proposed, extending the theory to settings such as vector-valued integration, non-predictable integrands, and infinite-dimensional spaces.

The development of stochastic integration began with early attempts to describe random processes mathematically. Norbert Wiener (1894–1964) constructed an integral for Brownian motion with smooth deterministic integrands, using integration by parts [160, 189, 190]. His formulation imposed strong regularity assumptions and did not yield a general probabilistic calculus. Around the same time, Wolfgang Doeblin (1915–1940) developed a theory of stochastic integration that allowed non-differentiable random integrands. His work anticipated many features of modern stochastic calculus but remained unknown until rediscovered much later [17].

A systematic and rigorous treatment appeared in the 1940s through the work of Kiyosi Itô (1915–2008) [84, 85, 86, 87, 88, 89]. He defined integration with respect to Brownian motion by taking limits of adapted Riemann sums. His construction introduced what is now known as the Itô integral and led to the Itô formula, which generalizes the chain rule in the stochastic setting.

Efforts to extend the integral beyond Brownian motion led to integration with respect to martingales. Doob observed that both B_t and $B_t^2 - t$ are martingales and suggested a general martingale-based theory [51]. This idea required a decomposition of submartingales into a martingale and a predictable process of finite variation. A general version of this decomposition was proved in continuous time by Paul-André Meyer in the early 1960s [139, 140], providing the foundation for stochastic integration with a broader class of processes.

In subsequent work, Kunita, Watanabe, Itô, and Meyer developed the theory for square-integrable and locally square-integrable martingales [121, 154, 91, 143, 144, 145, 146]. They introduced the predictable quadratic variation and clarified the Hilbert space structure of square-integrable martingales. These ideas made it possible to extend the stochastic integral by means of isometries and closure arguments.

By the 1970s, semimartingales had been identified as the natural generalization of both martingales and finite variation processes. Doléans-Dade and Meyer formalized the class of semimartingales and proved that locally bounded predictable integrands can be integrated with respect to them [49]. This result allowed the theory to move beyond Brownian motion and Markov processes and to adopt a general martingale-based formulation.

Further generalizations removed the assumption that the integrand is locally bounded. Jacod developed a version of the stochastic integral that allows unbounded predictable integrands, based on an analysis of semimartingale jumps [92, 93]. Chou later proposed a different approach, using topological tools such as the semimartingale topology [27], which had been introduced by Émery [57]. These developments also showed that the space of stochastic integrals with respect to vector semimartingales is not closed under componentwise summation. A more general construction of the vector-valued stochastic integral was then introduced by Jacod [94], and its closure properties were studied by Mémin [135] and Shiryaev and Cherny [178].

In parallel with the classical approach, a functional analytic method was developed. This approach was motivated by the problem of characterizing the class of processes for which the stochastic integral is continuous on simple predictable integrands. Results by Bichteler [11, 12] and Dellacherie [42] showed that semimartingales form the largest such class. These findings were further elaborated by Meyer [149], Lenglart [127], and Protter [164, 165]. The stochastic integral is first defined for càglàd integrands and then extended to general predictable processes. The functional analytic viewpoint leads to the Bichteler–Dellacherie–Mokobodzki theorem, which characterizes semimartingales as the maximal class of integrators for which the stochastic integral is well-behaved under convergence.

Although the functional analytic approach simplifies some aspects of the theory, it still relies on decomposition theorems when extending the integral to general predictable

integrands. Moreover, it does not provide a separate construction for the vector-valued case. For this extension, one still needs to resort to the classical approach.

Three concrete research questions guide this thesis:

1. Can the general stochastic vector integral be constructed directly via the semimartingale topology, without recourse to the decomposition theorems on which the classical and functional analytic approaches depend, and, more importantly, can all the fundamental results be established in a self-contained manner within this approach, so that the construction is usable for the full range of applications of stochastic integration?

2. Can the Ansel–Stricker Lemma be reproved and strictly generalized by replacing the constant lower bound with a càdlàg process minorant with locally integrable supremum process, and does this generalized result provide the precise tool needed to pass from nominal admissibility to discounted local-martingale and supermartingale properties in general semimartingale market models?

3. Can the Jacod–Yor representation theorem be generalized from local martingales to sigma martingales, and does this generalization yield a proof of the Second Fundamental Theorem of Asset Pricing in the sigma martingale setting, under the same hypotheses as the First Fundamental Theorem?

The remainder of this introduction sketches our answers to these questions; the chapter-by-chapter strategy by which they are reached is laid out at the end of the section.

This thesis develops a new, topology-based approach to defining the general stochastic vector integral. By leveraging the semimartingale topology introduced by Émery [57], we construct the integral directly via closure and continuity arguments. This eliminates the need for decompositions and leads to a formulation that is both conceptually elegant and practically useful.

While this approach is sufficient for defining the integral and for establishing the fundamental theorems of asset pricing, we go further. We also develop the Itô formula in this setting and demonstrate key applications. Even though this is not technically necessary for the definition of the general stochastic integral nor for the proof of the Fundamental Theorems of Asset Pricing, Itô calculus is indispensable in the practice of finance. Many pricing models, such as Black–Scholes, rely on Itô's formula to derive dynamics, hedging strategies, and partial differential equations. Our version of the Itô formula applies directly to our topological stochastic integral and confirms that this framework retains all the core tools of classical stochastic calculus.

Another central object in modern financial mathematics is the sigma martingale. Introduced by Chou [26] and further developed by Émery [58], sigma martingales generalize local martingales and naturally arise in financial models with unbounded jumps and discontinuities, or constraints. Sigma martingales are especially important when discussing admissible trading strategies and when applying the fundamental theorems of asset pricing in full generality.

Despite their importance, sigma martingales remain underrepresented in the literature. When they are discussed, it is typically in the context of general semimartingale models or via the highly technical theory of semimartingale characteristics [95]. Even fewer treatments explore their structure deeply or provide accessible constructions.

In this thesis, we provide a simplified, self-contained approach to sigma martingales. Our framework is independent of semimartingale characteristics and instead builds upon core properties of stochastic integrals and local martingales. This culminates in a new proof of the Ansel–Stricker Lemma [3], which characterizes one-sided bounded sigma martingales as local martingales. We go further and present a novel generalization: if a sigma martingale is bounded from below by the sum of a local martingale and a càdlàg minorant whose supremum process is locally integrable, then it is itself a local martingale. This generalization plays a crucial role in financial modeling, especially in transforming a market model from nominal to discounted form, where the integrability properties of price processes must be preserved.

The subsequent part investigates martingale representation theorems. These results are essential in mathematical finance, as they allow every integrable claim to be expressed as the terminal value of a stochastic integral, hence admitting replication. While such representation theorems exist for Brownian motion, Lévy processes, and local martingales, there has not been, until now, a corresponding theory for sigma martingales. This is surprising because in models with no arbitrage, the equivalent measure under which discounted prices are martingales is in general only a sigma martingale measure.

To fill this gap, we generalize the celebrated Jacod–Yor Theorem [96] to sigma martingales. Our result shows that if a sigma martingale X admits no non-trivial orthogonal sigma martingales, then every sigma martingale is a stochastic integral with respect to X. This gives a complete characterization of the martingale representation property in the sigma martingale setting and allows us to establish the Second Fundamental Theorem of Asset Pricing in a fully general semimartingale model, under the same assumptions used to prove the First Fundamental Theorem. To the best of our knowledge, this is the first such proof in the literature.

The thesis is structured as follows.

- **Chapter 2** introduces the mathematical background for stochastic processes, filtra-

tions, stopping times, local martingales, and the semimartingale framework.

- **Chapter 3** reviews the classical and functional analytic constructions of the stochastic integral, including their limitations and generalizations.
- **Chapter 4** presents our topological construction of the general vector stochastic integral using the semimartingale topology.
- **Chapter 5** establishes the Itô formula for our stochastic integral and shows that classical applications remain valid.
- **Chapter 6** develops a general theory of sigma martingales, including a new version of the Ansel–Stricker Lemma and the Jacod–Yor representation theorem for sigma martingales.
- **Chapter 7** applies the framework to mathematical finance: the general semimartingale market model with dividends, the Fundamental Theorems of Asset Pricing for sigma martingales, fair pricing bounds, and a worked example through stochastic differential equations.
- **Chapter 8** discusses conclusions and possible extensions to infinite-dimensional or constrained market settings.

Chapter 2

Preliminaries

This chapter establishes the mathematical foundation for the stochastic analysis developed in this thesis. We present the fundamental concepts, definitions, and classical results that underpin the theory of stochastic integration and semimartingales. The material is organized to provide a self-contained treatment, starting from basic probabilistic concepts and progressing to more specialized notions required for the advanced topics in subsequent chapters.

2.1 Basic Framework and Stochastic Processes

Throughout this thesis, we work within the framework of a complete filtered probability space. This setting provides the necessary structure for modeling the evolution of random phenomena over time while ensuring technical conditions that facilitate rigorous mathematical analysis.

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*:

(a) $\mathcal{F}_0$ includes all $\mathbf{P}$-null sets from $\mathcal{F}$ (completeness);

(b) 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 $t \geq 0$.

The completeness condition ensures that null sets can be handled uniformly, while right-continuity provides technical advantages in the study of stopping times and stochastic processes. These conditions, introduced by Meyer [139], have become standard in modern stochastic analysis.

For the analysis of random variables and stochastic processes, we employ the classical L^p spaces. For $p \in [1, \infty)$, we define the seminorm

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

and the corresponding space

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

The *essential supremum* of a random variable X, defined as

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

plays a crucial role in defining L^∞ and in various boundedness conditions throughout the thesis.

The *indicator function* $\mathbb{1}_A$ of a subset A is defined by

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

A set $E \subseteq \Omega \times \mathbb{R}_+$ is called *evanescent* if the set $\{\omega \in \Omega : \exists t \geq 0 \text{ such that } (t, \omega) \in E\}$ is a $\mathbf{P}$-null set. This concept is crucial when working with stochastic processes, as we typically consider processes modulo evanescent sets, meaning two processes that differ only on an evanescent set are considered equivalent.

A *stochastic process* X on $(\Omega, \mathcal{F}, \mathbf{P})$ is a collection of $\mathbb{R}$-valued or $\mathbb{R}^d$-valued random variables $(X_t)_{t\geq 0}$. The behavior and properties of these processes form the core objects of study in this thesis.

A process X is said to be *adapted* to the filtration $\mathbb{F}$ if X_t is $\mathcal{F}_t$-measurable for each t. This condition ensures that the value of the process at time t depends only on information available up to time t, capturing the non-anticipative nature of real-world phenomena.

Let $(\Omega, \mathcal{F}, \mathbf{P})$ be a probability space and let $T \subseteq \mathbb{R}$ be an index set (typically $T = [0, \infty)$). Two stochastic processes $X = (X_t)_{t\in T}$ and $Y = (Y_t)_{t\in T}$ defined on this space are said to be *modifications* of each other if $X_t = Y_t$ almost surely for every $t \in T$, while they are *indistinguishable* if $X_t(\omega) = Y_t(\omega)$ for all $t \in T$ and for almost every $\omega \in \Omega$. The distinction becomes important when dealing with uncountable index sets.

Regarding path regularity, a process X is said to be *càdlàg* if its paths are right-continuous with left limits, and *càglàd* if its paths are left-continuous with right limits. Throughout this thesis, unless stated otherwise, all processes are assumed to have càdlàg paths.

For a càdlàg process X, we define:

(a) The left-limit process X_- by $(X_-)_t = X_{t-} := \lim_{s\uparrow t} X_s$, with the convention $X_{0-} = X_0$;

(b) The jump process ΔX by $\Delta X_t := X_t - X_{t-}$, with the convention $\Delta X_0 = 0$.

A *stopping time* T is a non-negative random variable defined on Ω such that $\{T \leq t\} \in \mathcal{F}_t$ for each $t \geq 0$. This condition ensures that the decision to stop at time T depends only on information available up to that time. For a stopping time T, the σ-algebra $\mathcal{F}_T$ of events occurring up to time T consists of those events $A \in \mathcal{F}$ satisfying

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

Given a process X and a stopping time T, the *stopped process* is defined by $X_t^T := X_{t \wedge T}$. For stopping times S, T, we define the *stochastic intervals*:

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

A property is said to hold *locally* for a stochastic process X if there exists a sequence of stopping times $(T^n)_{n \in \mathbb{N}}$ increasing to ∞ almost surely such that the property holds for each stopped process X^{T^n}. Similarly, a property holds *prelocally* if it holds for each $X_-^{T^n}$. Common examples include:

(a) A process X is *locally bounded* if there exists a localizing sequence (T^n) such that each X^{T^n} is bounded;

(b) A process X is *locally integrable* if there exists a localizing sequence (T^n) such that $\sup_{t \leq T^n} |X_t| \in L^1$ for each n.

2.2 Martingales and Their Properties

Martingales are fundamental objects in probability theory, representing fair games and providing powerful tools for stochastic analysis. A real-valued, adapted process $X = (X_t)_{0 \leq t < \infty}$ is called a *martingale* with respect to the filtration $\mathbb{F}$ if:

(a) $X_t \in L^1(\mathbf{P})$ for all t;

(b) For $s \leq t$, we have $\mathbf{E}[X_t | \mathcal{F}_s] = X_s$ almost surely.

Similarly, X is a *supermartingale* (resp. *submartingale*) if the equality in condition (b) is replaced by "$\leq$" (resp. "$\geq$").

A process X is a *local martingale* if there exists a sequence of stopping times $(T^n)_{n\in\mathbb{N}}$ increasing to ∞ almost surely such that each stopped process X^{T^n} is a martingale. This localization concept allows us to work with processes that may not be integrable globally but exhibit martingale behavior locally.

We denote by $\mathcal{M}^d$ the space of d-dimensional martingales, by $\mathcal{M}^d_\infty$ the space of bounded d-dimensional martingales, and by $\mathcal{M}^d_{\text{loc}}$ the space of d-dimensional local martingales. A subscript 0, as in $\mathcal{M}_{\text{loc},0}$, indicates that the process starts at 0, that is, $M_0 = 0$ almost surely for all $M \in \mathcal{M}_{\text{loc},0}$.

A family of random variables $(U_\alpha)_{\alpha\in A}$ is *uniformly integrable* if

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

This concept is crucial for ensuring the convergence of expectations and plays a central role in martingale limit theorems.

Theorem 2.2.1 (Doob's Convergence Theorem, [166, Theorem I.12])**.** *Let X be a right-continuous martingale which is uniformly integrable. Then $Y = \lim_{t\to\infty} X_t$ exists almost surely, $\mathbf{E}[|Y|] < \infty$, and Y is said to* close *X as a martingale.*

This theorem establishes that uniformly integrable martingales converge almost surely to an integrable limit. The limit Y provides a natural extension of the martingale to time infinity, and the martingale property extends to include this limiting value. This result is fundamental in many applications, particularly in proving representation theorems and in the study of optimal stopping problems.

Theorem 2.2.2 ([166, Theorem I.13])**.** *Let X be a right-continuous martingale with $X_0 = 0$. Then $(X_t)_{0\leq t<\infty}$ is uniformly integrable if and only if $Y = \lim_{t\to\infty} X_t$ exists almost surely, $\mathbf{E}[|Y|] < \infty$, and $(X_t)_{0\leq t\leq\infty}$ is a martingale, where $X_\infty = Y$.*

This characterization theorem provides a convenient criterion for checking uniform integrability of martingales through their long-term behavior. It shows that uniform integrability is equivalent to the existence of a finite limit that maintains the martingale property.

Lemma 2.2.3 ([79, 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.

This transformation lemma is a powerful tool in constructing new local martingales from existing ones. It shows that multiplying the increment of a local martingale after a stopping time by a bounded random variable measurable at that stopping time preserves the local martingale property. This result is extensively used in the construction of stochastic integrals and in proving various properties of semimartingales.

Theorem 2.2.4 (Martingale Characterization, [31, Theorem 5.4.6])**.** *An adapted càdlàg process X with closure X_∞ is a uniformly integrable martingale if and only if:*

(a) $X_T \in L^1$ *for each stopping time* T*;*

(b) $\mathbf{E}[X_T] = \mathbf{E}[X_0]$ *for each stopping time* T*.*

This characterization provides a practical way to verify the martingale property without checking the conditional expectation condition directly. It is particularly useful in proving that certain constructed processes are martingales.

Corollary 2.2.5 (Optional Stopping, [166, Theorem I.18])**.** *Let X be a right-continuous martingale, and let T be a stopping time. Then X^T is also a right-continuous martingale. If T is bounded or X is uniformly integrable, then $X^T \in L^1$.*

The optional stopping theorem is one of the most useful tools in martingale theory. It ensures that the martingale property is preserved under stopping, which is essential for localization arguments.

For martingale spaces, we introduce the $\mathcal{H}^p$ spaces which play a crucial role in the theory:

Definition 2.2.6. For M a martingale and $p \in [1, \infty)$, write

$$\|M\|_{\mathcal{H}^p} := \|M^*_\infty\|_p = \mathbf{E}\left[\sup_t |M_t|^p\right]^{\frac{1}{p}}.$$

Here $\|\cdot\|_p$ denotes the norm in L^p. Then $\mathcal{H}^p$ is the space of martingales such that

$$\|M\|_{\mathcal{H}^p} < \infty.$$

Remark 2.2.7. If we identify processes that are indistinguishable, $\|\cdot\|_{\mathcal{H}^p}$ becomes a norm and $\mathcal{H}^p$ a normed vector space.

Theorem 2.2.8 ([31, Theorem 10.1.7]). *The space of bounded martingales (and hence $\mathcal{H}^p$ for any $p > 1$) is dense in $\mathcal{H}^1$.*

This density result is crucial for approximation arguments. It allows us to first prove results for bounded martingales, where calculations are simpler, and then extend them to the larger space $\mathcal{H}^1$ by continuity.

Theorem 2.2.9 ([31, Lemma 10.1.9]). *Let M be a local martingale. Then M is locally in $\mathcal{H}^1$, that is, there exists a sequence of stopping times $(T^n)_{n\in\mathbb{N}}$ with $T^n \uparrow \infty$ almost surely such that each stopped process M^{T^n} belongs to $\mathcal{H}^1$. Equivalently,*

$$\mathbf{E}\left[\sup_{t\geq 0}|M_{t\wedge T^n}|\right] < \infty \quad \text{for all } n \in \mathbb{N}.$$

Theorem 2.2.10. *Let $0 < p \leq \infty$. Then, a càdlàg process X is locally uniformly bounded in L^p if and only if ΔX is locally uniformly bounded in L^p.*

Proof. For $t \geq 0$, write $X_t^* := \sup_{0\leq s\leq t}|X_s|$ and $(\Delta X)_t^* := \sup_{0\leq s\leq t}|\Delta X_s|$.

If X is locally uniformly bounded in L^p, then the same is true of ΔX, since

$$(\Delta X)_t^* = \sup_{0\leq s\leq t}|X_s - X_{s-}| \leq \sup_{0\leq s\leq t}|X_s| + \sup_{0\leq s\leq t}|X_{s-}| \leq 2X_t^*.$$

Localization on X therefore also localizes ΔX, which proves one implication.

Conversely, suppose first that ΔX is uniformly bounded in L^p on compact intervals, not merely locally so. Define

$$T_n := \inf\{t \geq 0 \,:\, |X_t| \geq n\}.$$

Since X has càdlàg paths, it is pathwise bounded on compact intervals, and hence $T_n \uparrow \infty$ almost surely. For every fixed $t \geq 0$, the running supremum of the stopped process satisfies

$$(X^{T_n})_t^* \leq n + \mathbb{1}_{\{T_n\leq t\}}|\Delta X_{T_n}| \leq n + (\Delta X)_t^*,$$

where the first inequality uses that immediately before T_n the process is strictly below n in absolute value and any excess at the entry time itself is absorbed by the jump ΔX_{T_n}. Thus X^{T_n} is uniformly bounded in L^p on compact intervals, and X is locally uniformly bounded in L^p.

In the general local case, choose a localizing sequence $(R_m)_{m\in\mathbb{N}}$ for ΔX. Applying the preceding step to each stopped process X^{R_m} (whose jump process ΔX^{R_m} is uniformly

bounded in L^p on compact intervals by the choice of R_m) and combining the resulting stopping times by a standard diagonal localization argument yields a single localizing sequence for X. This proves the converse implication. □

The space $\mathcal{BMO}$ (bounded mean oscillation) plays a special role in martingale theory, particularly as the dual space of $\mathcal{H}^1$:

Definition 2.2.11. Let $M \in \mathcal{H}^2$. We say $M \in \mathcal{BMO}$, if there exists a constant c such that for any stopping time T we have

$$\mathbf{E}\left[(M_\infty - M_{T-})^2 \mid \mathcal{F}_T\right] \leq c^2 \quad \text{a.s.} \tag{2.1}$$

The smallest such c is defined to be the *$\mathcal{BMO}$-Norm* of M, and it is written $\|M\|_{\mathcal{BMO}}$. If the constant c does not exist, or if M is not in $\mathcal{H}^2$, then we set $\|M\|_{\mathcal{BMO}} = \infty$.

Theorem 2.2.12 ([31, Lemma A.8.7 and Lemma A.8.5])**.** *Each $\mathcal{BMO}$-martingale is locally bounded, and each bounded martingale is an element of $\mathcal{BMO}$.*

Theorem 2.2.13 ([31, Theorem A.8.14])**.** *Any continuous linear function $\varphi : \mathcal{H}^1 \to \mathbb{R}$ can be written*

$$\varphi_N(M) := \mathbf{E}\left(M_\infty N_\infty\right)$$

for some $N \in \mathcal{BMO}$.

This duality theorem shows that $\mathcal{BMO}$ is the dual space of $\mathcal{H}^1$, in analogy with classical analysis where L^∞ embeds continuously into the dual of L^1.

Lemma 2.2.14 ([31, Lemma 10.1.5])**.** *Identifying indistinguishable martingales, $\mathcal{H}^1$ is a Banach space under the $\mathcal{H}^1$ norm.*

The completeness of $\mathcal{H}^1$ is essential for many approximation arguments and for establishing the duality between $\mathcal{H}^1$ and $\mathcal{BMO}$.

The spaces $\mathcal{H}^1$ and $\mathcal{BMO}$ will be used in Section 6.2 in the proof of the Jacod–Yor representation theorem for sigma martingales.

2.3 The Predictable σ-Algebra and Related Concepts

The predictable σ-algebra provides the natural class of integrands for stochastic integration with respect to semimartingales.

Definition 2.3.1. The *predictable σ-algebra* $\mathcal{P}$ on $\Omega \times [0, \infty)$ is the smallest σ-algebra making all left-continuous, adapted processes measurable.

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 $\mathcal{P}$ and the Borel σ-algebra on $\mathbb{R}$.

Theorem 2.3.2 ([142, Section 2])**.** *The predictable σ-algebra $\mathcal{P}$ can be characterized equivalently as:*

(a) The smallest σ-algebra such that all left-continuous, adapted processes are measurable;

(b) The σ-algebra generated by all continuous, adapted processes.

This characterization shows that predictable processes are those that can be approximated by continuous adapted processes, highlighting their non-anticipative nature.

A càdlàg process A is a *finite variation process* (FV-process) if almost all of its paths have finite variation on each compact interval. These processes can be decomposed into their continuous and purely discontinuous parts, providing a fundamental building block for semimartingales.

The class of semimartingales, which includes both local martingales and finite variation processes, forms the most general class of integrators for stochastic integration. The precise definition and properties of semimartingales are developed in detail in subsequent chapters of this thesis.

Brownian motion serves as the canonical example of a continuous martingale and plays a fundamental role in stochastic analysis:

Definition 2.3.3. An adapted process $B = (B_t)_{0 \leq t < \infty}$ taking values in $\mathbb{R}^n$ is called an n-dimensional *Brownian motion* if:

(a) For $0 \leq s < t < \infty$, $B_t - B_s$ is independent of $\mathcal{F}_s$;

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

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

Theorem 2.3.4 ([166, Theorem I.26])**.** *Let B be a Brownian motion. Then there exists a modification of B which has continuous paths almost surely.*

Throughout this thesis, we assume that all Brownian motions have continuous paths and, unless stated otherwise, the covariance matrix C is the identity matrix (standard Brownian motion).

These Brownian-motion preliminaries will be used in Section 7.3, where dividend-paying stock prices are modeled by stochastic differential equations driven by Brownian motion.

Given two probability measures $\mathbf{P}$ and $\mathbf{Q}$ on $(\Omega, \mathcal{F})$, we say that $\mathbf{Q}$ is *absolutely continuous* with respect to $\mathbf{P}$, denoted $\mathbf{Q} \ll \mathbf{P}$, if $\mathbf{P}(A) = 0$ implies $\mathbf{Q}(A) = 0$ for all $A \in \mathcal{F}$. If both $\mathbf{Q} \ll \mathbf{P}$ and $\mathbf{P} \ll \mathbf{Q}$, we say that $\mathbf{P}$ and $\mathbf{Q}$ are *equivalent*, denoted $\mathbf{P} \sim \mathbf{Q}$.

The *Radon-Nikodym derivative* $\frac{\mathrm{d}\mathbf{Q}}{\mathrm{d}\mathbf{P}}$ is the unique (up to $\mathbf{P}$-null sets) $\mathcal{F}$-measurable function such that

$$\mathbf{Q}(A) = \int_A \frac{\mathrm{d}\mathbf{Q}}{\mathrm{d}\mathbf{P}} \, \mathrm{d}\mathbf{P}$$

for all $A \in \mathcal{F}$. When restricted to a sub-σ-algebra $\mathcal{G} \subset \mathcal{F}$, the conditional expectation $\mathbf{E}_{\mathbf{P}}\left[\frac{\mathrm{d}\mathbf{Q}}{\mathrm{d}\mathbf{P}} \mid \mathcal{G}\right]$ serves as the Radon-Nikodym derivative on $\mathcal{G}$.

2.4 Quasimartingales

In preparation for the Bichteler-Dellacherie theorem, we introduce the notion of quasimartingales, which provide a natural generalization of martingales and play a crucial role in characterizing semimartingales.

Definition 2.4.1. Let X be a d-dimensional adapted process such that $X_t \in L^1$ for all $t \in [0,1]$. For a partition $\pi = \{0 = t_0 < t_1 < \cdots < t_n = 1\}$ of $[0,1]$, the *mean variation* of X along π is defined as

$$\mathrm{MV}(X, \pi) = \mathbf{E}\left[\sum_{i=0}^{n-1} \left\| \mathbf{E}[X_{t_{i+1}} - X_{t_i} \mid \mathcal{F}_{t_i}] \right\|_{\mathbb{R}^d}\right].$$

The process X is called a *quasimartingale* if

$$\mathrm{MV}(X) := \sup_{\pi} \mathrm{MV}(X, \pi) < \infty,$$

where the supremum is taken over all finite partitions π of $[0,1]$.

Remark 2.4.2. The mean variation is monotone with respect to refinement of partitions: if π' refines π, then $\mathrm{MV}(X, \pi) \leq \mathrm{MV}(X, \pi')$. This follows from the conditional Jensen inequality.

The importance of quasimartingales lies in their characterization as differences of submartingales, providing the stochastic analogue of functions of bounded variation being differences of increasing functions.

Theorem 2.4.3 (Rao's Theorem, [166, Chapter III, Theorem 17]). *A càdlàg process X is a quasimartingale if and only if it has a decomposition $X = Y - Z$ where Y and Z are càdlàg submartingales.*

This theorem, originally due to Rao [167], provides a fundamental decomposition result that will be crucial in our proof of the Bichteler-Dellacherie theorem.

2.5 Measurability, Inequalities, and Supporting Results

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 [53] and Chung and Doob [28], 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 Courrège [34, 33] and Meyer [139, 140, 143, 144, 145, 146].

We start with the basic definitions.

Definition 2.5.1. 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 [44, p 120 ff].

Verifying progressive measurability is 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 ([110, 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 ([110, 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 2.5.2 (Chung-Doob-Meyer [28, 141]). *Let X be a real-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, [28] proved the existence of a progressively measurable modification under the assumption of separability. Meyer [141] 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 [44], using advanced tools from functional analysis. The theorem plays a pivotal role in the creation of the stochastic integral as presented in [110].

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 thesis, 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.

Several fundamental inequalities play crucial roles in the development of stochastic integration theory:

Theorem 2.5.3 (Burkholder's Inequality, [157, 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. Define the discrete stochastic integral*

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

Then for any $T \in \mathbb{N}$ and $\alpha > 0$,

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

If M is positive, then 18 can be replaced by 9.

This inequality, originally due to Burkholder [19], provides crucial estimates for martingale transforms and extends to continuous-time settings. It is fundamental in proving the continuity of the stochastic integral operator and establishing the Bichteler-Dellacherie theorem.

Theorem 2.5.4 (Generalized Fatou's Lemma, [193, Theorem 9.19])**.** *Let $(f_n)_{n\in\mathbb{N}}$ be a sequence of measurable functions on a measure space $(\Omega, \mathcal{F}, \mu)$. Suppose there exists an integrable function $g \in L^1(\mu)$ such that $f_n \geq g$ for all n. Then*

$$\int \liminf_{n\to\infty} f_n \, d\mu \leq \liminf_{n\to\infty} \int f_n \, d\mu.$$

This generalized version of Fatou's lemma is essential in our development, particularly for proving the generalized Ansel–Stricker Lemma in Chapter 6. It allows us to handle sequences of random variables that are bounded below by an integrable function rather than by zero, which is crucial when dealing with sigma martingales that may take negative values.

Theorem 2.5.5 ([31], Lemma 1.3.38)**.** *Suppose $\{f_n\}_{n\in\mathbb{N}}$ is a sequence of measurable functions that converges in measure. Then there exists a subsequence $\{f_{n_k}\}_{k\in\mathbb{N}}$ which converges almost everywhere.*

This result is crucial for passing from convergence in probability to almost sure convergence, which is often needed in the construction of stochastic integrals and in proving limit theorems for stochastic processes.

Theorem 2.5.6 (Mazur's Lemma, [175, Theorem 3.13])**.** *Let X be a Banach space and let $(x_n)_{n\in\mathbb{N}}$ be a sequence in X that converges weakly to some $x \in X$. Then there exists a sequence of convex combinations of the x_n that converges to x in norm. More precisely, for each $n \in \mathbb{N}$, there exist $N_n \geq n$ and non-negative numbers $\lambda_{n,k}$ for $k = n, \ldots, N_n$ with $\sum_{k=n}^{N_n} \lambda_{n,k} = 1$ such that*

$$y_n := \sum_{k=n}^{N_n} \lambda_{n,k} x_k \to x \quad \text{in norm.}$$

For more details, refer to [175] or [196].

Remark 2.5.7. In the context of $L^2(\mathbf{P})$ spaces, this theorem is particularly useful as it allows us to pass from weak convergence to strong convergence through convex combinations. This is utilized in the proof of the Accumulation Lemma (Lemma 4.2.8).

Theorem 2.5.8 ([125, Theorem 4.2.3.])**.** *Let $(X, \|\cdot\|)$ be a normed space with dual space X', $M \subset X$ a subspace, and $y_0 \in X \setminus M$ with*

$$d := \inf_{x \in M} \|y_0 - x\| > 0.$$

Then there exists an $f \in X'$ satisfying:

(1) $f(x) = 0$ *for all* $x \in M$,

(2) $f(y_0) = d$,

(3) $\|f\|_{X'} = 1$.

Chapter 3

The General Stochastic Vector Integral – Historical Perspectives

This chapter reviews the development of stochastic integration from its early origins in the 1920s and 1930s through the foundational work of Itô in the 1940s, and into the general theory of semimartingales and vector-valued stochastic integrals. We describe the main constructions, examine their relations, and explain the differences between the two leading approaches. We also summarize the extensions that have been introduced for specific purposes or within certain mathematical frameworks. While we follow the historical progression of the theory, we adapt the terminology to align with modern conventions. This avoids confusion arising from inconsistent or outdated vocabulary in the original sources. In addition, we make several minor adjustments to definitions and formulations to match the contemporary treatment. For instance, although Itô's original construction used optional integrands, we consistently work with predictable integrands throughout, as this reflects the current standard and ensures compatibility with general semimartingale theory.

This chapter is organized as follows. In Section 3.1, we examine the classical approach, following its historical development from the Brownian motion case to the general semimartingale setting. Section 3.2 discusses the functional analytic approach, which builds on continuity properties of the integral operator and characterizations of semimartingales. In Section 3.3, we present the general construction of the vector stochastic integral. Section 3.4 gives an overview of how the subject is treated in major textbooks.

While the focus lies on the historical development of the theory, we state the main constructions in a form compatible with the notation and conventions of the rest of the thesis. The required preliminaries are collected in Chapter 2. In addition, we provide a short section on the measurability of stochastic processes, as this aspect has often been overlooked in the literature and has led to incomplete or imprecise results.

3.1 The Classical Approach

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

For further details on the historical progression of stochastic processes and integration theory, see [100, 123, 150, 188].

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.

3.1.1 The Beginning: Wiener and Doeblin

Norbert Wiener (1894 – 1964), an American mathematician, is generally regarded as one of 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" [160, 189, 190]. 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), known in France as Vincent Doblin, 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 [17].

3.1.2 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 [84, 85, 86, 87, 88, 89], which are extensively summarized in [132]. 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 chapter, 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 [51] because it has proven to be irrelevant for most practical applications.

In addition, we have standardized the notation and terminology to align with modern conventions and recent approaches, which helps to avoid inconsistent definitions and notations. We have 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 2.5).

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 ($\mathbf{P}$-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, [166, 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 3.1.1. An *elementary predictable process* $H : \Omega \times [0,\infty) \to \mathbb{R}$ 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 $0 = t_1 \leq t_2 \leq \cdots \leq t_{n+1} < \infty$. The space of elementary predictable processes is denoted by $\mathcal{E}$.

Remark 3.1.2. 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 2.3.2 or, for more elaborations, [170, 29]). 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 3.1.3. 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

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

and

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

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 3.1.4. 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 3.1.5. 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

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

σ-algebra, the use of right-continuous step functions would lead to optional processes. Even though there has also been a stochastic integration theory developed for optional processes, it is generally considered less important. Itô also adjusted his approach in later publications and switched to left-continuous elementary predictable processes [90].

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 reappears in many later extensions of the stochastic integral and is, therefore, extremely crucial.

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

Definition 3.1.6. (a) For any jointly measurable and adapted process H, we define the seminorm $\|\cdot\|_{\text{Itô}}$ by

$$\|H\|^2_{\text{Itô}} := \mathbf{E}\left[\int_0^\infty H^2(s,\omega)\mathrm{d}s\right]. \tag{3.1}$$

(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 3.1.7 ([84], Theorem 2.1)**.** *The mapping*

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

is an isometry. Specifically,

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

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 3.1.8. 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 [132, 29, 100, 103].

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

Theorem 3.1.9 ([90], Theorem 5.2.1)**.** *Let $\mathscr{L}^2$ be the space of jointly measurable, adapted processes with*

$$\mathbf{E}\left[\int_0^\infty H^2(s,\omega)\,\mathrm{d}s\right] < \infty. \tag{3.2}$$

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 3.1.9, which is often combined with the proof of the Itô isometry Theorem 3.1.10 (see below), formal measurability arguments are important but are often neglected (see Section 2.5). 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 3.1.9, 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 [132] but also in the more recent [159], see also the discussions in [29, 100].

One proposed solution to ensure the measurability is to assume that the process X is separable [66], 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 [103]. With Theorem 2.5.2, one obtains the general result.

By Theorem 3.1.9, the Itô Isometry can be naturally extended to $\overline{\mathcal{E}} = \mathscr{L}^2$. For completeness, we mention the result again.

Theorem 3.1.10 (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 [114, p. 642 ff] or Kuo [122, 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.

3.1.3 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 is a brief outline of these developments.

In his book [51] 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

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

   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 Eq. (3.2) to

$$\int_0^\infty \mathbf{E}\left[{H_t}^2\right] \mathrm{d}F(t) < \infty, \tag{3.3}$$

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

Doob extended Itô's isometry Theorem 3.1.10 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 [51].

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, especially through his papers on Markov processes [80, 81, 82]. 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 [139, 140]. 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 [139]. He later even proved the uniqueness of this decomposition in [140].
- 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 [91, 154, 121].

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 [101].

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

Definition 3.1.11. A càdlàg supermartingale Z with $Z_0 = 0$ is of *Class (D)* if the collection $\{Z_T\,;\, T \text{ a finite valued stopping time}\}$ is uniformly integrable.

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

Theorem 3.1.12 (Doob-Meyer Decomposition [166], 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 [46, 47], who demonstrated that an increasing process is natural if it is predictable. Alternative proofs of Theorem 3.1.12 were presented by Rao [168] and Bass [6] and Jakubowski [97].

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 Komlós' Lemma [116].

A concise and modern proof can be found in the recent paper by [8], which combines ideas from Jakubowski [97] and Beiglböck, Schachermayer, and Veliyev [7] 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 [34, 33]. He focused 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,

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

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.

3.1.4 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 [91], Motoo and Watanabe [154], and Kunita and Watanabe [121], who extended the class of possible integrators, and

Meyer [143, 144, 145, 146], who explored the potential class of integrands and relaxed some assumptions on the filtration.

The first three works ([91, 154, 121]) not only marked a significant milestone in the evolution of stochastic integration but also contributed notably to the theory of martingales and stochastic processes: Itô'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 [91], 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 3.1.13. 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).

Theorem 3.1.14 ([143], 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*

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

Remark 3.1.15. 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 3.1.16. From the definition, it is easy to see that for $X, Y \in \mathcal{M}_0^2$, we have

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

The existence of the sharp-bracket process made it possible 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 Motoo and Watanabe [154] and Kunita and Watanabe [121] and later Meyer [143, 144, 145, 146] 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 3.1.25). 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 $L \in \mathcal{M}_0^2$ that satisfies

$$\mathbf{E}\left[L_\infty N_\infty\right] = \mathbf{E}\left[\int_0^\infty H_s \, d\langle M, N\rangle_s\right] \tag{3.4}$$

for every square-integrable martingale $N \in \mathcal{M}_0^2$ (see Theorem 3.1.25).

This approach was later extended to more general stochastic integrals [49] and even

the vector-valued stochastic integral [31]. 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 3.1.17. 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.

Throughout this subsection we assume $M_0 = 0$. For a general square-integrable martingale, one applies the construction to $M - M_0$.

Definition 3.1.18. Let $M \in \mathcal{M}_0^2$ 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

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

and

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

Since $M_0 = 0$, this integral starts at zero and belongs to $\mathcal{M}_0^2$.

Similar to the other approaches, the definition of the stochastic integral for square-integrable martingales is straightforward. 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 3.1.19. Let $M \in \mathcal{M}_0^2$.

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

$$\|H\|_M^2 := \mathbf{E}\left[\int_0^\infty H_s^2 \mathrm{d} \langle M, M \rangle_s \right]. \tag{3.5}$$

(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 3.1.20 ([31], Lemma 12.1.4)**.** *By endowing $\mathcal{M}_0^2$, the space of square-integrable martingales starting at zero, with the norm $\|M\|_{\mathcal{M}_0^2} := (\mathbf{E}\,[M_\infty^2])^{1/2} = (\mathbf{E}\,[\langle M, M\rangle_\infty])^{1/2}$, the mapping*

$$J_\infty^M : \quad \begin{cases} \mathcal{E}_M & \to & \mathcal{M}_0^2 \\ H & \mapsto & H \bullet M \end{cases}$$

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 $\mathcal{M}_0^2$.

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,

$$\|H \bullet M\|_{\mathcal{M}_0^2}^2 = \mathbf{E}\,[\langle H \bullet M, H \bullet M\rangle_\infty] = \mathbf{E}\left[\int_0^\infty H_s^2 \mathrm{d}\,\langle M, M\rangle_s\right],$$

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

The seminorm $\|\cdot\|_M$ is the L^2-seminorm associated with the measure on the predictable σ-field induced by the increasing process $\langle M, M\rangle$. After quotienting by its null space, the resulting Hausdorff space is a Hilbert space, which we continue to denote by $L^2(M)$ for brevity. 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, $\mathcal{E}_M$ is dense in $L^2(M)$:

Lemma 3.1.21 ([31], 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 3.1.22. The continuous linear mapping $\overline{J_\infty^M} : L^2(M) \to \mathcal{M}_0^2$, obtained by extending $J_\infty^M : \mathcal{E}_M \to \mathcal{M}_0^2$ 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 3.1.23 ([31], Lemma 12.1.4)**.** *After quotienting $L^2(M)$ by the null space of $\|\cdot\|_M$, the map*

$$J^M : L^2(M) \to \mathcal{M}_0^2, \qquad H \mapsto H \bullet M,$$

is an isometry. Specifically, for any $H \in L^2(M)$, it holds that

$$\|H\|_M^2 = \mathbf{E}\left[\int_0^\infty H_s^2 \,\mathrm{d}\,\langle M, M\rangle_s\right] = \mathbf{E}\left[\left(J_\infty^M(H)\right)^2\right] = \left\|J_\infty^M(H)\right\|_{\mathcal{M}_0^2}^2.$$

Motoo and Watanabe [154] and Kunita and Watanabe [121] 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-Watanabe inequality, which is due to Courrège and Meyer [151] but based on a result by Kunita and Watanabe [121].

Theorem 3.1.24 (Kunita-Watanabe Inequality, [151] Théorème 21)**.** *If M and N are two square-integrable martingales and H, K are predictable processes for which the right-hand side below is finite, then we have almost surely:*

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

The Kunita-Watanabe inequalities provide crucial insights into the most important 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 3.1.25 (Kunita and Watanabe, [173] Theorem 28.1)**.** *Let $M \in \mathcal{M}_0^2$ and $H \in L^2(M)$. Then the stochastic integral $H \bullet M$ is the unique element of $\mathcal{M}_0^2$ with the property that, for every $N \in \mathcal{M}_0^2$,*

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

With the help of Theorem 3.1.24, it is straightforward to show that the novel approach by Motoo and Watanabe [154] and Kunita and Watanabe [121] is equivalent to the one outlined above. Fix $N \in \mathcal{M}_0^2$. The mapping

$$H \mapsto \mathbf{E}\Big[(H \bullet M)_\infty \, N_\infty\Big] - \mathbf{E}\left[\int_0^\infty H_s \, \mathrm{d}\langle M, N\rangle_s\right]$$

defines a linear functional on $L^2(M)$. The Kunita-Watanabe inequality implies that this functional is continuous, and elementary computation shows that it vanishes on simple processes. By continuity of the extension, it vanishes on the whole space $L^2(M)$, and the identity

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

therefore holds for every $H \in L^2(M)$.

Furthermore, the right-hand side of the equation in Theorem 3.1.25 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 $\mathcal{M}_0^2$, equivalently in $L^2(\Omega, \mathcal{F}, \mathbf{P})$ through terminal values. By the Riesz-Fréchet representation theorem, for given H and M, there exists a unique element $H \bullet M \in \mathcal{M}_0^2$ that corresponds to this linear functional.

3.1.5 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 [91]. With this, it was possible to define the square-bracket process for local L^2-martingales (see [144, 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 3.1.14 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 3.1.26 ([79], Lemma 7.28)**.** *Let M be a locally square-integrable martingale with $M_0 = 0$. 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. This process is again called the* sharp-bracket process *or* predictable quadratic variation.

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

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

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 3.1.27 ([45], 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 [143, 144, 145, 146]. These works extended the stochastic integral to predictable integrands and locally square-integrable martingales.

In the original publication by Itô [91], the definition of the stochastic integral via Eq. (3.4) 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 [144] 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 34. 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 [44] and later by Protter [166] 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 $H \in L^2(M^{T^n})$ for all n.

Theorem 3.1.28 (Definition of the Stochastic Integral, [144] p. 98)**.** *Let $M \in \mathcal{M}^2_{\text{loc}}$ and $H \in L^2_{\text{loc}}(M)$. Then there exists a unique process L such that*

$$L^{T^n} = H \bullet M^{T^n}$$

for every localizing sequence (T^n) with $M^{T^n} \in \mathcal{M}^2$ and $H \in L^2(M^{T^n})$, where the preceding construction is applied to $M^{T^n} - M_0$ when necessary. 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 [154, 121] (see Theorems 3.1.25 and 3.1.29). This

method allows the direct definition of the stochastic integral for localized processes, which simplifies the application of Hilbert space methods.

Theorem 3.1.29 ([121], Theorem 2.1)**.** *For $M \in \mathcal{M}^2_{\mathrm{loc}}$ and $H \in L^2(M)_{\mathrm{loc}}$, the integral $H \bullet M$ is uniquely characterized by the relation*

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

Besides the definition of the stochastic integral for locally square-integrable martingales, one of the key results of Meyer's work [143, 144, 145, 146] 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 [16] and the textbooks by Dellacherie and Meyer [44] and He, Wang, and Yan [79] 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 3.1.30 ([144], 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

$$M_t = M_t^c + M_t^d.$$

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 3.1.26 that $\langle M^c, M^c \rangle$ exists for any local martingale M.

Definition 3.1.31. 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

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

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

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

Theorem 3.1.32 ([79], 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 [121, 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 3.1.33 ([110], 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 $\|P\| = \max_{0 \le k < n}(t_{k+1} - t_k)$ *is the mesh of the partition. The limit is taken as the mesh of the partition tends to zero, with convergence in probability for each fixed* t*.*

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 [166].

Theorem 3.1.34 ([144], 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 3.1.35. 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 3.1.36. 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 Karatzas and Shreve [110].

Remark 3.1.37. Even though the predictable quadratic variation as well as the optional quadratic variation generally coincide only for continuous square-integrable martingales, the seminorms they naturally induce are equivalent (see [31, 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.

3.1.6 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 [49], 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 3.1.38. 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 [49] 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 3.1.29 but used the quadratic variation as was done by Meyer [144].

The key idea of the approach by Doléans-Dade and Meyer [49] 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 [49] is the following:

Proposition 3.1.39 ([49], 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 $1 \leq p < \infty$.*
- 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 3.1.40 (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

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

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 finite-variation local martingale agrees with the pathwise Lebesgue-Stieltjes integral. In the continuous case such a local martingale is constant, but in the presence of jumps the compatibility of the two notions is nontrivial and depends on predictability of the integrand; 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 [49] also demonstrated the following crucial result, which allows for the unique characterization of the stochastic integral with local martingale integrators.

Proposition 3.1.41 ([49], Proposition 5)**.** *Let M be a local martingale and H a locally bounded, predictable process. Then $H \bullet M$ is the unique local martingale Y with $Y_0 = 0$ such that*

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

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.

3.1.7 The General Real-Valued Stochastic Integral

The stochastic integral was further generalized to accommodate unbounded predictable integrands by Jacod [92, 93] and Chou [27], 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 [27] 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 the publication by Meyer [147]) and Doléans-Dade [48] independently.

Theorem 3.1.42 (*Fundamental Theorem of Local Martingales* [147])**.** *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 [92, 93].

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

$$\int_0^t |H_s| \mathrm{d}|X|_s < \infty \text{ for } t < \infty, \tag{3.6}$$

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

$$\text{the increasing process } (H^2 \bullet [X, X])^{1/2} \text{ is locally integrable.} \tag{3.7}$$

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 3.1.43. 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 3.1.44. 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 [31, Theorem 11.6.10].

Special semimartingales were initially introduced by Meyer [151] and Yoeurp [194]. 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 3.1.45 ([92], Théorème 6)**.** *Let X be a classical semimartingale and $H \in \mathcal{P}$. Then*

$$Y_t = \sum_{0 \le s \le t} \Delta X_s \mathbb{1}_{\{|\Delta X_s| \ge 1\} \cup \{|H_s \Delta X_s| \ge 1\}} \tag{3.8}$$

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

Now, the stochastic integral can be defined as follows:

Definition 3.1.46. Let X be a classical semimartingale and $X = Y + Z$ be a decomposition as given by Theorem 3.1.45, 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

$$H \bullet X := H \bullet Y + H \bullet Z = H \bullet Y + H \bullet M + H \bullet A. \tag{3.9}$$

The stochastic integral is well-defined.

Remark 3.1.47. 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 Eq. (3.9), we take $J = \{s : |\Delta X_s| \vee |H_s||\Delta X_s| \vee |K_s||\Delta X_s| > 1\}$; Theorem 3.1.45 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 3.1.48. 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 3.1.45 by taking $J = \{s\,;\, |\Delta X_s| \vee |H_s||\Delta X_s| > 1\}$ in Eq. (3.9), reducing it to a dominated convergence theorem for usual stochastic integrals.

Jacod laid the foundation for general stochastic integration [92, 93]. 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 [94], and Mémin examined the connections between these integrals and the topology of classical semimartingales [135]. We briefly outline these approaches here.

The ansatz by Chou, Meyer, and Stricker is based on the work of Jacod [94] and Mémin [135], as well as Meyer [147]. 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, Meyer, and Stricker [27] 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

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

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 3.2.1). 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 3.1.49 (Sophisticated Definition of Stochastic Integral)**.** Let X be a classical semimartingale, H be a predictable process and $H^n = H\mathbb{1}_{\{|H|\leq n\}}$. 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 [49] and presented in Section 3.1.6 is crucial.

Remark 3.1.50. 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 3.1.51. 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 3.1.52. The sophisticated definition of the stochastic integral is equivalent to the definition as given in Definition 3.1.46, and $H \in L(X)$ is equivalent to H being X-integrable. For details, refer to [27], page 130 d).

Remark 3.1.53. 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, [31]).

When Doléans-Dade and Meyer [49] 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 raises 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 [27] via the following theorem:

Theorem 3.1.54 ([27], 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 [79, 31] 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 3.1.30), 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 3.1.55. 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, Davis, and Gundy [20] and Garsia [67]), one can show the equivalence of the norms $M \mapsto \mathbf{E}\left[[M,M]_\infty^{\frac{1}{2}}\right]$ and $M \mapsto \mathbf{E}\left[|M_\infty^*|\right]$ 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 3.1.55.

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

Definition 3.1.56. 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:

- $H \in L^2_{\text{loc}}(M)$,
- 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 3.1.55, and $H \bullet A$ as a Lebesgue-Stieltjes integral.

The original publication by Jacod [92, 93] introduced the condition that, similar to the definition in Definition 3.1.55, 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 fulfill 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.

3.2 The Functional Analytic Approach

The approach to the general stochastic integral for classical semimartingales described in Section 3.1 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 [151].

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, [165]).

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 one to define stochastic integrals with respect to left continuous, adapted processes very quickly and to prove 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 Meyer [151], Bichteler

[11, 12], and Dellacherie [42]. 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 [149] and Dellacherie [42] were the first to propose and describe the functional analytic technique as it is understood today. Then, Lenglart [127] and Protter [164, 165] built upon it. Protter [166] 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 [128] came up with a concept similar to the functional analytic approach discussed here.

3.2.1 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. 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 3.2.1. A *simple predictable process* $H : \Omega \times [0, \infty)$ is defined as

$$H(t, \omega) = H_0(\omega)\mathbb{1}_{\{0\}}(t) + \sum_{i=1}^{n} H_i(\omega)\mathbb{1}_{]\!]T^i, T^{i+1}]\!]}(t). \tag{3.10}$$

with stopping times $0 \leq T^1 \leq \ldots \leq T^n \leq T^{n+1} < \infty$ 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:

$$L^q(\Sigma_p, \mu) = \left\{ f \,;\, \int_{\Omega \times \mathbb{R}_+} |f|^q \, \mathrm{d}\mu < \infty \right\}.$$

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

For a stochastic process X and a simple predictable process H with presentation as in Eq. (3.10), let

$$J_\infty^X(H) = H_0 X_0 + \sum_{i=1}^{n} H^i \left(X_{T^{i+1}} - X_{T^i} \right) \tag{3.11}$$

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 a suitable continuity property. Eq. (3.11) is clearly linear. In order to define or check continuity, one needs to define suitable 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 3.2.2. A càdlàg, adapted process X is called a *good integrator* if, for each $t \geq 0$, the map

$$I_t^X : \Lambda_{\mathrm{u}} \to L^0, \qquad H \mapsto J_t^X(H),$$

is continuous.

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, [109, 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 3.2.3 (Bichteler-Dellacherie-Mokobodzki Theorem, [12] 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 Meyer [151], Bichteler [11, 12], and Dellacherie [42]. The original proofs were later simplified by Beiglböck and Siorpaes [9], which can also be found in the appendix of [31]. 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 3.2.4. Let X be a good integrator and $H \in \Lambda$ represented as

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

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

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

and

$$J_\infty^X(H) = H_0 X_0 + \sum_{i=1}^{n} H_i \left(X_{T^{i+1}} - X_{T^i}\right)$$

Then $\Lambda \subset \mathbb{L}$, and for each $H \in \Lambda$ the stochastic integral process $J^X(H)$ belongs to $\mathbb{D}$.

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 3.2.5. (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 $\mathbb{L}_{\text{ucp}}$, $\mathbb{D}_{\text{ucp}}$ accordingly.

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

Lemma 3.2.6 ([166], Theorem 10)**.** *The space* Λ_{ucp} *is dense in* $\mathbb{L}_{\text{ucp}}$.

Lemma 3.2.7 ([109], Theorem 2.71)**.** *The space* $\mathbb{D}_{\text{ucp}}$ *is complete.*

The continuity extends from I_X to J^X.

Lemma 3.2.8 ([166], 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 3.2.9. The continuous linear mapping $\overline{J^X} : \mathbb{L}_{\text{ucp}} \to \mathbb{D}_{\text{ucp}}$, obtained by extending $J^X : \Lambda_{\text{ucp}} \to \mathbb{D}_{\text{ucp}}$ 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 $H \bullet X$, 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 3.1.31) or through approximating sums (Theorem 3.1.33). 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 3.2.10. 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 Lévy's stochastic area formula. Nevertheless, for more advanced applications, a more general stochastic integral is required.

3.2.2 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 [166] and was first published in the initial edition of the cited textbook. However, similar techniques were already discussed in [164, 165], Jacod [93], and Chou, Meyer, and Stricker [27].

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 integrands, 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 subspace 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 3.1.43). 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 [151] and Yoeurp [194]).

We follow here closely the presentation in [166].

Definition 3.2.11. Let X be a special semimartingale with canonical decomposition $X = N + A$. The $\mathcal{H}^2$-*norm* of X is defined as

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

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

Remark 3.2.12. 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,

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

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

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

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

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

Theorem 3.2.14 ([166], 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 $\mathcal{N} = \{H : d_X(H,0) = 0\}$ is the null space of the seminorm $H \mapsto d_X(H,0)$, is a normed 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

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

Lemma 3.2.15 ([166], 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 [49] defined the stochastic integral first for bounded predictable integrands as the monotone class theorem could be utilized for that (also see [151] for a nice presentation).

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

$$\|H \bullet X - J \bullet X\|_{\mathcal{H}^2} = d_X(H, J). \tag{3.12}$$

Theorem 3.2.16 ([166], Theorem IV.3). *If $(H^n)_{n\in\mathbb{N}}$ is a Cauchy sequence in $\mathbf{b}\mathbb{L}$ 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 3.2.17. Let $X \in \mathcal{H}^2, H \in \mathbf{b}\mathcal{P}$ and $(H^n)_{n\in\mathbb{N}}$ be a sequence in $\mathbf{b}\mathbb{L}$ 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 3.2.18 ([166], 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, a suitable localization procedure (similar to how it was done in Section 3.1.5) 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. [95, 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 behavior is the following theorem.

Theorem 3.2.19 ([166], 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 3.2.20. 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

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

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.

3.3 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 further generalized. Jacod [94] presented this generalization when he developed the stochastic vector integral with regard to a multidimensional semimartingale. Jean Mémin showed in [135] that this construction yields a closed, hence complete, space of stochastic integrals. This concept of the stochastic vector integral is also briefly treated in C. Dellacherie and Paul-André Meyer's book [44].

In [94], 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 [177] and then shown in detail in [178]. In [108] and [109], 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 [31].

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 $H \in L^p_{\mathrm{loc}}(M)$ if H is locally in $L^p(M)$.

It is straightforward to demonstrate that $\|H\|_{L^p(M)}$ is a seminorm on predictable integrands, and a norm after quotienting by its null space, since it is composed of a Hilbert space seminorm and an L^p-seminorm. A dominated 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 $\left(\left(H^\top \theta H\right) \bullet C\right)^{p/2}$ 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 $L^2_{\mathrm{loc}}(M)$ perfectly corresponds to the space of predictable processes H, for which $\left(H^\top \theta H\right) \bullet C$ is locally integrable.

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

Definition 3.3.1. For a local martingale M with $M_0 = 0$ and a predictable process $H \in L^1_{\mathrm{loc}}(M)$, 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 $[M^i, N] = K^i \bullet C$.

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 3.3.2. Let $A = (A^1, \dots, A^d)$ be an $\mathbb{R}^d$-valued càdlàg process with components of finite variation. Then there exists a càdlàg increasing process V and an optional, or suitably measurable, $\mathbb{R}^d$-valued process $v = (v^1, \dots, v^d)$ such that each component A^i admits the representation

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

If A is predictable, then V and v may be chosen 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 3.3.3. 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. The corresponding density v is then automatically predictable when A is predictable.

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 X is both a vector local martingale and a vector finite-variation process, and if both constructions apply, then the resulting integrals agree.

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

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

Definition 3.3.4. A vector process H is integrable with respect 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 chosen decomposition, and the resulting space of X-integrable processes is denoted by $L(X)$.

Remark 3.3.5. 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 $H \in L^{\text{var}}(A) \cap L^1_{\text{loc}}(M)$. Thus, $L(X)$ includes all locally bounded predictable processes.

Remark 3.3.6. If $H \in L(X)$, then H may not belong to $L_{\text{var}}(A) \cap L^1_{\text{loc}}(M)$ for all decompositions $X = A + M$ with A a d-dimensional process of finite variation and M a d-dimensional local martingale, see [178, Example 5.2].

The following holds with this formulation of the stochastic integral:

Theorem 3.3.7 ([135], Théorème II.7)**.** *In the semimartingale topology, the space*

$$\{H \bullet X\}_{H \in L(X)}$$

is complete for every semimartingale X.

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

Kunita [119], Métivier and Pistone [138], Meyer [148], Métivier and Pellaumail [137], Grigelionis and Mikulevičius [71], and Korezlioglu and Martaias [117], 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 Mikulevičius and Rozovskii [153] and Mikulevičius and Rozovskii [152] is one method for approaching this topic. In Donno and Pratelli [50] and De Donno, Guasoni, and Pratelli [37], a particular application of this idea was refined and applied to finance. Alternatively, the theory of Random Measures, which is presented in detail in [13], can be extended to cover infinite-dimensional

stochastic integration. If X takes values in a Hilbert space, the construction described in [136] proceeds in a form closely analogous to the finite-dimensional case, although additional topological and measurability issues have to be handled.

3.4 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 and Doléans-Dade [5], Bingham and Kiesel [14], Choe [25], Ebenfeld [54], Elliott and Kopp [55], Klenke [114], Musiela and Rutkowski [156], and Øksendal [159]. This stochastic integral provides a foundation for most applications in biology, physics, and finance. Notably, Øksendal [159] and Musiela and Rutkowski [156] 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. Cont and Tankov [32] and Tankov [185] and Applebaum [4] 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 modeling jumps in financial markets.

Several works extend the stochastic integral to continuous local martingales. Chung and Williams [29] present the stochastic integral for continuous local martingales using integration with respect to random measures, and furthermore extend the set of potential integrands beyond predictable ones. Additionally, Meintrup and Schäffler [134], Émery [60], Irle [83], Karatzas and Shreve [110], Kallianpur [105], Revuz and Yor [170], Kuo [122], and Durrett [52], and Le Gall [126] 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 [60] provides an appendix discussing the general semimartingale integral, the proofs are omitted. Similarly, Dellacherie and Meyer [45] offer a rigorous treatment of semimartingale theory and stochastic integration but focus more on the probabilistic potential side.

In some texts, like Kallenberg [104] and Gihman and Skorohod [69], 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 [192], Klebaner [113], and Irle [83] follow the classical approach,

starting with the Itô integral for Brownian motions and extend it to local martingales and semimartingales. However, they often only sketch or omit proofs.

Some works, such as Shiryaev [177], 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 [173] define the general stochastic integral by starting with Itô processes and extending to continuous local martingales, martingales, and finally semimartingales. Unlike earlier works, Rogers and Williams use 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 [31] provide 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, Wang, and Yan [79] 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 [107] follow a similar approach. A detailed treatment of general semimartingale theory is covered in Jacod and Shiryaev [95], which extensively discusses limit theorems and the full semimartingale framework.

An alternative approach to the classical treatment is provided by Medvegyev [133], 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 [12] develops a comprehensive L^p theory for semimartingales, covering similar topics in stochastic integration.

Infinite-dimensional stochastic integration is treated by Métivier [136] and Métivier and Pellaumail [137], 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 Kunita [120], Gyöngy and Krylov [75], and Walsh [186]. The theory of stochastic integration in infinite dimensions, particularly in the context of stochastic partial differential equations, is further discussed in Gawarecki and Mandrekar [68], 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 [166] and Bichteler [13] take a different path. Though the functional analytic approach is mentioned in several textbooks, such as Rogers and Williams [173], where it is used to justify certain results, detailed development is often limited to appendices, as seen in Émery [60]. Bichteler [12] and Protter [166] provide more in-depth treatments of this approach, particularly with respect to general semimartingales.

Additionally, textbooks such as Grigoriu [72], Koller [115], Mürmann [155], and Prigent [163] closely follow Protter's approach, although they often omit most proofs and detailed explanations.

In Karandikar and Rao [109], the classical integral for Brownian motions is first defined, and then the integral is extended via the functional analytic approach as introduced by Protter [166].

Rüschendorf [176] employs the random measures approach to define the integral for general semimartingales in one dimension. This approach is straightforward and easy to understand, although some details require care.

Rao [169] 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.

3.5 Comparison with Recent Developments

The main results of stochastic integration were developed between the 1940s and the 1980s, and the most active period of research on sigma martingales took place between the late 1970s and the late 1990s. Although both topics have been studied thoroughly and their main results are already well established, they have continued to attract research interest up to the present day. Several recent strands bear directly on the topological construction of this thesis or on its principal applications. We organize the discussion into four themes: topological constructions of the stochastic integral, sigma martingales in mathematical finance, frameworks for simplified stochastic calculus, and cylindrical or infinite-dimensional extensions.

3.5.1 Topological Constructions of the Stochastic Integral

Several recent works construct or use stochastic integrals through topological closure or continuity arguments, in the general spirit of our approach.

Cuchiero and Teichmann [35] exploit the semimartingale topology of Émery, the same topology used in our Chapter 4, to give a variant of the Delbaen–Schachermayer FTAP

proof in the setting of admissible portfolio wealth processes. They show that the no-unbounded-profit-with-bounded-risk condition implies a predictable uniform tightness property in the Émery topology and combine this with the closedness results of Mémin and Słominski to obtain an alternative proof of the First FTAP, which is also extended to large financial markets. Their work uses the topology as a tool inside the FTAP machinery rather than as the foundation of the integral itself, whereas our construction uses the same topology to define and study the general stochastic vector integral from first principles.

Kardaras [112] defines stochastic integrals with respect to arbitrary, possibly uncountable, collections of continuous semimartingales by finite-dimensional approximation. He characterizes the resulting space of integrals as the closure, in the semimartingale topology, of the set of integrals against finite sub-collections, and identifies the corresponding integrand class through the covariation structure. The work covers the continuous setting and develops applications to large financial markets. Our framework operates in the càdlàg setting, treats integrators with jumps directly, and characterizes the integral by continuity of the elementary integral on a topological vector space rather than by approximation across an index set. The two constructions are complementary: the present thesis provides a foundation for finitely many càdlàg integrators with jumps, while Kardaras [112] handles arbitrary collections of continuous integrators.

The author has published a survey-style precursor of the present approach in [181], where the topological construction is sketched alongside the classical and functional analytic approaches. Chapter 4 of the present thesis develops that construction in full, with self-contained proofs of all properties used in subsequent chapters and with the vector-valued case treated directly.

3.5.2 Sigma Martingales in Mathematical Finance

Although sigma martingales are not new, their systematic study has accelerated in the last decade, driven largely by mathematical finance.

The closest result in the literature to our generalized Ansel–Stricker Lemma (Theorem 6.1.22) is contained in the monograph [74], where a related statement is given as Theorem 3.21 and obtained through the structural theory of special semimartingales and canonical decompositions. Our proof avoids semimartingale characteristics, local Doob–Meyer decompositions, and the auxiliary criterion that sigma local supermartingales are local martingales. Instead, it is based on the De Donno–Pratelli approximation criterion and the direct local-integrability characterization of sigma martingales. The result also includes equivalent jump-minorant criteria. We discuss this comparison in greater detail in Chapter 6.

The work of Karandikar and Rao [108] proves the Second FTAP in the sigma martingale

setting through a direct argument on stable subspaces, which bypasses a Jacod–Yor representation theorem altogether. The hypotheses of their proof differ from those of the First FTAP, and this leads to a logical asymmetry between the two Fundamental Theorems. Our route, presented in Chapter 6, is to extend the Jacod–Yor representation theorem itself to sigma martingales and to derive the Second FTAP as a consequence. This brings the hypotheses of the Second FTAP in line with those of the First and recovers the structure familiar from the local-martingale case. The two approaches yield the same conclusion in the cases where they overlap, but the route through Jacod–Yor is what justifies posing the question under the same hypotheses as the First FTAP.

The recent paper Acciaio, Backhoff-Veraguas, and Pammer [1] provides a quantitative version of the FTAP, with the no-arbitrage condition replaced by a quantification of the size of arbitrage and with the corresponding pricing measures required only to make asset prices close to martingales. The work is concerned with model uncertainty and robust pricing. Their work is complementary: it quantifies approximate absence of arbitrage and approximate martingale measures, whereas Chapter 7 remains in the qualitative NFLVR and equivalent-sigma martingale-measure framework.

The author has published the relevant parts of Chapter 6 in two papers. The foundational treatment of sigma martingales together with the new proof of the Ansel–Stricker Lemma appears in:

- Moritz Sohns. "σ-Martingales: Foundations, Properties, and a New Proof of the Ansel–Stricker Lemma". In: *Mathematics* 13.4 (2025), p. 682. ISSN: 2227-7390. DOI: 10.3390/math13040682

The Jacod–Yor extension together with its application to the Second FTAP appears in:

- Moritz Sohns. "The Jacod–Yor Theorem for Sigma Martingales and the Second Fundamental Theorem of Asset Pricing". In: *Journal of Stochastic Analysis* 6.2 (2025), Article 1. DOI: 10.31390/josa.6.2.01

The thesis presents both lines together, with the integration framework of Chapter 4 as the underlying setting and with stronger versions of some auxiliary results.

3.5.3 Frameworks for Simplified Stochastic Calculus

A different but closely related agenda is pursued by Černý and Ruf [22] and follow-up works. They introduce a calculus of *semimartingale representations*, in which semimartingales are represented by predictable triplets, and they track how this triplet transforms under predictable changes of variable. This bookkeeping makes drift computations, changes of variable, and changes of measure substantially easier in practice. The follow-up [23] extends

the framework to multiplicative compensation and changes of measure for complex-valued semimartingales, and [21] develops a taxonomy of pure-jump semimartingales within the same framework.

The simplification offered by Černý and Ruf [22] is complementary to ours rather than competing. They simplify the bookkeeping for predictable transformations of semimartingales; the existence and the properties of the stochastic integral itself are taken for granted. We replace the construction of the stochastic integral by a direct topological argument; standard predictable transformations are taken for granted. The two simplifications target different layers of the theory and could, in principle, be combined.

3.5.4 Cylindrical and Infinite-Dimensional Approaches

A substantial body of recent work develops stochastic integration in infinite-dimensional spaces, often through topological methods on the space of semimartingales. Fonseca-Mora [63] constructs a stochastic integral with respect to cylindrical semimartingales valued in the strong dual of a nuclear space, through tensor products of topological vector spaces and a notion of *good integrator* adapted to that setting. Related work of the same author [62] develops UCP convergence and a topology on the corresponding space of good integrators, and the follow-up [64] studies vector-valued stochastic integration for operator-valued integrands with respect to semimartingales in duals of nuclear spaces. The work of Kosmala and Riedle [118] develops stochastic integration with respect to cylindrical Lévy processes by p-summing operators, building on the Hilbert-space cylindrical Lévy integration of Riedle [171] and Jakubowski and Riedle [98]. Banach-space integration with jump noise and Poisson random measures, represented for instance by Brzeźniak, Hausenblas, and Zhu [18], is closely related but not identical to the cylindrical-Lévy programme.

Very recent work of Bodó and Riedle [15] develops stochastic integration with respect to arbitrary cylindrical Lévy processes in Hilbert spaces. Since such cylindrical Lévy processes need not possess a semimartingale decomposition, their construction uses decoupled tangent sequences rather than the semimartingale topology. This further illustrates that infinite-dimensional integration often requires tools quite different from those used in the finite-dimensional semimartingale setting of this thesis.

These works generalize stochastic integration to infinite dimensions and keep a topological perspective on the space of integrators. The present thesis remains finite-dimensional in both the integrator and the integrand, since this suffices for the financial applications targeted. Whether the construction of Chapter 4 extends naturally to the infinite-dimensional setting of nuclear or Banach-space-valued integrators is a possible direction for future work, briefly mentioned in Chapter 8.

Chapter 4

The Topological Stochastic Vector Integral

This chapter presents a novel approach to defining the stochastic vector integral through the topology of semimartingales. While traditional approaches rely heavily on the decomposition of semimartingales into local martingale and finite-variation components as well as continuous and purely discontinuous parts, we demonstrate that by utilizing the semimartingale topology, we can bypass these decompositions entirely and provide a direct, streamlined definition of the integral.

The key innovation lies in using the completeness of the semimartingale topology and the density of simple predictable processes to extend the integral operator continuously in a single step. This method avoids the technical decomposition arguments that characterize both the classical and functional analytic approaches when dealing with unbounded predictable integrands. Specifically, our approach eliminates the need for the Doob–Meyer decomposition and special semimartingale structure, bypasses excursions through $\mathcal{H}^2$ spaces or quasimartingale machinery typically required for unbounded integrands, and extends naturally from the scalar to the vector case without componentwise reconstruction. As a result, this topological approach not only simplifies several arguments in the general theory but also shortens proofs while retaining all standard properties of the general stochastic integral, making it particularly suitable for applications in mathematical finance and accessible for university lectures.

4.1 Semimartingales

Definition 4.1.1. (a) For a d-dimensional stochastic process

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

we put

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

(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

$$H_t(\omega) = H_0(\omega)\mathbb{1}_{\{0\}}(t) + \sum_{i=1}^{n} H_i(\omega)\mathbb{1}_{]\!]T_i,T_{i+1}]\!]}(\omega,t), \tag{4.1}$$

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

The set of $\mathbb{R}^d$-valued simple predictable processes will be denoted by Λ^d.

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

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

and we let Λ_u^d denote Λ^d equipped with the topology induced by $\|\cdot\|_u$ (i.e., uniform convergence in the essential supremum norm): we write $H^n \to H$ in Λ_u^d if $H, H^n \in \Lambda^d$ and $\|H^n - H\|_u \to 0$.

Remark 4.1.2. A stochastic process is a d-dimensional simple predictable process if and only if each component is a one-dimensional simple predictable process.

Remark 4.1.3. 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, [31, Corollary 7.4.3]).

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

Definition 4.1.4. Let X be an $\mathbb{R}^d$-valued stochastic process with càdlàg paths and $H \in \Lambda^d$ as the one given in (Eq. (4.1)). We define the *stochastic integral* $J_X : \Lambda^d \to \mathbb{D}$ by

$$J_X(H) := H \bullet X := H_0^\top X_0 + \sum_{i=1}^{n} H_i^\top (X^{T_{i+1}} - X^{T_i}). \tag{4.2}$$

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 4.1.5. (a) For any left- or right-continuous process X, we put

$$\|X\|_{\text{ucp}} := \sum_{n=1}^{\infty} \frac{1}{2^n} \mathbf{E}\left[X_n^* \wedge 1\right]. \tag{4.3}$$

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

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

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

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 $H^n \xrightarrow{\text{ucp}} H$ if it converges to a process H in the topology of ucp convergence. We use $X^n \xrightarrow{\mathcal{S}} X$ analogously.

Remark 4.1.6. Even though the name suggests otherwise, the Émery norm is not a norm but only a metric (see, for example, [31] page 278 for elaborations).

Remark 4.1.7. 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

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

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

Remark 4.1.8. 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, which ensures 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. [162, 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 in our own definition to highlight the different definition.

Definition 4.1.9. 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 $\mathcal{S}^d_e$ for the space $\mathcal{S}^d$ endowed with the semimartingale topology.

Remark 4.1.10. At first sight, Item (c) may look redundant: for a genuine norm one would have $\|\lambda X\| = |\lambda|\,\|X\| \to 0$ as $\lambda \to 0$. The point is that the so-called Émery "norm" $\|\cdot\|_{\mathcal{S}}$ is not a norm (indeed, not even a homogeneous seminorm); it is a metric generating the semimartingale (Émery) topology.[1] Consequently, continuity of scalar multiplication at 0 is not automatic and must be imposed; this is exactly what Item (c) does. Together with the (translation-invariant) metric structure, Item (c) ensures that $\mathcal{S}^d_e$ is a topological vector space. In particular, without Item (c) the map $(\lambda, X) \mapsto \lambda X$ need not be (jointly) continuous in the product topology.

The necessity of Item (c) is not merely formal: there exist càdlàg processes X such that $\|\lambda_n X\|_{\mathcal{S}} \not\to 0$ for some $\lambda_n \downarrow 0$; see Example 4.1.15 for a stochastic example (fractional Brownian motion with $H < \frac{1}{2}$) and Example 4.1.13 for a deterministic pure-jump example. Moreover, in Theorem 4.1.17 we show that this seemingly minimal definition still recovers the functional-analytic characterization of semimartingales as "good integrators".

Remark 4.1.11. With our (standard) choice of the ucp metric using the truncation $y \mapsto y \wedge 1$, we always have $\|Y\|_{\mathrm{ucp}} \leq 1$ for any càdlàg process Y, and therefore $\|X\|_{\mathcal{S}} \leq 1$ for any càdlàg X. Hence Item (b) is automatically satisfied; we keep it only to stress that $\|\cdot\|_{\mathcal{S}}$ controls the topology rather than the magnitude. If one removes the truncation and works with the untruncated family

$$\|X\|^{\mathrm{untr}}_{\mathcal{S},t} := \sup_{\substack{K\in\Lambda^1 \\ \|K\|_u\leq 1}} \mathbf{E}\Big[\sup_{s\leq t}\big|(K\bullet X)_s\big|\Big], \qquad t > 0,$$

then one can construct càdlàg processes with $\|X\|^{\mathrm{untr}}_{\mathcal{S},t} = +\infty$; see the following example.

Example 4.1.12. Let $t_n := 2^{-n}$ and define a (deterministic, hence adapted) càdlàg pure-jump process

$$X_t := \sum_{n=1}^{\infty} \frac{(-1)^n}{n}\,\mathbb{1}_{\{t\geq t_n\}}, \qquad t \geq 0.$$

[1] Recall $\|X\|_{\mathrm{ucp}} = \sum_{n\geq 1} 2^{-n}\,\mathbf{E}[X^*_n \wedge 1]$ and $\|X\|_{\mathcal{S}} := \sup_{\|K\|_u\leq 1}\|K\bullet X\|_{\mathrm{ucp}}$. The truncation $\wedge 1$ breaks homogeneity.

Then X is finite for every t (the alternating harmonic series converges), but it has infinite total variation on $[0,1]$ since $\sum_{n\geq 1}|\Delta X_{t_n}| = \sum_{n\geq 1}\frac{1}{n} = +\infty$.

For each $N \in \mathbb{N}$, define the finite simple predictable integrand

$$K_t^N := \sum_{n=1}^{N} \operatorname{sgn}(\Delta X_{t_n})\, \mathbb{1}_{]\!]t_{n+1},t_n]\!]}(t),$$

which is genuinely simple predictable in the sense of Definition 4.1.1 (a finite linear combination of indicators of stochastic intervals). Since the indicator sets $]\!]t_{n+1},t_n]\!]$ are pairwise disjoint and the coefficients are in $\{-1,+1\}$, we have $\|K^N\|_u \leq 1$ for every N. For any $t \in (0,1]$, on the event $\{t_N \leq t\}$,

$$(K^N \bullet X)_t = \sum_{n=1}^{N} \mathbb{1}_{\{t_n\leq t\}} \operatorname{sgn}(\Delta X_{t_n})\, \Delta X_{t_n} = \sum_{n=1}^{N} \mathbb{1}_{\{t_n\leq t\}} \left|\Delta X_{t_n}\right| = \sum_{\{n\leq N:\, 2^{-n}\leq t\}} \frac{1}{n}.$$

In particular, evaluating at $t=1$ gives $(K^N \bullet X)_1 = \sum_{n=1}^{N}\frac{1}{n} \to +\infty$ as $N\to\infty$, so

$$\sup_{N\in\mathbb{N}} (K^N \bullet X)_1 = +\infty, \qquad \sup_{N\in\mathbb{N}} \|K^N\|_u \leq 1.$$

Hence the supremum over finite simple predictable integrands of unit uniform norm of the value of $K^N \bullet X$ at $t=1$ is already infinite, which gives $\|X\|_{\mathcal{S},1}^{\mathrm{untr}} = +\infty$ for the un-truncated version of the $\mathcal{S}$-norm without ever needing to invoke an infinite sum as a single simple predictable process. By contrast, under our truncated definition of $\|\cdot\|_{\mathcal{S}}$ used in this thesis one always has $\|X\|_{\mathcal{S}} \leq 1$, so the phenomenon $\|X\|_{\mathcal{S}} = +\infty$ cannot occur. However, it turns out that X is still not a semimartingale, as Item (c) is not satisfied (see, Example 4.1.13).

Example 4.1.13. Let X be the deterministic pure-jump process from Example 4.1.12, i.e.,

$$X_t = \sum_{n=1}^{\infty} \frac{(-1)^n}{n}\, \mathbb{1}_{\{t\geq t_n\}}, \qquad t_n := 2^{-n}.$$

For $m \in \mathbb{N}$ choose N_m so large that the harmonic sum

$$H_{N_m} := \sum_{n=1}^{N_m} \frac{1}{n} \;\geq\; m.$$

Define the simple predictable integrand

$$\widehat{K}_t^{(m)} \;:=\; \sum_{n=1}^{N_m} \operatorname{sgn}(\Delta X_{t_n})\, \mathbb{1}_{]\!]t_{n+1},t_n]\!]}(t), \qquad \text{so that} \quad \|\widehat{K}^{(m)}\|_u = 1,$$

and set $\lambda_m := 1/m$. Then, at $t = 1$,

$$\left(\widehat{K}^{(m)} \bullet (\lambda_m X)\right)_1 = \lambda_m \sum_{n=1}^{N_m} \operatorname{sgn}(\Delta X_{t_n})\, \Delta X_{t_n} = \frac{1}{m} \sum_{n=1}^{N_m} \left|\Delta X_{t_n}\right| = \frac{1}{m} \sum_{n=1}^{N_m} \frac{1}{n} \;\geq\; 1.$$

Hence $\sup_{s\leq 1} \left|(\widehat{K}^{(m)} \bullet (\lambda_m X))_s\right| \geq 1$ deterministically, and therefore

$$\|\lambda_m X\|_{\mathcal{S}} = \sup_{\|K\|_u \leq 1} \|K \bullet (\lambda_m X)\|_{\mathrm{ucp}} \;\geq\; \|\widehat{K}^{(m)} \bullet (\lambda_m X)\|_{\mathrm{ucp}} \;\geq\; \tfrac{1}{2}$$

(for the ucp metric, the $n = 1$ term contributes $2^{-1}\mathbf{E}[(\cdot)_1^* \wedge 1] = \frac{1}{2}$). Thus $\|\lambda_m X\|_{\mathcal{S}} \not\to 0$ as $m \to \infty$, so Item (c) is violated. In particular, X is not a topological semimartingale; by Theorem 4.2.11, it is not a classical semimartingale either.

Remark 4.1.14. More generally, a deterministic càdlàg process is a semimartingale if and only if it has finite variation on compacts; the above X has infinite total variation on $[0, 1]$, hence fails this criterion.

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, [10] 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 [131, Example 2 of Section 4.9.13], Lin [130, Corollary 2.2] or Rogers [172, 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 $H < \frac{1}{2}$.

Example 4.1.15. 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

$$\begin{aligned} H^{\pi_0,N} \bullet (\lambda X)_T &= -\frac{\lambda}{N} \sum_{j=0}^{n-1} X_{t_j}(X_{t_{j+1}} - X_{t_j}) \\ &= \frac{\lambda}{N} \sum_{j=0}^{n-1} X_{t_j}^2 - X_{t_j} X_{t_{j+1}} \\ &= \frac{\lambda}{2N} \left(\sum_{j=0}^{n-1} (X_{t_{j+1}} - X_{t_j})^2 - X_T^2 \right) \\ &> \frac{\lambda}{2N} \left(\frac{2N}{\lambda} + N^2 - N^2 \right) = 1. \end{aligned}$$

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, [56, Lemma 4.1.2]). Hence

$$c := \lim_{\lambda \to 0} \sup_{\pi} \mathbf{P} \left(\sum_{j=0}^{n-1} \left(X_{t_{j+1}} - X_{t_j} \right)^2 > \frac{1}{\lambda} \right)$$

exists and is greater than 0.

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

$$\mathbf{P} \left(\sum_{j=0}^{n-1} \left(X_{t_{j+1}} - X_{t_j} \right)^2 > \frac{2N}{\lambda} + N^2 \right) > \frac{c}{2}. \tag{4.4}$$

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

$$\mathbf{P} \left(X_T^* > N \right) < \frac{c}{4}. \tag{4.5}$$

With Eq. (4.5) and Eq. (4.4), we conclude

$$\mathbf{P} \left(A^{\pi_0,N,\lambda} \right) > \frac{c}{4} > 0.$$

Example 4.1.16. Following a similar strategy as above, one can show that for a Brownian motion B, $|B|^\alpha$ with $0 < \alpha < 1$ is also not a semimartingale. A proof of this fact in the classical semimartingale framework can be found in Wang [187] or Yor [195].

In the functional analytic approach, as presented by Protter [166], 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 4.1.17. *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.*

For the proof, we will use a Lemma, that will also be instrumental in the proof of the stochastic Dominated Convergence Theorem.

Lemma 4.1.18 (Stopping-time compression)**.** *Let X be a d-dimensional adapted càdlàg process and let $t > 0$. Then, for every $\theta > 0$,*

$$\sup_{\substack{G\in\Lambda^d \\ \|G\|_u\leq 1}} \mathbf{P}\Big((G\bullet X)_t^* > \theta\Big) = \sup_{\substack{K\in\Lambda^d \\ \|K\|_u\leq 1}} \mathbf{P}\Big(\,|(K\bullet X)_t| > \theta\Big). \tag{4.6}$$

Consequently, for every $M > 0$,

$$\sup_{\substack{G\in\Lambda^d \\ \|G\|_u\leq 1}} \mathbf{E}\Big[(G\bullet X)_t^* \wedge M\Big] \leq \int_0^M \sup_{\substack{K\in\Lambda^d \\ \|K\|_u\leq 1}} \mathbf{P}\Big(\,|(K\bullet X)_t| > \theta\Big)\,\mathrm{d}\theta. \tag{4.7}$$

Proof. We first prove the probability-tail equality. The inequality "$\geq$" in Eq. (4.6) is immediate. Indeed, for every $K \in \Lambda^d$ with $\|K\|_u \leq 1$,

$$|(K\bullet X)_t| \leq (K\bullet X)_t^* \qquad \text{a.s.}$$

Hence

$$\mathbf{P}\Big(\,|(K\bullet X)_t| > \theta\Big) \leq \mathbf{P}\Big((K\bullet X)_t^* > \theta\Big) \leq \sup_{\substack{G\in\Lambda^d \\ \|G\|_u\leq 1}} \mathbf{P}\Big((G\bullet X)_t^* > \theta\Big).$$

Taking the supremum over all $K \in \Lambda^d$ with $\|K\|_u \leq 1$ gives one direction.

For the converse, fix $G \in \Lambda^d$ with $\|G\|_u \leq 1$, and put

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

Then Y is adapted and càdlàg. Let $a > \theta$, and define

$$\tau_a := \inf\{s \geq 0 : |Y_s| > a\}, \qquad \inf\emptyset := \infty.$$

Since Y is adapted and càdlàg, and since the filtration is right-continuous by our standing assumptions, τ_a is a stopping time. If $Y_t^* > a$, then there is some $s \leq t$ such that $|Y_s| > a$. Hence $\tau_a \leq t$, and therefore

$$\{Y_t^* > a\} \subseteq \{\tau_a \leq t\}.$$

By right-continuity of Y and the definition of τ_a, we have

$$|Y_{\tau_a}| \geq a \qquad \text{on } \{\tau_a < \infty\}.$$

In particular,

$$|Y_{\tau_a}| > \theta \qquad \text{on } \{Y_t^* > a\}.$$

Set

$$\sigma_a := \tau_a \wedge t$$

and define

$$K := G\mathbb{1}_{[\![0,\sigma_a]\!]}.$$

Since simple predictable processes are stable under stopping, $K \in \Lambda^d$. Moreover,

$$\|K\|_u \leq \|G\|_u \leq 1.$$

The elementary stopping identity gives

$$K \bullet X = Y^{\sigma_a}.$$

Therefore, on the event $\{Y_t^* > a\}$, we have $\tau_a \leq t$ and hence

$$(K \bullet X)_t = Y_{\sigma_a} = Y_{\tau_a}.$$

Consequently,

$$\{Y_t^* > a\} \subseteq \Big\{ |(K \bullet X)_t| > \theta \Big\}.$$

It follows that

$$\mathbf{P}\Big(Y_t^* > a\Big) \leq \mathbf{P}\Big(|(K \bullet X)_t| > \theta \Big) \leq \sup_{\substack{K \in \Lambda^d \\ \|K\|_u \leq 1}} \mathbf{P}\Big(|(K \bullet X)_t| > \theta \Big).$$

Letting $a \downarrow \theta$, the events $\{Y_t^* > a\}$ increase to $\{Y_t^* > \theta\}$. Hence, by continuity of probability from below,

$$\mathbf{P}\Big(Y_t^* > \theta\Big) \leq \sup_{\substack{K \in \Lambda^d \\ \|K\|_u \leq 1}} \mathbf{P}\Big(|(K \bullet X)_t| > \theta \Big).$$

Taking the supremum over all $G \in \Lambda^d$ with $\|G\|_u \leq 1$ proves the opposite inequality in Eq. (4.6). Thus the probability-tail equality is proved.

It remains to prove the integrated consequence. For each fixed $G \in \Lambda^d$ with $\|G\|_u \leq 1$,

the layer-cake formula gives

$$\mathbf{E}\Big[(G\bullet X)_t^* \wedge M\Big] = \int_0^M \mathbf{P}\Big((G\bullet X)_t^* > \theta\Big)\,\mathrm{d}\theta.$$

By Eq. (4.6), for every $\theta > 0$,

$$\mathbf{P}\Big((G\bullet X)_t^* > \theta\Big) \leq \sup_{\substack{K\in\Lambda^d \\ \|K\|_u\leq 1}} \mathbf{P}\Big(\,|(K\bullet X)_t| > \theta\Big).$$

The right-hand side is decreasing as a function of θ, hence measurable. Integrating over $(0, M)$ and then taking the supremum over all $G \in \Lambda^d$ with $\|G\|_u \leq 1$ yields

$$\sup_{\substack{G\in\Lambda^d \\ \|G\|_u\leq 1}} \mathbf{E}\Big[(G\bullet X)_t^* \wedge M\Big] \leq \int_0^M \sup_{\substack{K\in\Lambda^d \\ \|K\|_u\leq 1}} \mathbf{P}\Big(\,|(K\bullet X)_t| > \theta\Big)\,\mathrm{d}\theta.$$

This proves Eq. (4.7). □

Proof of Theorem 4.1.17. Assume first that X is a topological semimartingale, and let $(H^n)_{n\in\mathbb{N}} \subset \Lambda^d$ satisfy $\|H^n\|_u \to 0$. Set

$$\lambda_n := \|H^n\|_u\,.$$

If $\lambda_n = 0$, then $H^n\bullet X = 0$. Otherwise, put

$$K^n := \lambda_n^{-1} H^n,$$

so that $\|K^n\|_u \leq 1$. By linearity of the elementary stochastic integral in the integrator,

$$H^n\bullet X = K^n\bullet(\lambda_n X).$$

Hence

$$\|H^n\bullet X\|_{\mathrm{ucp}} \leq \sup_{\substack{K\in\Lambda^d \\ \|K\|_u\leq 1}} \|K\bullet(\lambda_n X)\|_{\mathrm{ucp}} = \|\lambda_n X\|_{\mathcal{S}}\,.$$

Since X is a topological semimartingale and $\lambda_n \to 0$, Item (c) gives $\|\lambda_n X\|_{\mathcal{S}} \to 0$. Thus $H^n\bullet X \to 0$ in ucp, and in particular $(H^n\bullet X)_t \to 0$ in probability for every $t \geq 0$.

Conversely, assume that for each sequence $(H^n)_{n\in\mathbb{N}} \subset \Lambda^d$ with $\|H^n\|_u \to 0$ one has $(H^n\bullet X)_t \to 0$ in probability for every $t \geq 0$. We verify Item (a)–Item (c). The first condition holds by assumption, and Item (b) is automatic for the truncated metric used here. It remains to show Item (c).

Let $\lambda_n \to 0$. For an adapted càdlàg process Y and $r \in \mathbb{N}$, write

$$q_r(Y) := \sup_{\substack{K \in \Lambda^d \\ \|K\|_u \le 1}} \mathbf{E}\Big[(K \bullet Y)_r^* \wedge 1\Big].$$

We first show that

$$q_r(\lambda_n X) \longrightarrow 0$$

for every fixed $r \in \mathbb{N}$.

Fix $r \in \mathbb{N}$ and $\theta > 0$. We claim that

$$\sup_{\substack{K \in \Lambda^d \\ \|K\|_u \le 1}} \mathbf{P}\Big(|(K \bullet (\lambda_n X))_r| \ge \theta \Big) \longrightarrow 0. \tag{4.8}$$

Indeed, if Eq. (4.8) failed, there would exist $\varepsilon > 0$, a subsequence $(n_j)_{j \in \mathbb{N}}$, and $K^j \in \Lambda^d$ with $\|K^j\|_u \le 1$ such that

$$\mathbf{P}\Big(\Big|(K^j \bullet (\lambda_{n_j} X))_r\Big| \ge \theta \Big) \ge \varepsilon$$

for all j. But $G^j := \lambda_{n_j} K^j$ satisfies $\|G^j\|_u \le \big|\lambda_{n_j}\big| \to 0$, while

$$(G^j \bullet X)_r = (K^j \bullet (\lambda_{n_j} X))_r.$$

This contradicts the assumed fixed-time convergence applied to the sequence (G^j). Thus Eq. (4.8) holds.

By Lemma 4.1.18,

$$q_r(\lambda_n X) \le \int_0^1 \sup_{\substack{K \in \Lambda^d \\ \|K\|_u \le 1}} \mathbf{P}\Big(|(K \bullet (\lambda_n X))_r| \ge \theta \Big) \mathrm{d}\theta.$$

For every $\theta > 0$, the integrand tends to 0 by Eq. (4.8), and it is bounded by 1. Dominated convergence gives $q_r(\lambda_n X) \to 0$ for every fixed $r \in \mathbb{N}$.

Finally,

$$\begin{aligned} \|\lambda_n X\|_{\mathcal{S}} &= \sup_{\substack{K \in \Lambda^d \\ \|K\|_u \le 1}} \sum_{r=1}^{\infty} 2^{-r} \mathbf{E}\Big[(K \bullet (\lambda_n X))_r^* \wedge 1\Big] \\ &\le \sum_{r=1}^{\infty} 2^{-r} q_r(\lambda_n X). \end{aligned}$$

Let $\varepsilon > 0$. Choose $R \in \mathbb{N}$ such that $\sum_{r>R} 2^{-r} < \varepsilon/2$. For $r = 1, \dots, R$, choose n_0 so large

that $q_r(\lambda_n X) < \varepsilon/(2R)$ for all $n \geq n_0$. Then, for $n \geq n_0$,

$$\|\lambda_n X\|_{\mathcal{S}} \leq \sum_{r=1}^{R} 2^{-r} \frac{\varepsilon}{2R} + \sum_{r>R} 2^{-r} < \varepsilon.$$

Hence $\|\lambda_n X\|_{\mathcal{S}} \to 0$, which is Item (c). Therefore X is a topological semimartingale. □

Remark 4.1.19. Theorem 4.1.17 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 4.1.17 that would yield an equivalent definition of semimartingales. We are going to show one of them in our version of the Bichteler-Dellacherie-Mokobozki Theorem. For equivalent statements, we refer to [109, Theorem 5.89].

4.2 Properties of Semimartingales

Theorem 4.2.1. *(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.*

(b) Local topological semimartingales and processes that are prelocally topological semimartingales are topological semimartingales.

(c) 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 4.1.17. Let X be a local topological semimartingale with localizing sequence $(T^k)_{k\in\mathbb{N}}$, so that for every k the stopped process X^{T^k} is a topological semimartingale and $T^k \to \infty$ a.s. Fix $t > 0$ and $\varepsilon > 0$. Let $(H^n)_{n\in\mathbb{N}} \subset \Lambda^d$ with $\|H^n\|_u \to 0$; we must show $(H^n \bullet X)_t \to 0$ in probability.

For every $k \in \mathbb{N}$ and every $n \in \mathbb{N}$, since $(H^n \bullet X)^{T^k} = H^n \bullet X^{T^k}$ on $\{T^k \geq t\}$,

$$\mathbf{P}(|(H^n \bullet X)_t| \geq \varepsilon) \;\leq\; \mathbf{P}\Big(|(H^n \bullet X^{T^k})_t| \geq \varepsilon\Big) \;+\; \mathbf{P}(T^k \leq t).$$

Order of limits. For each fixed k, the stopped integrator X^{T^k} is a topological semimartingale (by hypothesis), so by Theorem 4.1.17 applied to X^{T^k},

$$\mathbf{P}\Big(|(H^n \bullet X^{T^k})_t| \geq \varepsilon\Big) \longrightarrow 0 \qquad \text{as } n \to \infty, \quad \text{with } k \text{ fixed.}$$

Hence

$$\limsup_{n\to\infty} \mathbf{P}(|(H^n \bullet X)_t| \geq \varepsilon) \;\leq\; \mathbf{P}(T^k \leq t).$$

Now let $k \to \infty$. Since $T^k \to \infty$ a.s., $\mathbf{P}(T^k \leq t) \to 0$, hence

$$\limsup_{n\to\infty} \mathbf{P}(|(H^n \bullet X)_t| \geq \varepsilon) = 0.$$

This shows $(H^n \bullet X)_t \to 0$ in probability, so by Theorem 4.1.17, X is a topological semimartingale.

If X is prelocally a topological semimartingale, the same argument applies with the prelocalizing sequence, using X^{T^k-} in place of X^{T^k}.

(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 4.1.17). In fact, these definitions are equivalent, which was proved in the late 1970s independently of each other by Klaus Bichteler [11, 12] and Claude Dellacherie with contributions by Paul-André Meyer and Gabriel Mokobodzki [43]. In addition to the literature mentioned above, the theorem is also discussed and proven in modern textbooks such as in Cohen and Elliott [31, Theorem 12.3.26], in Karandikar and Rao [109, Theorem 5.59], and in Protter [166, Theorem III.47]. While the latter two employ classical methods, the former adopts an approach by Beiglböck, who has presented alternative proofs in [7, 8, 9].

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

The final form of the Bichteler-Dellacherie-Mokobodzki Theorem, including the bounded pointwise continuity criterion used later in the proof of the Dominated Convergence Theorem, is stated and proved after the auxiliary estimates below.

For the proofs of Proposition 4.2.6 and Proposition 4.2.3, and thus also for Theorem 4.2.11, 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 4.2.2. *Let M be a d-dimensional càdlàg process and $H \in \Lambda^d$.*

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

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

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

$$\alpha \mathbf{P}\left[(H \bullet M)_t^* \geq \alpha\right] \leq 18 \left\|H\right\|_u d \sum_{j=1}^{d} \mathbf{E}\left[\left|M_t^j\right|\right].$$

Proof. (a) By Lemma 2.2.3, an application of the tower rule and the Cauchy-Schwarz inequality, we get

$$\begin{aligned}
\mathbf{E}\left[(H \bullet M_t)^2\right] &= \mathbf{E}\left[\left(H_0^\top M_0 + \sum_{i=1}^{n} H_i^\top \left(M_t^{T_{i+1}} - M_t^{T_i}\right)\right)^2\right] \\
&= \mathbf{E}\left[\left(H_0^\top M_0\right)^2 + \sum_{i=1}^{n} \left(H_i^\top \left(M_t^{T_{i+1}} - M_t^{T_i}\right)\right)^2\right] \\
&\leq \mathbf{E}\left[\left\|H_0\right\|_{\mathbb{R}^d}^2 \left\|M_0\right\|_{\mathbb{R}^d}^2 + \sum_{i=1}^{n} \left\|H_i\right\|_{\mathbb{R}^d}^2 \left\|M_t^{T_{i+1}} - M_t^{T_i}\right\|_{\mathbb{R}^d}^2\right] \\
&\leq d\|H\|_u^2 \mathbf{E}\left[\left\|M_0\right\|_{\mathbb{R}^d}^2 + \sum_{i=1}^{n} \left\|M_t^{T_{i+1}} - M_t^{T_i}\right\|_{\mathbb{R}^d}^2\right] \\
&\leq d \left\|H\right\|_u^2 \mathbf{E}\left[\left\|M_t\right\|_{\mathbb{R}^d}^2\right],
\end{aligned}$$

where the last step uses that the orthogonal martingale increments telescope, so that $\mathbf{E}\big[\left\|M_0\right\|^2 + \sum_{i=1}^{n} \left\|M_t^{T_{i+1}} - M_t^{T_i}\right\|^2\big] = \mathbf{E}\big[\left\|M_{T_{n+1} \wedge t}\right\|^2\big] \leq \mathbf{E}\big[\left\|M_t\right\|^2\big]$, the final inequality being the submartingale property of $\left\|M\right\|^2$ (with equality whenever $T_{n+1} \geq t$).

(b) The idea is to transform the time-continuous statement into a time-discrete one and then apply Theorem 2.5.3. We first assume $d = 1$ and $\sup_t |H| < 1$. Since $H \in \Lambda$,

there is a representation

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

with stopping times T_i. By the tower rule and Doob's optional stopping theorem, one easily sees that $(M^T_{T_i})_{i\in\mathbb{N}}$ 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 2.5.3. The discrete maximal inequality Theorem 2.5.3 thus controls $\sup_i |(H\bullet M)_{T_i\wedge t}|$. To pass from this discrete maximum to the genuine running supremum $(H\bullet M)^*_t = \sup_{0\le s\le t} |(H\bullet M)_s|$, note that on each interval $]\!]T_i, T_{i+1}]\!]$ the integrand H is constant, so refining the grid $\{T_i\}$ by additional time points leaves H unchanged — and hence also $\|H\|_u$ and the right-hand side of Theorem 2.5.3 — while $(H\bullet M)$ evaluated along the refined grid is still the time-discrete transform of M with the same bound. Choosing a sequence of refinements whose mesh tends to 0, the right-continuity of $s \mapsto (H\bullet M)_s$ gives $\sup_{T_i \text{ in refined grid}} |(H\bullet M)_{T_i\wedge t}| \uparrow (H\bullet M)^*_t$ almost surely. Since the bound is uniform over all refinements, it passes to the limit and yields the result for $d = 1$. For general d and H, we obtain

$$\begin{aligned}
\alpha \mathbf{P}\left[(H\bullet M)^*_t \ge \alpha\right] &= \alpha \mathbf{P}\left[\left|\frac{H}{\|H\|_u}\bullet M\right|^*_t \ge \frac{\alpha}{\|H\|_u}\right] \\
&\le \alpha \sum_{j=1}^{d} \mathbf{P}\left[\left|\frac{H^j}{\|H\|_u}\bullet M^j\right|^*_t \ge \frac{\alpha}{d\|H\|_u}\right] \\
&\le 18\|H\|_u d \sum_{j=1}^{d} \mathbf{E}\left[|M^j_t|\right].
\end{aligned}$$

□

We now continue with the first part of the Bichteler-Dellacherie-Mokobodzki Theorem.

Proposition 4.2.3. *Every classical semimartingale is a topological semimartingale.*

Proof. We 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

$$\begin{aligned}
\mathbf{Var}[H\bullet A] &= \mathbf{Var}\left[\sum_{i=1}^{n} H_i^\top \left(A^{T_{i+1}} - A^{T_i}\right)\right] \\
&\le \sum_{i=1}^{n}\sum_{j=1}^{d} \left|h_i^j\right| \mathbf{Var}\left[A^{j,T_{i+1}} - A^{j,T_i}\right] \\
&\le \|H\|_u \sum_{j=1}^{d} \mathbf{Var}\left[A^j\right].
\end{aligned}$$

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 4.1.17. For any martingale M, we can proceed analogously. By applying Theorem 4.2.2, we see that M is a topological semimartingale. Via localization, we get that any local martingale is locally a topological semimartingale; thus, by Theorem 4.2.1, any local martingale is a topological semimartingale. By Theorem 4.2.1 $M+A$ is a topological semimartingale as well. □

Remark 4.2.4. 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 4.2.2 would suffice to demonstrate one direction of the Bichteler-Dellacherie-Mokobodzki Theorem.

Example 4.2.5.

- 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.

We have already shown that classical semimartingales are topological semimartingales. It remains to prove the converse:

Proposition 4.2.6. *Every topological semimartingale is a classical semimartingale.*

For this statement, there are multiple proofs available in the literature. Most of them require deep knowledge and technical results about the decomposition of semimartingales and stochastic integration. We now follow an ansatz by [9] to give a shorter alternative proof with basic tools.

We first establish several key lemmas. The main idea is to show that topological semimartingales are locally quasimartingales, and then apply the Doob-Meyer decomposition.

We start with the key technical result:

Lemma 4.2.7. *Let X be a bounded càdlàg adapted topological semimartingale. Then for every $\varepsilon > 0$, there exist a constant C and a sequence of stopping times $(\varrho_n)_{n\in\mathbb{N}}$ with*

[2] see Cohen and Elliott [31] Definition 5.5.16
[3] see Cohen and Elliott [31] Theorem 5.5.18
[4] see Cohen and Elliott [31] Definition 13.5.1
[5] This was initially shown in Lévy [129], see also Itô [90] for an English proof, Kunita and Watanabe [121] for a proof in the classical semimartingale setting or Applebaum [4] Proposition 2.7.1 for a modern textbook treatment.

values in $[0,1] \cup \{\infty\}$ *such that* $\mathbf{P}(\varrho_n = \infty) \geq 1 - \varepsilon$ *and* $\mathrm{MV}(X^{\varrho_n}, D_n) \leq C$*, where* $D_n = \{0, 1/2^n, 2/2^n, \ldots, 1\}$ *is the* n*-th dyadic partition.*

Proof. Since X is a topological semimartingale, by Theorem 4.1.17, for given $\varepsilon > 0$ there exists $C > 0$ such that for all $H \in \Lambda^d$ with $\|H\|_u \leq 1$, we have

$$\mathbf{P}\left((H \bullet X)_1 \geq C - 2d \|X\|_\infty\right) \leq \varepsilon,$$

where $\|X\|_\infty := \operatorname{ess\,sup}_{\omega,t} \|X_t(\omega)\|_{\mathbb{R}^d}$.

For each n, define the simple predictable process

$$H^n := \sum_{t_i \in D_n} \mathbb{1}_{(t_i, t_{i+1}]} \cdot \frac{\mathbf{E}[X_{t_{i+1}} - X_{t_i} \mid \mathcal{F}_{t_i}]}{\left\|\mathbf{E}[X_{t_{i+1}} - X_{t_i} \mid \mathcal{F}_{t_i}]\right\|_{\mathbb{R}^d}},$$

with the convention that a summand is set to 0 on the set $\{\mathbf{E}[X_{t_{i+1}} - X_{t_i} \mid \mathcal{F}_{t_i}] = 0\}$. By construction each coefficient is therefore either a unit vector or zero. We further define the stopping time

$$\varrho_n := \inf\left\{t \in D_n \,;\, (H^n \bullet X)_t \geq C - 2d \|X\|_\infty\right\}.$$

Note that $\|H^n\|_u \leq 1$ by construction. On the set $\{\varrho_n < \infty\}$, we have $(H^n \bullet X)_1^{\varrho_n} \geq C - 2d \|X\|_\infty$, hence $\mathbf{P}(\varrho_n = \infty) \geq 1 - \varepsilon$.

Moreover, by the definition of the stochastic integral and the fact that jumps are bounded by $2\|X\|_\infty$, and writing $v_i := \mathbf{E}[X_{t_{i+1}} - X_{t_i} \mid \mathcal{F}_{t_i}]$,

$$\begin{aligned}
C &\geq \mathbf{E}[(H^n \bullet X)_1^{\varrho_n}] \\
&= \mathbf{E}\left[\sum_{t_i \in D_n} \mathbb{1}_{\{t_i < \varrho_n\}} \frac{(X_{t_{i+1}} - X_{t_i})^\top v_i}{\|v_i\|_{\mathbb{R}^d}}\right] \\
&= \mathbf{E}\left[\sum_{t_i \in D_n} \mathbb{1}_{\{t_i < \varrho_n\}} \frac{v_i^\top v_i}{\|v_i\|_{\mathbb{R}^d}}\right] \\
&= \mathbf{E}\left[\sum_{t_i \in D_n} \mathbb{1}_{\{t_i < \varrho_n\}} \|v_i\|_{\mathbb{R}^d}\right] \\
&= \mathrm{MV}(X^{\varrho_n}, D_n),
\end{aligned}$$

where the second equality is the tower rule (the indicator $\mathbb{1}_{\{t_i < \varrho_n\}}$ and the normalizing factor $1/\|v_i\|_{\mathbb{R}^d}$ are $\mathcal{F}_{t_i}$-measurable, and $\mathbf{E}[X_{t_{i+1}} - X_{t_i} \mid \mathcal{F}_{t_i}] = v_i$), and the third uses $v_i^\top v_i = \|v_i\|_{\mathbb{R}^d}^2$ together with the convention $0/0 := 0$. This is exactly where the unit-vector normalization is needed: had we instead divided by $\|v_i\|_{\mathbb{R}^d} \vee 1$, the summand would equal $\|v_i\|_{\mathbb{R}^d}^2 / (\|v_i\|_{\mathbb{R}^d} \vee 1)$, which is strictly smaller than $\|v_i\|_{\mathbb{R}^d}$ whenever $\|v_i\|_{\mathbb{R}^d} < 1$, and the

bound $\mathrm{MV}(X^{\varrho_n}, D_n) \leq C$ would fail. □

Lemma 4.2.8. *Let $(\varrho_n)_{n\in\mathbb{N}}$ be a sequence of stopping times with values in $[0,1]\cup\{\infty\}$ such that $\mathbf{P}(\varrho_n = \infty) \geq 1-\varepsilon$ for all n. Then there exists a stopping time ϱ and for each $n \geq 1$ convex weights $\mu_n^n, \ldots, \mu_{N_n}^n$ such that $\mathbf{P}(\varrho = \infty) \geq 1 - 3\varepsilon$ and*

$$\mathbb{1}_{[0,\varrho]} \leq 2\sum_{k=n}^{N_n} \mu_k^n \mathbb{1}_{[0,\varrho_k]} \quad \textit{for all } n \geq 1.$$

Proof. Apply Mazur's lemma 2.5.6 in $L^2(\mathbf{P})$ to the sequence $(\mathbb{1}_{\{\varrho_n=\infty\}})_{n\in\mathbb{N}}$ to obtain convex combinations that converge. The details follow as in [9]. □

Lemma 4.2.9. *If X is locally a classical semimartingale, then X is a classical semimartingale.*

Proof. There exists a sequence $(\sigma_n)_n$ of stopping times with $\mathbf{P}(\sigma_n < \infty) \leq 2^{-n}$ such that each $X^{\sigma_n} = M^n + A^n$ with M^n a martingale and A^n an FV-process. Setting $\tau_k := \inf_{n\geq k} \sigma_n$, we have

$$X = X^{\tau_1} + \sum_{k=1}^{\infty}(X^{\tau_{k+1}} - X^{\tau_k}) = \left[\sum_{k=1}^{\infty}(M^{k+1,\tau_{k+1}} - M^{k,\tau_k})\right] + \left[\sum_{k=1}^{\infty}(A^{k+1,\tau_{k+1}} - A^{k,\tau_k})\right],$$

which gives the desired decomposition. □

We can now prove the other direction of the Bichteler-Dellacherie-Mokobodzki Theorem.

Proof of Proposition 4.2.6. Now we complete the proof. Let X be a topological semimartingale. By taking a suitable decomposition (extracting large jumps), we may assume without loss of generality that X is bounded.

Given $\varepsilon > 0$, apply Lemma 4.2.7 to obtain C and $(\varrho_n)_n$, then apply Lemma 4.2.8 to obtain ϱ with $\mathbf{P}(\varrho = \infty) \geq 1 - 3\varepsilon$.

For any $n \geq 1$, using the convex combination from Lemma 4.2.8:

$$\begin{aligned}
\mathrm{MV}(X^{\varrho}, D_n) &\leq 2\sum_{k=n}^{N_n} \mu_k^n \mathrm{MV}(X^{\varrho_k}, D_n) + 2d\,\|X\|_\infty \\
&\leq 2\sum_{k=n}^{N_n} \mu_k^n \mathrm{MV}(X^{\varrho_k}, D_k) + 2d\,\|X\|_\infty \\
&\leq 2C + 2d\,\|X\|_\infty .
\end{aligned}$$

Since this holds for all n, we have $\mathrm{MV}(X^{\varrho}) \leq 2C + 2d\,\|X\|_\infty < \infty$. By Theorem 2.4.3, X^{ϱ} is componentwise the difference of submartingales, hence by the Doob-Meyer decomposition, X^{ϱ} is a classical semimartingale.

Since this holds for arbitrarily small $\varepsilon > 0$, X is locally a classical semimartingale. By Lemma 4.2.9, X is a classical semimartingale. □

Instead of the uniform convergence used in Theorem 4.1.17, one may also characterize semimartingales by bounded pointwise convergence of simple predictable integrands. We prove the required bounded-pointwise implication in a form that will be used below. The proof uses only the martingale facts collected in Chapter 2 and the estimates established in the present chapter.

Proposition 4.2.10 (Bounded pointwise convergence for classical semimartingales)**.** *Let X be a real-valued classical semimartingale and let $T > 0$. Let $(H^n)_{n\in\mathbb{N}} \subset \Lambda^1$ satisfy*

$$\sup_{n\in\mathbb{N}} \|H^n\|_u < \infty$$

and suppose that

$$H^n \longrightarrow 0$$

pointwise outside an evanescent set. Then

$$(H^n \bullet X)^*_T \longrightarrow 0$$

in probability.

Proof. Put

$$C := \sup_{n\in\mathbb{N}} \|H^n\|_u < \infty.$$

Step 1: square-integrable martingales. Let N be a square-integrable martingale. We first show that

$$(H^n \bullet N)^*_T \longrightarrow 0$$

in probability.

The initial term satisfies

$$H^n_0 N_0 \longrightarrow 0$$

in probability, since $H^n_0 \to 0$ a.s. and $|H^n_0| \le C$. Hence it is enough to consider $N - N_0$, and we may assume without loss of generality that $N_0 = 0$.

We construct the finite measure associated with N on the predictable σ-algebra restricted to $\Omega \times (0, T]$. For stochastic rectangles of the form $A \cap]\!]S, R]\!]$, where $S \le R \le T$ are bounded stopping times and $A \in \mathcal{F}_S$, define

$$\mu^T_N\big(A \cap]\!]S, R]\!]\big) := \mathbf{E}\big[\mathbb{1}_A (N_R - N_S)^2\big].$$

By Corollary 2.2.5, martingale increments over disjoint predictable stochastic rectangles are orthogonal in L^2. Hence, after writing a simple predictable set as a finite disjoint union of such rectangles, the value obtained by summing the preceding expression is independent of the chosen representation. This gives a finite premeasure on the algebra of finite disjoint unions of predictable stochastic rectangles. Its total mass is $\mathbf{E}[N_T^2]$, and therefore it extends to a finite measure, still denoted by μ_N^T, on the predictable σ-algebra restricted to $\Omega \times (0,T]$.

For every $K \in \Lambda^1$, the elementary martingale-transform calculation and Corollary 2.2.5 give

$$\mathbf{E}\Big[((K\bullet N)_T)^2\Big] = \int_{\Omega\times(0,T]} K^2 \mathrm{d}\mu_N^T. \tag{4.9}$$

Indeed, after writing K on a common stopping-time grid, the cross terms vanish by orthogonality of martingale increments, and the diagonal terms are precisely the values of μ_N^T on the corresponding predictable stochastic rectangles.

The evanescent exceptional set on which H^n may fail to converge is μ_N^T-null. Since $|H^n| \le C$ and μ_N^T is finite, ordinary dominated convergence yields

$$\int_{\Omega\times(0,T]} (H^n)^2 \mathrm{d}\mu_N^T \longrightarrow 0.$$

Now apply Lemma 4.1.18 to the adapted càdlàg process $H^n\bullet N$. For $K \in \Lambda^1$ with $\|K\|_u \le 1$, the elementary associativity identity following from Eq. (4.2) gives

$$K\bullet(H^n\bullet N) = (KH^n)\bullet N.$$

Therefore, for every $\varepsilon > 0$,

$$\begin{aligned}
\mathbf{P}\Big((H^n\bullet N)_T^* > \varepsilon\Big) &\le \sup_{\substack{K\in\Lambda^1\\ \|K\|_u\le 1}} \mathbf{P}\Big(\,|((KH^n)\bullet N)_T| \ge \varepsilon\Big) \\
&\le \frac{1}{\varepsilon^2} \sup_{\substack{K\in\Lambda^1\\ \|K\|_u\le 1}} \mathbf{E}\Big[((KH^n)\bullet N)_T^2\Big] \\
&\le \frac{1}{\varepsilon^2} \int_{\Omega\times(0,T]} (H^n)^2 \mathrm{d}\mu_N^T \longrightarrow 0.
\end{aligned}$$

This proves the square-integrable martingale case.

Step 2: martingales. Let M be a martingale. For $r \in \mathbb{N}$, define

$$\xi^r := (-r) \vee (M_T \wedge r),$$

and set

$$M_t^r := \mathbf{E}[\xi^r \mid \mathcal{F}_t], \qquad 0 \le t \le T,$$

using a càdlàg version and extending M^r by $M^r_t := M^r_T$ for $t > T$. Then M^r is bounded and hence square-integrable. By Step 1,

$$(H^n \bullet M^r)^*_T \longrightarrow 0$$

in probability for every fixed r.

Put $R^r := M - M^r$. Then R^r is a martingale on $[0, T]$, and

$$R^r_T = M_T - \xi^r.$$

By Theorem 4.2.2 (b), for every n, r and every $\varepsilon > 0$,

$$\mathbf{P}\Big((H^n \bullet R^r)^*_T > \varepsilon\Big) \le \frac{18C}{\varepsilon}\mathbf{E}\Big[\,|R^r_T|\,\Big] = \frac{18C}{\varepsilon}\mathbf{E}\Big[\,|M_T - \xi^r|\,\Big].$$

Since $\xi^r \to M_T$ in L^1, the right-hand side tends to 0 as $r \to \infty$, uniformly in n. Hence

$$\begin{aligned}\limsup_{n\to\infty} \mathbf{P}\Big((H^n \bullet M)^*_T > 2\varepsilon\Big) &\le \limsup_{n\to\infty} \mathbf{P}\Big((H^n \bullet M^r)^*_T > \varepsilon\Big) \\ &\quad + \sup_{n\in\mathbb{N}} \mathbf{P}\Big((H^n \bullet R^r)^*_T > \varepsilon\Big) \\ &\le \frac{18C}{\varepsilon}\mathbf{E}\Big[\,|M_T - \xi^r|\,\Big].\end{aligned}$$

Letting $r \to \infty$ proves

$$(H^n \bullet M)^*_T \longrightarrow 0$$

in probability.

Step 3: local martingales. Let M be a local martingale, and let $(\tau_k)_{k\in\mathbb{N}}$ be a localizing sequence such that M^{τ_k} is a martingale for every k. On the set $\{\tau_k > T\}$, the processes $H^n \bullet M$ and $H^n \bullet M^{\tau_k}$ coincide on $[0, T]$. Therefore, for every $\varepsilon > 0$,

$$\mathbf{P}\Big((H^n \bullet M)^*_T > \varepsilon\Big) \le \mathbf{P}(\tau_k \le T) + \mathbf{P}\Big((H^n \bullet M^{\tau_k})^*_T > \varepsilon\Big).$$

For fixed k, Step 2 yields

$$(H^n \bullet M^{\tau_k})^*_T \longrightarrow 0$$

in probability. Hence

$$\limsup_{n\to\infty} \mathbf{P}\Big((H^n \bullet M)^*_T > \varepsilon\Big) \le \mathbf{P}(\tau_k \le T).$$

Letting $k \to \infty$ gives

$$(H^n \bullet M)^*_T \longrightarrow 0$$

in probability.

Step 4: classical semimartingales. Let

$$X = X_0 + M + A,$$

where $M_0 = A_0 = 0$, M is a local martingale and A is an FV-process.

The initial term satisfies $H_0^n X_0 \to 0$ in probability. By Step 3,

$$(H^n \bullet M)_T^* \longrightarrow 0$$

in probability.

It remains to handle the FV-part. Let V denote the total variation process of A. Pathwise,

$$(H^n \bullet A)_T^* \leq \int_{(0,T]} |H_s^n| \, \mathrm{d}V_s.$$

For almost every ω, the measure $\mathrm{d}V(\omega)$ is finite on $(0, T]$, and $H^n(\omega, \cdot) \to 0$ pointwise while $|H^n| \leq C$. Ordinary dominated convergence for finite measures gives

$$\int_{(0,T]} |H_s^n| \, \mathrm{d}V_s \longrightarrow 0 \qquad \text{a.s.}$$

Therefore $(H^n \bullet A)_T^* \to 0$ in probability. Combining the initial, local-martingale and FV-parts proves the claim. □

Theorem 4.2.11 (Bichteler-Dellacherie-Mokobodzki Theorem with bounded pointwise criterion)**.** *Let $X = (X^1, \ldots, X^d)$ be an adapted 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;

(iv) For every sequence $(H^n)_{n\in\mathbb{N}} \subset \Lambda^d$ such that

$$\sup_{n\in\mathbb{N}} \|H^n\|_u < \infty$$

and

$$H^n \longrightarrow 0$$

pointwise outside an evanescent set, one has

$$H^n \bullet X \longrightarrow 0$$

in the ucp topology.

Proof. The equivalence of (i), (ii), and (iii) follows from Theorem 4.1.17, Theorem 4.2.1 (a), Proposition 4.2.3, and Proposition 4.2.6. It remains to connect these conditions with (iv).

Assume first that (iv) holds. Let $(K^n)_{n\in\mathbb{N}} \subset \Lambda^d$ satisfy

$$\|K^n\|_u \to 0.$$

Then (K^n) is uniformly bounded and converges pointwise outside an evanescent set to 0. Hence, by (iv),

$$K^n \bullet X \to 0$$

in ucp, and therefore $(K^n \bullet X)_t \to 0$ in probability for every $t \geq 0$. By Theorem 4.1.17, X is a topological semimartingale. Thus (iv) implies (i).

Conversely, assume that X is a topological semimartingale. By the equivalence of (i) and (ii), each component X^i is a classical semimartingale. Let

$$H^n = (H^{n,1}, \dots, H^{n,d}) \in \Lambda^d$$

be uniformly bounded and suppose that $H^n \to 0$ pointwise outside an evanescent set. For each i, $(H^{n,i})_{n\in\mathbb{N}} \subset \Lambda^1$ is uniformly bounded and converges pointwise outside an evanescent set to 0. Since X^i is a real-valued classical semimartingale, Proposition 4.2.10 gives

$$H^{n,i} \bullet X^i \to 0$$

in ucp. As $d < \infty$,

$$H^n \bullet X = \sum_{i=1}^{d} H^{n,i} \bullet X^i \to 0$$

in ucp. Hence (i) implies (iv).

This proves the equivalence of all four statements. □

4.3 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 [57, 59] and Mémin [135]. 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 Mémin [135] that these definitions are equivalent.

Definition 4.3.1. 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^d_X 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 4.3.2. 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 4.3.3. *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 $\|cX\|_{\mathcal{S}} \le \varepsilon$. Without loss of generality, we assume $H^n \xrightarrow{\text{ucp}} 0$ and hence, we can always find a $N \in \mathbb{N}$ such that $\mathbf{P}((H^n)^*_t > c) \le \varepsilon$ for all $n \ge N$.

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 $n \ge N$ and $K \in \Lambda^1$ with $\|K\|_u \le 1$

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

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

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

□

Remark 4.3.4. The converse of these statements does not hold (see Example 4.3.5). 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, [27]).

Example 4.3.5. Let $X_t(\omega) = t$ and

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

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

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

Hence, we have $H^n \bullet X \xrightarrow{\text{ucp}} 0$ 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 $H^n \not\xrightarrow{\text{ucp}} 0$.

The following results were initially shown by Mémin [135]. In our setting, we can provide a simplified proof.

Theorem 4.3.6. *The space of semimartingales is complete under the semimartingale topology.*

Proof. Let $(X^n)_{n\in\mathbb{N}}$ be a Cauchy sequence in the semimartingale topology, i.e. $\|X^n - X^m\|_{\mathcal{S}} \to 0$ as $m, n \to \infty$.

Step 1: ucp limit. Taking $K = \mathbb{1}_{]\!]0,\infty[\![}\,\mathbf{1}$ with $\|K\|_u \leq 1$, we have $K \bullet (X^n - X^m) = X^n - X^m - X_0^n + X_0^m$, hence $\|X^n - X^m\|_{\text{ucp}} \to 0$ (after absorbing the constant initial

differences, which themselves form a Cauchy sequence in $\mathbb{R}^d$). By Theorem 4.3.3, $X^n \to X$ in ucp for some càdlàg adapted process X.

Step 2: $X^n \to X$ in the semimartingale topology. Let $K \in \Lambda^d$ with $\|K\|_u \leq 1$ be a fixed simple predictable integrand, and let $t > 0$. Since K is simple, $K \bullet (X^n - X^m)$ is a fixed linear combination of values of $X^n - X^m$, so the ucp convergence $X^n \to X$ gives $K \bullet X^n \to K \bullet X$ in ucp. Furthermore, by the definition of the $\mathcal{S}$-norm at horizon t,

$$\|K \bullet (X^n - X^m)\|_{\mathrm{ucp},t} \;\leq\; \|X^n - X^m\|_{\mathcal{S},t}\,.$$

Letting $m \to \infty$ on the left-hand side gives

$$\|K \bullet (X^n - X)\|_{\mathrm{ucp},t} \;\leq\; \liminf_{m\to\infty} \|X^n - X^m\|_{\mathcal{S},t}\,.$$

Taking the supremum over $K \in \Lambda^d$, $\|K\|_u \leq 1$,

$$\|X^n - X\|_{\mathcal{S},t} \;\leq\; \liminf_{m\to\infty} \|X^n - X^m\|_{\mathcal{S},t}\,,$$

and the right-hand side tends to 0 as $n \to \infty$ by the Cauchy property. Since $t > 0$ was arbitrary, $\|X^n - X\|_{\mathcal{S}} \to 0$, so $X^n \to X$ in the semimartingale topology.

Step 3: X is a topological semimartingale. It remains to verify $X \in \mathcal{S}^d$. The càdlàg and integrability properties Item (a),Item (b) hold by construction (cf. the discussion before Example 4.1.12). For Item (c), by Theorem 4.1.17 it suffices to show that for every $(H^k)_{k\in\mathbb{N}} \subset \Lambda^d$ with $\left\|H^k\right\|_u \to 0$ and every $t > 0$, $(H^k \bullet X)_t \to 0$ in probability.

Fix $\varepsilon > 0$ and $t > 0$. For any $n \in \mathbb{N}$ and any $k \in \mathbb{N}$, by linearity and the triangle inequality,

$$\left\|H^k \bullet X\right\|_{\mathrm{ucp},t} \;\leq\; \left\|H^k \bullet (X - X^n)\right\|_{\mathrm{ucp},t} \;+\; \left\|H^k \bullet X^n\right\|_{\mathrm{ucp},t}\,.$$

For k large enough that $\left\|H^k\right\|_u \leq 1$, the first term is bounded by $\|X - X^n\|_{\mathcal{S},t}$, which converges to 0 as $n \to \infty$ by Step 2. For each *fixed* n, X^n is a topological semimartingale, so by Theorem 4.1.17 the second term tends to 0 as $k \to \infty$. Choose n so large that $\|X - X^n\|_{\mathcal{S},t} < \varepsilon/2$, then let $k \to \infty$ to make the second term less than $\varepsilon/2$. This shows $\limsup_{k\to\infty} \left\|H^k \bullet X\right\|_{\mathrm{ucp},t} \leq \varepsilon$, and since ε was arbitrary, $\left\|H^k \bullet X\right\|_{\mathrm{ucp},t} \to 0$, hence $(H^k \bullet X)_t \to 0$ in probability. □

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 $\mathcal{S}_e^1$:

$$J_X : \quad \begin{cases} \Lambda_X^d & \to & \mathcal{S}_e^1 \\ H & \mapsto & H \bullet X \end{cases}$$

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 [99, 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 motivates the following definition.

Definition 4.3.7. 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

$$\overline{J}_X : \quad \begin{cases} L(X) & \to & \mathcal{S}_e^1 \\ H & \mapsto & H \bullet X := \overline{J}_X(H) \end{cases} .$$

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 P.-A. Meyer in the classical approach towards stochastic integration (see [144, page 97 ff] for more details). However, this is not the case, as the following theorem demonstrates.

Theorem 4.3.8. *Let $H \in \mathcal{P}$ be locally in $L(X)$. Then $H \in L(X)$. In other words:*

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

We utilize the following lemma for the proof.

Lemma 4.3.9. *Let $(X^n)_{n\in\mathbb{N}} \subset \mathcal{S}^d$, $X \in \mathcal{S}^d$, and $(T^m)_{m\in\mathbb{N}}$ an increasing sequence of stopping times with $T^m \uparrow \infty$ a.s. Assume that for every m,*

$$(X^n)^{T^m} \xrightarrow{\mathcal{S}} X^{T^m} \quad \text{in } \mathcal{S}^d \text{ as } n \to \infty.$$

Then $X^n \xrightarrow{\mathcal{S}} X$ in $\mathcal{S}^d$.

Proof. Suppose not. Then there exist $\varepsilon > 0$, a subsequence n_k, and $K^{(k)} \in \Lambda^d$ with $\|K^{(k)}\|_u \le 1$ such that

$$\|K^{(k)} \bullet (X^{n_k} - X)\|_{\mathrm{ucp}} \;\ge\; \varepsilon \quad \text{for all } k.$$

Write $Z^{(k)} := K^{(k)} \bullet (X^{n_k} - X)$. By the definition $\|Z^{(k)}\|_{\mathrm{ucp}} = \sum_{r=1}^{\infty} 2^{-r}\mathbf{E}[(Z_r^{(k)})^* \wedge 1]$. Choose N so large that $\sum_{r>N} 2^{-r} < \varepsilon/4$. Then for every k,

$$\sum_{r=1}^{N} 2^{-r}\mathbf{E}[(Z_r^{(k)})^* \wedge 1] \;\ge\; \varepsilon - \tfrac{\varepsilon}{4} = \tfrac{3\varepsilon}{4}.$$

Hence, for each k there exists an index $r_k \in \{1, \dots, N\}$ with

$$\mathbf{E}[(Z_{r_k}^{(k)})^* \wedge 1] \;\ge\; \frac{3\varepsilon}{4N}.$$

Passing to a further subsequence (not relabeled), we may assume $r_k \equiv t$ is constant. Thus

$$\mathbf{E}\big[(Z_t^{(k)})^* \wedge 1\big] \;\ge\; c_\varepsilon := \frac{3\varepsilon}{4N} \quad \text{for all } k. \tag{4.10}$$

Fix this $t \in \{1, \dots, N\}$. By Item (d), for any m,

$$\big(Z^{(k)}\big)^{T^m} = \big(K^{(k)}\mathbb{1}_{[\![0,T^m]\!]}\big) \bullet (X^{n_k} - X) = K^{(k)} \bullet \big((X^{n_k} - X)^{T^m}\big).$$

Therefore, for each fixed m,

$$\lim_{k\to\infty} \mathbf{E}\Big[\big(Z^{(k)}\big)_t^{T^m,*} \wedge 1\Big] \le \lim_{k\to\infty} \sup_{\|K\|_u \le 1} \Big\|K \bullet \big((X^{n_k} - X)^{T^m}\big)\Big\|_{\mathrm{ucp}} = 0,$$

by the assumed convergence in $\mathcal{S}^d$ of $(X^n)^{T^m}$ to X^{T^m}.

Next, note that truncation at T^m can only change the path on $\{T^m < t\}$, hence

$$\mathbf{E}\big[(Z_t^{(k)})^* \wedge 1\big] \;\le\; \mathbf{E}\Big[\big(Z^{(k)}\big)_t^{T^m,*} \wedge 1\Big] \;+\; \mathbf{P}(T^m < t).$$

Since $T^m \uparrow \infty$ a.s., choose m so large that $\mathbf{P}(T^m < t) < c_\varepsilon/2$, and then k large so that $\mathbf{E}[(Z^{(k)})_t^{T^m,*} \wedge 1] < c_\varepsilon/2$. This gives $\mathbf{E}[(Z_t^{(k)})^* \wedge 1] < c_\varepsilon$, contradicting (4.10). The

contradiction proves $X^n \xrightarrow{\mathcal{S}} X$ in $\mathcal{S}^d$. □

Proof of Theorem 4.3.8. Since $H \in L(X)_{\mathrm{loc}}$, 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 4.3.9, we have $H^{\tilde{n}} \bullet X \xrightarrow{\mathcal{S}} H \bullet X$, and thus $H \in L(X)$. □

Remark 4.3.10. 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 Mémin [135]. An alternative proof is presented in [31, 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 4.3.11. *Let $X \in \mathcal{S}^d$, then*

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

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 4.4.10). With Theorem 4.3.8, we get $\mathbf{b}\mathcal{P}_{\mathrm{loc}} \subset L(X)$. □

Remark 4.3.12. Usually, we even have $\mathbb{L} \subsetneq L(X) \subsetneq \mathcal{P}$. In Example 4.5.2, 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 4.3.13. Corollary 4.4.16 and Theorem 4.4.15 will provide a more detailed understanding of $L(X)$.

Remark 4.3.14. 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 4.1.3), 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.

4.4 The Dominated Convergence Theorem

The Dominated Convergence 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.

Our proof is based on two auxiliary seminorms on integrands, a bounded-convergence step via a monotone-class argument, and a bootstrap from bounded to dominated convergence. The overall structure mirrors the "define-close-extend" pattern of Carathéodory-type arguments, but is carried out entirely within the predictable σ-algebra and the semimartingale topology.

Theorem 4.4.1 (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 $\left|\tilde{H}^{i,m}\right| \leq G^i$ 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).*

Remark 4.4.2. The condition requiring G to be componentwise integrable is crucial, as demonstrated in Example 4.7.11. 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 4.7.9.

The proof of the Dominated Convergence Theorem is self-contained and does not use any of the properties proved later in Section 4.5. The argument proceeds by a bootstrap in four transparent steps:

(A) *Auxiliary seminorms on integrands.* We introduce two auxiliary seminorms on simple predictable, respectively bounded predictable, integrands (parts (a)-(b) of Definition 4.4.3):

$$\|H\|_{\Lambda,X} := \sup\{\|G \bullet X\|_{\mathrm{ucp}}\,;\, G \in \Lambda^1,\ |G| \leq |H|\},$$
$$\|H\|_{\mathbf{b}\mathcal{P},X} := \inf\left\{\sum_n \|G^n\|_{\Lambda,X}\,;\, |H| \leq \sum_n |G^n|\right\}.$$

They are designed so that convergence in these seminorms implies ucp convergence of the resulting stochastic integrals.

(B) *Bounded Convergence.* Using these seminorms plus a monotone-class argument, we prove a bounded convergence theorem for bounded predictable integrands (Theorem 4.4.10), without invoking any of the later properties in Section 4.5.

(C) *Dominated Convergence on $L^1(X)$.* We define $L^1(X)$ (Definition 4.4.11), prove dominated convergence on $L^1(X)$ (Theorem 4.4.13), and use only: (i) associativity for simple integrands (Theorem 4.4.12), and (ii) the continuity of J_X in the X-seminorm.

(D) *Identification $L^1(X) = L(X)$.* Finally, we show $L^1(X) = L(X)$ (Theorem 4.4.15) by a telescoping/series argument and truncation; hence dominated convergence holds for all integrands in $L(X)$.

As the Dominated Convergence Theorem requires componentwise integrability, we can, without loss of generality, assume for its proof that $d = 1$.

Definition 4.4.3. Let X be a one-dimensional semimartingale.

(a) For $H \in \Lambda^1$, we define

$$\|H\|_{\Lambda,X} = \sup\{\|G \bullet X\|_{\mathrm{ucp}}\,;\, G \in \Lambda^1,\ |G| \le |H|\}.$$

(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 4.4.4. We will see later in the proof of the bounded convergence theorem that $\mathbb{A}$ coincides with $\mathbf{b}\mathcal{P}$.

Lemma 4.4.5. *Let X be a one-dimensional semimartingale.*

(i) If $H \in \Lambda^1$ and $G \in \Lambda^1$ satisfy $|G| \le |H|$, then $\|G\|_{\Lambda,X} \le \|H\|_{\Lambda,X}$. In particular, $|H_1| \le |H_2|$ implies $\|H_1\|_{\Lambda,X} \le \|H_2\|_{\Lambda,X}$.

(ii) For all $H_1, H_2 \in \Lambda^1$,

$$\|H_1 + H_2\|_{\Lambda,X} \le \|H_1\|_{\Lambda,X} + \|H_2\|_{\Lambda,X}.$$

(iii) Consequently, after identifying indistinguishable processes, $\|\cdot\|_{\Lambda,X}$ is a seminorm on Λ^1, and $\|\cdot\|_{\mathbf{b}\mathcal{P},X}$ is a metric on $\mathcal{P}$.

Proof. (i) is immediate from the definition of the supremum in $\|\cdot\|_{\Lambda,X}$.

(ii) Fix $G \in \Lambda^1$ with $|G| \le |H_1 + H_2|$. Set

$$\alpha := \frac{|H_1|}{|H_1| + |H_2|} \mathbb{1}_{\{|H_1|+|H_2|>0\}}, \qquad G_1 := \alpha G, \quad G_2 := (1-\alpha)G.$$

On the common refinement of the simple partitions, α is predictable and bounded by 1, hence $G_1, G_2 \in \Lambda^1$. Moreover $|G_1| \le |H_1|$ and $|G_2| \le |H_2|$, and $G = G_1 + G_2$. Therefore

$$\|G \bullet X\|_{\mathrm{ucp}} \le \|G_1 \bullet X\|_{\mathrm{ucp}} + \|G_2 \bullet X\|_{\mathrm{ucp}} \le \|H_1\|_{\Lambda,X} + \|H_2\|_{\Lambda,X}.$$

Taking the supremum over all G with $|G| \le |H_1 + H_2|$ yields the claim.

(iii) follows from (i)–(ii) and the definitions. □

Lemma 4.4.6. *Let $X \in \mathcal{S}^1$, and let $(H^n)_{n\in\mathbb{N}} \subset \Lambda^1$ be such that*

$$H^n \le H^{n+1} \le K \qquad \text{for all } n, \text{ for some deterministic constant } K.$$

Then, for every $\varepsilon > 0$ there exists n_0 with

$$\|H^m - H^n\|_{\Lambda,X} \le \varepsilon \qquad \text{for all } m, n \ge n_0.$$

Proof. Let

$$H := \lim_{n\to\infty} H^n.$$

Since (H^n) is increasing and bounded above by the deterministic constant K, the limit exists pointwise and is finite. Since $H^1 \in \Lambda^1$ is bounded, there is a deterministic constant

$$C := |K| + \|H^1\|_u$$

such that, for all $m > n$,

$$0 \le H^m - H^n \le H - H^n \le K - H^1 \le C.$$

For $r \in \mathbb{N}$ and $F \in \Lambda^1$, put

$$q_r(F) := \sup_{\substack{G\in\Lambda^1 \\ |G|\le|F|}} \mathbf{E}\left[((G\bullet X)^*_r) \wedge 1\right].$$

We first show that, for every fixed $r \in \mathbb{N}$,

$$q_r(H^m - H^n) \longrightarrow 0 \qquad \text{as } m, n \to \infty.$$

By symmetry, it is enough to consider $m > n$. Suppose the claim is false. Then there exist $r \in \mathbb{N}$, $\varepsilon > 0$, and sequences $m_j > n_j \geq j$ such that

$$q_r(H^{m_j} - H^{n_j}) > \varepsilon \qquad \text{for all } j.$$

By the definition of q_r, we may choose $G^j \in \Lambda^1$ with

$$|G^j| \leq H^{m_j} - H^{n_j}$$

and

$$\mathbf{E}\left[\left((G^j \bullet X)_r^*\right) \wedge 1\right] \geq \frac{\varepsilon}{2} \qquad \text{for all } j.$$

But

$$|G^j| \leq H^{m_j} - H^{n_j} \leq H - H^{n_j} \longrightarrow 0$$

pointwise outside an evanescent set, because $n_j \to \infty$. Moreover,

$$|G^j| \leq C \qquad \text{for all } j.$$

Thus $(G^j)_{j\in\mathbb{N}} \subset \Lambda^1$ is uniformly bounded and converges pointwise outside an evanescent set to 0. Since $X \in \mathcal{S}^1$, Theorem 4.2.11 (iv) gives

$$G^j \bullet X \longrightarrow 0$$

in ucp. In particular,

$$\mathbf{E}\left[\left((G^j \bullet X)_r^*\right) \wedge 1\right] \longrightarrow 0,$$

contradicting the lower bound above. Hence

$$q_r(H^m - H^n) \longrightarrow 0 \qquad \text{for every fixed } r \in \mathbb{N}.$$

It remains to pass from fixed horizons to the full seminorm. By definition,

$$\begin{aligned} \|H^m - H^n\|_{\Lambda,X} &= \sup_{\substack{G\in\Lambda^1 \\ |G|\leq|H^m-H^n|}} \sum_{r=1}^{\infty} 2^{-r}\mathbf{E}\left[((G\bullet X)_r^*) \wedge 1\right] \\ &\leq \sum_{r=1}^{\infty} 2^{-r} q_r(H^m - H^n). \end{aligned}$$

Let $\delta > 0$. Choose $R \in \mathbb{N}$ such that

$$\sum_{r>R} 2^{-r} < \frac{\delta}{2}.$$

For each $r = 1, \dots, R$, the fixed-horizon convergence just proved gives $n_r \in \mathbb{N}$ such that

$$q_r(H^m - H^n) < \frac{\delta}{2R} \qquad \text{for all } m, n \geq n_r.$$

Set

$$n_0 := \max_{1 \leq r \leq R} n_r.$$

Then, for all $m, n \geq n_0$,

$$\begin{aligned} \|H^m - H^n\|_{\Lambda,X} &\leq \sum_{r=1}^{R} 2^{-r} \frac{\delta}{2R} + \sum_{r>R} 2^{-r} \\ &< \frac{\delta}{2} + \frac{\delta}{2} = \delta. \end{aligned}$$

Therefore $(H^n)_{n\in\mathbb{N}}$ is Cauchy in $\|\cdot\|_{\Lambda,X}$. □

Lemma 4.4.7. *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

$$\left\| \sum_{j=1}^{\infty} H^j \right\|_{\mathbf{b}\mathcal{P},X} \leq \sum_{j=1}^{\infty} \left\| H^j \right\|_{\mathbf{b}\mathcal{P},X}.$$

Proof. We have

$$\left\| \sum_{j=1}^{n} H^j \right\|_{\mathbf{b}\mathcal{P},X} \leq \sum_{j=1}^{n} \left\| H^j \right\|_{\mathbf{b}\mathcal{P},X} \leq \sum_{j=1}^{\infty} \left\| H^j \right\|_{\mathbf{b}\mathcal{P},X}.$$

The result follows from the continuity of the metric. □

Lemma 4.4.8. *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 4.4.9. *Let $H \in \mathbf{b}\mathcal{P}$ and $(H^n)_{n\in\mathbb{N}} \in \mathbb{A}$ with $|H^n| \leq |H^{n+1}| \leq |H|$ for all $n \in \mathbb{N}$ 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 $\|G^{n,m} - H^n\|_{\mathbf{b}\mathcal{P},X} < \varepsilon$ for all $m \geq m_n(\varepsilon)$.

We put

$$\overline{m}_n(\varepsilon) := \max\left\{\overline{m}_{n-1}\left(\frac{\varepsilon}{3}\right) + 1, m_n\left(\frac{\varepsilon}{3}\right)\right\} \quad \text{and} \quad \tilde{G}^n := G^{n,\overline{m}_n(\varepsilon)}.$$

As $\tilde{G}^m - \tilde{G}^n \in \Lambda^1$, we have

$$\left\|\tilde{G}^m - \tilde{G}^n\right\|_{\mathbf{b}\mathcal{P},X} \leq \left\|\tilde{G}^m - \tilde{G}^n\right\|_{\Lambda,X}$$

and by Lemma 4.4.6, there exists an $\tilde{n}(\varepsilon) \in \mathbb{N}$, such that $\left\|\tilde{G}^m - \tilde{G}^n\right\|_{\Lambda,X} < \frac{\varepsilon}{3}$ for all $n, m \geq \tilde{n}(\varepsilon)$.

We obtain for $m, n \geq \max\{\tilde{n}(\varepsilon), \overline{m}_n(\varepsilon)\}$

$$\begin{aligned}
&\|H^m - H^n\|_{\mathbf{b}\mathcal{P},X} \\
&= \left\|H^m - \tilde{G}^m + (\tilde{G}^n - H^n) + (\tilde{G}^m - \tilde{G}^n)\right\|_{\mathbf{b}\mathcal{P},X} \\
&\leq \underbrace{\left\|H^m - \tilde{G}^m\right\|_{\mathbf{b}\mathcal{P},X}}_{\leq\frac{\varepsilon}{3}} + \underbrace{\left\|\tilde{G}^n - H^n\right\|_{\mathbf{b}\mathcal{P},X}}_{\leq\frac{\varepsilon}{3}} + \underbrace{\left\|\tilde{G}^m - \tilde{G}^n\right\|_{\mathbf{b}\mathcal{P},X}}_{\leq\frac{\varepsilon}{3}}.
\end{aligned}$$

Hence, for each ε, there exists an $n(\varepsilon) \in \mathbb{N}$, such that $\|H^m - H^n\|_{\mathbf{b}\mathcal{P},X} \leq \varepsilon$ for all $m, n \geq n(\varepsilon)$.

Now we put $n^k = n\left(2^{-k}\right)$ and obtain a sequence $(H^{n^k})_{k\in\mathbb{N}}$ such that

$$\left\|H^{n^{k+1}} - H^{n^k}\right\|_{\mathbf{b}\mathcal{P},X} \leq 2^{-k}.$$

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 4.4.7

$$\left\|H - H^{n^k}\right\|_{\mathbf{b}\mathcal{P},X} \leq \sum_{m=k}^{\infty} \left\|H^{n^{m+1}} - H^{n^m}\right\|_{\mathbf{b}\mathcal{P},X} \leq 2^{-k+1}.$$

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 4.4.8, this proves the result and implies $H \in \mathbb{A}$. □

Theorem 4.4.10 (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 $|H^n| \leq K$ 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 $(G^n - G)^* \leq a_n$ and $a_n \to 0$. Now we have

$$\|G^n - G\|_{\mathbf{b}\mathcal{P},X} \leq \|(G^n - G)^*\|_{\mathbf{b}\mathcal{P},X} \leq \|a_n\|_{\mathbf{b}\mathcal{P},X} = |a_n| \, \|X\|_{\mathrm{ucp}} \to 0.$$

By Lemma 4.4.8, 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 4.4.9, 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 $\tilde{G}^n := \sup_{m\geq n} |H^m - H|$ and $\hat{G}^n := \tilde{G}^1 - \tilde{G}^n$ and have $\left|\hat{G}^n\right| \leq \left|\tilde{G}^1\right|$ and $\hat{G}^n \to \tilde{G}^1$ pointwise almost surely. By Lemma 4.4.9, we obtain

$$\|H^n - H\|_{\mathbf{b}\mathcal{P},X} \leq \left\|\tilde{G}^n\right\|_{\mathbf{b}\mathcal{P},X} = \left\|\hat{G}^n - \tilde{G}^1\right\|_{\mathbf{b}\mathcal{P},X} \to 0.$$

Hence, we deduce $\|H^n - H\|_{\mathbf{b}\mathcal{P},X} \to 0$.

Furthermore, it is easy to check that we have

$$\|H^n - H\|_X \leq \|H^n - H\|_{\mathbf{b}\mathcal{P},X}.$$

Thus, $\|H^n - H\|_X \to 0$, which implies $H^n \xrightarrow{X} H$. □

Definition 4.4.11. 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 $|H^n| \leq |G|$ and $H^n \to 0$ pointwise a.s., we have $\|H^n\|_X \to 0$.

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 4.4.18 and Theorem 4.5.1.

Theorem 4.4.12. *Let X be a one-dimensional semimartingale, $H \in \Lambda^1$ and $K \in L(X)$. Then*

$$H\bullet(K\bullet X) = (HK)\bullet X.$$

Proof. This is a simple calculation for $K \in \Lambda^1$ (see, Theorem 4.5.1). 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 $K^n \xrightarrow{X} K$. □

Theorem 4.4.13 (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 a.s. to H with $|H^n| \le |G|$ for some $G \in L^1(X)$. Then $\|H^n - H\|_X \to 0$.*

Proof. Set $Y^n := (H^n - H)\bullet X$. Recall that

$$\|Y^n\|_{\mathcal{S}} = \sup_{\substack{K\in\Lambda^1\\ \|K\|_u\le 1}} \|K\bullet Y^n\|_{\text{ucp}} = \sup_{\substack{K\in\Lambda^1\\ \|K\|_u\le 1}} \|(K(H^n - H))\bullet X\|_{\text{ucp}}$$

where the second equality follows from Theorem 4.4.12, and $\|H^n - H\|_X = \|Y^n\|_{\mathcal{S}}$ by definition.

Let $(K^n)_{n\in\mathbb{N}} \subset \Lambda^1$ with $\|K^n\|_u \le 1$. Then $K^n(H^n - H) \to 0$ pointwise a.s. and $|K^n(H^n - H)| \le 2\,|G|$. Since $G \in L^1(X)$ implies $2G \in L^1(X)$, Theorem 4.4.10 yields

$$\|(K^n(H^n - H))\|_X \;\to\; 0.$$

Because $\|Z\|_{\text{ucp}} \le \|Z\|_{\mathcal{S}}$ (take the outer integrand $J \equiv 1$ in the definition of $\|\cdot\|_{\mathcal{S}}$), we get by Theorem 4.4.12

$$\|K^n\bullet Y^n\|_{\text{ucp}} = \|(K^n(H^n - H))\bullet X\|_{\text{ucp}} \;\longrightarrow\; 0. \qquad (*)$$

Now suppose $\|Y^n\|_{\mathcal{S}} \not\to 0$. Then there exist $\varepsilon > 0$, a subsequence n_k, and $K^{(k)} \in \Lambda^1$ with $\|K^{(k)}\|_u \le 1$ such that

$$\|K^{(k)}\bullet Y^{n_k}\|_{\text{ucp}} \;\ge\; \varepsilon \quad \text{for all } k.$$

Define the sequence $\widetilde{K}^n$ by $\widetilde{K}^{n_k} := K^{(k)}$ and choosing $\widetilde{K}^n$ arbitrarily (with $\|\widetilde{K}^n\|_u \leq 1$) for other n. Then $(*)$ applied to the sequence $(\widetilde{K}^n)_n$ gives $\|\widetilde{K}^n \bullet Y^n\|_{\mathrm{ucp}} \to 0$, contradicting the choice of $(n_k, K^{(k)})$. Hence $\|Y^n\|_{\mathcal{S}} \to 0$.

Since $\|H^n - H\|_X = \|Y^n\|_{\mathcal{S}}$, we conclude $\|H^n - H\|_X \to 0$. □

Lemma 4.4.14. *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 $|H^n| \leq |G|$.

(a) Let $(\lambda^n)_{n\in\mathbb{N}}$ be a sequence of positive numbers such that $|\lambda^n H^n| \leq |G|$. 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 $\lambda\,|H| \leq |G|$, $\lambda(H \wedge K^n) \vee (-K^n)$ converges pointwise almost surely to λH and we have $\lambda(H \wedge K^n) \vee (-K^n) \leq |\lambda H|$.

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

$$\lambda(H \wedge K^n) \vee (-K^n) \leq \sum_{k=1}^{n} |\lambda H| \wedge \left|G^k\right|$$

and thus

$$\begin{aligned}
\|\lambda H\|_X &= \lim_{n\to\infty} \|\lambda(H \wedge K^n) \vee (-K^n)\|_X \\
&\leq \lim_{n\to\infty} \sum_{k=1}^{n} \left\| |\lambda H| \wedge \left|G^k\right| \right\|_X \\
&= \sum_{n=1}^{\infty} \| |\lambda H| \wedge |G^n| \|_X\,,
\end{aligned}$$

where we used the triangle inequality of the metric $\|\cdot\|_X$.

Hence, we obtain

$$\sup_{m\in\mathbb{N}} \|\lambda^m H^m\|_X \leq \sup_{m\in\mathbb{N}} \sum_{n=1}^{\infty} \| |\lambda^m H^m| \wedge G^n \|_X \leq \sum_{n=1}^{\infty} \|G^n\|_X < \infty.$$

That proves Item (a). For Item (b), we proceed analogously and obtain, by putting $\lambda^m = 1$,

$$\lim_{m\to\infty} \|H^m\|_X \leq \lim_{m\to\infty} \sum_{n=1}^{\infty} \| |H^m| \wedge G^n \|_X ,$$

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

$$\lim_{m\to\infty} \|H^m\|_X \leq \sum_{n=1}^{\infty} \lim_{m\to\infty} \| |H^m| \wedge G^n \|_X .$$

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 Item (c) we use Item (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 $\|\mathbb{1}_A H\|_X = 0$. □

Theorem 4.4.15. *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 4.4.14, $\mathbb{1}_{A^c} G \in L^1(X)$. 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 $\tilde{H} \in L^1(X)$ and by dominated convergence, we obtain $\left\|\mathbb{1}_{A^c} H^n - \tilde{H}\right\|_X \to 0$. By Lemma 4.4.14, 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 $H \in L^1(X)$. □

The Dominated Convergence Theorem has a useful corollary that can help comprehend the space of integrands more thoroughly.

Corollary 4.4.16. *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 $H^n \xrightarrow{X} H$.*

Proof. Without loss of generality, take the localizing sequence $(T^n)_{n\in\mathbb{N}}$ to be increasing with $T^n \uparrow \infty$ a.s. . Set $T^0 := 0$ and define the predictable (vector) process $G = (\tilde{G}^1, \ldots, \tilde{G}^d)$ by

$$\tilde{G}^i := \sum_{n=1}^{\infty} \widetilde{K}_n \mathbb{1}_{\rrbracket T^{n-1}, T^n \rrbracket}, \qquad \widetilde{K}_n := \max_{j\leq n} K_j.$$

Then on each slice $\rrbracket T^{n-1}, T^n \rrbracket$ we have $|H^m_t| \leq K_n \leq \widetilde{K}_n$, hence $|H^m| \leq G$ componentwise for all m and $|H| \leq G$ as well. Moreover, G is locally bounded (indeed, $G^{T^n} \leq \widetilde{K}_n$), so by Theorem 4.3.11 we have $\tilde{G}^i \in L(X)$ for every component i.

Since $H^n_t \to H_t$ a.s. for each t and $|H^n| \leq G$ componentwise with $G \in L(X)$ (componentwise), the Dominated Convergence Theorem for the topological stochastic integral (Theorem 4.4.1) yields

$$H \in L(X) \qquad \text{and} \qquad H^n \xrightarrow{X} H.$$

□

As there is no version of the Dominated Convergence Theorem that guarantees the integrability of a process that is not componentwise integrable, the following Theorem 4.4.19 is very useful. The proof requires some preparation.

Lemma 4.4.17. *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 4.1.17. □

For the proof of the important Theorem 4.4.19, we need to show the associativity of the stochastic integral first.

Theorem 4.4.18. *Let $X = (X^1, \dots, X^d)^\top \in \mathcal{S}^d$ and $H = (H^1, \dots, H^d)^\top$ with $H^i \in L(X^i)$ for all i. Furthermore, let $K = (K^1, \dots, K^d)^\top$ be a predictable d-dimensional process and we set $Y^i := H^i \bullet X^i$, $Y := (Y^1, \dots, Y^d)^\top$, $G^i := K^i H^i$ and $G = (G^1, \dots, G^d)^\top$. 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 $i \in \{1, \dots, d\}$. By Remark 4.1.3, there exists a $\tilde{K}^n = (\tilde{K}^{1,n}, \dots, \tilde{K}^{d,n})^\top \in \Lambda^d$, such that $\left|\tilde{K}^{i,n}\right| \leq |K^i|$ and $\tilde{K}^{i,n}_t \to K^i_t$ for all t almost surely. Now, by the first part and by the Dominated Convergence Theorem 4.4.1, 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 $i \in \{1, \dots, d\}$, 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

$$\begin{aligned} &\|K \bullet Y - G \bullet X\|_{\mathcal{S}} \\ &\quad \leq \left\|K \bullet Y - \tilde{K}^n \bullet Y\right\|_{\mathcal{S}} + \left\|(\tilde{K}^n) \bullet Y - \tilde{G}^n \bullet X\right\|_{\mathcal{S}} + \left\|\tilde{G}^n \bullet X - G \bullet X\right\|_{\mathcal{S}} \end{aligned}$$

For the middle term, we can apply bounded convergence and see that it vanishes. The first and last tend to 0 (Lemma 4.4.17). □

Theorem 4.4.19. *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 $H^n = H\mathbb{1}_{\{\|H\| \leq n\}}$ converges in the semimartingale topology.*

Proof. The one direction has already been shown in Lemma 4.4.17. For the other direction, assume $H^n \bullet X$ converges in the semimartingale topology, and let $Z := \lim_{n\to\infty} H^n \bullet X$ denote its limit. Following [178], Remark 4.3.10 (completeness of the space of stochastic integrals) 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 4.4.18 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 4.4.18 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 4.4.18 states that G^{-1} also belongs to $L((GK) \bullet X)$, which is identical to $L((GH) \bullet X)$. A subsequent application of Theorem 4.4.18 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. □

4.5 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 4.5.1. *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

$$(\alpha H + \beta J) \bullet X = \alpha H \bullet X + \beta J \bullet X. \tag{4.11}$$

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 $(\Delta(H \bullet X)_s)_{s \geq 0}$ is indistinguishable from $(H_s^\top(\Delta X_s))_{s \geq 0}$.

(d) We have

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

(e) Assume $X = (X^1, \dots, X^d)^\top \in \mathcal{S}^d$ and $H = (H^1, \dots, H^d)^\top$ with $H^i \in L(X^i)$ for all i. Furthermore, let $K = (K^1, \dots, K^d)^\top$ be a predictable d-dimensional process and we set $Y^i := H^i \bullet X^i$, $Y := (Y^1, \dots, Y^d)^\top$, $G^i := K^i H^i$ and $G = (G^1, \dots, G^d)^\top$. 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 Item (a) holds. It is clear that $\Lambda^d \subset \mathbb{A}$. With the help of the bounded convergence theorem 4.4.10, we can show $\mathbf{b}\mathcal{P} \subset \mathbb{A}$ and an application of Theorem 4.4.19 yields the result.

(b) Analogous to Item (a).

(c) Let $\mathbb{A}$ be the set of processes for which Item (c) 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 4.1.1.

Hence we have

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

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

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

and thus

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

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

Now let $(H^n)_{n\in\mathbb{N}}$ be a uniformly bounded, increasing sequence in Λ^d converging pointwise to a predictable process H. By the Bounded Convergence Theorem 4.4.10, $H^n\bullet X \to H\bullet X$ in ucp. Since ucp convergence implies that some subsequence (n_k) converges almost surely uniformly on compacts, there is a measurable set Ω' of full probability such that on Ω' and for every fixed $t > 0$,

$$\sup_{s\in[0,t]} \left| (H^{n_k}\bullet X)_s - (H\bullet X)_s \right| \;\to\; 0 \qquad (k \to \infty).$$

Uniform convergence on $[0, t]$ preserves left limits, so on Ω' and for every $s \in [0, t]$,

$$(H^{n_k}\bullet X)_{s-} \;\to\; (H\bullet X)_{s-}, \qquad \Delta(H^{n_k}\bullet X)_s \;\to\; \Delta(H\bullet X)_s.$$

On the integrand side, $H^n \to H$ pointwise on $\Omega \times \mathbb{R}_+$, so $(H^{n_k})_s^\top \Delta X_s \to H_s^\top \Delta X_s$ pointwise. Combined with the simple-predictable identity $\Delta(H^n \bullet X)_s = (H^n)_s^\top \Delta X_s$ already established (and which holds indistinguishably), we obtain on Ω', for every s,

$$\Delta(H \bullet X)_s \;=\; H_s^\top \Delta X_s.$$

This is an identity of càdlàg-valued (in the case of $H \bullet X$) and jump-valued processes that holds simultaneously for all $s \in [0,t]$ on a set of full measure; since $t > 0$ was arbitrary, the identity holds indistinguishably for all $s \geq 0$. The monotone class theorem yields $\mathbf{b}\mathcal{P} \subset \mathbb{A}$, and the case $H \in L(X)$ follows by truncation as in part (a).

(d) 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 proceed as before with the monotone class theorem and then with Theorem 4.4.19.

(e) See Theorem 4.4.18.

(f) 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.

(g) 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.

□

Example 4.5.2. Let $X_t := t$ and define

$$H_t := \frac{1}{t}\,\mathbb{1}_{(0,\infty)}(t) \qquad (\text{so } H_0 := 0).$$

Then X is a semimartingale of finite variation and H is predictable. If $H \in L(X)$, then by Item (f) the stochastic integral coincides pathwise with the Lebesgue-Stieltjes integral:

$$(H \bullet X)_t = \int_{(0,t]} \frac{1}{s}\,\mathrm{d}s = +\infty \quad \text{for every } t > 0,$$

which is impossible for a càdlàg process. Hence $H \notin L(X)$.

Remark 4.5.3. For the same integrator $X_t = t$, the bounded predictable integrand $H \equiv 1$ *is* in $L(X)$, and

$$(1 \bullet X)_t = \int_0^t 1 \, \mathrm{d}s = t$$

by Item (f). Thus Example 4.5.2 does not say that "$\int 1 \, \mathrm{d}s$ does not exist"; it shows that not *every* predictable process is X-integrable. The obstruction for $H_t = 1/t$ is the divergence of the pathwise Lebesgue-Stieltjes integral at 0.

Example 4.5.4. 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 4.7.9 demonstrates that an integrand that is integrable does not necessarily have 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 4.5.5. *Let X be a d-dimensional semimartingale and*

$$H = (H^1, \dots, H^d)^\top$$

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. □

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 [31, Example 12.5.1] or [178, Example 6.4].

Example 4.5.6. 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.

4.6 Stability of Local Martingales Under Stochastic Integration

The following elementary closure criterion will be used below. It is tailored to approximation arguments for stochastic integrals: instead of trying to dominate the whole running supremum of the approximating martingales, we localize that supremum and only require a locally integrable bound for the jumps.

Lemma 4.6.1. *Let $(Y^n)_{n\in\mathbb{N}}$ be a sequence of real-valued local martingales such that*

$$Y^n \xrightarrow{ucp} Y.$$

Assume that the process

$$J_t := \sup_{n\in\mathbb{N}} \sup_{0\le s\le t} |\Delta Y^n_s|, \qquad t \ge 0,$$

is locally integrable. Then Y is a local martingale.

Proof. After replacing Y by a càdlàg modification if necessary and passing to a subsequence, we may assume that

$$\sup_{0\le s\le t} |Y^n_s - Y_s| \longrightarrow 0 \qquad \text{a.s. for every } t < \infty.$$

This does not change the limit process and is sufficient for proving the martingale property.

Set

$$Z_t := \sup_{n\in\mathbb{N}} (Y^n)^*_t, \qquad (Y^n)^*_t := \sup_{0\le s\le t} |Y^n_s|.$$

Then $Z_t < \infty$ a.s. for every $t < \infty$, because the convergence is uniform on compact time intervals and Y is càdlàg. Choose stopping times $(R_k)_{k\in\mathbb{N}}$ with $R_k \uparrow \infty$ such that

$$J_{t\wedge R_k} \in L^1 \qquad \text{for every } t < \infty.$$

For $m \in \mathbb{N}$, define

$$S_m := \inf\{t \ge 0 : Z_t > m\}, \qquad T_m := R_m \wedge S_m.$$

Since Z is adapted and increasing, S_m is a stopping time; indeed $\{S_m < t\} = \bigcup_{q\in\mathbb{Q}_+,\, q<t} \{Z_q > m\} \in \mathcal{F}_t$. Moreover, $S_m \uparrow \infty$ on compact time intervals, hence $T_m \uparrow \infty$.

Fix m and $t < \infty$. For every n,

$$\left((Y^n)^{T_m}\right)^*_t \le m + J_{t\wedge R_m}.$$

Indeed, before S_m all processes Y^n are bounded by m, and at S_m the possible overshoot is bounded by the jump size. The right-hand side is integrable by the choice of R_m. Thus $(Y^n)^{T_m}$ is a true martingale on every finite horizon.

Let $0 \le s \le t < \infty$ and $A \in \mathcal{F}_s$. Since

$$Y^n_{t\wedge T_m} \to Y_{t\wedge T_m}, \qquad Y^n_{s\wedge T_m} \to Y_{s\wedge T_m} \qquad \text{a.s.},$$

and since all variables are dominated by $m + J_{t\wedge R_m}$, dominated convergence yields

$$\begin{aligned}\mathbf{E}\Big[\mathbb{1}_A\Big(Y_{t\wedge T_m} - Y_{s\wedge T_m}\Big)\Big] &= \lim_{n\to\infty} \mathbf{E}\Big[\mathbb{1}_A\Big(Y^n_{t\wedge T_m} - Y^n_{s\wedge T_m}\Big)\Big] \\ &= 0.\end{aligned}$$

Hence Y^{T_m} is a martingale on every finite horizon. Since $T_m \uparrow \infty$, the process Y is a local martingale. □

Theorem 4.6.2. *Let X be an $\mathbb{R}^d$-valued local martingale and let $H \in \mathcal{P}$ be $\mathbb{R}^d$-valued and locally bounded. Then $H \bullet X$ is a local martingale.*

If X is locally square-integrable, then $H \bullet X$ is locally square-integrable.

Proof. By Theorem 4.5.5,

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

Thus it suffices to prove the assertion in the one-dimensional case.

Let first X be a real-valued local martingale and let H be real-valued, predictable and locally bounded. By Theorem 2.2.9, choose stopping times $(\tau_n)_{n\in\mathbb{N}}$ with $\tau_n \uparrow \infty$ such that

$$X^{\tau_n} \in \mathcal{H}^1.$$

Since H is locally bounded, choose stopping times $(\sigma_n)_{n\in\mathbb{N}}$ with $\sigma_n \uparrow \infty$ such that

$$|H_t| \le n \qquad \text{on } \{t \le \sigma_n\}.$$

Put

$$\rho_n := \tau_n \wedge \sigma_n \wedge n.$$

Then $\rho_n \uparrow \infty$. It is enough to show that $(H \bullet X)^{\rho_n}$ is a local martingale for every n.

Fix n. Set

$$M := X^{\rho_n}, \qquad K := H \mathbb{1}_{[\![0,\rho_n]\!]}.$$

Then $M \in \mathcal{H}^1$, and K is bounded by n. By Remark 4.1.3, followed by truncation at level n, choose $K^m \in \Lambda$ such that

$$|K^m| \leq n, \qquad K^m \to K \quad \text{pointwise.}$$

By the stochastic dominated convergence theorem, in the form of Corollary 4.4.16,

$$K^m \bullet M \xrightarrow{\text{ucp}} K \bullet M.$$

For every m, $K^m \bullet M$ is a local martingale by Lemma 2.2.3. Moreover, by Item (c),

$$\Delta(K^m \bullet M) = K^m \Delta M,$$

and therefore

$$\sup_m \sup_{0 \leq s \leq t} |\Delta(K^m \bullet M)_s| \leq n \sup_{0 \leq s \leq t} |\Delta M_s| \leq 2n M_t^*.$$

Since $M \in \mathcal{H}^1$, the process M^* is integrable. Hence Lemma 4.6.1 implies that $K \bullet M$ is a local martingale. By the stopping property of the stochastic integral,

$$K \bullet M = \left(H \mathbb{1}_{[\![0,\rho_n]\!]}\right) \bullet X^{\rho_n} = (H \bullet X)^{\rho_n}.$$

Thus $(H \bullet X)^{\rho_n}$ is a local martingale for every n, and since $\rho_n \uparrow \infty$, the process $H \bullet X$ is a local martingale.

Now suppose that X is locally square-integrable. Choose a localizing sequence $(\tau_n)_{n \in \mathbb{N}}$ such that X^{τ_n} is a square-integrable martingale for every n, and define $\rho_n = \tau_n \wedge \sigma_n \wedge n$ as before. Fix n, and set again

$$M := X^{\rho_n}, \qquad K := H \mathbb{1}_{[\![0,\rho_n]\!]}.$$

Then M is a square-integrable martingale and $|K| \leq n$. Choose $K^m \in \Lambda$ with $|K^m| \leq n$, $K^m \to K$ pointwise, and

$$K^m \bullet M \xrightarrow{\text{ucp}} K \bullet M.$$

For every m, the process $K^m \bullet M$ is a square-integrable martingale. Indeed, it is a martingale

by Lemma 2.2.3, and by Theorem 4.2.2 (a), for every $t \geq 0$,

$$\mathbf{E}\left[\left((K^m \bullet M)_t\right)^2\right] \leq n^2 \mathbf{E}\left[M_t^2\right] < \infty.$$

Since $(K^m \bullet M)_t \to (K \bullet M)_t$ in probability for every fixed t, passing to an a.s. convergent subsequence and applying Fatou's lemma gives

$$\mathbf{E}\left[\left((K \bullet M)_t\right)^2\right] \leq \liminf_{m\to\infty} \mathbf{E}\left[\left((K^m \bullet M)_t\right)^2\right] < \infty.$$

The same uniform L^2-bound implies that $\{(K^m \bullet M)_t : m \in \mathbb{N}\}$ is uniformly integrable for each fixed t. Hence the convergence holds in L^1 at each fixed time. Therefore, for $0 \leq s \leq t$ and $A \in \mathcal{F}_s$,

$$\begin{aligned}\mathbf{E}\left[\mathbb{1}_A\Big((K \bullet M)_t - (K \bullet M)_s\Big)\right] &= \lim_{m\to\infty} \mathbf{E}\left[\mathbb{1}_A\Big((K^m \bullet M)_t - (K^m \bullet M)_s\Big)\right] \\ &= 0.\end{aligned}$$

Thus $K \bullet M = (H \bullet X)^{\rho_n}$ is a square-integrable martingale for every n. Since $\rho_n \uparrow \infty$, this means precisely that $H \bullet X$ is locally square-integrable. □

Remark 4.6.3. 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 4.6.4. For a semimartingale $X = M + A$ with $M \in \mathcal{M}_{\text{loc}}$ and an FV-process A and $H \in L(M)$, one cannot assume that $H \bullet M \in \mathcal{M}_{\text{loc}}$. However, there is always a decomposition $X = \overline{M} + \overline{A}$ with $\overline{M} \in \mathcal{M}_{\text{loc}}$ and an FV-process A such that $H \bullet \overline{M} \in \mathcal{M}_{\text{loc}}$. This is a direct consequence of the Fundamental Theorem of Local Martingales and Theorem 4.6.2. For details, we refer to Shiryaev and Cherny [178, Remarks after Definition 3.9].

4.7 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 [94, 178]). 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 [136].

The structure of this section is as follows. After defining the quadratic covariation $[X, Y]$ and the quadratic variation $[X] = [X, X]$, we establish their basic algebraic and analytical properties, including the approximation by Riemann-type sums along refining random partitions and the polarization identity. We then derive the behavior of the quadratic (co-)variation under stochastic integration, which is the crucial input to the Itô formula of Chapter 5. Finally, we record the relationship between the quadratic variation and the classical predictable quadratic variation for square-integrable martingales.

Definition 4.7.1. Let $X, Y \in \mathcal{S}^d$ and

$$XY^\top = \begin{pmatrix} X^1Y^1 & \cdots & X^1Y^d \\ \vdots & \ddots & \vdots \\ X^dY^1 & \cdots & X^dY^d \end{pmatrix}.$$

Then we define *quadratic covariation* of X and Y as

$$[X, Y] := \begin{pmatrix} X^1Y^1 - X^1_- \bullet Y^1 - Y^1_- \bullet X^1 & \cdots & X^1Y^d - X^1_- \bullet Y^d - Y^d_- \bullet X^1 \\ \vdots & \ddots & \vdots \\ X^dY^1 - X^d_- \bullet Y^1 - Y^1_- \bullet X^d & \cdots & X^dY^d - X^d_- \bullet Y^d - Y^d_- \bullet X^d \end{pmatrix}.$$

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, [113, Section 8.5]), which is an equivalent definition. There are even more ways to define the quadratic variation, especially for martingales (see, for example, [31, Chapter 11], [166, Chapter II Corollary 2 of Theorem 27] or [2, Definition 1.15]).

One easily sees that the $(X, Y) \mapsto [X, Y]$ is bilinear and symmetric. Therefore, we get the *polarization identity*

$$[X, Y] = \frac{1}{2}([X + Y, X + Y] - [X, X] - [Y, Y]) \tag{4.12}$$

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 4.7.2. A finite set σ of stopping times, which satisfies $0 = T_1 \leq T_2 \leq \cdots \leq T_k < \infty$ 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 4.7.3. *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

$$X_0 Y_0 + \sum_{i=1}^{k_n} (X^{T_{i+1}^n} - X^{T_i^n})(Y^{T_{i+1}^n} - Y^{T_i^n}) \to [X, Y], \quad n \to \infty, \tag{4.13}$$

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 forms an algebra, it is easy to see that each component is a semimartingale, and by Item (d), it is an FV-process.

We can, for all properties, without loss of generality, assume $d = 1$.

(a) By definition, $[X,Y]_0 = X_0Y_0 - X_- \bullet Y_0 - Y_- \bullet X_0 = X_0Y_0$, and the first equation is clear. For the second equation, using Theorem 4.5.1,

$$\begin{aligned}\Delta X \Delta Y &= (X - X_-)(Y - Y_-) = XY - X_-Y - Y_-X + X_-Y_- \\ &= (XY - X_-Y_-) - X_-Y - Y_-X + 2X_-Y_- \\ &= \Delta(XY) - X_-(Y - Y_-) - Y_-(X - X_-) \\ &= \Delta(XY) - X_-\Delta Y - Y_-\Delta X \\ &= \Delta(XY) - \Delta(X_- \bullet Y) - \Delta(Y_- \bullet X),\end{aligned}$$

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:

$$\begin{aligned}&(X^R)_t^2 \\ &= (X_0^R)^2 + 2\sum_i X_{T^i}^R (X_{T^{i+1}\wedge t}^R - X_{T^i\wedge t}^R) + \sum_i (X_{T^{i+1}\wedge t}^R - X_{T^i\wedge t}^R)^2 \\ &= (X_0^R)^2 + 2\left((X_{\sigma_n}\mathbb{1}_{[\![0,R]\!]}) \bullet X\right)_t + \sum_i (X_{T^{i+1}\wedge t\wedge R} - X_{T^i\wedge t\wedge R})^2. \qquad (4.14)\end{aligned}$$

By the assumptions on σ_n, we note that only finitely many terms of this sum are nonzero. Define $Z(\sigma_n, t) := \sum_i (X_{T_{i+1}^n\wedge t} - X_{T_i^n\wedge t})^2$. By the Bounded Convergence Theorem (Theorem 4.4.10), applied to the integrands $X^{\sigma_n}\mathbb{1}_{[\![0,R]\!]}$, which are uniformly bounded by N and converge pointwise to $X_-\mathbb{1}_{[\![0,R]\!]}$, 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 (4.14) implies $Z = [X, X]$, and therefore $Z = [X, X]$ in general.

(c) We assume $d = 1$. By Theorem 4.5.1, 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.

(d) This follows with $X = Y$ directly from the summation representation in Item (b).

(e) 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 4.7.4. *For $X, Y \in \mathcal{S}^1, H \in L(X)$ and $K \in L(Y)$, we have*

$$[H \bullet X, K \bullet Y]_t = \int_0^t H_s K_s \mathrm{d}[X, Y]_s \quad (t \ge 0). \tag{4.15}$$

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 4.7.3, 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 4.4.19 provides the result for $H \in L(X)$. □

Remark 4.7.5. Theorem 4.7.4 has a multi-dimensional version as well ([178, Theorem 4.19]). However, the multidimensional variant requires some more technical definitions, and we, therefore, stick to the much simpler version in one dimension.

To apply standard arguments, for instance those used in the proof of Itô's formula, it is essential to decompose the quadratic variation into its continuous and purely discontinuous

parts. For completeness, we state this result below. With this decomposition at hand, the proofs of the usual results follow directly.

Proposition 4.7.6. *Let $X, Y \in \mathcal{S}^1$ be (one-dimensional) semimartingales and recall that $[X,Y]$ is càdlàg and of finite variation with $\Delta[X,Y]_t = \Delta X_t\,\Delta Y_t$ and $[X,Y]_0 = X_0Y_0$ by Theorem 4.7.3. Define*

$$[X,Y]^d_t := [X,Y]_0 \;+\; \sum_{0<s\le t} \Delta X_s\,\Delta Y_s, \qquad [X,Y]^c := [X,Y] - [X,Y]^d.$$

Then:

(i) The sum in the definition of $[X,Y]^d$ converges absolutely on compacts; hence $[X,Y]^d$ is a càdlàg finite-variation process which is purely discontinuous. Moreover,

$$\Delta[X,Y]^d_t = \Delta X_t\,\Delta Y_t \quad \text{for all } t \ge 0, \qquad [X,Y]^d_0 = X_0Y_0.$$

Equivalently, $[X,Y]^d_t = X_0Y_0 + \sum_{0<s\le t} \Delta X_s\,\Delta Y_s$.

(ii) The process $[X,Y]^c$ is continuous, of finite variation, and satisfies

$$[X,Y]^c_0 = 0.$$

(iii) The decomposition is unique and behaves well under stopping: for every stopping time T,

$$[X,Y]^{c,T} = [X^T, Y^T]^c, \qquad [X,Y]^{d,T} = [X^T, Y^T]^d.$$

(iv) (Compatibility with stochastic integration.) *For every $H \in L(X)$, $K \in L(Y)$,*

$$[H\bullet X,\, K\bullet Y]^c = \int_0^{\cdot} H_sK_s\,\mathrm{d}[X,Y]^c_s,$$
$$[H\bullet X,\, K\bullet Y]^d_t = (H\bullet X)_0\,(K\bullet Y)_0 + \sum_{0<s\le t} H_sK_s\,\Delta X_s\,\Delta Y_s.$$

Equivalently,

$$[H\bullet X,\, K\bullet Y]_t = [H\bullet X,\, K\bullet Y]_0 + \int_0^t H_sK_s\,\mathrm{d}[X,Y]^c_s + \sum_{0<s\le t} H_sK_s\,\Delta X_s\,\Delta Y_s.$$

All assertions hold componentwise for $X, Y \in \mathcal{S}^d$.

Proof. By Theorem 4.7.3 the bracket $[X,Y]$ is (componentwise) a finite-variation semimartingale with $\Delta[X,Y]_t = \Delta X_t\,\Delta Y_t$ and $[X,Y]_0 = X_0Y_0$. For any càdlàg finite-variation process V the total variation on $[0,t]$ dominates the sum of absolute jumps, hence

$\sum_{0<s\le t}|\Delta[X,Y]_s| \le \mathbf{Var}([X,Y])_t < \infty$. Thus the series in $[X,Y]^d$ converges absolutely (on compacts), so $[X,Y]^d$ is of finite variation and purely discontinuous, with the stated jumps and initial value. Setting $[X,Y]^c := [X,Y] - [X,Y]^d$ gives a càdlàg finite-variation process with no jumps; hence it is continuous and $[X,Y]^c_0 = 0$. Uniqueness follows from equality of jumps and initial value.

The stopping identities follow from Theorem 4.7.3: since $[X^T,Y^T] = [X,Y]^T$, the same algebra gives $[X^T,Y^T]^d = [X,Y]^{d,T}$ (because both have the same initial value and jumps), and hence also $[X^T,Y^T]^c = [X,Y]^{c,T}$.

For the integration statement, Theorem 4.7.4 yields

$$[H\bullet X,\ K\bullet Y] = \int HK\,\mathrm{d}[X,Y].$$

Decomposing $[X,Y] = [X,Y]^c + [X,Y]^d$ and using linearity of the (Lebesgue–)Stieltjes integral for finite-variation integrators (cf. Item (f)) gives $[H\bullet X,\ K\bullet Y] = \int HK\,\mathrm{d}[X,Y]^c + \int HK\,\mathrm{d}[X,Y]^d$. Since $[X,Y]^d$ is purely discontinuous with jumps $\Delta[X,Y]^d_t = \Delta X_t\,\Delta Y_t$, the pathwise Stieltjes integral against $[X,Y]^d$ reduces to the sum of atoms, $\int_0^t H_sK_s\,\mathrm{d}[X,Y]^d_s = \sum_{0<s\le t} H_sK_s\,\Delta X_s\,\Delta Y_s$, while the integral against $[X,Y]^c$ yields the claimed continuous part. The displayed formulas follow. □

Remark 4.7.7. Two immediate consequences of Proposition 4.7.6 align with Theorem 4.7.3:

- If one of X or Y has finite variation, then $[X,Y]^c \equiv 0$ and $[X,Y]_t = X_0Y_0 + \sum_{0<s\le t}\Delta X_s\,\Delta Y_s$.
- If X and Y are continuous, then $[X,Y]^d \equiv X_0Y_0$ and the evolving part $[X,Y] - X_0Y_0$ is continuous.

Remark 4.7.8. 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 4.6.2, it is established that $MN^\top - [M,N]$ behaves as a local martingale in each component. Furthermore, according to Theorem 4.7.3, 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 4.7.9. 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 4.7.3, $[H^1 \bullet X^1, H^1 \bullet X^1]$ 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 4.7.10. There are several criteria for determining when an integrand is componentwise integrable. One such criterion is discussed in Remark 4.7.12. Moreover, all locally bounded predictable integrands are necessarily integrable (see Theorem 4.3.11). 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 Dominated Convergence Theorem cannot be dropped.

Example 4.7.11. 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 4.7.9, we also have $H \in L(X)$ but we have $K \notin L(X)$.

Remark 4.7.12. 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,

$$\mathbf{E}\left[\int_0^\infty H_s^2 \, \mathrm{d}[N, N]_s\right] + \mathbf{E}\left[\left(\int_0^\infty |H_s| \, \mathrm{d}|A|_s\right)^2\right] < \infty$$

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 [166, Theorem IV.14]). Next, we demonstrate that $H^n \bullet X$ converges in the semimartingale topology ([166, Theorem VI.4.19]). The rest follows with the equivalence of the $\mathcal{H}^2$ and X norms ([166, Corollary of Theorem IV.24 and Theorem V.14]).

Chapter 5

The Itô Formula

In this chapter we develop Itô's formula within the semimartingale framework. Building on the topological definition of the stochastic integral introduced earlier, we give a concise proof of the change-of-variables formula for càdlàg semimartingales, record its continuous and multi-dimensional versions, and collect the corollaries needed later (integration by parts, chain rules, and the Doléans–Dade exponential). We also indicate precisely where the quadratic-variation term enters and compare with the classical Lebesgue–Stieltjes calculus.

5.1 Statement and Proof

In classical calculus, the chain rule plays a central role. We now want to derive a corresponding rule for semimartingales. It is natural to assume that the chain rule can be directly transferred to our integration calculus, since after all, the stochastic integral coincides with the Lebesgue-Stieltjes integral in some cases (cf. Item (f)). If A is a process with continuous paths of finite variation, one can show that for $f \in \mathcal{C}^1$, the chain rule for the Lebesgue-Stieltjes integral yields:

$$f(A_t) - f(A_0) = \int_0^t f'(A_s)\mathrm{d}A_s.$$

However, we have already seen that for a Brownian motion B,

$$\int_0^t B_s \mathrm{d}B_s = \frac{1}{2}B_t^2 - \frac{1}{2}t,$$

which for $f(x) = \frac{x^2}{2}$ corresponds to

$$f(B_t) - f(B_0) = \int_0^t f'(B_s)\mathrm{d}B_s + \frac{1}{2}\int_0^t f''(B_s)\mathrm{d}s,$$

and therefore does not coincide with the formula for the chain rule of Lebesgue-Stieltjes integration. Indeed, due to the possibly unbounded variation, we will need to add an additional term, as the above calculation already suggests.

The formula analogous to the chain rule in stochastic analysis is called the *Itô formula*. It represents a natural generalization of the fundamental theorem of calculus to stochastic integrals and has far-reaching applications, particularly in mathematical finance.

5.1.1 The One-Dimensional Itô Formula

Theorem 5.1.1 (Itô Formula)**.** *Let X be a semimartingale and $f : \mathbb{R} \to \mathbb{R}$ be a twice continuously differentiable function (i.e., $f \in \mathcal{C}^2$). Then $f(X)$ is also a semimartingale and we have*

$$f(X_t) - f(X_0) = f'(X_-) \bullet X_t + \frac{1}{2} f''(X_-) \bullet [X, X]^c_t + \sum_{0<s\leq t} \left(f(X_s) - f(X_{s-}) - f'(X_{s-}) \Delta X_s \right). \tag{5.1}$$

We could also write Eq. (5.1) as

$$\begin{aligned} f(X_t) - f(X_0) = &\int_{0+}^{t} f'(X_{s-}) \mathrm{d}X_s + \frac{1}{2} \int_{0+}^{t} f''(X_{s-}) \mathrm{d}[X, X]^c_s \\ &+ \sum_{0<s\leq t} \left(f(X_s) - f(X_{s-}) - f'(X_{s-}) \Delta X_s \right). \end{aligned}$$

Remark 5.1.2. (a) Semimartingales are stable under $\mathcal{C}^2$-transformations, but in general not under merely continuous transformations.

(b) The first term of the Itô formula

$$\int_{0+}^{t} f'(X_{s-}) \mathrm{d}X_s$$

exists as a stochastic integral, since $f'(X_-) \in \mathcal{L}$. According to the fundamental theorem for Riemann-Stieltjes integrals, this term is the only non-vanishing expression on the right-hand side if the integrator X is continuous and of bounded variation on compact intervals.

If we also allow continuous integrators of unbounded variation, we must consider the second term of the Itô formula,

$$\frac{1}{2} \int_{0+}^{t} f''(X_{s-}) \mathrm{d}[X, X]^c_s.$$

Here $[X, X]^c$ is a continuous and increasing process, hence the integral can be

understood as a path-wise Riemann-Stieltjes or Lebesgue-Stieltjes integral.

If $X = B$ is a Brownian motion, the formula simplifies to

$$f(B_t) = f(B_0) + \int_0^t f'(B_s)\mathrm{d}B_s + \frac{1}{2}\int_0^t f''(B_s)\mathrm{d}s.$$

5.1.2 Proof of the Itô Formula

From now on, let X be a semimartingale with $X_0 = 0$ and $f \in \mathcal{C}^2$. This is a notational convenience: the general case is obtained by carrying X_0 through the same telescoping argument, since the stochastic integral term is unaffected by a constant initial shift. The quadratic variation of X decomposes into a continuous part and a jump part (cf. Proposition 4.7.6), so that

$$\int_{0+}^t f''(X_{s-})\mathrm{d}[X,X]_s = \int_{0+}^t f''(X_{s-})\mathrm{d}[X,X]^c_s + \sum_{0<s\leq t} f''(X_{s-})(\Delta X_s)^2.$$

Since $[X,X]^c_s$ is continuous and increasing and, therefore, of bounded variation on compact sets, we can understand the integral term as a Lebesgue-Stieltjes integral by Item (f) from Theorem 4.5.1. Furthermore, since X is càdlàg it is bounded on compact intervals pathwise, and f'' is therefore bounded on the range of X up to time t. Together with $\sum_{0<s\leq t}(\Delta X_s)^2 < \infty$, this shows that $\sum_{0<s\leq t} f''(X_{s-})(\Delta X_s)^2$ is well-defined.

Using the above decomposition, we formulate the following statement equivalent to the Itô formula, which is, however, easier to handle in the proof.

Variant of the Itô Formula. *Let X be a semimartingale and $f : \mathbb{R} \to \mathbb{R}$ be a $\mathcal{C}^2$-function. Then $f(X)$ is also a semimartingale and we have*

$$\begin{aligned} f(X_t) - f(X_0) = &\int_{0+}^t f'(X_{s-})\mathrm{d}X_s + \frac{1}{2}\int_{0+}^t f''(X_{s-})\mathrm{d}[X,X]_s \\ &+ \sum_{0<s\leq t}\left(f(X_s) - f(X_{s-}) - f'(X_{s-})\Delta X_s - \frac{1}{2}f''(X_{s-})(\Delta X_s)^2\right). \end{aligned}$$

The remaining part of this section is devoted to the proof of this variant of the Itô formula. A central role is played by the following version of Taylor's theorem.

Theorem 5.1.3 (Taylor's Theorem)**.** *Let $I = [a,b]$ be an interval, $f \in \mathcal{C}^2(I)$ and $x \in I$. Then*

$$f(y) - f(x) = (y-x)f'(x) + \frac{1}{2}(y-x)^2 f''(x) + R(x,y)$$

for every $y \in I$, where there exists a monotonically increasing function $r : \mathbb{R} \to \mathbb{R}$ with

$r(b-a) \leq \sup_{x\in I} |f''(x)|$ *and* $r(u) \searrow 0$ *as* $u \searrow 0$, *such that*

$$|R(x,y)| \leq (y-x)^2 r(|x-y|).$$

We first prove the variant of the Itô formula for continuous semimartingales. We then present the relatively technical generalization to arbitrary semimartingales.

Proposition 5.1.4. *The variant of the Itô formula holds for every continuous semimartingale.*

Proof. If we write the Itô formula in abbreviated form as

$$f(X_t) - f(X_0) = I(f,X)_t,$$

then one easily verifies with Item (d) and Theorem 4.7.3 that

$$I(f,X)_t^{R_m} = I(f, X^{R_m})_t \tag{5.2}$$

holds.

Define $R_m := \inf\{t > 0 : |X_t| \geq m\}$. Then R_m is a stopping time with $R_m \nearrow \infty$ and $|X^{R_m}| \leq m$. If we can prove that the Itô formula holds for each m, then by Eq. (5.2) it also holds in general. Hence it suffices to prove the claim for bounded continuous semimartingales.

We can therefore assume that X takes values only in a compact interval I. Let $t > 0$ be fixed and σ_n be a sequence of stopping times $0 = T_0^n \leq T_1^n \leq \ldots \leq T_{k_n}^n = t$ that converges to the identity and refines the sequence of stopping times R_m. Thus

$$\begin{aligned} f(X_t) - f(X_0) &= \sum_{1\leq i\leq k_n} \left(f(X_{T_i^n}) - f(X_{T_{i-1}^n}) \right) \\ &= \sum_{1\leq i\leq k_n} \left((X_{T_i^n} - X_{T_{i-1}^n}) f'(X_{T_{i-1}^n}) + \frac{1}{2}(X_{T_i^n} - X_{T_{i-1}^n})^2 f''(X_{T_{i-1}^n}) \right) \\ &\quad + \sum_{1\leq i\leq k_n} R(X_{T_{i-1}^n}, X_{T_i^n}). \end{aligned}$$

Since f is twice continuously differentiable, we have $f'(X_-), f''(X_-) \in \mathcal{L}$ and thus by Theorem 4.5.1 and Item (b), the first sum converges in probability to

$$\int_{0+}^t f'(X_{s-}) \mathrm{d}X_s + \frac{1}{2} \int_{0+}^t f''(X_{s-}) \mathrm{d}[X,X]_s.$$

We only need to consider the remainder term. By assumption, X is continuous, hence

uniformly continuous on $[0,t]$ for each ω, and consequently

$$|X_{T_i^n} - X_{T_{i-1}^n}| \to 0 \quad \text{a.s.}$$

as $n \to \infty$. Thus

$$\begin{aligned}\sum_{1\le i\le k_n} |R(X_{T_{i-1}^n}, X_{T_i^n})| &\le \sum_{1\le i\le k_n} (X_{T_i^n} - X_{T_{i-1}^n})^2 r(|X_{T_i^n} - X_{T_{i-1}^n}|) \\ &\le \left(\sup_{1\le i\le k_n} r(|X_{T_i^n} - X_{T_{i-1}^n}|)\right) \sum_{1\le i\le k_n} (X_{T_i^n} - X_{T_{i-1}^n})^2 \\ &\xrightarrow{\mathbf{P}} 0,\end{aligned}$$

since $\sum_{1\le i\le k_n}(X_{T_i^n} - X_{T_{i-1}^n})^2 \xrightarrow{\mathbf{P}} [X,X]_t$, which is almost surely finite, and r goes to 0 by assumption. The limit in probability is unique only almost surely, hence for each $t \ge 0$

$$f(X_t) - f(X_0) \overset{\text{a.s.}}{=} \int_{0+}^t f'(X_{s-})\mathrm{d}X_s + \frac{1}{2}\int_{0+}^t f''(X_{s-})\mathrm{d}[X,X]_s.$$

The exceptional set on which the identity does not hold is a t-dependent null set. Using the continuity of X, however, the t-dependence can be resolved, so that the identity holds for all t simultaneously except on a null set. □

Proposition 5.1.5. *The variant of the Itô formula holds for arbitrary semimartingales.*

Proof. Roadmap of the proof. We localize the integrand by stopping X at the first exit from $[-m,m]$, so we may assume X takes values in a compact interval $I = [-m,m]$ and is pathwise bounded. Following this localization, the proof has three blocks. (1) For each $\varepsilon > 0$, partition the at-most-countable random jump set D of X on $(0,t]$ into a *finite* "large-jump" subset D_ε collecting all jumps with $|\Delta X_s| > \varepsilon$ and a "small-jump" tail $D \setminus D_\varepsilon$ with $\sum_{s\in D\setminus D_\varepsilon}(\Delta X_s)^2 \le \varepsilon$. (2) Along a sequence of stopping-time partitions σ_n that refine the stopping times capturing D_ε, split the telescoping sum into intervals containing a large jump (finitely many) and the rest. (3) Apply Taylor's theorem on the small-jump-only intervals, control the remainder uniformly in n via the oscillation of X, and pass to the limit in the order $n \to \infty$, $\varepsilon \downarrow 0$, then $m \to \infty$. The quadratic-variation convergence input is Item (b) from Theorem 4.7.3, and the stochastic-integral convergence input is Theorem 4.5.1.

Fix $t > 0$. Since each càdlàg path is bounded on $[0,t]$, the following estimates may be carried out pathwise on the events $\{\sup_{s\le t}|X_s| \le m\}$, which increase to Ω, and then we let $m \to \infty$. For $t > 0$, we define a random set

$$D := \{s \in (0,t]\,;\, |\Delta X_s| > 0\}.$$

As a semimartingale, X is in particular càdlàg and hence D is at most countable. By Item (d) from Theorem 4.5.1 it follows that

$$\sum_{0<s\leq t} (\Delta X_s)^2 \leq [X,X]_t < \infty.$$

The sum is thus in particular convergent. For each $\varepsilon > 0$ choose a finite set $D_\varepsilon \subset D$ containing all jumps with $|\Delta X_s| > \varepsilon$ and enough smaller jumps so that

$$\sum_{s\in D\setminus D_\varepsilon} (\Delta X_s)^2 \leq \varepsilon.$$

Then all jumps outside D_ε have absolute size at most ε. Choose the family $\{D_\varepsilon\}$ monotonically along a sequence $\varepsilon_n \downarrow 0$ so that $D_{\varepsilon_n} \uparrow D$ and $\sum_{s\in D\setminus D_{\varepsilon_n}} (\Delta X_s)^2 \to 0$.

Let σ_n again be a sequence of stopping times on $[0,t]$ that converges to the identity, and

$$A^n = \{1 \leq i \leq k_n : \exists s \in D_\varepsilon \text{ with } T^n_{i-1} < s \leq T^n_i\},$$

so A^n depends on randomness and $\#A^n \leq \#D_\varepsilon < \infty$. We consider

$$\begin{aligned} f(X_t) - f(X_0) &= \sum_{1\leq i\leq k_n} \left(f(X_{T^n_i}) - f(X_{T^n_{i-1}})\right) \\ &= \sum_{i\in A^n} \left(f(X_{T^n_i}) - f(X_{T^n_{i-1}})\right) + \sum_{i\notin A^n} \left(f(X_{T^n_i}) - f(X_{T^n_{i-1}})\right), \end{aligned}$$

where we have decomposed the telescoping sum so that the first sum collects the finitely many partition intervals containing the selected large jumps in D_ε, and the second sum collects the remaining intervals, which contain only small jumps of size at most ε together with the continuous part of X.

Since the cardinality of all A^n is bounded, we can pass to the limit in the first sum, so that

$$\lim_{n\to\infty} \sum_{i\in A^n} \left(f(X_{T^n_i}) - f(X_{T^n_{i-1}})\right) = \sum_{s\in D_\varepsilon} (f(X_s) - f(X_{s-})).$$

To the summands of the second sum we apply Taylor's theorem and obtain

$$\begin{aligned}
&\sum_{i\notin A^n}\Big(f(X_{T_i^n})-f(X_{T_{i-1}^n})\Big)\\
&=\sum_{i\notin A^n}\left(f'(X_{T_{i-1}^n})(X_{T_i^n}-X_{T_{i-1}^n})+\frac{1}{2}f''(X_{T_{i-1}^n})(X_{T_i^n}-X_{T_{i-1}^n})^2\right)\\
&\quad+\sum_{i\notin A^n}R(X_{T_{i-1}^n},X_{T_i^n})\\
&=\sum_{1\le i\le k_n}\left(f'(X_{T_{i-1}^n})(X_{T_i^n}-X_{T_{i-1}^n})+\frac{1}{2}f''(X_{T_{i-1}^n})(X_{T_i^n}-X_{T_{i-1}^n})^2\right) \qquad (5.3)\\
&\quad-\sum_{i\in A^n}\left(f'(X_{T_{i-1}^n})(X_{T_i^n}-X_{T_{i-1}^n})+\frac{1}{2}f''(X_{T_{i-1}^n})(X_{T_i^n}-X_{T_{i-1}^n})^2\right) \qquad (5.4)\\
&\quad+\sum_{i\notin A^n}R(X_{T_{i-1}^n},X_{T_i^n}). \qquad (5.5)
\end{aligned}$$

Analogous to the proof of Proposition 5.1.4, the term in (5.3) converges in probability by Theorem 4.5.1 and Item (b) of Theorem 4.5.1 to

$$\int_{0+}^{t}f'(X_{s-})\mathrm{d}X_s+\frac{1}{2}\int_{0+}^{t}f''(X_{s-})\mathrm{d}[X,X]_s.$$

The term in (5.4) is also convergent due to the finiteness of A^n, namely to

$$\sum_{s\in D_\varepsilon}\left(f'(X_{s-})\Delta X_s+\frac{1}{2}f''(X_{s-})(\Delta X_s)^2\right).$$

It remains to control the remainder term $\sum_{i\notin A^n}R(X_{T_{i-1}^n},X_{T_i^n})$. By assumption $|X_s|\le m$ for some $m\ge 1$, so f'' is bounded on $I=[-m,m]$ by $M:=\sup_{x\in I}|f''(x)|$ and the increasing function r from Theorem 5.1.3 satisfies $r(u)\le M$ for all u and $r(u)\searrow 0$ as $u\searrow 0$. Fix $\eta>0$. By the standard oscillation property of càdlàg functions, after removing the finitely many jumps in D_ε, the oscillation of X on sufficiently fine partition intervals not meeting D_ε is bounded by $\eta+\varepsilon$. Consequently $r(|X_{T_i^n}-X_{T_{i-1}^n}|)\le r(\eta+\varepsilon)$ on these intervals, once n is large enough. Therefore

$$\sum_{i\notin A^n}|R(X_{T_{i-1}^n},X_{T_i^n})|\le r(\eta+\varepsilon)\sum_{i\notin A^n}(X_{T_i^n}-X_{T_{i-1}^n})^2.$$

As $n\to\infty$, $\sum_{1\le i\le k_n}(X_{T_i^n}-X_{T_{i-1}^n})^2\xrightarrow{\mathbf{P}}[X,X]_t$, while the contribution of the finitely many intervals in A^n converges to $\sum_{s\in D_\varepsilon}(\Delta X_s)^2$. Hence

$$\sum_{i\notin A^n}(X_{T_i^n}-X_{T_{i-1}^n})^2\xrightarrow{\mathbf{P}}[X,X]_t^c+\sum_{s\notin D_\varepsilon}(\Delta X_s)^2\le[X,X]_t^c+\varepsilon.$$

Taking limits in the order $n\to\infty$, $\eta\to 0$, and finally $\varepsilon\to 0$, and noting that $r(\eta+\varepsilon)\to 0$,

the remainder vanishes in probability.

In summary, for fixed ε, after letting $n \to \infty$ we obtain

$$\begin{aligned} f(X_t) - f(X_0) = &\int_{0+}^{t} f'(X_{s-})\mathrm{d}X_s + \frac{1}{2}\int_{0+}^{t} f''(X_{s-})\mathrm{d}[X,X]_s \\ &+ \sum_{s\in D_\varepsilon}\Big(f(X_s) - f(X_{s-}) - f'(X_{s-})\Delta X_s - \frac{1}{2}f''(X_{s-})(\Delta X_s)^2\Big) \\ &+ \mathcal{R}_{\eta,\varepsilon}, \end{aligned}$$

where the residual $\mathcal{R}_{\eta,\varepsilon}$ has lim sup in probability bounded by

$$r(\eta+\varepsilon)\Big([X,X]_t^c + \varepsilon\Big) + C\sum_{s\in D\setminus D_\varepsilon}(\Delta X_s)^2.$$

Letting $\eta \downarrow 0$ and then $\varepsilon \downarrow 0$ gives zero: $r(\eta+\varepsilon) \to 0$, the square-tail tends to zero by construction, and $D_{\varepsilon_n} \uparrow D$ along the chosen exhaustion. It remains to show that the jump sum converges absolutely as $\varepsilon \downarrow 0$.

We show the absolute convergence of the series for $\varepsilon \searrow 0$. Since the term in the corrected variant of the Itô formula is exactly the Taylor remainder $R(X_{s-}, X_s)$, Taylor's theorem gives

$$\Big|f(X_s) - f(X_{s-}) - f'(X_{s-})\Delta X_s - \tfrac{1}{2}f''(X_{s-})(\Delta X_s)^2\Big| = |R(X_{s-}, X_s)| \le C(\Delta X_s)^2,$$

where $C = \sup_{x\in I}|f''(x)| < \infty$ since X is bounded. Summing directly over all jumps,

$$\begin{aligned} &\sum_{s\in D}\Big|f(X_s) - f(X_{s-}) - f'(X_{s-})\Delta X_s - \frac{1}{2}f''(X_{s-})(\Delta X_s)^2\Big| \\ &\quad\le C\sum_{s\in D}(\Delta X_s)^2 \le C\,[X,X]_t < \infty. \end{aligned}$$

Hence the partial sums over D_ε converge absolutely to the full jump correction as $\varepsilon \downarrow 0$. Finally, letting $m \to \infty$ removes the bounded-path restriction, since the events $\{\sup_{s\le t}|X_s| \le m\}$ increase to Ω. This completes the proof. □

5.1.3 The Multi-Dimensional Itô Formula

So far we have only dealt with one-dimensional integration. The Itô formula also holds in higher dimensions. We present here a version for d-dimensional stochastic processes.

Theorem 5.1.6 (Multi-Dimensional Itô Formula)**.** *Let $X = (X^1, \dots, X^d)$ be a d-tuple of semimartingales and $f : \mathbb{R}^d \to \mathbb{R}$ have continuous second partial derivatives. Then the*

stochastic process $f(X)$ is again a semimartingale and we have, up to indistinguishability,

$$\begin{aligned} f(X_t) = f(X_0) &+ \sum_{i=1}^{d} \left(\frac{\partial f}{\partial x_i}(X_-) \right) \bullet X_t^i \\ &+ \frac{1}{2} \sum_{1 \le i,j \le d} \left(\frac{\partial^2 f}{\partial x_i \partial x_j}(X_-) \right) \bullet [X^i, X^j]_t^c \\ &+ \sum_{0<s\le t} \left(f(X_s) - f(X_{s-}) - \sum_{j=1}^{d} \frac{\partial f}{\partial x_j}(X_{s-}) \Delta X_s^j \right). \end{aligned} \tag{5.6}$$

The proof proceeds analogously to the proof of Theorem 5.1.1, except that the multi-dimensional Taylor formula is used.

Remark 5.1.7. The potentially infinite sum in Eq. (5.6) (over all jumps) is absolutely convergent. To see this, fix ω and t: on each compact interval, the càdlàg path of X is bounded, so the Taylor estimate

$$f(x+h) = f(x) + h^\top \nabla f(x) + O(\|h\|^2), \quad h \to 0,$$

is uniform on the range of X up to time t. Combined with

$$\sum_{0<s\le t} \|\Delta X_s\|^2 < \infty, \quad \mathbf{P}\text{-a.s.}$$

(cf. Item (d)), the jump sum

$$\sum_{0<s\le t} \left(f(X_s) - f(X_{s-}) - \sum_{j=1}^{d} \frac{\partial f}{\partial x_j}(X_{s-}) \Delta X_s^j \right)$$

is absolutely convergent.

For $d = 1$ and a real-valued càdlàg process A of finite variation, the Itô formula reduces to

$$f(A_t) = f(A_0) + \int_0^t f'(A_s) \mathrm{d}A_s^c + \sum_{0<s\le t} (f(A_s) - f(A_{s-})),$$

where A^c denotes the continuous part of the process A, which can be obtained analogously to Definition 3.1.31 due to the finite variation of A. Note that

$$\int_0^t f'(A_{s-}) \mathrm{d}A_s^c = \int_0^t f'(A_s) \mathrm{d}A_s^c,$$

since a modification of the integrand at countably many points does not change the integral for a continuous integrator.

The Itô formula is thus a generalization of the fundamental theorem of calculus,

according to which the integral over the derivative of a function yields the function itself (minus its initial value).

In applications, one of the semimartingales is often time, i.e., $X_t^1 = t$. Then $[X^1, X^j] = 0$ for $j = 1, 2, \ldots, d$ (due to $[X^1, X^1] = 0$ and the Cauchy-Schwarz inequality) and $\Delta X^1 = 0$.

Remark 5.1.8. So far, we have proven that for a semimartingale X, $f(X)$ is also a semimartingale when $f \in \mathcal{C}^2$. The question arises whether the "relatively strong" condition $f \in \mathcal{C}^2$ can be weakened. This is indeed possible, although the resulting class is still quite structured. It holds that:

A real function f maps every real semimartingale X to a semimartingale $f(X)$ if and only if f is locally the difference of two convex functions.

A proof and more detailed exposition of this statement can be found, for example, in [30]. The Itô formula can also be generalized to this case. This leads to the concept of *local times*, which we will not discuss further here. From the above result, it follows directly that if X is a semimartingale, so are

$$|X|, X^+, X^-, X \vee Y \text{ and } X \wedge Y.$$

5.2 Applications of the Itô Formula

5.2.1 The Stochastic Exponential

Theorem 5.2.1. *Let X be a semimartingale with $X_0 = 0$. There exists a unique semimartingale Z that satisfies the equation:*

$$Z_t = 1 + Z_- \bullet X_t, \quad \text{for all } t \geq 0 \tag{5.7}$$

The stochastic process Z is given by

$$Z_t = \exp\left(X_t - \frac{1}{2}[X, X]_t^c\right) \prod_{0<s\leq t} (1 + \Delta X_s)\, e^{-\Delta X_s}, \tag{5.8}$$

where the product is well-defined on compact time intervals and the associated logarithmic series is absolutely convergent.

Note that Eq. (5.8) simplifies considerably for continuous processes of finite variation on compact sets:

$$Z_t = \exp(X_t),$$

and for merely continuous processes, the formula simplifies to

$$Z_t = \exp\left(X_t - \frac{1}{2}[X,X]_t\right).$$

Furthermore, note that the restriction $X_0 = 0$ is not a real restriction. If $\tilde{X}$ is a semimartingale with $\tilde{X}_0 \neq 0$, then $X_t = \tilde{X}_t - \tilde{X}_0$ satisfies the above conditions.

We first verify that the product is well-defined. Fix $\varepsilon \in (0,1)$. Since X is càdlàg, on every compact interval $[0,t]$ there are only finitely many jumps with $|\Delta X_s| \geq \varepsilon$; in particular, finitely many jumps with $\Delta X_s \leq -1$. Removing these finitely many factors does not affect the convergence of the product. On the remaining factors $|\Delta X_s| < \varepsilon < 1$, so $1 + \Delta X_s > 0$ and the Taylor expansion

$$\log\Big((1+\Delta X_s)e^{-\Delta X_s}\Big) = \log(1+\Delta X_s) - \Delta X_s = -\tfrac{1}{2}(\Delta X_s)^2 + O\Big((\Delta X_s)^3\Big)$$

is valid. Hence

$$\sum_{0<s\leq t,\ |\Delta X_s|<\varepsilon} \Big|\log\Big((1+\Delta X_s)e^{-\Delta X_s}\Big)\Big| \leq C \sum_{0<s\leq t} (\Delta X_s)^2 \leq C\,[X,X]_t < \infty,$$

which proves the absolute convergence of the corresponding logarithmic series, and hence of the product.

If a jump $\Delta X_s = -1$ occurs at some time s, then the factor $(1+\Delta X_s)e^{-\Delta X_s}$ equals zero, so the product vanishes from time s onward. Jumps $\Delta X_s < -1$ produce negative factors, which change the sign of finitely many factors on each compact interval but do not affect convergence.

Remark 5.2.2. Equation (5.7) is an example of a so-called *stochastic differential equation* (SDE). The term "differential equation" comes from the fact that Eq. (5.7) is usually written in shorthand as

$$\mathrm{d}Z_t = Z_{t-}\mathrm{d}X_t \quad \text{and} \quad Z_0 = 1.$$

However, even in the differential notation, what is meant is that Z satisfies the integral equation (5.7), i.e., $\mathrm{d}X_t$ is not an independent quantity.

Proof of Theorem 5.2.1. We first establish uniqueness. Suppose Z and $\tilde{Z}$ are two semimartingale solutions of Eq. (5.7). Their difference $D = Z - \tilde{Z}$ satisfies $D_0 = 0$ and $D = D_- \bullet X$. A standard application of the local boundedness of D_- and Gronwall-type estimates for linear stochastic integral equations gives $D \equiv 0$ (see, e.g., [166, Theorem V.7]). Hence the solution is unique.

We now construct the solution. We first consider the case where X is a continuous

semimartingale. Define

$$K_t := X_t - \frac{1}{2}[X,X]_t, \qquad Z_t := \exp(K_t).$$

Since X is continuous, so is K. By Theorem 4.7.4, we have

$$[K,K] = [X,X],$$

since $[X,X]$ is a continuous finite variation process and thus contributes zero to the quadratic variation.

Applying the Itô formula (Theorem 5.1.1) to the function $f(x) = \exp(x)$, we obtain:

$$\begin{aligned} Z_t = \exp(K_t) &= \exp(K_0) + \int_0^t \exp(K_{s-})\,\mathrm{d}K_s + \frac{1}{2}\int_0^t \exp(K_{s-})\,\mathrm{d}[K,K]_s \\ &= 1 + \int_0^t \exp(K_{s-})\left(\mathrm{d}X_s - \frac{1}{2}\mathrm{d}[X,X]_s\right) + \frac{1}{2}\int_0^t \exp(K_{s-})\,\mathrm{d}[X,X]_s \\ &= 1 + \int_0^t \exp(K_{s-})\,\mathrm{d}X_s. \end{aligned}$$

Since X is continuous, $K_{s-} = K_s$ and $\exp(K_{s-}) = \exp(K_s) = Z_s$. Therefore,

$$Z_t = 1 + \int_0^t Z_s\,\mathrm{d}X_s = 1 + Z_- \bullet X_t,$$

which shows that $Z = \mathcal{E}(X)$ in the continuous case.

Now consider the general case. Set

$$Y_t := \exp\left(X_t - \tfrac{1}{2}[X,X]_t^c\right), \qquad U_t := \prod_{0<s\le t} (1+\Delta X_s)\, e^{-\Delta X_s}, \qquad Z := YU,$$

so that Z is the right-hand side of (5.8).

We first show that U is well-defined and is a purely discontinuous càdlàg adapted process of finite variation on compact intervals. Since $(1+x)e^{-x} - 1 = O(x^2)$ as $x \to 0$ and $\sum_{s\le t}(\Delta X_s)^2 < \infty$, the logarithmic series associated to the product converges absolutely after removing the finitely many large-jump factors, and the jump variation of U is finite on each compact interval. Therefore U is a finite-variation semimartingale with $U^c \equiv 0$.

By integration by parts, $Z = YU = Y_0U_0 + Y_-\bullet U + U_-\bullet Y + [Y,U]$; since U is of finite variation with $U^c \equiv 0$, the covariation $[Y,U] = \sum_{0<s\le\cdot} \Delta Y_s\,\Delta U_s$ is purely discontinuous and $[Y,U]^c = 0$. We verify $Z = 1 + Z_-\bullet X$ by matching the jump parts and the continuous parts separately.

At a jump time s, the exponential factor contributes $e^{\Delta X_s}$ and U gains the factor

$(1 + \Delta X_s)\, e^{-\Delta X_s}$, so

$$Z_s = Z_{s-}\, e^{\Delta X_s} \cdot (1 + \Delta X_s)\, e^{-\Delta X_s} = Z_{s-}(1 + \Delta X_s),$$

which gives the jump identity $\Delta Z_s = Z_{s-}\Delta X_s$.

For the continuous part of the semimartingale differential, the jump product U is constant and the continuous Itô formula applied to $Y_t = \exp(X_t - \frac{1}{2}[X, X]^c_t)$ gives

$$\mathrm{d}Y^c_t = Y_{t-}\, \mathrm{d}X^c_t,$$

where the superscript c denotes the continuous contribution; the jump contribution has already been handled by the identity $\Delta Z_s = Z_{s-}\Delta X_s$. The $-\frac{1}{2}\mathrm{d}[X, X]^c_t$ term in the exponent cancels the second-order term from Itô's formula. Since $U^c \equiv 0$, the continuous quadratic covariation $[Y, U]^c$ also vanishes, and the integration-by-parts formula gives $\mathrm{d}Z^c_t = U_{t-}\, \mathrm{d}Y^c_t = U_{t-}Y_{t-}\, \mathrm{d}X^c_t = Z_{t-}\, \mathrm{d}X^c_t$.

Combining the continuous part $\mathrm{d}Z^c_t = Z_{t-}\, \mathrm{d}X^c_t$ with the jump identity $\Delta Z_s = Z_{s-}\Delta X_s$ yields $\mathrm{d}Z_t = Z_{t-}\, \mathrm{d}X_t$, that is, $Z = 1 + Z_- \bullet X$, as required. □

The initial value problem $z' = z$ with $z_0 = 1$ can be used to define the exponential function e^t. By the fundamental theorem of calculus, this is the unique solution of the integral equation

$$e^t = 1 + \int_0^t e^s \mathrm{d}s.$$

The following definition generalizes this idea to stochastic semimartingales.

Definition 5.2.3. Let X be a semimartingale with $X_0 = 0$. The *unique* solution of the integral equation (5.7) is called the *stochastic exponential* of X. We write $\mathcal{E}(X) := Z$, i.e.,

$$\mathcal{E}(X)_t = 1 + \mathcal{E}(X)_- \bullet X_t.$$

$\mathcal{E}(X)$ is also called the *Doléans-Dade exponential*.

Remark 5.2.4. If instead of Eq. (5.7) we consider the stochastic integral equation

$$Z_t = Z_0 + Z_- \bullet X_t \tag{5.9}$$

with an $\mathcal{F}_0$-measurable random variable Z_0, then one easily shows that $Z_0\mathcal{E}(X)$ is a solution of Eq. (5.9).

It should be noted that $\mathcal{E}(X)$ generally differs from the "ordinary" exponential $\exp(X)$.

Example 5.2.5 (Geometric Brownian Motion). Let B be the standard Brownian motion and c a real number. Since B is continuous, the jump product in Eq. (5.8) is trivial, and

$$\mathcal{E}(cB)_t = \exp\left(cB_t - \tfrac{1}{2}[cB, cB]_t^c\right) = \exp\left(cB_t - \tfrac{c^2 t}{2}\right).$$

As a stochastic exponential, $Z = \mathcal{E}(cB)$ solves the stochastic integral equation $\mathrm{d}Z = cZ_-\mathrm{d}B$, and by Theorem 4.6.2, Z is a local martingale. In fact, Z is even a martingale (this can be shown using Novikov's condition).

Remark 5.2.6. (i) The solution of Eq. (5.7) is positive with probability 1 if and only if $\Delta X > -1$.

(ii) If X is a local martingale, then $\mathcal{E}(X)$ is also a local martingale. This follows from Theorem 4.6.2.

(iii) If X is a Lévy martingale and $\Delta X > -1$, then $\mathcal{E}(X)$ is a true martingale; this follows from the independent-increment structure, or from the standard exponential-martingale criterion for Lévy processes (see, e.g., [166]).

The exponential function can also be defined as a solution of the functional equation

$$f(x+y) = f(x)f(y), \quad f(0) = 1.$$

The following theorem transfers this to the stochastic exponential.

Theorem 5.2.7. *Let X and Y be two semimartingales with $X_0 = Y_0 = 0$. Then*

$$\mathcal{E}(X)\mathcal{E}(Y) = \mathcal{E}(X + Y + [X, Y]).$$

Proof. Let $U = \mathcal{E}(X)$ and $V = \mathcal{E}(Y)$, which means in particular $U = 1 + U_- \bullet X$ and thus $\mathrm{d}U = \mathrm{d}(U_- \bullet X)$. Using the integration by parts formula (Theorem 4.5.1) and Theorem 4.7.4, we obtain

$$\begin{aligned}
U_t V_t - U_0 V_0 &= \int_{0+}^t U_{s-}\mathrm{d}V_s + \int_{0+}^t V_{s-}\mathrm{d}U_s + [U,V]_t - [U,V]_0 \\
&= \int_{0+}^t U_{s-}\mathrm{d}(V_- \bullet Y)_s + \int_{0+}^t V_{s-}\mathrm{d}(U_- \bullet X)_s + \int_{0+}^t \mathrm{d}[U,V]_s \\
&= \int_{0+}^t U_{s-}V_{s-}\mathrm{d}Y_s + \int_{0+}^t U_{s-}V_{s-}\mathrm{d}X_s + \int_{0+}^t U_{s-}V_{s-}\mathrm{d}[X,Y]_s \\
&= \int_{0+}^t U_{s-}V_{s-}\mathrm{d}(X + Y + [X,Y])_s.
\end{aligned}$$

Setting $W_t = U_t V_t$, we have, since $U_0 V_0 = 1$,

$$W_t = 1 + \int_{0+}^{t} W_{s-} \mathrm{d}(X + Y + [X, Y])_s,$$

and therefore $W = \mathcal{E}(X + Y + [X, Y])$, which was to be shown. □

Corollary 5.2.8. *Let X be a continuous semimartingale with $X_0 = 0$. Then*

$$\frac{1}{\mathcal{E}(X)} = \mathcal{E}(-X + [X, X]).$$

Proof. By Theorem 5.2.7 with $A = X$ and $B = -X + [X, X]$,

$$\mathcal{E}(X)\mathcal{E}(-X + [X, X]) = \mathcal{E}\Big(X + (-X + [X, X]) + [X, -X + [X, X]]\Big).$$

Since X is continuous and $[X, X]$ is of finite variation, $[X, [X, X]] = 0$, so the cross term is $[X, -X + [X, X]] = [X, -X] = -[X, X]$. Hence the argument of $\mathcal{E}$ equals $X + (-X + [X, X]) - [X, X] = 0$, and since $\mathcal{E}(0) = 1$, the result follows. □

5.2.2 Lévy's Characterization Theorem

To conclude this chapter, we examine the standard Brownian motion once more. It is continuous and a local martingale. Furthermore, the quadratic variation of a standard Brownian motion B is given by $[B, B]_t = t$. It turns out that this property already uniquely characterizes the standard Brownian motion.

Theorem 5.2.9 (Lévy's Theorem)**.** *A stochastic process X is a standard Brownian motion if and only if $X_0 = 0$, X is a continuous local martingale, and*

$$[X, X]_t = t, \quad \textit{for all } t \geq 0.$$

Proof. $\Rightarrow$ By definition, Brownian motion is continuous and a martingale, hence in particular a local martingale. By direct calculation, $[B, B]_t = t$.

$\Leftarrow$ By assumption, X is already continuous and adapted. By definition, we need to show two things:

(i) $X_t - X_s$ is independent of $\mathcal{F}_s$, and

(ii) $X_t - X_s$ is $\mathcal{N}(0, t - s)$ distributed.

Let $u \in \mathbb{R}$ be fixed and $F(x,t) = \exp\left(\mathrm{i}ux + \frac{u^2 t}{2}\right)$. Then $F : \mathbb{R}^2 \to \mathbb{C}$ is twice continuously differentiable with

$$\begin{aligned}
\frac{\partial F}{\partial x} &= \mathrm{i}u \exp\left(\mathrm{i}ux + \frac{u^2 t}{2}\right) \\
\frac{\partial F}{\partial t} &= \frac{u^2}{2} \exp\left(\mathrm{i}ux + \frac{u^2 t}{2}\right) \\
\frac{\partial^2 F}{\partial x^2} &= -u^2 \exp\left(\mathrm{i}ux + \frac{u^2 t}{2}\right).
\end{aligned}$$

With $Z_t = F(X_t, t)$, by Theorem 4.7.4 we have

$$[X,t]^c = 0.$$

We apply the multi-dimensional Itô formula (Theorem 5.1.6) to the two-dimensional semimartingale (X_t, t), whose second component is continuous and of finite variation, separately to the real and imaginary parts of F:

$$\begin{aligned}
Z_t - Z_0 &= \mathrm{i}u \int_{0+}^{t} Z_s \mathrm{d}X_s + \frac{u^2}{2} \int_{0+}^{t} Z_s \mathrm{d}s + \frac{1}{2}(\mathrm{i}u)^2 \int_{0+}^{t} Z_s \mathrm{d}[X,X]_s \\
&= \mathrm{i}u \int_{0+}^{t} Z_s \mathrm{d}X_s.
\end{aligned}$$

From Theorem 4.6.2, it follows that Z is a continuous complex-valued local martingale, equivalently its real and imaginary parts are local martingales.

If we define a stopping time $T = t_0$ for some $t_0 > 0$, then Z^T is also a continuous local martingale, and

$$|Z_{t\wedge T}| = \left|\exp\left(\mathrm{i}uX_{t\wedge T} + (t \wedge t_0)\frac{u^2}{2}\right)\right| = \exp\left((t \wedge t_0)\frac{u^2}{2}\right) \leq \exp\left(\frac{t_0 u^2}{2}\right).$$

Since Z^T is bounded on $[0, t_0]$, it is a martingale on $[0, t_0]$, and as t_0 is arbitrary, Z is a martingale on every finite time interval. Therefore $\mathbf{E}[Z_t | \mathcal{F}_s] = Z_s$, which means

$$\mathbf{E}\left[\exp\left(\mathrm{i}uX_t + t\frac{u^2}{2}\right) \mid \mathcal{F}_s\right] = \exp\left(\mathrm{i}uX_s + s\frac{u^2}{2}\right).$$

Thus

$$\mathbf{E}\left[\exp(\mathrm{i}u(X_t - X_s)) \mid \mathcal{F}_s\right] = e^{-\frac{u^2(t-s)}{2}},$$

so $X_t - X_s$ is independent of $\mathcal{F}_s$. Furthermore, by the uniqueness theorem for

characteristic functions,

$$X_t - X_s \sim \mathcal{N}(0, t-s).$$

Therefore, X is a Brownian motion.

□

An analogous statement also exists for multi-dimensional Brownian motions.

Theorem 5.2.10 (Lévy's Theorem)**.** *An n-dimensional stochastic process $X = (X^1, \dots, X^n)$ is a standard Brownian motion if and only if $X_0 = 0$, X is a continuous local martingale, and*

$$[X^i, X^j]_t = t\delta_{ij}. \tag{5.10}$$

For this, we first show:

Lemma 5.2.11. *If B^1, B^2 are two independent Brownian motions, then $[B^1, B^2] = 0$.*

Proof. By the independence, $B^1 B^2$ is a martingale. Thus

$$[B^1, B^2] = B^1 B^2 - B^1_- \bullet B^2 - B^2_- \bullet B^1$$

is a continuous local martingale by Theorem 4.6.2. But $[B^1, B^2]$ is also a process of bounded variation, and thus by Theorem 4.7.3,

$$[B^1, B^2]_t = [B^1, B^2]_0 = B^1_0 B^2_0 = 0.$$

□

Proof of Theorem 5.2.10. $\Rightarrow$ We have already shown that the standard Brownian motion is a martingale, and with Lemma 5.2.11, it is clear that its quadratic variations/covariations satisfy condition (5.10).

$\Leftarrow$ Let $X = (X^1, \dots, X^n)$ be a continuous local martingale satisfying (5.10). For fixed $u \in \mathbb{R}^n$, we define the function $(x,t) = (x_1, \dots, x_n, t) \mapsto f(x_1, \dots, x_n, t) := \exp\left(iu^\top x + \frac{u^\top u}{2} t\right)$, and apply the real-valued Itô formula separately to the real and imaginary parts of f applied to the process $Z_t := f(X^1_t, \dots, X^n_t, t)$. With the

notation $\partial_k f = \frac{\partial f}{\partial x_k}$, we obtain

$$\begin{aligned}
Z_t &= Z_0 + \sum_{k=1}^{n} \int_0^t \partial_k f(X_s^1, \dots, X_s^n, s) \mathrm{d}X_s^k + \int_0^t \partial_{n+1} f(X_s^1, \dots, X_s^n, s) \mathrm{d}s \\
&\quad + \frac{1}{2} \sum_{1 \le l,k \le n} \int_0^t \partial_{kl} f(X_s^1, \dots, X_s^n, s) \mathrm{d}[X^k, X^l]_s \\
&= 1 + \sum_{k=1}^{n} \int_0^t \mathrm{i}u_k f(X_s^1, \dots, X_s^n, s) \mathrm{d}X_s^k + \frac{u^\top u}{2} \int_0^t f(X_s^1, \dots, X_s^n, s) \mathrm{d}s \\
&\quad + \frac{1}{2} \sum_{k=1}^{n} \int_0^t (\mathrm{i}u_k)^2 f(X_s^1, \dots, X_s^n, s) \mathrm{d}[X^k, X^k]_s \\
&= 1 + \mathrm{i} \sum_{k=1}^{n} \int_0^t u_k f(X_s^1, \dots, X_s^n, s) \mathrm{d}X_s^k, \quad \forall t \ge 0.
\end{aligned}$$

Since X is a local martingale, Z is also a local martingale as an integral with respect to a local martingale. Since $|Z_t| = \exp(\frac{1}{2} u^\top u\, t)$ is deterministic and bounded on each finite interval, Z is a martingale on every finite time interval, i.e.,

$$\mathbf{E}\left[\exp\left(\mathrm{i}u^\top X_t + \frac{u^\top u}{2} t\right) | \mathcal{F}_s\right] = \exp\left(\mathrm{i}u^\top X_s + \frac{u^\top u}{2} s\right), \quad \forall s \le t < \infty$$

and thus

$$\mathbf{E}\left[\exp(\mathrm{i}u^\top (X_t - X_s)) | \mathcal{F}_s\right] = \exp\left(-\frac{u^\top u}{2}(t-s)\right), \quad \text{for all } s \le t < \infty. \tag{5.11}$$

Since this holds for all $u \in \mathbb{R}^n$, $X_t - X_s$ is stochastically independent of the σ-algebra $\mathcal{F}_s$. Since the right-hand side of Eq. (5.11) is the characteristic function of an n-dimensional normal distribution with mean vector $(0, \dots, 0)$ and variance/covariance matrix $(t-s)I$, we have $(X_t - X_s) \sim \mathcal{N}((0, \dots, 0), (t-s)I)$, where I is the identity matrix. This proves the claim.

□

5.2.3 The Stochastic Logarithm

A special property of the exponential function is that it bijectively maps the space of real numbers onto the set of multiplicatively invertible elements of $\mathbb{R}$. A similar result holds for the stochastic exponential.

Lemma 5.2.12. *Let Z be a semimartingale such that $\inf_{s \in [0,t]} Z_s > 0$ for all $t \in \mathbb{R}_+$. Then there exists a semimartingale X, unique up to indistinguishability, with $X_0 = 0$ and*

$Z = Z_0\mathcal{E}(X)$. *It is given by*

$$X = \left(\frac{1}{Z}\right)_- \bullet Z.$$

Proof. Since $\inf_{s\in[0,t]} Z_s > 0$ for all t, the process $1/Z$ is càdlàg and locally bounded, hence $(1/Z)_-$ is predictable and locally bounded. In particular, $(1/Z)_- \in \mathcal{L}$, so $X = (1/Z)_- \bullet Z$ is well-defined.

Define $\tilde{Z} = \frac{Z}{Z_0}$. By Theorem 5.2.1, we only need to show that

$$\tilde{Z} = 1 + \tilde{Z}_- \bullet X \tag{5.12}$$

holds.

We have

$$\begin{aligned} 1 + \tilde{Z}_- \bullet X &= 1 + \frac{1}{Z_0}\left(Z_- \bullet \left(\left(\frac{1}{Z}\right)_- \bullet Z\right)\right) \\ &= 1 + \frac{1}{Z_0}\left(\left(\frac{Z_-}{Z_-}\right) \bullet Z\right) \\ &= 1 + \frac{1}{Z_0}(Z_t - Z_0) = \tilde{Z}_t, \end{aligned}$$

where the last step follows from stochastic integration of the constant integrand 1 on $(0, t]$. This proves Eq. (5.12). □

Chapter 6

Sigma Martingales

Sigma martingales are a class of stochastic processes that extend the concept of local martingales by broadening the localization procedure. While local martingales are processes that can be localized to uniformly integrable martingales using stopping times, sigma martingales generalize this idea through sequences of predictable sets. They exhibit distinctive properties, making them particularly valuable for addressing challenges in theoretical probability and financial mathematics. Notably, unlike the subset of local martingales, sigma martingales are closed under stochastic integration – a property that underscores their natural emergence in financial mathematics and is often used as an alternative definition for sigma martingales, distinct from the one adopted in this thesis.

In finance, sigma martingales play a central role in markets with no arbitrage opportunities and in modeling self-financing trading strategies. The First Fundamental Theorem of Asset Pricing [40] guarantees that with "No Free Lunch with Vanishing Risk" (NFLVR) satisfied, there exists an Equivalent sigma martingale Measure (ESMM), under which discounted price processes are sigma martingales. Even when price processes are modeled as local martingales (e.g., due to positivity constraints or continuity), stochastic integrals like $H \bullet M$, arising in self-financing trading strategies, are typically only sigma martingales and not local martingales.

The concept of sigma martingales was introduced as processes "de la classe (Σ_m)" in the pioneering work of Chou [26] and was further explored by Émery [58]. Despite their fundamental role in financial mathematics, sigma martingales have received relatively limited attention in the broader literature. For instance, the treatment in [166] is concise but somewhat superficial, whereas works such as [95] and [106] offer a more in-depth exploration. However, they heavily rely on the framework of semimartingale characteristics, which, while highly useful for establishing important properties, can also be mathematically demanding.

In this chapter we seek to bridge the gap in the literature by developing a comprehensive,

self-contained and accessible framework for sigma martingales without relying on the useful but technical apparatus of semimartingale characteristics. Instead, we demonstrate that all fundamental results about sigma martingales can be derived from first principles. While most major results from the first part are already known, the approach presented here is novel, more straightforward, and avoids the complex concept of semimartingale characteristics. This streamlined method leads to all major properties of sigma martingales and a new proof of the crucial Ansel–Stricker lemma, which states that one-sided bounded sigma martingales are local martingales. We use this new approach to prove a generalization, which, to the best of our knowledge, is so far only mentioned in a slightly different version in [74]. In the context of mathematical finance, this can be used to show that if a trading strategy is admissible (i.e., it satisfies specific boundedness conditions on the nominal price process), the discounted stochastic integral $H \bullet M$ is not merely a sigma martingale but also a local martingale, significantly simplifying the theoretical framework.

Additionally, this work elaborates on the major examples of sigma martingales that are widely known in the literature but are often only briefly mentioned without explicitly demonstrating their important properties or making use of more complex concepts (such as the above-mentioned characteristics of semimartingales), whereas the present approach avoids such complexity. By focusing on clear, intuitive arguments, we not only provide a new perspective on these examples but also ensure a more accessible and didactic exposition of sigma martingales.

In the second part, we prove the important Jacod–Yor theorem for sigma martingales, which is currently not available in the literature and thus fills a significant gap. In mathematical finance, this theorem has a wide range of applications, including the Second Fundamental Theorem of Asset Pricing and, more generally, the existence and construction of hedging strategies.

The content of this chapter is presented in the following two publications by the same author as this thesis:

- Moritz Sohns. "σ-Martingales: Foundations, Properties, and a New Proof of the Ansel–Stricker Lemma". In: *Mathematics* 13.4 (2025), p. 682. ISSN: 2227-7390. DOI: 10.3390/math13040682

- Moritz Sohns. "The Jacod–Yor Theorem for Sigma Martingales and the Second Fundamental Theorem of Asset Pricing". In: *Journal of Stochastic Analysis* 6.2 (2025), Article 1. DOI: 10.31390/josa.6.2.01

6.1 Definition and Properties

The conventions used throughout this chapter are those of Chapter 2: a complete filtered probability space $(\Omega, \mathcal{F}, \mathbb{F}, \mathbf{P})$ whose filtration $\mathbb{F} = (\mathcal{F}_t)_{t\geq 0}$ satisfies the usual conditions, all stochastic processes are càdlàg and considered modulo evanescence, $\mathcal{P}$ denotes the predictable σ-algebra, d_{ucp} the metric of uniform convergence on compacts in probability, and $\|\cdot\|_{\mathcal{S}}$ the Émery semimartingale metric of Chapter 4. The classes $\mathcal{M}^d$, $\mathcal{M}^d_\infty$, $\mathcal{M}^d_{\mathrm{loc}}$, the subscript convention $\mathcal{M}_{\mathrm{loc},0}$ for processes starting at 0, and the Hardy spaces $\mathcal{H}^p$ are as defined in Chapter 2. Absolute continuity and equivalence of measures are denoted $\mathbf{Q} \ll \mathbf{P}$ and $\mathbf{P} \sim \mathbf{Q}$. The stochastic integral $H \bullet X$ is the topological vector integral of Chapter 4, with integrand class $L(X)$.

The structure of this section is as follows. We start from the definition of a sigma martingale via a predictable localizing sequence, establish the basic stability properties of the class (closedness under stopping, stochastic integration, and orthogonal projection), and characterize sigma martingales as stochastic integrals against local martingales (Theorem 6.1.13). We establish an important result about sigma martingales: a sigma martingale is a local martingale if and only if it is locally integrable. The Ansel–Stricker Lemma 6.1.20 and the Generalized Ansel–Stricker Lemma 6.1.22, including equivalent jump-minorant criteria, follow from these results. This generalization is the result that allows the nominal-to-discounted transition in the market models of Chapter 7. The section closes with examples that show the limits of the result, including a sigma martingale with no equivalent local-martingale measure.

There are several equivalent definitions of sigma martingales in the literature. The definition adopted in this work was originally proposed by Goll, Kallsen, et al. [70] and later refined by other authors (e.g., [95, Definition III.6.33]). This definition emphasizes how sigma martingales extend local martingales through a broader localization framework. In contrast, works such as [40] and [166] define sigma martingales as processes that can be represented as stochastic integrals with respect to martingales. While this perspective underscores their crucial role in mathematical finance, it makes the connection to their generalization of local martingales less immediately apparent. From a didactic perspective, we find that the earlier definition - adopted here - provides a clearer and more intuitive introduction to the concept. In Theorem 6.1.13, we establish the equivalence of our definition with that of [40] and [166].

Definition 6.1.1. A one-dimensional semimartingale S is called a *sigma martingale* if there exists a sequence of sets $D_n \in \mathcal{P}$ such that:

(i) $D_n \subset D_{n+1}$ for all n;

(ii) $\bigcup_{n=1}^{\infty} D_n = \Omega \times \mathbb{R}_+$;

(iii) for any $n \geq 1$, the process $\mathbb{1}_{D_n} \bullet S$ is a uniformly integrable martingale.

Such a sequence $(D_n)_{n\in\mathbb{N}}$ is called a *σ-localizing sequence*. A d-dimensional semimartingale is called a *sigma martingale* if each of its components is a one-dimensional sigma martingale. By $\mathcal{M}_\sigma^d$, we denote the set of all d-dimensional sigma martingales.

First, observe that by setting $D_n := [\![0, T^n]\!]$, all local martingales are sigma martingales.

Theorem 6.1.2. *Every local martingale is a sigma martingale.*

Proof. Let M be a local martingale and $(T^n)_{n\in\mathbb{N}}$ a localizing sequence. Define $D_n := [\![0, T^n]\!]$. Since $\mathbb{1}_{[\![0,T^n]\!]} \bullet M = M^{T^n}$, it follows that M is also a sigma martingale. □

In discrete time, any sigma martingale is a local martingale. This follows from the fact that, in discrete time, any predictable set D_n can be expressed as a finite union of intervals of the form $]\!]T_k, T_n]\!]$, combined with the property that the set of local martingales forms a vector space. Alternatively, this result can also be derived using Theorem 6.1.13 and the observation that, in discrete time, the predictable integrand can be assumed to be locally bounded. The conclusion then follows directly from results such as [31, Theorem 12.3.3], which states that a stochastic integral with a locally bounded integrand and a local martingale integrator is again a local martingale, or [191, p. 10.7], which states that in discrete time any stochastic integral with a bounded integrand and a martingale integrator is again a martingale.

It turns out that $\mathcal{M}_{\mathrm{loc}}^d \subsetneq \mathcal{M}_\sigma^d$, as we will illustrate in the following example, which can also be found in [95, Example 6.40]. We revisit the example with a detailed demonstration and elaboration on the claimed properties.

Example 6.1.3. Let $(Y_n)_{n\in\mathbb{N}}$ be a sequence of independent random variables with

$$\mathbf{P}(Y_n = n) = \frac{1}{2n^2}, \quad \mathbf{P}(Y_n = -n) = \frac{1}{2n^2}, \quad \mathbf{P}(Y_n = 0) = 1 - \frac{1}{n^2}.$$

We put

$$X_t := \sum_{\{n\,;\,1-t\leq\frac{1}{n}\}} Y_n.$$

Then X is a well-defined sigma martingale but no local martingale with respect to the filtration created by X.

First, we have to show that X is well-defined. Therefore, we define $A_n = \{\omega \in \Omega\,;\, Y_n \neq 0\}$. Clearly, we have $\mathbf{P}(A_n) = \frac{1}{n^2}$ and hence $\sum_{n=1}^{\infty} \mathbf{P}(A_n) < \infty$. By the Borel-Cantelli lemma, we conclude $\mathbf{P}(\{Y_n \neq 0 \text{ for infinitely many } n\}) = 0$ and thus X is well-defined.

By setting

$$D_n := \Omega \times \left(\left[0, 1 - \frac{1}{n} \right] \cup [1, \infty) \right),$$

we obtain $\mathbb{1}_{D_n} \bullet X = \sum_{i=1}^{n} Y_i$ for each n. Since the sum is finite and all Y_n are symmetric and integrable, it is easy to see that D_n is a localizing sequence and X a sigma martingale.

Furthermore, X is not a local martingale. In order to show that, we assume $X \in \mathcal{M}^1_{\mathrm{loc}}$. Since X is a process with independent increments, we even have $X \in \mathcal{M}^1$ (see, for example, Medvegyev [133] Theorem 7.97). We put

$$A_n := \{Y_n = n\} \cap \bigcap_{\substack{k \geq 2 \\ k \neq n}} \{Y_k = 0\}, \qquad n \geq 2,$$

and by the independence of the random variables Y_k, we obtain for every $n \geq 2$

$$\mathbf{P}(A_n) = \mathbf{P}(Y_n = n) \prod_{\substack{k \geq 2 \\ k \neq n}} \mathbf{P}(Y_k = 0) \geq \mathbf{P}(Y_n = n) \prod_{k \geq 2} \mathbf{P}(Y_k = 0) = \frac{c}{2n^2}$$

with $c := \prod_{k \geq 2} \left(1 - \frac{1}{k^2}\right) = \frac{1}{2}$, [1] where the inequality uses that dropping the factor $\mathbf{P}(Y_n = 0) \leq 1$ only increases the product.

On A_n all Y_k with $k \geq 2$ and $k \neq n$ vanish, while $Y_n = n$ and $Y_1 \in \{-1, 1\}$, so $X_1 = Y_1 + n$ and hence $|X_1| \geq n - 1$. Since the sets A_n $(n \geq 2)$ are pairwise disjoint, monotone convergence gives

$$\mathbf{E}\left[|X_1|\right] \geq \mathbf{E}\left[|X_1| \sum_{n \geq 2} \mathbb{1}_{A_n} \right] = \sum_{n \geq 2} \mathbf{E}\left[|X_1| \, \mathbb{1}_{A_n}\right] \geq \sum_{n \geq 2} (n-1) \mathbf{P}(A_n) \geq \frac{c}{2} \sum_{n \geq 2} \frac{n-1}{n^2} = \infty,$$

where the last sum diverges because $\frac{n-1}{n^2} \geq \frac{1}{2n}$ for $n \geq 2$. This is a contradiction, and we conclude that X is not a local martingale.

The following example is a variant of the most prominent example for a sigma martingale that is not a local martingale. It is from Émery [58] and mentioned in most publications about sigma martingales (see, for example, [109, Example 9.29], [166] the example preceding Theorem IV.34 or [178, Example 5.2]).

Example 6.1.4. Let τ, ξ be independent random variables with $\mathbf{P}(\tau > t) = \exp(-2t)$ and $\mathbf{P}(\xi = 1) = \mathbf{P}(\xi = -1) = \frac{1}{2}$. We put $X_t = \mathbb{1}_{[\![\tau, \infty[\![} \frac{\xi}{\tau}$. Then, X is a sigma martingale but not a local martingale with respect to the filtration created by X.

[1] The product telescopes: $\prod_{k=2}^{N} (1 - \frac{1}{k^2}) = \prod_{k=2}^{N} \frac{(k-1)(k+1)}{k^2} = \frac{N+1}{2N} \to \frac{1}{2}$. The factor $k = 1$ must be excluded here, since $\mathbf{P}(Y_1 = 0) = 0$ would otherwise make the product vanish; the random variable $Y_1 \in \{-1, 1\}$ is instead accounted for separately below.

By putting $D_n = \{0\} \cup \left[\frac{1}{n}, \infty\right)$, we obtain

$$\mathbb{1}_{D_n} \bullet \left(\frac{\xi}{\tau} \mathbb{1}_{[\![\tau, \infty[\![}\right) = \begin{cases} \frac{\xi}{\tau} & \text{for } \tau \geq \frac{1}{n} \text{ and } t \geq \tau \\ 0 & \text{else} \end{cases}.$$

And it is easy to see that D_n is a localizing sequence.

But X is not a local martingale as we encounter integrability problems. We assume $X \in \mathcal{M}^1_{\text{loc}}$, and hence, there exists a stopping time $T > 0$ such that X^T is a uniformly integrable martingale.

Since X is constant on $[\![0, \tau[\![$, we deduce that T is constant on $\{\omega \in \Omega\,;\, T < \tau\}$. There exists an $\varepsilon > 0$ such that $T \geq \tau$ on $\{\omega \in \Omega\,;\, \tau < \varepsilon\}$. Hence we have $T \geq \tau \wedge \varepsilon$ and thus

$$\mathbf{E}\left[\left|X^T_\varepsilon\right|\right] = \int_0^\varepsilon \left|\frac{\xi}{t}\right| \mathrm{d}\mathbf{P}(\tau \leq t) = \int_0^\varepsilon \frac{1}{t} 2\exp(-2t)\mathrm{d}t = \infty.$$

So X^T is not a martingale, and thus X is not a local martingale.

Remark 6.1.5. It turns out that for both of the above examples, there exists an equivalent probability measure $\mathbf{Q}$ such that X is a $\mathbf{Q}$-martingale.

For the first example, assume a probability measure $\mathbf{Q}$ such that Y_n are independent random variables with

$$\mathbf{Q}(Y_n = n) = \frac{1}{2n^3}, \quad \mathbf{Q}(Y_n = -n) = \frac{1}{2n^3}, \quad \mathbf{Q}(Y_n = 0) = 1 - \frac{1}{n^3}.$$

Then we have

$$\mathbf{E}\,|X_t| \leq \mathbf{E}\,|X_1| = \mathbf{E}\left|\sum_{n=1}^\infty Y_n\right| \leq \mathbf{E}\left[\sum_{n=1}^\infty |Y_n|\right] = \sum_{n=1}^\infty n \cdot \frac{1}{n^3} = \sum_{n=1}^\infty \frac{1}{n^2} < \infty.$$

Furthermore, it is easy to see that we have $\mathbf{E}\,[X_t \mid \mathcal{F}_s] = X_s$. Hence, X is a martingale.

The equivalence of the original probability measure $\mathbf{P}$ and $\mathbf{Q}$ can be seen by constructing the Radon–Nikodym derivative. For each $n \in \mathbb{N}$, the laws of Y_n under $\mathbf{P}$ and $\mathbf{Q}$ are mutually absolutely continuous. In fact, defining

$$Z_n := \frac{\mathrm{d}\mathbf{Q}}{\mathrm{d}\mathbf{P}}(Y_n) = \begin{cases} \frac{\frac{1}{2n^3}}{\frac{1}{2n^2}} = \frac{1}{n}, & \text{if } Y_n \in \{n, -n\}, \\ \frac{1 - \frac{1}{n^3}}{1 - \frac{1}{n^2}}, & \text{if } Y_n = 0, \end{cases}$$

the overall Radon–Nikodym derivative is given by

$$\frac{\mathrm{d}\mathbf{Q}}{\mathrm{d}\mathbf{P}} = \prod_{n=1}^\infty Z_n.$$

Since by the Borel–Cantelli lemma only finitely many of the events $\{Y_n \neq 0\}$ occur $\mathbf{P}$-almost surely, the infinite product involves only finitely many factors differing from 1, and hence converges to a strictly positive random variable. This shows that $\frac{\mathrm{d}\mathbf{Q}}{\mathrm{d}\mathbf{P}}$ is well defined and strictly positive $\mathbf{P}$-almost surely, implying that $\mathbf{P}$ and $\mathbf{Q}$ are equivalent.

(Alternatively, one could verify the equivalence by applying Kakutani's theorem.)

For the second example, assume a probability measure $\mathbf{Q}$ such that τ and ξ are independent random variables and $\mathbf{P}(\xi \in A) = \mathbf{Q}(\xi \in A)$ for all $A \in \mathcal{B}$ and $\int_0^t \left|\frac{\xi}{s}\right| \mathrm{d}\mathbf{Q}(\tau \leq s) < \infty$ (for example, you can choose $\mathbf{Q}$ with $\mathbf{Q}(\tau > t) = \frac{1}{t+1}$ for all $t \geq 0$). Then we have $\mathbf{E_Q}\,|X_t| < \infty$ and $\mathbf{E_Q}[X_t \mid \mathcal{F}_s] = X_s$ and hence X is a martingale.

However, in general, such a probability measure does not necessarily exist. We will illustrate that in Example 6.1.26.

In the definition of sigma martingales, we referred to the localizing sequence D_n as a σ-localizing sequence.

To establish properties of sigma martingales, we prefer to work with a definition that is relatively "strong". However, when proving that a given process is a sigma martingale, it is more convenient to use criteria that appear "weak" or less restrictive but are nonetheless equivalent to the definition of sigma martingales.

We achieve this by potentially relaxing the conditions that the sequence $\mathbb{1}_{D_n} \bullet S$ must satisfy for a sigma martingale S. Consequently, we extend the notion of a σ-localizing sequence.

Definition 6.1.6. A sequence of sets $D_n \in \mathcal{P}$ is called a Σ-*localizing sequence* if:

(i) $D_n \subset D_{n+1}$ for all n;

(ii) $\bigcup_{n=1}^{\infty} D_n = \Omega \times \mathbb{R}_+$;

(iii) for any $n \geq 1$, the process $\mathbb{1}_{D_n} \bullet S$ is a local martingale.

The notion of the σ– (or Σ–) localizing sequence is new in the literature but is inspired by the procedure of σ-localization, which was first described by Jacod and Shiryaev [95] and Kallsen [106]. It does simplify some proofs since the following theorem holds.

Theorem 6.1.7. *Let S be a semimartingale. The following are equivalent:*

(i) The process S is a sigma martingale;

(ii) for S, there exists a σ-localizing sequence;

(iii) for S, there exists a Σ-localizing sequence;

(iv) for S, there exists a sequence $D_n \in \mathcal{P}$, such that $D_n \subset D_{n+1}$ and $\bigcup_{n=1}^{\infty} D_n = \Omega \times \mathbb{R}_+$ and for any $n \geq 1$ the process $\mathbb{1}_{D_n} \bullet S$ is a sigma martingale.

In order to prove this theorem, we need the following lemma.

Lemma 6.1.8. *Let $S \in \mathcal{S}^1$ and $D_n \in \mathcal{P}$ sets, which form a countable partition of $\Omega \times \mathbb{R}_+$ such that $\mathbb{1}_{D_n} \bullet S$ is a uniformly integrable martingale for any $n \geq 1$, then S is a sigma martingale.*

Proof. We put $\tilde{D}_n = \bigcup_{i=1}^{n} D_i$, then it is easy to see that $(\tilde{D}_n)_{n\in\mathbb{N}}$ is a σ-localizing sequence. □

Proof of Theorem 6.1.7. It suffices to prove the theorem for $S \in \mathcal{S}^1$. $(i) \Leftrightarrow (ii) \Rightarrow (iii) \Rightarrow (iv)$ is clear, so we just have to show $(iv) \Rightarrow (i)$.

Let S be a semimartingale, for which a sequence $(D_n)_{n\in\mathbb{N}}$ of predictable sets exists such that $D_n \subset D_{n+1}$ and $\bigcup_{n=1}^{\infty} D_n = \Omega \times \mathbb{R}_+$ and for which $S^n := \mathbb{1}_{D_n} \bullet S$ is a sigma martingale for any $n \geq 1$.

By assumption, for every $n \geq 1$ there exists a sequence $(D_{n,m})_{m\in\mathbb{N}}$, such that

$$S^{n,m} := \mathbb{1}_{D_{n,m}} \bullet S^n$$

is a uniformly integrable martingale for all m. By defining $\tilde{D}_{n,m} := D_{n,m} \cap (D_n \setminus D_{n-1})$, we get $\tilde{D}_{n,m} = D_{n,m} \cap D_n \cap (D_n \setminus D_{n-1})$ and

$$\begin{aligned}\tilde{S}^{n,m} &:= \mathbb{1}_{\tilde{D}_{n,m}} \bullet S = \left(\mathbb{1}_{D_{n,m}} \mathbb{1}_{D_n \setminus D_{n-1}} \mathbb{1}_{D_n}\right) \bullet S \\ &= \mathbb{1}_{D_n \setminus D_{n-1}} \bullet \left(\mathbb{1}_{D_{n,m}} \bullet S^n\right) = \mathbb{1}_{D_n \setminus D_{n-1}} \bullet S^{n,m}\end{aligned}$$

is a local martingale because stochastic integrals with bounded integrands and local martingale integrators are again local martingales. Thus for every pair (n, m) there exists a fundamental sequence $(T^{n,m,k})_{k\in\mathbb{N}\setminus 0}$. We put

$$\overline{D}_{n,m,1} = \tilde{D}_{n,m} \cap [\![0, T^{n,m,1}]\!], \quad \text{and} \quad \overline{D}_{n,m,k} = \tilde{D}_{n,m} \cap]\!]T^{n,m,k-1}, T^{n,m,k}]\!]$$

for $k = 2, 3, \ldots$ and all $(m, n) \in \mathbb{N}^2$. Now

$$\mathbb{1}_{\overline{D}_{n,m,1}} \bullet S = \left(\mathbb{1}_{\tilde{D}_{n,m}} \bullet S\right)^{T^{n,m,1}} = \left(\tilde{S}^{n,m}\right)^{T^{n,m,1}}$$

is a uniformly integrable martingale, and so is

$$\begin{aligned}\mathbb{1}_{\overline{D}_{n,m,k}} \bullet S &= \left(\mathbb{1}_{\tilde{D}_{n,m}} \bullet S\right)^{T^{n,m,k}} - \left(\mathbb{1}_{\tilde{D}_{n,m}} \bullet S\right)^{T^{n,m,k-1}} \\ &= \left(\tilde{S}^{n,m}\right)^{T^{n,m,k}} - \left(\tilde{S}^{n,m}\right)^{T^{n,m,k-1}}\end{aligned}$$

The sets $\overline{D}_{n,m,k}$ are predictable sets and form a countable partition of $\Omega \times \mathbb{R}_+$. Thus, by Lemma 6.1.8, S is a sigma martingale. □

The following corollary is immediate.

Corollary 6.1.9. *Every local sigma martingale X is a sigma martingale.*

The following result shows that the set of sigma martingales is closed under stochastic integration, as opposed to the set of local martingales.

Corollary 6.1.10. *Let $S \in \mathcal{M}_\sigma^d$ and $H \in L(S)$, then $H \bullet S$ is also a sigma martingale.*

Proof. Let $\tilde{S}^1, \dots, \tilde{S}^d$ be the components of S. Consider a Σ-localizing sequence $(D_n)_{n\in\mathbb{N}}$ and define $\tilde{D}_n = \{(\omega, t) \in \Omega \times \mathbb{R}_+ \,;\, \|H_t(\omega)\|_u \leq n\}$.

Since H is a predictable process, all of the $\tilde{D}_n$ lie in the predictable σ-algebra. Therefore, the sets $\overline{D}_n = D_n \cap \tilde{D}_n$ are predictable and we have $\overline{D}_n \subseteq \overline{D}_{n+1}$ and $\bigcup_{n=1}^\infty \overline{D}_n = \Omega \times \mathbb{R}_+$. By putting $H^n := H\mathbb{1}_{\tilde{D}_n}$ and $S^n := (\mathbb{1}_{D_n} \bullet \tilde{S}^1, \dots, \mathbb{1}_{D_n} \bullet \tilde{S}^d)^\top$, the process H^n is bounded and by the linearity of the integral, we obtain

$$\mathbb{1}_{\overline{D}_n} \bullet (H \bullet S) = (\mathbb{1}_{D_n}\mathbb{1}_{\tilde{D}_n}) \bullet (H \bullet S) = \left(\mathbb{1}_{\tilde{D}_n} H \mathbb{1}_{D_n}\right) \bullet S = H^n \bullet S^n$$

Hence, since S^n is a local martingale, $\mathbb{1}_{\overline{D}_n} \bullet (H \bullet S)$ is also a local martingale and thus $\overline{D}_n$ a Σ-localizing sequence. By Theorem 6.1.7, $H \bullet S$ is a sigma martingale. □

Corollary 6.1.11. *For a sigma martingale S with Σ-localizing sequence $(D_n)_{n\in\mathbb{N}}$ and a sequence of predictable sets $(\tilde{D}_n)_{n\in\mathbb{N}}$, which satisfies $D_n \subset D_{n+1}$ and $\bigcup_{n=1}^\infty D_n = \Omega \times \mathbb{R}_+$, $(\overline{D}_n)_{n\in\mathbb{N}} := (D_n \cap \tilde{D}_n)_{n\in\mathbb{N}}$ is also a Σ-localizing sequence.*

Proof. We have

$$\mathbb{1}_{\overline{D}_n} \bullet S = \left(\mathbb{1}_{D_n}\mathbb{1}_{\tilde{D}_n}\right) \bullet S = \left(\mathbb{1}_{\tilde{D}_n} \bullet (\mathbb{1}_{D_n} \bullet S)\right).$$

Since $(D_n)_{n\in\mathbb{N}}$ is a Σ-localizing sequence, $\mathbb{1}_{D_n} \bullet S$ is, by definition, a local martingale and since $\mathbb{1}_{\tilde{D}_n}$ is bounded, $\mathbb{1}_{\overline{D}_n} \bullet S$ is a local martingale for all n and $(\overline{D}_n)_{n\in\mathbb{N}}$ is a Σ-localizing sequence. □

Corollary 6.1.12. *The set of sigma martingales forms a vector space.*

Proof. Without loss of generality, we assume $d = 1$. Consider $\alpha, \beta \in \mathbb{R}$ and $X, Y \in \mathcal{M}_\sigma^1$ with Σ-localizing sequences $(D_n)_{n\in\mathbb{N}}$ for X and $(\tilde{D}_n)_{n\in\mathbb{N}}$ for Y. By Corollary 6.1.11 $(D_n \cap \tilde{D}_n)_{n\in\mathbb{N}}$ is a Σ-localizing sequence for both X and Y and we have

$$\mathbb{1}_{D_n\cap\tilde{D}_n} \bullet (\alpha X + \beta Y) = \alpha \left(\mathbb{1}_{D_n\cap\tilde{D}_n} \bullet X\right) + \beta \left(\mathbb{1}_{D_n\cap\tilde{D}_n} \bullet Y\right).$$

Since $\mathcal{M}_{\mathrm{loc}}^1$ is a vector space, $\mathbb{1}_{D_n\cap\tilde{D}_n} \bullet (\alpha X + \beta Y)$ is a local martingale and thus $D_n \cap \tilde{D}_n$ is a Σ-localizing sequence for $\alpha X + \beta Y$. Hence, $\alpha X + \beta Y$ is a sigma martingale. □

We now come to one of the main statements about sigma martingales. As mentioned earlier, $H \bullet M$ for $M \in \mathcal{M}_{\mathrm{loc}}$ is not necessarily a local martingale. Hence, we have a closer look at the class

$$\{H \bullet M \,;\, H \in L(M), M \in \mathcal{M}_{\mathrm{loc}}\}$$

and it turns out that this class corresponds exactly to the vector space of the sigma martingales. Furthermore, by proving this theorem, we also show that our definition of a sigma martingale is equivalent to the definition used in [40] and [166].

The theorem is mentioned in almost any publication about sigma martingales (for example, in [166, Theorem IV.89] or [95, Theorem 6.4.1]). Because of our different approach, the proof given here differs slightly from the one given in the above-mentioned literature.

Theorem 6.1.13. *Let $X = (X^1, \dots, X^d)$ be a d-dimensional semimartingale. The following are equivalent:*

(i) The process X is a sigma martingale.

(ii) There exists a strictly positive process $H \in \mathcal{P}$ and an $\mathcal{H}^1$-martingale

$$M = (M^1, \dots, M^d)$$

with

$$X^i = H \bullet M^i \quad \text{for all } i \in \{1, \dots, d\}.$$

(iii) There exists a strictly positive process $H \in \mathcal{P}$ and a martingale

$$M = (M^1, \dots, M^d)$$

with

$$X^i = H \bullet M^i \quad \text{for all } i \in \{1, \dots, d\}.$$

(iv) There exists a strictly positive process $H \in \mathcal{P}$ and a local martingale

$$M = (M^1, \dots, M^d)$$

with

$$X^i = H \bullet M^i \quad \text{for all } i \in \{1, \dots, d\}.$$

(v) There exists a local martingale $M = (M^1, \dots, M^d)$ and a predictable process $H = (H^1, \dots, H^d)$ with

$$H^i \in L(M^i) \quad \text{and} \quad X^i = H^i \bullet M^i \quad \text{for all } i \in \{1, \dots, d\}.$$

Proof. The implications $(ii) \Rightarrow (iii) \Rightarrow (iv) \Rightarrow (v)$ are clear.

$(i) \Rightarrow (ii)$. By assumption, there exists a Σ-localizing sequence $(D_n)_{n\in\mathbb{N}}$ for X, so that for every n and every $i \in \{1, \dots, d\}$ the process $\mathbb{1}_{D_n} \bullet X^i$ is a local martingale. By Theorem 2.2.9, every local martingale is locally in $\mathcal{H}^1$, so for each n and i there exists an increasing sequence of stopping times $(T^{n,i,m})_{m\in\mathbb{N}} \uparrow \infty$ a.s. with $(\mathbb{1}_{D_n} \bullet X^i)^{T^{n,i,m}} \in \mathcal{H}^1$.

Set $T^{n,m} := \min_{1\le i\le d} T^{n,i,m}$; then $(T^{n,m})_m$ is an increasing sequence of stopping times tending to infinity, and for every $i \in \{1, \dots, d\}$,

$$\left(\mathbb{1}_{D_n} \bullet X^i\right)^{T^{n,m}} \in \mathcal{H}^1 \quad \text{for all } n, m \in \mathbb{N}.$$

Put $T^{n,0} := 0$, and define for each $n, m \in \mathbb{N}$ and $i \in \{1, \dots, d\}$ the $\mathcal{H}^1$-martingale increment

$$M^{n,m,i} := \left(\mathbb{1}_{D_n} \bullet X^i\right)^{T^{n,m}} - \left(\mathbb{1}_{D_n} \bullet X^i\right)^{T^{n,m-1}} = \mathbb{1}_{]\!]T^{n,m-1}, T^{n,m}]\!] \cap D_n} \bullet X^i \in \mathcal{H}^1.$$

Choose constants $\alpha_{n,m} > 0$ so that

$$\max_{1\le i\le d} \left\| \alpha_{n,m}\, M^{n,m,i} \right\|_{\mathcal{H}^1} \le 2^{-(n+m)} \qquad \text{for all } n, m \in \mathbb{N}.$$

Define the strictly positive predictable process

$$\tilde{K} := \mathbb{1}_{[\![0]\!]} + \sum_{n=1}^{\infty} \sum_{m=1}^{\infty} \alpha_{n,m} \mathbb{1}_{]\!]T^{n,m-1}, T^{n,m}]\!] \cap D_n},$$
$$N^i := \tilde{K} \bullet X^i, \quad i = 1, \dots, d.$$

Since $\|\alpha_{n,m} M^{n,m,i}\|_{\mathcal{H}^1} \le 2^{-(n+m)}$, the double series $\sum_{n,m} \alpha_{n,m} M^{n,m,i}$ converges absolutely in $\mathcal{H}^1$, hence in $\mathcal{H}^1$ by completeness (Lemma 2.2.14). The limit is exactly N^i. Therefore $N = (N^1, \dots, N^d)$ is an $\mathcal{H}^1$-martingale. The process $\tilde{K}$ is strictly positive almost everywhere

on $\bigcup_n D_n = \Omega \times \mathbb{R}_+$, so $K := 1/\tilde{K}$ is well defined and predictable, and by associativity (Theorem 4.4.18) $X^i = K \bullet N^i$ for all i.

$(v) \Rightarrow (i)$. As every martingale is locally in $\mathcal{H}^1$, for each i, there exists an increasing sequence of stopping times $(T_i^n)_{n\in\mathbb{N}}$ tending to infinity, such that $(M^i)^{T_i^n} \in \mathcal{H}^1$. Hence, we can construct an increasing sequence of stopping times $(T^n)_{n\in\mathbb{N}}$ tending to infinity, such that $(M^i)^{T^n} \in \mathcal{H}^1$ for all $n \in \mathbb{N}$ and for all $i \in \{1, \dots, d\}$. We define

$$D_n := \bigcap_{i=1}^{d} \left\{ \left| H^i \right| \leq n \right\} \cap [\![0, T^n]\!] \subset \Omega \times \mathbb{R}_+$$

(an increasing sequence of predictable sets with $\bigcup_n D_n = \Omega \times \mathbb{R}_+$). Hence, we have

$$\mathbb{1}_{D_n} \bullet X^i = \left(\mathbb{1}_{D_n} H^i \right) \bullet M^i = \left(\mathbb{1}_{D_n} H^i \right) \bullet (M^i)^{T^n}.$$

Because of $(\mathbb{1}_{D_n} H^i) \leq n$ and because the process $(M^i)^{T^n}$, stopped at T^n, is a $\mathcal{H}^1$ martingale, $\mathbb{1}_{D_n} \bullet X^i$ is for all n a local martingale, as a stochastic integral with a locally bounded integrand and a local martingale integrator is again a local martingale by Theorem 4.6.2. Therefore, (D_n) is a Σ-localizing sequence for X, and thus X is a sigma martingale. □

In the following, we are going to examine how sigma martingales behave when they are bounded from one-side by a sequence of random variables. The following lemma will be instrumental. It is based on ideas by De Donno and Pratelli [38].

Lemma 6.1.14. *Let $T < \infty$ be deterministic. Let X be an adapted càdlàg process with $X_0 \in L^1$, and let $(M^n)_{n\in\mathbb{N}}$ be a sequence of càdlàg martingales satisfying $M_0^n = X_0$ for all n. Assume that*

$$\sup_{0\leq s\leq T} |M_s^n - X_s| \longrightarrow 0 \quad \textit{in probability.}$$

Assume further that there exist stopping times $\eta_m \uparrow T$ almost surely and integrable random variables $\theta_m \in L^1$ such that

$$X_{s\wedge\eta_m} \geq \theta_m, \qquad 0 \leq s \leq T.$$

Finally, assume that for every stopping time $\tau \leq T$ and every $n \in \mathbb{N}$,

$$(\Delta M_\tau^n)^- \leq (\Delta X_\tau)^-.$$

Then X is a local supermartingale on $[0, T]$.

Proof. By passing to a subsequence and relabelling it, we may assume that

$$\sum_{n=1}^{\infty} \mathbf{P}\left(\sup_{0\le s\le T} |M_s^n - X_s| > \frac{1}{n} \right) < \infty.$$

For $n \in \mathbb{N}$, define the error time

$$\varepsilon_n := \inf\left\{ s \le T : \sup_{0\le u\le s} |M_u^n - X_u| > \frac{1}{n} \right\} \wedge T.$$

By the Borel–Cantelli lemma,

$$\inf_{n\ge m} \varepsilon_n \uparrow T \quad \text{almost surely as } m \to \infty.$$

For $m \in \mathbb{N}$, define

$$\rho_m := \inf\{ s \le T : \ X_s > m \} \wedge T$$

and

$$\sigma_m := \eta_m \wedge \rho_m \wedge \inf_{n\ge m} \varepsilon_n.$$

Since X is càdlàg, it is pathwise bounded on $[0,T]$, and hence $\rho_m \uparrow T$ almost surely. Together with $\eta_m \uparrow T$ and the preceding convergence of the error times, this gives

$$\sigma_m \uparrow T \quad \text{almost surely.}$$

Fix $m \in \mathbb{N}$ and $n \ge m$. For every $s < \sigma_m$, we have

$$|M_s^n - X_s| \le \frac{1}{n} \le 1, \qquad X_s \le m, \qquad X_s \ge \theta_m.$$

Hence

$$M_s^n \ge \theta_m - 1, \qquad s < \sigma_m.$$

At the terminal time σ_m, the lower bound for X and the pre-stopping upper bound give

$$X_{\sigma_m} \ge \theta_m, \qquad X_{\sigma_m -} \le m.$$

Consequently

$$(\Delta X_{\sigma_m})^- = (X_{\sigma_m -} - X_{\sigma_m})^+ \le (m - \theta_m)^+.$$

By the negative-jump domination assumption,

$$(\Delta M_{\sigma_m}^n)^- \le (\Delta X_{\sigma_m})^- \le (m - \theta_m)^+.$$

Therefore

$$M^n_{\sigma_m} \geq M^n_{\sigma_m -} - (\Delta M^n_{\sigma_m})^- \geq \theta_m - 1 - (m - \theta_m)^+.$$

Set

$$L_m := \theta_m - 1 - (m - \theta_m)^+.$$

Then $L_m \in L^1$, and for every $n \geq m$,

$$M^n_{s\wedge\sigma_m} \geq L_m, \qquad 0 \leq s \leq T.$$

We first show integrability. Since $M^n_{s\wedge\sigma_m} - L_m \geq 0$, Fatou's lemma and the martingale property of $(M^n)^{\sigma_m}$ give

$$\begin{aligned}\mathbf{E}[X_{s\wedge\sigma_m} - L_m] &\leq \liminf_{n\to\infty} \mathbf{E}\Big[M^n_{s\wedge\sigma_m} - L_m\Big] \\ &= \liminf_{n\to\infty} \mathbf{E}[M^n_0 - L_m] \\ &= \mathbf{E}[X_0 - L_m] < \infty.\end{aligned}$$

Since $L_m \in L^1$, it follows that $X_{s\wedge\sigma_m} \in L^1$ for all $s \leq T$.

Now let $0 \leq r \leq s \leq T$. By the construction of σ_m, $M^n_{\cdot\wedge\sigma_m} \to X_{\cdot\wedge\sigma_m}$ uniformly on $[0,T]$ almost surely along the selected subsequence. The conditional Fatou lemma applied to the non-negative random variables $M^n_{s\wedge\sigma_m} - L_m$ gives

$$\mathbf{E}[X_{s\wedge\sigma_m} - L_m \mid \mathcal{F}_r] \leq \liminf_{n\to\infty} \mathbf{E}\Big[M^n_{s\wedge\sigma_m} - L_m \mid \mathcal{F}_r\Big].$$

Using the martingale property of $(M^n)^{\sigma_m}$, the right-hand side equals

$$\liminf_{n\to\infty} \Big(M^n_{r\wedge\sigma_m} - \mathbf{E}[L_m \mid \mathcal{F}_r]\Big) = X_{r\wedge\sigma_m} - \mathbf{E}[L_m \mid \mathcal{F}_r].$$

Cancelling $\mathbf{E}[L_m \mid \mathcal{F}_r]$ yields

$$\mathbf{E}[X_{s\wedge\sigma_m} \mid \mathcal{F}_r] \leq X_{r\wedge\sigma_m}.$$

Therefore X^{σ_m} is a supermartingale on $[0,T]$. Since $\sigma_m \uparrow T$ almost surely, X is a local supermartingale on $[0,T]$. □

The following statement is, to the best of our knowledge, not mentioned in the literature. It proves to be quite valuable to avoid a detour into the special semimartingale theory.

Theorem 6.1.15. *A one-dimensional sigma martingale X is a local martingale if and only if it is locally integrable in the following sense: there exists a sequence of stopping times $(T_n)_{n\in\mathbb{N}}$ with $T_n \uparrow \infty$ almost surely such that, for every $n \in \mathbb{N}$ and every finite horizon*

$t \geq 0$,

$$(X^{T_n})^*_t := \sup_{0 \leq s \leq t} |X_{s \wedge T_n}| \in L^1.$$

Equivalently, it is enough to require this condition for every $n \in \mathbb{N}$ and every integer horizon $m \in \mathbb{N}$.

Proof. Every local martingale is locally in $\mathcal{H}^1$ by Theorem 2.2.9, hence locally integrable.

Conversely, let X be a one-dimensional locally integrable sigma martingale. Fix a deterministic horizon $T < \infty$. By local integrability, there are stopping times $\eta_m \uparrow T$ almost surely and random variables $K_m \in L^1$ such that

$$\sup_{0 \leq s \leq T} |X_{s \wedge \eta_m}| \leq K_m, \qquad m \in \mathbb{N}.$$

In particular, $X_0 \in L^1$.

Let $(D_n)_{n \in \mathbb{N}}$ be a Σ-localizing sequence for X. By replacing D_n with $D_n \cup (\Omega \times \{0\})$, we may and do assume that $\Omega \times \{0\} \subset D_n$ for every n. This only adds the integrable $\mathcal{F}_0$-measurable constant X_0 at the initial time and therefore preserves the local-martingale property. Put

$$X^n := \mathbb{1}_{D_n} \bullet X.$$

Then X^n is a local martingale and $X^n_0 = X_0$ for every n. Since $\mathbb{1}_{D_n} \to 1$ pointwise and the integrands are bounded, dominated convergence for stochastic integrals gives

$$X^n \longrightarrow X \qquad \text{ucp on } [0, T].$$

For every n, choose a stopping time ρ_n such that $(X^n)^{\rho_n}$ is a martingale on $[0, T]$ and

$$\mathbf{P}(\rho_n < T) \leq \frac{1}{n}.$$

Set

$$M^n := (X^n)^{\rho_n}.$$

Then M^n is a càdlàg martingale on $[0, T]$, $M^n_0 = X_0$, and

$$\sup_{0 \leq s \leq T} |M^n_s - X_s| \longrightarrow 0 \qquad \text{in probability.}$$

Indeed, on $\{\rho_n \geq T\}$ one has $M^n = X^n$ on $[0, T]$, while $\mathbf{P}(\rho_n < T) \to 0$.

For every stopping time $\tau \leq T$,

$$\Delta M^n_\tau = \mathbb{1}_{\{\tau \leq \rho_n\}} \mathbb{1}_{D_n}(\tau) \Delta X_\tau.$$

Consequently,

$$(\Delta M^n_\tau)^- \le (\Delta X_\tau)^-.$$

The lower-bound hypothesis in Lemma 6.1.14 is satisfied with $\theta_m := -K_m$, because

$$X_{s\wedge\eta_m} \ge -K_m, \qquad 0 \le s \le T.$$

Thus Lemma 6.1.14 implies that X is a local supermartingale on $[0,T]$.

Apply the same argument to $-X$, using the martingales $-M^n$. For every stopping time $\tau \le T$,

$$\Delta(-M^n)_\tau = \mathbb{1}_{\{\tau\le\rho_n\}}\mathbb{1}_{D_n}(\tau)\Delta(-X)_\tau,$$

and hence

$$(\Delta(-M^n)_\tau)^- \le (\Delta(-X)_\tau)^-.$$

The lower bound is again provided by local integrability, since

$$(-X)_{s\wedge\eta_m} \ge -K_m, \qquad 0 \le s \le T.$$

Therefore $-X$ is a local supermartingale on $[0,T]$.

Choose localizing sequences for X and $-X$ as local supermartingales and stop them simultaneously. Then there are stopping times $\gamma_m \uparrow T$ such that both X^{γ_m} and $(-X)^{\gamma_m}$ are supermartingales on $[0,T]$. For $0 \le r \le s \le T$ this gives

$$\mathbf{E}[X_{s\wedge\gamma_m} \mid \mathcal{F}_r] \le X_{r\wedge\gamma_m}$$

and, from the supermartingale property of $-X^{\gamma_m}$,

$$\mathbf{E}[X_{s\wedge\gamma_m} \mid \mathcal{F}_r] \ge X_{r\wedge\gamma_m}.$$

Hence equality holds, so X^{γ_m} is a martingale on $[0,T]$. Thus X is a local martingale on $[0,T]$. Since $T<\infty$ was arbitrary, a diagonal localization over integer horizons yields that X is a local martingale on $\mathbb{R}_+$. The multidimensional case follows componentwise. □

As every continuous semimartingale is locally integrable, the following corollary is immediate.

Corollary 6.1.16. *Every continuous sigma martingale is a local martingale.*

Remark 6.1.17. A similar but better known criterion than Theorem 6.1.15 is given through the theory of special semimartingales: any sigma martingale that is also a special semi-

martingale[2] is a local martingale ([31, Corollary 12.3.20] or [166, Theorem IV.91]), and a semimartingale is a special semimartingale if and only if its supremum process is locally integrable (see, for example, [31, Theorem 11.6.10] or [79, Theorem 8.6]). Composing these two results yields a similar criterion.

Lemma 6.1.18. *Let $T < \infty$ be deterministic. Let Y be a càdlàg local supermartingale on $[0, T]$. Assume that there exist stopping times $\eta_m \uparrow T$ almost surely and random variables $\theta_m \in L^1$ such that*

$$Y_{s \wedge \eta_m} \geq \theta_m, \qquad 0 \leq s \leq T.$$

Then Y is locally integrable on $[0, T]$.

Proof. Let $(R_j)_{j \in \mathbb{N}}$ be a localizing sequence such that Y^{R_j} is a supermartingale on $[0, T]$. Fix $m, j \in \mathbb{N}$, and put

$$U := Y^{\eta_m \wedge R_j}.$$

Then U is a supermartingale and $U_s \geq \theta_m$ for $0 \leq s \leq T$. Let

$$L_s := \mathbf{E}[\theta_m \mid \mathcal{F}_s], \qquad 0 \leq s \leq T,$$

where L is chosen càdlàg. Since $\theta_m \leq U_s$ and U_s is $\mathcal{F}_s$-measurable, we have $L_s \leq U_s$. Thus

$$Z := U - L$$

is a non-negative supermartingale.

For $p \in \mathbb{N}$, define

$$\alpha_p := \inf\{s \leq T : Z_s \geq p\} \wedge T, \qquad \beta_p := \inf\{s \leq T : |L_s| \geq p\} \wedge T.$$

Since Z is càdlàg and non-negative,

$$\sup_{0 \leq s \leq T} Z_{s \wedge \alpha_p} \leq p + Z_{\alpha_p}.$$

The stopped value Z_{α_p} is integrable by optional sampling for non-negative supermartingales. Similarly,

$$\sup_{0 \leq s \leq T} |L_{s \wedge \beta_p}| \leq p + |L_{\beta_p}|,$$

and L_{β_p} is integrable because L is a uniformly integrable martingale on the finite horizon $[0, T]$.

[2] see, for example, [31, Definition 11.6.9] for the definition of a special semimartingale

Consequently,

$$\sup_{0\le s\le T} |U_{s\wedge\alpha_p\wedge\beta_p}| \le 2p + Z_{\alpha_p} + |L_{\beta_p}| \in L^1.$$

To make the last localization explicit, choose triples (m_q, j_q, p_q) so that the corresponding stopping times

$$\delta_q := \eta_{m_q} \wedge R_{j_q} \wedge \alpha_{p_q}^{m_q,j_q} \wedge \beta_{p_q}^{m_q,j_q}$$

satisfy $\mathbf{P}(\delta_q < T) \le 2^{-q}$. This is possible because, for fixed m, j, the stopping times $\alpha_p^{m,j} \wedge \beta_p^{m,j}$ increase to T as $p \to \infty$, and because $\eta_m \uparrow T$ and $R_j \uparrow T$. By Borel–Cantelli, $\delta_q \to T$ almost surely. Setting $\gamma_n := \inf_{q\ge n} \delta_q$ gives an increasing sequence $\gamma_n \uparrow T$ a.s.; moreover $\gamma_n \le \delta_n$, so the integrable supremum bound for Y^{δ_n} also bounds Y^{γ_n}. Hence Y is locally integrable on $[0, T]$. □

Theorem 6.1.19 (Lower-minorant sigma martingale theorem)**.** *Let X be a one-dimensional sigma martingale. Suppose there exist stopping times $T_k \uparrow \infty$ and, for every k and every finite horizon $t \ge 0$, an integrable random variable $G_{k,t} \in L^1$ such that*

$$X_{s\wedge T_k} \ge -G_{k,t}, \qquad 0 \le s \le t.$$

Then X is a local martingale.

Proof. Fix $k \in \mathbb{N}$ and a deterministic horizon $t < \infty$. It is enough to prove that $Y := X^{T_k}$ is a local martingale on $[0, t]$.

Let $(D_n)_{n\in\mathbb{N}}$ be a Σ-localizing sequence for X. As in the proof of Theorem 6.1.15, we may assume $\Omega \times \{0\} \subset D_n$ for all n. Put

$$X^n := \mathbb{1}_{D_n} \bullet X.$$

Then X^n is a local martingale, $X_0^n = X_0$, and

$$X^n \longrightarrow X \qquad \text{ucp on } [0, t].$$

For each n, choose a stopping time ρ_n such that $(X^n)^{\rho_n}$ is a martingale on $[0, t]$ and

$$\mathbf{P}(\rho_n < t) \le \frac{1}{n}.$$

Define

$$M^n := (X^n)^{T_k\wedge\rho_n}.$$

Then M^n is a càdlàg martingale on $[0,t]$, $M^n_0 = Y_0 = X_0$, and

$$\sup_{0\le s\le t} |M^n_s - Y_s| \longrightarrow 0 \qquad \text{in probability.}$$

Indeed, on $\{\rho_n \ge t\}$ one has $M^n = (X^n)^{T_k}$ on $[0,t]$, while $\mathbf{P}(\rho_n < t) \to 0$.

For every stopping time $\tau \le t$,

$$\Delta M^n_\tau = \mathbb{1}_{\{\tau \le T_k \wedge \rho_n\}} \mathbb{1}_{D_n}(\tau) \Delta X_\tau,$$

whereas

$$\Delta Y_\tau = \mathbb{1}_{\{\tau \le T_k\}} \Delta X_\tau.$$

Thus ΔM^n_τ is either 0 or ΔY_τ, and consequently

$$(\Delta M^n_\tau)^- \le (\Delta Y_\tau)^-.$$

The lower-bound assumption gives

$$Y_s = X_{s\wedge T_k} \ge -G_{k,t}, \qquad 0 \le s \le t.$$

Therefore the hypotheses of Lemma 6.1.14 are satisfied on $[0,t]$ with $\eta_m \equiv t$ and $\theta_m = -G_{k,t}$. Hence $Y = X^{T_k}$ is a local supermartingale on $[0,t]$.

By Lemma 6.1.18, the same lower bound implies that Y is locally integrable on $[0,t]$. Since $Y = X^{T_k} = \mathbb{1}_{[\![0,T_k]\!]} \bullet X$ is again a sigma martingale by Corollary 6.1.10, the finite-horizon form of Theorem 6.1.15 implies that Y is a local martingale on $[0,t]$.

Since k and t were arbitrary, a diagonal localization over k and over integer horizons gives that X is a local martingale on $\mathbb{R}_+$. □

The following theorem is of principal importance in financial mathematics. It can be found in many publications on financial mathematics using the semimartingale terminology (not only is it mentioned in almost all of the publications we mentioned frequently in this work, such as [39, 40, 41] and [178] but it is also mentioned in many textbooks dealing with the different aspects of financial mathematics such as [161, 111, 61, 156]). However, to our knowledge, the only published proofs are the French-language original publication [3, Corollaire 3.5] and the more recent [38]. Theorem 6.1.15 enables us to give an alternative proof.

The proof of Theorem 6.1.20 below reduces the Ansel–Stricker Lemma to the lower-minorant sigma martingale theorem 6.1.19 proved above. The constant-minorant case of the latter is recorded here as a corollary for convenient reference.

Theorem 6.1.20 (Ansel–Stricker)**.** *A one-sided bounded sigma martingale X is a local martingale. If X is bounded from below (resp. above), it is also a supermartingale (resp. submartingale).*

Proof. It is enough to consider the case where X is bounded from below; the upper-bounded case follows by applying the lower-bounded case to $-X$, and the multidimensional case follows componentwise. Assume $X \geq -c$ for some deterministic constant $c \geq 0$.

By Theorem 6.1.19, X is a local martingale.

A lower-bounded local martingale is a supermartingale by the standard Fatou argument. Indeed, let (T_k) be a localizing sequence such that X^{T_k} is a martingale. For $0 \leq s \leq t$ and $A \in \mathcal{F}_s$,

$$\mathbf{E}\Big[\mathbb{1}_A\Big(X_{t\wedge T_k} - X_{s\wedge T_k}\Big)\Big] = 0.$$

Since $X \geq -c$, the random variables

$$\mathbb{1}_A\Big(X_{t\wedge T_k} - X_{s\wedge T_k}\Big) + \mathbb{1}_A(X_s + c)$$

are nonnegative. Fatou's lemma applied as $k \to \infty$ gives

$$\mathbf{E}[\mathbb{1}_A(X_t - X_s)] \leq 0.$$

Hence X is a supermartingale. The submartingale assertion for X bounded above follows by applying the preceding argument to $-X$. □

In finite, discrete-time, any non-negative local martingale that is bounded from below is even a martingale and not just a supermartingale. The difference in continuous time is that $\mathbf{E}[Y_t] < \infty$ for all t does not imply $\mathbf{E}[\sup_t Y_t] < \infty$ (not even on compacts). Thus, there is no integrable pointwise majorant which would be needed to prove the martingale property.

Remark 6.1.21. There are three other approaches to proving the Ansel–Stricker Lemma in the literature.

The original proof by Ansel and Stricker [3]: This approach relies on the classification and control of jumps and the theory of special semimartingales. It involves intricate technical calculations and constructs a sequence that converges to the integral $H \bullet M$, ensuring that this sequence is bounded from below by integrable random variables.

The proof by De Donno and Pratelli [38]: This method categorizes jumps into positive and negative components to manage the stochastic integral's behavior. It leverages Fatou's Lemma and Lebesgue's Theorem to control approximation sequences, thereby circumventing the complexities associated with special semimartingales.

The proof by Gushchin [74]: Here, the lemma by Ansel–Stricker appears as a corollary, a more general result that is very close to Theorem 6.1.22. This approach relies extensively on the theory of special semimartingales, utilizing their structural properties to establish the result.

As mentioned earlier, the Ansel–Stricker Lemma is widely referenced in the literature on mathematical finance, particularly in the context of semimartingale models. However, our proof shows that with minimal additional effort, the result can be further generalized. This generalization is crucial for the analysis of general semimartingale market models. Surprisingly, this result does not seem to be widely known. To the best of our knowledge, only in [74, Theorem 3.21] a similar result is mentioned.

Theorem 6.1.22 (Generalized Ansel–Stricker Lemma)**.** *Let X be a one-dimensional sigma martingale. The following statements are equivalent:*

(i) *X is a local martingale.*

(ii) *There exist a local martingale M and a càdlàg adapted process A with A^* locally integrable such that*
$$X \geq M + A.$$

(iii) *There exist a local martingale M and a càdlàg adapted process A of locally integrable variation such that*
$$X \geq M + A.$$

(iv) *There exist a local martingale M and a càdlàg adapted process A with A^* locally integrable such that*
$$\Delta X \geq \Delta M + \Delta A.$$

(v) *There exist a local martingale M and a càdlàg adapted process A of locally integrable variation such that*
$$\Delta X \geq \Delta M + \Delta A.$$

Proof. The implications $(i) \Rightarrow (ii)$, $(i) \Rightarrow (iii)$, $(i) \Rightarrow (iv)$ and $(i) \Rightarrow (v)$ are immediate by taking $M = X$ and $A \equiv 0$. Moreover, $(iii) \Rightarrow (ii)$ and $(v) \Rightarrow (iv)$ follow because a càdlàg process of locally integrable variation has locally integrable supremum process: if $\mathbf{Var}(A)$ denotes the total-variation process of A, then

$$A_t^* \leq |A_0| + \mathbf{Var}(A)_t.$$

We prove $(ii) \Rightarrow (i)$. Assume

$$X \geq M + A,$$

where M is a local martingale and A^* is locally integrable. Set

$$Y := X - M.$$

Since sigma martingales form a vector space and every local martingale is a sigma martingale, Y is a sigma martingale. Moreover, $Y \geq A$. Choose stopping times $T_k \uparrow \infty$ such that

$$A^*_{t\wedge T_k} \in L^1 \qquad \text{for every finite horizon } t \geq 0.$$

Then, for $0 \leq s \leq t$,

$$Y_{s\wedge T_k} \geq A_{s\wedge T_k} \geq -A^*_{t\wedge T_k} =: -G_{k,t},$$

where $G_{k,t} \in L^1$. Thus Y satisfies the hypotheses of Theorem 6.1.19, and hence Y is a local martingale. Since M is a local martingale, $X = Y + M$ is a local martingale.

It remains to prove $(iv) \Rightarrow (i)$. Assume

$$\Delta X \geq \Delta M + \Delta A,$$

where M is a local martingale and A^* is locally integrable. Since A is adapted, the process introduced below is adapted and càdlàg. Put

$$Z := X - M - A.$$

Then Z has no negative jumps, i.e. $\Delta Z \geq 0$. For $n \in \mathbb{N}$ define

$$\tau_n := \inf\{t \geq 0 : Z_t < -(n + |Z_0|)\}.$$

Since Z has no negative jumps,

$$Z_{s\wedge\tau_n} \geq -(n + |Z_0|), \qquad s \geq 0.$$

Therefore

$$X^{\tau_n} \geq M^{\tau_n} + A^{\tau_n} - (n + |Z_0|).$$

The process $A^{\tau_n} - (n + |Z_0|)$ has locally integrable supremum process: indeed, A^* is locally integrable by assumption, and $Z_0 = X_0 - M_0 - A_0$ is locally integrable because X_0 and M_0 are integrable and A_0 is locally integrable whenever A^* is locally integrable. Applying the already proved implication $(ii) \Rightarrow (i)$ to the stopped sigma martingale X^{τ_n} shows that

X^{τ_n} is a local martingale for every n. Finally, $\tau_n \uparrow \infty$ almost surely because Z is càdlàg and therefore pathwise bounded on compact intervals. Hence X is a local martingale. □

Remark 6.1.23 (Use in financial applications). The form used in the financial application is condition (ii). Nominal admissibility gives the discounted lower bound $\widehat{V} \geq -\alpha/N$. With the strengthened numéraire definition of Definition 7.1.1, the process $A = -\alpha/N$ has essentially bounded supremum on every finite horizon and hence locally integrable supremum process. Therefore Theorem 6.1.22 converts the discounted sigma martingale into a local martingale, and the finite-horizon essential bound on $1/N$ then gives the supermartingale property by the standard Fatou argument.

Remark 6.1.24. The jump-minorant criteria (iv) and (v) are not needed for the financial application below, where the process-minorant form (ii) is the relevant statement. They are included to make the theorem structurally parallel to Gushchin's special-semimartingale criterion and to show that the same lower-minorant principle also captures one-sided jump control.

Remark 6.1.25 (Comparison with Gushchin's Theorem 3.21). The closest statement to Theorem 6.1.22 in the literature is [74, Theorem 3.21]. Gushchin's result is formulated through special semimartingales, whereas the present theorem is stated directly in the sigma martingale language used in this chapter. The process-minorant form is the form needed in the financial application below, where nominal admissibility yields the discounted lower bound $\widehat{V} \geq -\alpha/N$. The jump-minorant criteria are included for structural completeness and parallel the corresponding one-sided jump conditions in Gushchin's formulation.

To conclude the study of sigma martingales, let us illustrate that there are sigma martingales where no equivalent probability measure exists under which it is a local martingale. The example is inspired by [40, Example 2.3] and [178, Example 5.3].

Example 6.1.26. Assume a two-dimensional sigma martingale $Z = (X, Y)$ with X being the process from Example 6.1.4 and

$$Y_t = \begin{cases} -2t & \text{for } t < \tau, \\ 1 - 2\tau & \text{else.} \end{cases}$$

Let $\mathcal{F}$ be the σ-algebra created by τ and ξ (from Example 6.1.4) and, furthermore, let $\mathcal{F}_t$ be the σ-algebra generated by $(X_s)_{s \leq t}$. We examine the processes with respect to the corresponding filtered probability space. Obviously, Y is a stopped compensated Poisson Process. As the compensated Poisson Process is a martingale (see, for example, [31, Theorem 5.5.18]) and a stopped martingale is again a martingale by the Optional Stopping Theorem, Y is also a martingale. Hence, Z is a sigma martingale that is not a

local martingale, and we are going to show that no probability measure $\mathbf{Q} \sim \mathbf{P}$ exists such that Z is a local martingale under $\mathbf{Q}$.

To that end, let $\mathbf{Q}$ be an equivalent probability such that Y is a sigma martingale under $\mathbf{Q}$. For each $t > 0$, the stopped process Y^t is bounded and since, by Theorem 6.1.20, any bounded sigma martingale is a martingale, we conclude that Y^t is a martingale. Thus, we have $\mathbf{E}_{\mathbf{Q}}[Y_t] = 0$ for all $t \geq 0$. With f being the density function of τ under $\mathbf{Q}$ and F being the cumulative distribution function of τ (that means $F(t) = \mathbf{Q}(\tau \leq t) = \int_0^t f(s)\mathrm{d}s$), we get

$$\begin{aligned}
0 = \mathbf{E}_{\mathbf{Q}}[Y_t] &= -2t\mathbf{Q}(\tau > t) + \int_0^t (1-2s)f(s)\mathrm{d}s \\
&= -2t\mathbf{Q}(\tau > t) + \int_0^t f(s)\mathrm{d}s - 2\int_0^t -sf(s)\mathrm{d}s \\
&= -2t\mathbf{Q}(\tau > t) + F(t) - 2\left([sF(s)]_0^t - \int_0^t F(s)\mathrm{d}s\right) \\
&= -2t\mathbf{Q}(\tau > t) - 2t\mathbf{Q}(\tau \leq t) + F(t) + 2\int_0^t F(s)\mathrm{d}s \\
&= -2t + F(t) + 2\int_0^t F(s)\mathrm{d}s
\end{aligned}$$

We derive with respect to t and obtain $0 = -2 + F'(t) + 2F(t)$. By putting $G(t) := 1 - F(t)$, we get $G'(t) = -F'(t)$ and $F'(t) = 2G(t)$, thus

$$G'(t) = -2G(t) \text{ and } G(0) = 1.$$

Hence, we obtain $G(t) = \exp(-2t)$ and thus

$$\mathbf{Q}(\tau > t) = 1 - \mathbf{Q}(\tau \leq t) = 1 - F(t) = G(t) = \exp(-2t) = \mathbf{P}(\tau > t).$$

Now, we turn to X. Our first goal is to show that we have $\mathbf{E}_{\mathbf{Q}}[\xi \mid \tau] = 0$ or equivalently, $\int_A \xi = 0$ for all $A \in \sigma(\tau)$. We proceed in a sequence of steps.

1. We have $\mathbf{E}_{\mathbf{Q}}\left[\frac{\xi}{\tau}\mathbb{1}_{\{s<\tau\leq t\}}\right] = 0$.

   For any $s > 0$, we define $\tilde{X}_t^{(s)} := X_t - X_t^s$ where $X_t^s = X_{s\wedge t}$. Then $\tilde{X}^{(s)}$ is clearly a sigma martingale. Furthermore, we have

$$\tilde{X}_t^{(s)} = \begin{cases} 0 & \text{for } t \leq s \\ X_t - X_s, & \text{for } t > s, \end{cases} = \begin{cases} 0 & \text{for } t \leq s, \\ 0 & \text{for } t > s \text{ and } t < \tau, \\ \frac{\xi}{\tau} & \text{for } t > s \text{ and } t \geq \tau \text{ and } s < \tau, \\ 0 & \text{for } t > s \text{ and } s \geq \tau. \end{cases}$$

And since $\left|\frac{\xi}{\tau}\right| < \left|\frac{\xi}{s}\right|$ for $s < \tau$ and $|\xi| \leq 1$, we conclude that $\tilde{X}_t^{(s)}$ is a bounded sigma martingale for all $s > 0$. Hence, by Theorem 6.1.20, it is a martingale and we obtain

$$0 = \mathbf{E}_{\mathbf{Q}}\left[\tilde{X}_t^{(s)}\right] = \mathbf{E}_{\mathbf{Q}}\left[X_t \mathbb{1}_{\{s<\tau\leq t\}}\right] = \mathbf{E}_{\mathbf{Q}}\left[\frac{\xi}{\tau}\mathbb{1}_{\{s<\tau\leq t\}}\right].$$

2. We have $\int_A \xi \mathrm{d}\mathbf{Q} = 0$ for all $A \in \{\{\tau \in (s,t]\}\,;\, 0 < s < t\}$.

   Let $A \in \{\{\tau \in (s,t]\}\,;\, 0 < s < t\}$. Then we have $\mathbf{E}_{\mathbf{Q}}\left|\frac{\xi}{\tau}\mathbb{1}_A\right| < \infty$ and we get

$$\begin{aligned}\frac{\mathbb{1}_A}{\tau}\int_A \xi \mathrm{d}\mathbf{Q} = \frac{\mathbb{1}_A}{\tau}\int_A \mathbf{E}_{\mathbf{Q}}\left[\xi \mid \tau\right]\mathrm{d}\mathbf{Q} &= \int_A \mathbf{E}_{\mathbf{Q}}\left[\frac{\mathbb{1}_A}{\tau}\xi \mid \tau\right]\mathrm{d}\mathbf{Q} \\ &= \int_A \frac{\mathbb{1}_A}{\tau}\xi \mathrm{d}\mathbf{Q} = 0.\end{aligned}$$

   As $\frac{\mathbb{1}_A}{\tau}$ is not constantly zero, we conclude $\int_A \xi \mathrm{d}\mathbf{Q} = 0$.

3. We have $\int_A \xi \mathrm{d}\mathbf{Q} = 0$ for all $A \in \sigma(\tau)$.

   By definition, we have $\int_A \xi \mathrm{d}\mathbf{Q} = \int_A \xi^+ \mathrm{d}\mathbf{Q} - \int_A \xi^- \mathrm{d}\mathbf{Q}$. Both summands can be seen as measures on Ω and since the half-open intervals are an $\cap$-closed generator of the Borel σ algebra,[3] these measures are uniquely determined by its values on $\{\{\tau \in (s,t]\}\,;\, 0 < s < t\}$, [4] we conclude $\int_A \xi \mathrm{d}\mathbf{Q} = 0$ for all $A \in \sigma(\tau)$ and hence

$$\mathbf{E}_{\mathbf{Q}}\left[\xi \mid \tau\right] = 0. \tag{6.1}$$

Since $\mathbf{Q}$ and $\mathbf{P}$ are equivalent probability measures, we have

$$0 < \mathbf{Q}(\xi = 1 \mid \tau), \mathbf{Q}(\xi = -1 \mid \tau) < 1 \quad \Longleftrightarrow \quad 0 < \mathbf{P}(\xi = 1 \mid \tau), \mathbf{P}(\xi = -1 \mid \tau) < 1.$$

Thus with (Eq. (6.1)), we obtain

$$\mathbf{Q}(\xi = -1 \mid \tau) = \mathbf{Q}(\xi = 1 \mid \tau) = \frac{1}{2}$$

almost surely and hence ξ and τ are $\mathbf{Q}$-independent. We conclude $\mathbf{Q} = \mathbf{P}$.

6.2 The Jacod–Yor Theorem for Sigma Martingales

The question of determining criteria for a class of processes $\mathcal{A}$ to be representable as a stochastic integral with respect to a given semimartingale X is fundamental to many applications, particularly in modeling, risk management, and decision-making frameworks.

[3] see, for example, [114, Theorem 1.23]
[4] see, for example, [114, Lemma 1.42]

For example, martingale representation theorems enable the expression of portfolio wealth in self-financing trading strategies, directly linking the hedgeability of a claim to its representation as a stochastic integral. Beyond finance, these theorems are indispensable in filtering theory for constructing optimal filters for stochastic signals and in stochastic control for characterizing solutions to stochastic optimal control problems.

Most commonly, the martingale representation theorem is studied in the context of Brownian motion. Originally developed by Itô [87], the theory was extended to square-integrable martingales by Kunita and Watanabe [121] and Davis and Varaiya [36]. Occasionally, representation theorems are also developed for processes other than Brownian motion. For instance, Klebaner [113] explores the Poisson process, while Applebaum [4] and Nualart and Schoutens [158] extend the discussion to Lévy processes. These broader classes highlight the versatility of representation theorems in accommodating diverse stochastic dynamics. For local martingales, representation theory is notably sparse, with only a few significant contributions, such as those by Cohen and Elliott [31], Kussmaul [124], and Dellacherie and Meyer [44]. Many of these works depend on the Jacod–Yor Theorem, introduced by Jacod and Yor [96], which provides necessary and sufficient conditions for the existence of a martingale representation. While this theorem has been extensively studied for L^2-martingales and, in some cases, local martingales, it has not been extended to sigma martingales.

No general martingale representation theorem exists for sigma martingales. In particular, the Jacod–Yor Theorem has not been established for this class of processes. The lack of such a result leaves a significant gap in the literature. [108] explicitly acknowledge this shortfall:

> *"To the best of our knowledge, the martingale representation property in the framework of sigma martingales is not available in the literature. Indeed, most treatments deal with square-integrable martingales where the notion of orthogonality of martingales is available, which simplifies the treatment."*

Sigma martingales generalize local martingales and are particularly important in financial mathematics. For example, the First Fundamental Theorem of Asset Pricing (FTAP) asserts the existence of an equivalent sigma martingale measure under the condition of No Free Lunch with Vanishing Risk (NFLVR), as shown in [40]. This result highlights that, under no-arbitrage conditions, price processes are typically sigma martingales rather than local martingales.

The lack of a martingale representation theory for sigma martingales also affects the Second Fundamental Theorem of Asset Pricing (Second FTAP), which concerns market completeness. The classical Second FTAP, attributed to [76, 77, 78], is limited to cases

where discounted price processes are local martingales. While versions of the Second FTAP for sigma martingales are presented in [178] and [108], these works deviate from the economically justified approaches commonly associated with the First FTAP.

This thesis aims to fill these gaps by providing a version of the Jacod–Yor Theorem for sigma martingales. As an application, we prove the Second Fundamental Theorem of Asset Pricing in a general form, thereby laying the foundation for a unified theory that bridges sigma martingales, their representation properties, and their role in financial mathematics.

The predictable representation property (PRP) is usually defined for L^2-martingales; see, for example, [166]. Generalizations exist for local martingales as in [79, 178]. We extend this concept further to sigma martingales.

Definition 6.2.1. For $X \in \mathcal{M}^d_{\sigma,0}$, define:

$$\mathcal{I}(X) := \{H \bullet X \,;\, H \in L(X)\}.$$

We say that X has the *predictable representation property (PRP)* if

$$\mathcal{I}(X) = \mathcal{M}^1_{\sigma,0}.$$

Furthermore, define

$$\mathcal{I}^1(X) := \mathcal{I}(X) \cap \mathcal{H}^1.$$

Remark 6.2.2. The set $\mathcal{I}(X)$ depends on the probability measure $\mathbf{P}$. However, for the sake of simplicity, we only write $\mathcal{I}(X)$ instead of $\mathcal{I}(X, \mathbf{P})$.

The importance of orthogonality in stochastic analysis traces back to the foundational work by P. A. Meyer in the 1960s. In his seminal papers [139, 140], Meyer laid the groundwork for understanding martingales using Hilbert-space methods. This perspective was further developed by Itô [91], Motoo and Watanabe [154], and Kunita and Watanabe [121], who demonstrated that the space of L^2-martingales could be treated as a Hilbert space. These techniques and tools allowed for significant advancements in the theory of stochastic analysis.

Inspired by this classical notion of orthogonality, we extend the concept to sigma martingales.

Definition 6.2.3. Let $X \in \mathcal{M}^d_\sigma$. We say that a process $Z \in \mathcal{M}^1_\sigma$ is *orthogonal* to X if $XZ \in \mathcal{M}^d_\sigma$, and we write $X \perp Z$. Furthermore, we define the set of all processes orthogonal to X as

$$X^\perp = \{Z \in \mathcal{M}^1_\sigma \,;\, X \perp Z\}.$$

Remark 6.2.4. Clearly, $X^\perp$ contains all constant processes and for $d = 1$, by partial integration, two sigma martingales X, Y are orthogonal if and only if $[X, Y] \in \mathcal{M}^1_\sigma$.

Definition 6.2.5. Let $X \in \mathcal{M}^d_\sigma$. Then $\mathfrak{P}_\sigma(X)$ denotes the set of all probability measures $\mathbf{Q}$ on $(\Omega, \mathcal{F})$ such that X is a sigma martingale under $\mathbf{Q}$. A measure $\mathbf{P} \in \mathfrak{P}_\sigma(X)$ is *extremal* if, for any $\mathbf{Q}, \mathbf{R} \in \mathfrak{P}_\sigma(X)$ and $\lambda \in (0, 1)$ such that $\mathbf{P} = \lambda\mathbf{Q} + (1 - \lambda)\mathbf{R}$, we have $\mathbf{P} = \mathbf{Q} = \mathbf{R}$.

The main result of this chapter is a generalization of the classical Jacod–Yor Theorem to sigma martingales. This theorem is critical in understanding the uniqueness of probability measures and the structure of sigma martingales.

Theorem 6.2.6 (Jacod–Yor Theorem for Sigma Martingales)**.** *Let $X \in \mathcal{M}^d_\sigma$. The following are equivalent:*

(i) The process X has the predictable representation property.

(ii) The set $X^\perp$ only contains the constant processes.

(iii) For any $\mathbf{Q} \in \mathfrak{P}_\sigma(X)$ with $\mathbf{Q} \ll \mathbf{P}$, we have $\mathbf{P} = \mathbf{Q}$.

(iv) The probability measure $\mathbf{P}$ is an extremal point of $\mathfrak{P}_\sigma(X)$.

As before, we denote by $\mathcal{M}^d$ the space of d-dimensional martingales, by $\mathcal{M}^d_\infty$ the space of bounded d-dimensional martingales, and by $\mathcal{M}^d_{\mathrm{loc}}$ the space of d-dimensional local martingales. A subscript 0 as in $\mathcal{M}_{\mathrm{loc},0}$ will further indicate that the process starts at 0, that is $M_0 = 0$ almost surely for all $M \in \mathcal{M}_{\mathrm{loc},0}$.

Lemma 6.2.7. *Let $X \in \mathcal{M}^d_{\sigma,0}, Y \in \mathcal{M}^1_{\sigma,0}$ and let Y be locally in $\mathcal{I}(X)$. Then we have $Y \in \mathcal{I}(X)$.*

Proof. Let Y be locally in $\mathcal{I}(X)$. That means there exists an increasing sequence of stopping times $(T^n)_{n\in\mathbb{N}}$, tending to infinity, such that there are $H^n \in L(X)$ with $Y^{T^n} = H^n \bullet X$ for all $n \in \mathbb{N}$. By putting $T^0 := 0$ and $H := \sum_{n=1}^\infty H^n \mathbb{1}_{\rrbracket T^{n-1}, T^n \rrbracket}$, we obtain $H \in L(X)$ and locally we have $Y = H \bullet X$. But then we also have $Y = H \bullet X$ and $Y \in \mathcal{I}(X)$ in general. □

Definition 6.2.8. Let V be a vector space of stochastic processes. We say that V is *stable* if $X \in V$ implies $X^T \in V$.

In the literature, it is often required that a stable subspace is also closed. In such a case, we will speak of a stable, closed vector space.

Lemma 6.2.9. *Let $X \in \mathcal{M}^d_{\sigma,0}$. Then, $\mathcal{I}^1(X)$ is a stable and closed subspace of $\mathcal{H}^1$.*

Proof. Because of the linearity of the stochastic integral and since $\mathcal{H}^1$ is a vector space, $\mathcal{I}^1(X)$ is also a vector space. We have for $H \bullet X \in \mathcal{I}^1(X)$ and a stopping time T

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

where

$$[\![0,T]\!] := \{(\omega, t) \in \Omega \times \mathbb{R}_+ \,;\, 0 \leq t \leq T(\omega)\}.$$

Hence $(H \bullet X)^T \in \mathcal{I}(X)$. Since the quadratic variation is an increasing process, we further have

$$\left\| (H \bullet X)^T \right\|_{\mathcal{H}^1} \leq \|H \bullet X\|_{\mathcal{H}^1}$$

and thus $(H \bullet X)^T \in \mathcal{H}^1$. We conclude that $\mathcal{I}^1(X)$ is stable.

We still have to show the closedness. For $H \in L(X)$, we put

$$\|H\|_{\mathcal{I}^1(X)} := \mathbf{E}\left[\sqrt{[H \bullet X, H \bullet X]_\infty}\right].$$

We define $I^1(X)$ to be the quotient space of $\{H \in L(X) \,;\, \|H\|_{\mathcal{I}^1(X)} < \infty\}$ by $\{H \in L(X) \,;\, \|H\|_{\mathcal{I}^1(X)} = 0\}$. Analogously to the L^p theory,[5] it is easy to see that $I^1(X)$ is a Banach space under the quotient norm. Now let $H^n \bullet X$ be a Cauchy sequence in $\mathcal{H}^1$. For all $n, m \in \mathbb{N}$ we have

$$\|H^n - H^m\|_{\mathcal{I}^1(X)} \;=\; \mathbf{E}\left[\sqrt{[(H^n - H^m) \bullet X, (H^n - H^m) \bullet X]_\infty}\right] \;=\; \|(H^n - H^m) \bullet X\|_{\mathcal{H}^1},$$

so the corresponding equivalence classes $[H^n] \in I^1(X)$ form a Cauchy sequence in the quotient space $I^1(X)$. Since $I^1(X)$ is a Banach space, this sequence has a limit $[H] \in I^1(X)$, and any representative $H \in L(X)$ of this class satisfies

$$\|(H^n - H) \bullet X\|_{\mathcal{H}^1} \;=\; \|H^n - H\|_{\mathcal{I}^1(X)} \;\longrightarrow\; 0.$$

In particular, $H^n \bullet X \to H \bullet X$ in $\mathcal{H}^1$, and the limit $H \bullet X$ lies in $\mathcal{I}^1(X)$ since $H \in L(X)$. The space $\mathcal{I}^1(X)$ is therefore closed in $\mathcal{H}^1$. □

Theorem 6.2.10. *Let $X \in \mathcal{M}^d_{\sigma,0}$. The following are equivalent:*

(i) The process X has the predictable representation property;

(ii) $\mathcal{M}^1_{\infty,0} \subset \mathcal{I}^1(X)$;

[5] compare, for example, [174] Theorem 3.11

(iii) $\mathcal{H}^1_0 = \mathcal{I}^1(X)$;

(iv) $\mathcal{M}^1_{\mathrm{loc},0} \subset \mathcal{I}(X)$.

Proof. **(i) $\Rightarrow$ (ii).** By definition of the predictable representation property, we have $\mathcal{I}(X) = \mathcal{M}^1_{\sigma,0}$ and $\mathcal{I}^1(X) := \mathcal{I}(X) \cap \mathcal{H}^1$. Since $\mathcal{M}^1_{\infty,0} \subset \mathcal{H}^1$ and $\mathcal{M}^1_{\infty,0} \subset \mathcal{M}^1_{\sigma,0}$, we conclude $\mathcal{M}^1_{\infty,0} \subset \mathcal{I}^1(X)$.

(ii) $\Rightarrow$ (iii). Let $N \in \mathcal{H}^1_0$. By Theorem 2.2.8, $\mathcal{M}^\infty_0$ is dense in $\mathcal{H}^1_0$. Hence, there exists a sequence $(M^n)_{n\in\mathbb{N}} \subset \mathcal{M}^\infty_0$ that converges in $\mathcal{H}^1$ to N. By assumption, we have $M^n \in \mathcal{I}^1(X)$ for all n, and from Lemma 6.2.9, we know that $\mathcal{I}^1(X)$ is closed. Thus, $N \in \mathcal{I}^1(X)$.

(iii) $\Rightarrow$ (iv). By Theorem 2.2.9, each local martingale is locally in $\mathcal{H}^1$. Hence, by assumption, each local martingale is locally in $\mathcal{I}^1(X)$ and, therefore, in $\mathcal{I}(X)$. Lemma 6.2.7 yields $\mathcal{M}^1_{\mathrm{loc},0} \subset \mathcal{I}(X)$.

(iv) $\Rightarrow$ (i). Let $Y \in \mathcal{M}^1_{\sigma,0}$. By Theorem 6.1.7, there exists a one-dimensional martingale M and a process $H \in L(M)$ such that $Y = H \bullet M$. Since $\mathcal{M}^1_{\mathrm{loc},0} \subset \mathcal{I}(X)$, there also exists a process $K \in L(X)$ such that $M = K \bullet X$. Hence, we have

$$Y = H \bullet (K \bullet X) = (HK) \bullet X,$$

and $Y \in \mathcal{I}(X)$. □

The following lemma is often used in connection with Girsanov's theorem (see, for example, [166] Exercise III.21, [31] Lemma 15.2.1, [95] Proposition 3.8, [104] Lemma 18.16, [122] Lemma 8.9.3, [79] Lemma 12.11, [4] Lemma 5.2.11).

Lemma 6.2.11. *Suppose a semimartingale M, $\mathbf{Q} \ll \mathbf{P}$ and let $Z_\infty = \frac{\mathrm{d}\mathbf{Q}}{\mathrm{d}\mathbf{P}}$ be the Radon-Nikodym derivative with*

$$Z_t = \mathbf{E}_\mathbf{P}(Z_\infty \mid \mathcal{F}_t).$$

Then M is a $\mathbf{Q}$-(local) martingale if and only if MZ is a $\mathbf{P}$-(local) martingale.

We are going to investigate a possible generalization to sigma martingale measures.

Lemma 6.2.12. *Let X be a semimartingale, $\mathbf{Q} \ll \mathbf{P}$, $Z_\infty = \frac{\mathrm{d}\mathbf{Q}}{\mathrm{d}\mathbf{P}}$ the Radon-Nikodym derivative and $Z_t = \mathbf{E}_\mathbf{P}(Z_\infty \mid \mathcal{F}_t)$.*

(a) If X is a $\mathbf{Q}$ sigma martingale, then XZ is a $\mathbf{P}$-sigma martingale.

(b) If $\mathbf{Q} \sim \mathbf{P}$ then X is a $\mathbf{Q}$-sigma martingale if and only if XZ is a $\mathbf{P}$-sigma martingale.

Proof. First we note that by [166, Theorem IV.25], it is irrelevant whether we calculate the integral under $\mathbf{Q}$ or $\mathbf{P}$. We will therefore only write $\bullet$ instead of $\bullet_{\mathbf{Q}}$ or $\bullet_{\mathbf{P}}$. Without loss of generality, we assume X to be one-dimensional.

For any predictable set D_n, we have

$$\begin{aligned}
\mathbb{1}_{D_n}\bullet(XZ) &= \mathbb{1}_{D_n}\bullet\left(X_-\bullet Z + Z_-\bullet X + [X,Z]\right)\\
&= \mathbb{1}_{D_n}\bullet(X_-\bullet Z) + \mathbb{1}_{D_n}\bullet(Z_-\bullet X) + \mathbb{1}_{D_n}\bullet[X,Z]\\
&= (\mathbb{1}_{D_n}X_-)\bullet Z + (\mathbb{1}_{D_n}Z_-)\bullet X + [\mathbb{1}_{D_n}\bullet X, Z]\\
&= (\mathbb{1}_{D_n}X_-)\bullet Z + (\mathbb{1}_{D_n}Z_-)\bullet X + (\mathbb{1}_{D_n}\bullet X)\, Z\\
&\qquad - (\mathbb{1}_{D_n}\bullet X)_-\bullet Z - Z_-\bullet(\mathbb{1}_{D_n}\bullet X)\\
&= (\mathbb{1}_{D_n}X_-)\bullet Z + (\mathbb{1}_{D_n}Z_-)\bullet X + (\mathbb{1}_{D_n}\bullet X)\, Z\\
&\qquad - (\mathbb{1}_{D_n}\bullet X)_-\bullet Z - (Z_-\mathbb{1}_{D_n})\bullet X\\
&= (\mathbb{1}_{D_n}X_-)\bullet Z + (\mathbb{1}_{D_n}\bullet X)\, Z - (\mathbb{1}_{D_n}\bullet X)_-\bullet Z. \qquad (6.2)
\end{aligned}$$

Let X be a $\mathbf{Q}$-sigma martingale with localizing sequence $(D_n)_{n\in\mathbb{N}}$. By Lemma 6.2.11, $(\mathbb{1}_{D_n}\bullet X)Z$ is a local $\mathbf{P}$-martingale. Furthermore, $(\mathbb{1}_{D_n}X_-)$ and $(\mathbb{1}_{D_n}\bullet X)_-$ are locally bounded. Since Z is a $\mathbf{P}$-martingale, the first and the third summand on the right-hand side of (Eq. (6.2)) are local $\mathbf{P}$-martingales . Hence, $\mathbb{1}_{D_n}\bullet(XZ)$ is also a local $\mathbf{P}$-martingale and $(D_n)_{n\in\mathbb{N}}$ is a localizing sequence for XZ. Thus XZ is a $\mathbf{P}$-sigma martingale.

Now let XZ be a $\mathbf{P}$-sigma martingale and $\mathbf{Q}\sim\mathbf{P}$. Then we have $Z_\infty > 0$ and $\frac{\mathrm{d}\mathbf{P}}{\mathrm{d}\mathbf{Q}}$ is given by $\frac{1}{Z_\infty}$. Applying the already proven direction yields that $XZ\frac{1}{Z} = X$ is a $\mathbf{Q}$-sigma martingale. □

Lemma 6.2.13. *Let $X\in\mathcal{M}^d_\sigma$ and $H\in L(X)$, then there exists a process*

$$M = (M^1,\dots,M^d)^\top \text{ with } M^i\in\mathcal{H}^1$$

for all i and a process $K=(K^1,\dots,K^d)^\top$ with $K^i\in L(M^i)$ for all i with

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

Furthermore, if there exists a stochastic process $Z\in\mathcal{M}^1_\sigma$ such that $Z\perp X$, then we also have $Z\perp M$.

Proof. By Theorem 6.1.7, each X^i can be written as $X^i=\hat{K}\bullet N^i$ with a strictly positive $\hat{K}\in L(N^i)$ and $N^i\in\mathcal{H}^1$ for all i. We set $G^i := H^i\hat{K}$ and obtain $H\bullet X = G\bullet N$ for $G=(G^1,\dots,G^d)^\top$. Now we put $K := \|G\|\vee 1$ and $J=\frac{G}{K}$ and J is a bounded process

and thus $J^i \in L(X^i)$ for each i. We have

$$G \bullet N = K \bullet (J \bullet N) = K \bullet \left(\sum_{i=1}^{d} J^i \bullet N^i \right) = \sum_{i=1}^{d} K \bullet (J^i \bullet N^i).$$

By putting $M^i := J^i \bullet N^i$, we obtain the desired result.

Now let $Z \in \mathcal{M}_\sigma^1$ with $ZX \in \mathcal{M}_\sigma^d$. By partial integration, we get

$$ZX^i = Z(\hat{K} \bullet N^i) = Z_- \bullet (\hat{K} \bullet N^i) + (\hat{K} \bullet N^i)_- \bullet Z + [Z, \hat{K} \bullet N^i].$$

Hence, we obtain $\hat{K} \bullet [Z, N^i] \in \mathcal{M}_\sigma^1$. Since $\hat{K}$ is strictly positive, we also get

$$[Z, N^i] = \left(\frac{1}{\hat{K}} \hat{K} \right) \bullet [Z, N^i] = \frac{1}{\hat{K}} \bullet (\hat{K} \bullet [Z, N^i]) \in \mathcal{M}_\sigma^1.$$

Analogously, we obtain

$$ZM^i = Z_- \bullet M^i + M^i_- \bullet Z + [Z, M^i].$$

Since $[Z, M^i] = [Z, J \bullet N^i] = J \bullet [Z, N^i]$, all three summands on the right-hand side are sigma martingales. We conclude that ZM^i is a sigma martingale as well. □

For the proof of the Jacod–Yor Theorem for sigma martingales, we recall that $\mathcal{BMO}$ denotes the space of martingales of bounded mean oscillation.

Proof of Theorem 6.2.6. We are going to show (i) ⇒ (ii) ⇒ (iii) ⇒ (iv) ⇒ (i). The closure step (iv) ⇒ (i) is the only one that requires the full machinery of Theorem 6.2.10 and a Hahn–Banach/duality argument; the first three implications use only the lemmas just established.

(i) ⇒ (ii). Recall that (i) is the predictable representation property: every $Y \in \mathcal{M}_{\sigma,0}^1$ can be written as $Y = H \bullet X$ for some $H \in L(X)$. Recall also that $X^\perp$ consists of those $Z \in \mathcal{M}_\sigma^1$ for which the product XZ remains a sigma martingale. To say "$X^\perp$ contains only constant processes" is to say that any such Z must equal Z_0 a.s. for all t; "constant" here means constant as a function of t for almost every ω, equivalently a process whose increments vanish almost surely, equivalently $[Z, Z] \equiv 0$. The strategy of the implication is to show that, under (i), every $Z \in X^\perp$ has $[Z, Z] \equiv 0$, hence Z is constant.

Suppose $Z \in X^\perp$. Without loss of generality we may assume $Z_0 = 0$, so that $Z \in \mathcal{M}_{\sigma,0}^1$. By assumption (i), there exists an $H \in L(X)$ such that $Z = H \bullet X$. Now Lemma 6.2.13 yields a martingale $M = (M^1, \ldots, M^d)^\top$ and a process $K = (K^1, \ldots, K^d)^\top$

with $K^i \in L(M^i)$ such that

$$Z = K \bullet M \quad \text{and} \quad MZ \in \mathcal{M}^d_\sigma.$$

The point of this reduction is that M is an honest martingale (not merely a sigma martingale), so the integration-by-parts formula and the quadratic variation can be exploited cleanly. We define

$$N := [K \bullet M, K \bullet M] = \sum_{i=1}^{d} K^i \bullet [M^i, K \bullet M] = \sum_{i=1}^{d} K^i \bullet [M^i, Z],$$

where

$$N = \sum_{i=1}^{d} K^i \bullet (M^i Z - M^i_- \bullet Z - Z_- \bullet M^i).$$

Each summand in the last expression is a sigma martingale: $M^i Z$ is a sigma martingale because the components of MZ are sigma martingales by orthogonality $Z \in X^\perp$ combined with $Z = K \bullet M$ and the reduction in Lemma 6.2.13; the integrals $M^i_- \bullet Z$ and $Z_- \bullet M^i$ are sigma martingales by Corollary 6.1.10 since M^i_- and Z_- are locally bounded predictable. Hence N is a sigma martingale. However, as the quadratic variation of a semimartingale, N is also positive and increasing. By Theorem 6.1.20, a positive sigma martingale is a local martingale, and since N is increasing with $N_0 = 0$, the local-martingale property forces $N \equiv 0$ a.s. (every increasing local martingale starting at 0 is identically zero). Therefore $[Z, Z] = N = 0$, so Z is constant.

(ii) $\Rightarrow$ (iii). Let $\mathbf{Q} \in \mathfrak{P}_\sigma(X)$ with $\mathbf{Q} \ll \mathbf{P}$. Define

$$Z_t := \mathbf{E}\left[\frac{\mathrm{d}\mathbf{Q}}{\mathrm{d}\mathbf{P}} \mid \mathcal{F}_t\right].$$

Since X is a $\mathbf{Q}$-sigma martingale, XZ is a $\mathbf{P}$-sigma martingale by Lemma 6.2.12. By assumption, Z is constant. Hence, we have $\mathbf{P} = \mathbf{Q}$.

(iii) $\Rightarrow$ (iv). Let $\mathbf{Q}, \mathbf{R} \in \mathfrak{P}_\sigma(X)$ and $\lambda \in (0, 1)$ such that $\mathbf{P} = \lambda\mathbf{Q} + (1 - \lambda)\mathbf{R}$. Clearly, $\mathbf{Q} \ll \mathbf{P}$. By assumption, we have $\mathbf{Q} = \mathbf{P}$. Similarly, $\mathbf{R} = \mathbf{P}$. Thus, $\mathbf{P}$ is an extremal point in $\mathfrak{P}_\sigma(X)$.

(iv) $\Rightarrow$ (i). By Theorem 6.2.10, it suffices to show that $\mathcal{I}^1(X) = \mathcal{H}^1_0$.

Suppose this is not the case. Then, by Lemma 6.2.9, $\mathcal{I}^1(X)$ is a stable closed subspace of $\mathcal{H}^1$, and there exists an element $Y_0 \in \mathcal{H}^1$, $Y_0 \notin \mathcal{I}^1(X)$, such that

$$\inf_{Y \in \mathcal{I}^1(X)} \|Y_0 - Y\|_{\mathcal{H}^1} > 0.$$

By Theorem 2.5.8, there exists a bounded linear functional $\varphi : \mathcal{H}_0^1 \to \mathbb{R}$ with $\varphi\,|_{\mathcal{I}^1(X)} = 0$ and $\varphi(Y_0) = 1$. By Theorem 2.2.13, φ can be represented as

$$\varphi(I) = \mathbf{E}\left[I_\infty M_\infty\right],$$

where $M \in \mathcal{BMO}_0$.

Thus, $\mathbf{E}\left[I_\infty M_\infty\right] = 0$ for all $I \in \mathcal{I}^1(X)$. Since $\mathcal{I}^1(X)$ is stable, we have

$$\mathbf{E}\left[I_\infty M_\infty^T\right] = 0 \quad \text{for all } I \in \mathcal{I}^1(X). \tag{6.3}$$

By Theorem 2.2.12, M is locally bounded. Thus, there exists a sequence of stopping times $(T^n)_{n\in\mathbb{N}}$ such that $M^{T^n} \in \mathcal{M}_0^\infty$. Define

$$Z_t^n := \mathbf{E}\left[1 + \frac{M_\infty^{T^n}}{2c_n} \mid \mathcal{F}_t\right], \quad Z_t^{-n} := \mathbf{E}\left[1 - \frac{M_\infty^{T^n}}{2c_n} \mid \mathcal{F}_t\right],$$

where c_n is a constant larger than $\sup|M^{T^n}|$.

Clearly, all Z^n and Z^{-n} are bounded martingales. We define probability measures $\mathbf{Q}^n, \mathbf{Q}^{-n}$ by $Z_\infty^n = \frac{\mathrm{d}\mathbf{Q}^n}{\mathrm{d}\mathbf{P}}$ and $Z_\infty^{-n} = \frac{\mathrm{d}\mathbf{Q}^{-n}}{\mathrm{d}\mathbf{P}}$.

Since $XZ^n = X + \frac{XM^{T^n}}{2c_n}$ and $X \in \mathcal{I}^1(X)$, we conclude with (6.3) that XZ^n is a $\mathbf{P}$-sigma martingale. Similarly, XZ^{-n} is a $\mathbf{P}$-sigma martingale. By Lemma 6.2.12, $\mathbf{Q}^n, \mathbf{Q}^{-n} \in \mathfrak{P}_\sigma(X)$.

Because $\mathbf{P} = \frac{1}{2}(\mathbf{Q}^n + \mathbf{Q}^{-n})$, we get, by assumption, $\mathbf{P} = \mathbf{Q}^n = \mathbf{Q}^{-n}$. Thus,

$$1 + \frac{M_\infty^{T^n}}{2c_n} = \frac{\mathrm{d}\mathbf{Q}^n}{\mathrm{d}\mathbf{P}} = 1,$$

which implies $M_\infty^{T^n} = 0$ for all n. Taking $n \to \infty$, we get $M \equiv 0$, and hence $\varphi \equiv 0$.

Therefore, no process $Y_0 \in \mathcal{H}^1$ can exist with $Y_0 \notin \mathcal{I}^1(X)$. Thus, $\mathcal{I}^1(X) = \mathcal{H}_0^1$ follows. □

As an application of the theory developed in this chapter, we combine the stability of sigma martingales under stochastic integration (Corollary 6.1.10), the Ansel–Stricker Lemma (6.1.20), and (for the backward direction) the stochastic logarithm of Lemma 5.2.12 to characterize when a semimartingale and its Doléans-Dade exponential are simultaneously local martingales.

Theorem 6.2.14. *Let X be a semimartingale with $X_0 = 0$ and $\Delta X > -1$. If X or $\mathcal{E}(X)$ is a sigma martingale, then both X and $\mathcal{E}(X)$ are local martingales.*

Proof. Suppose first that X is a sigma martingale. Define the stopping times

$$T^n := \inf\{t \geq 0 : X_t \leq -n\},$$

which increase to ∞ a.s. because X is càdlàg. On $[\![0, T^n]\!]$ we have $X_{T^n-} \geq -n$ and $\Delta X_{T^n} > -1$, so $X^{T^n} \geq -(n+1)$ pathwise. The stopped process is again a sigma martingale by Corollary 6.1.10, and is bounded below, so the Ansel–Stricker Lemma (Theorem 6.1.20) yields that X^{T^n} is a local martingale. A diagonal localizing sequence then shows that X itself is a local martingale. Consequently, $\mathcal{E}(X) = 1 + \mathcal{E}(X)_- \bullet X$ is a local martingale by Theorem 4.6.2, since $\mathcal{E}(X)_-$ is locally bounded.

Conversely, suppose $\mathcal{E}(X)$ is a sigma martingale. Since $\Delta X > -1$, we have $\mathcal{E}(X) > 0$ by Remark 5.2.6, so by Ansel–Stricker (Theorem 6.1.20) $\mathcal{E}(X)$ is a local martingale. Define

$$T_n := \inf\{t \geq 0 : \mathcal{E}(X)_t \leq 1/n\};$$

since $\mathcal{E}(X) > 0$ is càdlàg, $T_n \to \infty$ a.s. and the integrand below is bounded by n. Applying Lemma 5.2.12 to the stopped process gives

$$X^{T_n} = \left(\tfrac{1}{\mathcal{E}(X)^{T_n}}\right)_- \bullet \mathcal{E}(X)^{T_n}.$$

Since the integrand is locally bounded and the integrator is a local martingale, X^{T_n} is a local martingale by Theorem 4.6.2; as $T_n \to \infty$, X is a local martingale. □

Chapter 7

Applications to Mathematical Finance

Mathematical finance is the principal application target of the theory developed in this thesis. The aim of this chapter is to apply the topological vector integral of Chapter 4, the Itô formula of Chapter 5, and the sigma martingale theory of Chapter 6 to financial markets in a self-contained way. We work with a general semimartingale market with dividends, define the numéraire and the discounted market, prove the transfer of the self-financing condition, and then derive the supermartingale property of admissible discounted wealth processes under equivalent sigma martingale measures. The chapter follows the unified treatment in

- Moritz Sohns. "The General Semimartingale Market Model". In: *AppliedMath* 5.3 (2025). ISSN: 2673-9909. DOI: 10.3390/appliedmath5030097;

further applications of the framework appear in [179, 182]. The chapter closes with a worked example, the dividend-paying geometric Brownian motion, in which the Itô formula of Chapter 5 is applied directly.

7.1 The General Semimartingale Market Model with Dividends

Throughout this chapter we work on the filtered probability space fixed in Chapter 2. The market consists of $d+1$ tradable securities. The ex-dividend price processes are collected in the $(d+1)$-dimensional semimartingale

$$S = (S^0, S^1, \dots, S^d),$$

and the cumulative cash-flow or dividend processes are collected in the $(d+1)$-dimensional semimartingale

$$D = (D^0, D^1, \dots, D^d).$$

We assume $D_0^i = 0$ for every $i = 0, \dots, d$. No monotonicity and no finite-variation assumption is imposed on D^i. Thus D^i may represent dividends, coupon payments, fees, taxes, storage costs, margin calls, capital calls, or other pass-through cash flows. Positive increments correspond to payments received by the holder of the security, and negative increments correspond to payments made by the holder.

The total-return process of security i is $S^i + D^i$. We assume the market is frictionless and competitive. Frictionless means that there are no strategy-dependent transaction costs, taxes, short-sale restrictions, borrowing limits, or margin requirements, and that assets are infinitely divisible. Competitive means that traders act as price takers. These idealizations are the standard assumptions under which the Fundamental Theorems of Asset Pricing are formulated.

7.1.1 Self-Financing Trading Strategies

A *trading strategy* is a predictable $(d+1)$-dimensional process $\varphi = (\varphi^0, \dots, \varphi^d) \in L(S+D)$. The component φ_t^i is the number of shares of security i held over the infinitesimal interval ending at t. The associated *wealth process* is

$$V_t(\varphi) := \sum_{i=0}^{d} \varphi_t^i \big(S_t^i + \Delta D_t^i\big), \qquad t \geq 0.$$

This convention values the portfolio at time t after the cash flow ΔD_t has been received or paid. Concretely, $V_t(\varphi) = \sum_i \varphi_t^i S_t^i + \sum_i \varphi_t^i \Delta D_t^i$ is the mark-to-market value $\sum_i \varphi_t^i S_t^i$ together with the contemporaneous cash flow $\sum_i \varphi_t^i \Delta D_t^i$ on the positions held across t, so V_t is the cum-dividend value of the portfolio; this bookkeeping convention follows [183], where its consistency with the gains process $\varphi \bullet (S+D)$ is examined in detail. The strategy φ is called self-financing in nominal units if

$$V(\varphi) = \varphi \bullet (S + D). \tag{7.1}$$

The equality is an identity of semimartingales. It says that changes in wealth are generated only by price changes and the cash flows attached to the securities, and not by external deposits or withdrawals.

7.1.2 The Numéraire and the Discounted Market

To compare values at different times, one fixes a strictly positive reference asset and expresses all other values in units of that reference asset. This reference asset is the numéraire. In the present chapter we use a slightly strengthened numéraire condition. It is strong enough to turn nominal admissibility into an integrable discounted lower bound on every finite horizon, and it therefore removes the need for any separate integrability assumption on the reciprocal numéraire later.

Definition 7.1.1. A predictable semimartingale N is called a *numéraire* if $N > 0$ up to evanescence and, for every finite horizon $T < \infty$,

$$c_T(N) := \left\| \sup_{0 \le s \le T} \frac{1}{N_s} \right\|_{L^\infty(P)} < \infty.$$

Equivalently, for every finite horizon $T < \infty$,

$$\operatorname*{ess\,inf}_{\omega \in \Omega} \inf_{0 \le s \le T} N_s(\omega) > 0.$$

Remark 7.1.2. The essential-supremum formulation is stronger than pathwise strict positivity on compact intervals. Its role is to make the lower bound obtained after discounting automatically integrable under every probability measure equivalent to $\mathbf{P}$. Since equivalent probability measures have the same null sets, the condition $c_T(N) < \infty$ is invariant under equivalent changes of measure.

Lemma 7.1.3. *Let N be a numéraire. Then $1/N$ is predictable and bounded on every deterministic finite horizon, up to evanescence. In particular, $1/N$ is locally bounded and belongs to $L(X)$ for every semimartingale X. Moreover, for every probability measure $\mathbf{Q} \sim \mathbf{P}$ and every finite horizon $T < \infty$,*

$$\mathbf{E_Q}\left[\sup_{0 \le s \le T} \frac{1}{N_s} \right] < \infty.$$

Proof. Predictability of $1/N$ follows from predictability of N and Borel measurability of $x \mapsto 1/x$ on $(0, \infty)$. By definition of a numéraire,

$$\sup_{0 \le s \le T} \frac{1}{N_s} \le c_T(N) \qquad \mathbf{P}\text{-a.s.}$$

for every finite horizon T. Hence $1/N$ is bounded on every deterministic finite horizon, up to evanescence, and therefore locally bounded. Since locally bounded predictable processes are integrable with respect to every semimartingale, $1/N \in L(X)$ for every semimartingale

X. If $\mathbf{Q} \sim \mathbf{P}$, the same essential bound holds $\mathbf{Q}$-a.s. because $\mathbf{P}$ and $\mathbf{Q}$ have the same null sets. Thus

$$\mathbf{E_Q}\left[\sup_{0\le s\le T}\frac{1}{N_s}\right] \le c_T(N) < \infty.$$

□

Given a numéraire N, define discounted ex-dividend prices, discounted cumulative dividend processes, and discounted wealth by

$$\widehat{S}^i := \frac{S^i}{N}, \qquad \widehat{D}^i := \frac{1}{N}\bullet D^i, \qquad \widehat{V}(\varphi) := \frac{V(\varphi)}{N}, \qquad i = 0, \dots, d.$$

The process $\widehat{D}^i$ is defined as a stochastic integral, not as the pointwise quotient D^i/N. This is the economically correct convention: cumulative dividends change only when a cash flow occurs, and the integral discounts each cash flow at the time it occurs. Since $1/N$ is locally bounded and predictable, $(1/N)\bullet D^i$ is well defined for arbitrary semimartingale dividend processes D^i.

Theorem 7.1.4. *Let N be a numéraire, and let S^i, D^i, $i = 0, \dots, d$, be semimartingales. Then*

$$\widehat{S}^i = \frac{S^i}{N}, \qquad \widehat{D}^i = \frac{1}{N}\bullet D^i$$

are semimartingales for every $i = 0, \dots, d$. If φ is a self-financing strategy, then $\widehat{V}(\varphi) = V(\varphi)/N$ is also a semimartingale.

Proof. Since N is a strictly positive semimartingale and $1/N$ is locally bounded, the Itô formula applied locally to $x \mapsto 1/x$ shows that $1/N$ is a semimartingale. Hence $\widehat{S}^i = S^i(1/N)$ is a product of semimartingales and is therefore a semimartingale. Moreover, $1/N$ is predictable and locally bounded, so $(1/N)\bullet D^i$ is a semimartingale for every semimartingale D^i. Finally, if φ is self-financing, then $V(\varphi) = \varphi\bullet(S + D)$ is a semimartingale, and $V(\varphi)/N = V(\varphi)(1/N)$ is again a product of semimartingales. □

7.1.3 Transferring the Self-Financing Condition

The semimartingale preservation result above does not by itself imply that self-financing is invariant under discounting. The next theorem gives the precise transfer statement. Since dividend processes are allowed to be arbitrary semimartingales, the theorem assumes the continuous-covariation condition actually needed in the calculation.

Theorem 7.1.5. *Let N be a numéraire and let S^i, D^i, $i = 0, \dots, d$, be semimartingales. Assume*

$$[N, D^i]^c = 0, \qquad i = 0, \dots, d.$$

Then a predictable strategy $\varphi \in L(S + D)$ is self-financing in nominal units if and only if its discounted wealth process satisfies

$$\widehat{V}(\varphi) = \sum_{i=0}^{d} \varphi^i \bullet (\widehat{S}^i + \widehat{D}^i). \tag{7.2}$$

Proof. By linearity it is enough to verify the identity for one security, so we suppress the index i. We write $\widehat{S} = S/N$ and $\widehat{D} = (1/N) \bullet D$. The integration-by-parts formula for $S = N\widehat{S}$ gives

$$S = N_- \bullet \widehat{S} + \widehat{S}_- \bullet N + [N, \widehat{S}].$$

For the dividend part, since $\widehat{D} = (1/N) \bullet D$, the jump identity gives

$$\Delta \widehat{D}_t = \frac{1}{N_t} \Delta D_t.$$

Furthermore, the bracket identity for stochastic integrals yields

$$[N, \widehat{D}]^c = \frac{1}{N} \bullet [N, D]^c = 0$$

by the assumption $[N, D]^c = 0$. The continuous part of $N_- \bullet \widehat{D}$ therefore coincides with the continuous part of D. At a jump time,

$$\Delta\Big(N_- \bullet \widehat{D} + [N, \widehat{D}]\Big)_t = N_{t-} \Delta \widehat{D}_t + \Delta N_t \Delta \widehat{D}_t = \frac{N_{t-} + \Delta N_t}{N_t} \Delta D_t = \Delta D_t.$$

Thus

$$N_- \bullet \widehat{D} + [N, \widehat{D}] = D.$$

Combining the price and dividend identities gives

$$N_- \bullet (\widehat{S} + \widehat{D}) + \widehat{S}_- \bullet N + [N, \widehat{S} + \widehat{D}] = S + D.$$

If the discounted self-financing identity (7.2) holds, then $V = N\widehat{V}$ and integration by parts, together with associativity of stochastic integration and the preceding identity, gives

$$V = \varphi \bullet (S + D).$$

This is the nominal self-financing condition. Conversely, starting from $V = \varphi \bullet (S + D)$

and applying the same calculation to $\widehat{V} = V/N$ gives (7.2). □

7.2 Admissibility and Equivalent Sigma Martingale Measures

The class of all self-financing strategies is too large for no-arbitrage analysis, because it includes doubling-type strategies. We therefore impose a nominal lower bound.

Definition 7.2.1. A self-financing strategy $\varphi \in L(S + D)$ is called *admissible* if there exists $\alpha \in \mathbb{R}_+$ such that

$$V_t(\varphi) \geq -\alpha, \qquad t \geq 0.$$

The set of admissible self-financing strategies is denoted by $\Phi(S, D, N)$, or simply by Φ when the market data are fixed.

The nominal lower bound is economically natural but mathematically delicate: after discounting it becomes $\widehat{V} \geq -\alpha/N$, a process-valued lower bound. The generalized Ansel–Stricker theorem of Chapter 6 is precisely the result that converts this process minorant into local martingality.

Definition 7.2.2. A probability measure $\mathbf{Q} \sim \mathbf{P}$ is called an *equivalent sigma martingale measure*, or *ESMM*, for (S, D, N) if

$$\widehat{S}^i + \widehat{D}^i$$

is a $\mathbf{Q}$-sigma martingale for every $i = 0, \ldots, d$.

Theorem 7.2.3. *Let φ be an admissible self-financing strategy, let N be a numéraire, assume*

$$[N, D^i]^c = 0, \qquad i = 0, \ldots, d,$$

and let $\mathbf{Q}$ be an ESMM for (S, D, N). Then

$$\widehat{V}(\varphi) = \varphi \bullet (\widehat{S} + \widehat{D})$$

is a $\mathbf{Q}$-local martingale. Moreover, $\widehat{V}(\varphi)$ is a $\mathbf{Q}$-supermartingale.

Proof. By Theorem 7.1.5, the self-financing condition transfers to the discounted setting, and

$$\widehat{V}(\varphi) = \sum_{i=0}^{d} \varphi^i \bullet (\widehat{S}^i + \widehat{D}^i) = \varphi \bullet (\widehat{S} + \widehat{D}).$$

Under $\mathbf{Q}$, each $\widehat{S}^i + \widehat{D}^i$ is a sigma martingale by the ESMM property. Since sigma martingales are stable under stochastic integration, $\widehat{V}(\varphi)$ is a $\mathbf{Q}$-sigma martingale.

Admissibility gives $V(\varphi) \geq -\alpha$ for some $\alpha \geq 0$. Hence

$$\widehat{V}(\varphi) = \frac{V(\varphi)}{N} \geq -\frac{\alpha}{N}.$$

Set $A := -\alpha/N$. For every finite horizon T, the numéraire condition gives

$$\sup_{0\leq s\leq T} |A_s| = \alpha \sup_{0\leq s\leq T} \frac{1}{N_s} \leq \alpha c_T(N)$$

almost surely under $\mathbf{Q}$. Hence A^* is locally integrable under $\mathbf{Q}$. The generalized Ansel–Stricker theorem, Theorem 6.1.22, applied with $M \equiv 0$, implies that $\widehat{V}(\varphi)$ is a $\mathbf{Q}$-local martingale.

Finally, on every finite horizon T,

$$\widehat{V}_s(\varphi) \geq -\alpha c_T(N), \qquad 0 \leq s \leq T.$$

Thus the local martingale $\widehat{V}(\varphi)$ is bounded from below by an integrable constant on every finite horizon. The standard Fatou argument for lower-bounded local martingales yields that $\widehat{V}(\varphi)$ is a $\mathbf{Q}$-supermartingale. □

7.2.1 The Fundamental Theorems of Asset Pricing

The First Fundamental Theorem links absence of arbitrage to the existence of an equivalent sigma martingale measure. The Second Fundamental Theorem links completeness to uniqueness of this measure. The sigma martingale formulation is essential because equivalent local-martingale measures need not exist in general semimartingale models.

Definition 7.2.4. Let K be the set of terminal discounted gains $(\varphi \bullet (\widehat{S} + \widehat{D}))_T$ obtained from admissible strategies with $\widehat{V}_0 = 0$, let $C = (K - L^0_+) \cap L^\infty$, and let $\overline{C}^{\|\cdot\|_\infty}$ denote its norm closure in L^∞. The market satisfies *No Free Lunch with Vanishing Risk*, or *NFLVR*, on $[0, T]$ if

$$\overline{C}^{\|\cdot\|_\infty} \cap L^\infty_+ = \{0\}.$$

Theorem 7.2.5 (First Fundamental Theorem of Asset Pricing, [40])**.** *Let S, D, N be as above and assume $[N, D^i]^c = 0$ for every $i = 0, \ldots, d$. The discounted market satisfies NFLVR on $[0, T]$ if and only if there exists an equivalent sigma martingale measure $\mathbf{Q}$ for (S, D, N).*

The proof is the Delbaen–Schachermayer Fundamental Theorem of Asset Pricing for unbounded semimartingales, applied to the discounted total-return process $\widehat{S} + \widehat{D}$, together with Theorem 7.2.3. The strengthened numéraire definition ensures that nominal admissibility yields true discounted supermartingales under every ESMM, so no further reciprocal-integrability hypothesis is needed.

The Second Fundamental Theorem characterizes completeness by uniqueness of the ESMM. In this section we absorb dividends into the discounted total-return process and write S for $\widehat{S} + \widehat{D}$ to lighten notation.

Theorem 7.2.6 (Second Fundamental Theorem of Asset Pricing)**.** *Let $(S_t)_{t\in[0,T]}$ be a $(d+1)$-dimensional sigma martingale under $\mathbf{P}$. The following are equivalent:*

(i) The market is complete: every contingent claim can be replicated by an admissible self-financing strategy.

(ii) Every uniformly integrable martingale admits a representation as a stochastic integral with respect to S.

(iii) Every local martingale admits a representation as a stochastic integral with respect to S.

(iv) The equivalent sigma martingale measure is unique.

This is the Jacod–Yor representation theorem for sigma martingales, Theorem 6.2.6, translated into financial language. The equivalence between completeness and representation is standard; the novelty here is that the representation theorem is formulated at the sigma martingale level, which is the natural level supplied by the First Fundamental Theorem.

Remark 7.2.7. In the replication statement one requires the wealth process of the replicating strategy to be a true martingale, not merely a local martingale, whenever pricing identities are written in expectation form. This is exactly where admissibility and Theorem 7.2.3 enter.

7.2.2 Fair Prices and Pricing Bounds

Let $X \geq 0$ be a contingent claim payable at time T. Under an ESMM $\mathbf{Q}$, the risk-neutral price in units of the numéraire is

$$\widetilde{\pi}_{\mathbf{Q}}(X) := \mathbf{E}_{\mathbf{Q}}\left[\frac{X}{N_T}\right],$$

and the corresponding nominal price at time 0 is

$$\pi_{\mathbf{Q}}(X) := N_0 \widetilde{\pi}_{\mathbf{Q}}(X).$$

The pricing bounds in this section are stated under the assumptions of Theorem 7.2.3. The strengthened numéraire definition ensures that nominal admissibility gives an integrable discounted lower bound on every finite horizon, so discounted wealth processes of admissible strategies are true supermartingales under every ESMM.

If an admissible strategy superreplicates X, i.e. $V_T(\varphi) \geq X$, then Theorem 7.2.3 gives

$$\widehat{V}_0(\varphi) \geq \mathbf{E}_{\mathbf{Q}}[\widehat{V}_T(\varphi)] \geq \mathbf{E}_{\mathbf{Q}}\left[\frac{X}{N_T}\right].$$

Thus every superhedging price dominates $\pi_{\mathbf{Q}}(X)$. In complete markets, the ESMM is unique and all no-arbitrage prices coincide with $\pi_{\mathbf{Q}^*}(X)$, where $\mathbf{Q}^*$ is the unique ESMM.

7.3 Application: Dividend-Paying Geometric Brownian Motion

We close the chapter with the canonical example: a stock paying discrete proportional dividends, with a geometric Brownian motion price process between dividend dates. This example uses the topological construction of the stochastic integral, the Itô formula, the numéraire transformation, the generalized Ansel–Stricker theorem, and the Jacod–Yor extension for sigma martingales.

Let $B = (B_t)_{t\geq 0}$ be a standard Brownian motion, let $r > 0$ be constant, and fix a finite horizon T. The bank account

$$N_t = e^{rt}$$

is predictable, strictly positive, and satisfies the numéraire condition since

$$\sup_{0\leq s\leq T} \frac{1}{N_s} = 1.$$

Let $\mu \in \mathbb{R}$ and $\sigma > 0$. Between dividend dates the stock price satisfies

$$\mathrm{d}S_t = \mu S_{t-}\mathrm{d}t + \sigma S_{t-}\mathrm{d}B_t, \qquad S_0 > 0. \tag{7.3}$$

If $0 < \tau_1 < \cdots < \tau_K \leq T$ are deterministic ex-dividend dates and $\delta_k \in (0,1)$ are

proportional dividend rates, then the cumulative dividend process is

$$D_t := \sum_{k:\tau_k \le t} \delta_k S_{\tau_k -}.$$

It is purely discontinuous, hence $[N, D]^c = 0$. The explicit solution is

$$S_t = S_0 \exp\Big(\Big(\mu - \frac{1}{2}\sigma^2\Big)\, t + \sigma B_t\Big) \prod_{k:\tau_k \le t} (1 - \delta_k).$$

Under the bank-account numéraire,

$$\widehat{S}_t = e^{-rt} S_t, \qquad \widehat{D}_t = (1/N) \bullet D_t = \sum_{k:\tau_k \le t} e^{-r\tau_k} \delta_k S_{\tau_k -}.$$

A direct Itô-formula calculation gives

$$\mathrm{d}(\widehat{S}_t + \widehat{D}_t) = \widehat{S}_{t-}\Big((\mu - r)\mathrm{d}t + \sigma \mathrm{d}B_t\Big).$$

With $\theta = (\mu - r)/\sigma$, the density

$$\frac{\mathrm{d}\mathbf{Q}}{\mathrm{d}\mathbf{P}} = \exp\Big(-\theta B_T - \frac{1}{2}\theta^2 T\Big)$$

makes $B_t^{\mathbf{Q}} := B_t + \theta t$ a Brownian motion under $\mathbf{Q}$, and then

$$\mathrm{d}(\widehat{S}_t + \widehat{D}_t) = \sigma \widehat{S}_{t-} \mathrm{d}B_t^{\mathbf{Q}}.$$

Thus the discounted total-return process is a $\mathbf{Q}$-martingale. By Theorem 7.2.3, every admissible discounted wealth process is a $\mathbf{Q}$-supermartingale. Since the Brownian filtration has the predictable representation property, the ESMM is unique, and by Theorem 7.2.6 the price of a contingent claim X is

$$\pi_{\mathbf{Q}}(X) = N_0 \mathbf{E}_{\mathbf{Q}}\Big[\frac{X}{N_T}\Big].$$

For a European call $X = (S_T - K)_+$, this gives the dividend-adjusted Black–Scholes price, with the multiplicative dividend correction $\prod_{k:\tau_k \le T}(1 - \delta_k)$ entering through the explicit solution for S.

7.4 Departures from the Classical Assumptions

The results of this chapter rest on the frictionless, competitive market assumed in Section 7.1. Before closing, we record briefly how the principal economic departures from that idealization sit with the framework. A first observation is that a substantial part of what is often called a market friction is already absorbed into the cumulative dividend process D^i at no cost: the semimartingale generality permits negative increments, so security-level pass-through taxes, management fees, storage costs, and capital calls can be encoded directly as cashflows of D^i. The strategy-dependent frictions discussed below cannot be absorbed in this way and require dedicated extensions.

Transaction costs and bid–ask spreads. The frictionless assumption is reasonable for highly liquid markets and small trades, but fails in retail markets, illiquid securities, and high-turnover strategies. The wealth equation then acquires a strategy-dependent cost term, and the set of self-financing strategies loses its vector-space structure. The standard recovery, due to Kabanov [102] and Guasoni [73], replaces the single price process by bid–ask processes or solvency cones; no-arbitrage is then characterized through consistent price systems or shadow prices. The sigma-martingale arguments of Chapter 6 apply to a selected consistent price process once it has been constructed.

Portfolio constraints. Regulatory capital requirements, position limits, investment mandates, long-only restrictions, and contractual covenants confine the agent to a strict subset $\Phi^c \subset \Phi(S)$ that is in general not a vector space, so the closure properties used in the proof of the First Fundamental Theorem fail on it. The Fundamental Theorems of Asset Pricing then take a one-sided form, in which the absence of constrained arbitrage implies the existence of a measure under which the constrained-strategy wealth process is a supermartingale; see Föllmer and Schied [61, Chapter 9] for a textbook treatment. The results of Section 7.2.1 apply on the unconstrained set and provide the baseline on which the constrained theory is calibrated.

Liquidity and market impact. The price-taker idealization is realistic for trades that are small relative to the depth of the order book, but fails when the agent's trades are large enough to move the market: the execution price then depends on the trade size and the strategy choice feeds back into the price process. Small-trader models, such as Çetin, Jarrow, and Protter [24], treat liquidity as a small perturbation and add a transaction-cost-like term to the wealth equation while keeping the underlying price process exogenous. Large-trader and equilibrium models, such as the structural framework of Fontana [65, Section 6], treat the price process as part of the equilibrium and require the theory to be developed afresh; in this regime the present framework is no longer the right setting.

Underlying all three of the above discussions is the additional robustness provided

by the sigma-martingale level at which the Fundamental Theorems of Section 7.2.1 are formulated: the existence of an equivalent sigma-martingale measure is required, which is strictly weaker than the existence of an equivalent local-martingale measure. The frictionless, competitive setting treated in this chapter is the limiting case against which the friction- and constraint-aware extensions cited above are calibrated.

Chapter 8

Conclusion

This thesis has constructed the general stochastic vector integral through the semimartingale topology of Émery [57], developed a self-contained theory of sigma martingales without recourse to semimartingale characteristics, and applied both to the Fundamental Theorems of Asset Pricing in a general semimartingale market with dividends. Chapter 7 also shows that the abstract results apply to numéraire changes, admissibility, pricing bounds, and an SDE-based dividend-paying model. The three research questions of Chapter 1 have been answered affirmatively; the remainder of this chapter restates those answers as a list of contributions, records the settings in which the approach is and is not the natural choice, and indicates directions for further work.

Topological construction of the integral. The classical and functional analytic constructions both depend on decomposition theorems and both require auxiliary machinery (special semimartingales, $\mathcal{H}^2$ spaces, quasimartingale techniques) to extend the integral to unbounded predictable integrands. Chapter 4 defines the integral as the continuous extension of the elementary integral from simple predictable integrands into the complete scalar semimartingale space $\mathcal{S}^1_e$. The vector-valued case is treated in the same step as the scalar case; bounded and unbounded integrands are handled by the same continuity argument; no decomposition is invoked. This answers Item 1.

Generalized Ansel–Stricker Lemma. The classical Ansel–Stricker Lemma (Theorem 6.1.20) states that a sigma martingale bounded from below by a constant is a local martingale. Theorem 6.1.22 weakens the lower bound to $X \geq M + A$, where M is a local martingale and A is a càdlàg minorant whose supremum process A^* is locally integrable. The role of this strengthening is delineated in Chapter 7: the strengthened numéraire condition of Definition 7.1.1 turns nominal admissibility into an integrable discounted lower bound on every finite horizon, after which Theorem 6.1.22 (applied with $M \equiv 0$) and a Fatou argument give the local-martingale and supermartingale properties of discounted wealth without any further integrability hypothesis on the reciprocal numéraire.

This answers Item 2.

Sigma martingale theory without characteristics. Chapter 6 develops the properties of sigma martingales through stochastic-integral and local-martingale arguments. The class is characterized via σ- and Σ-localizing sequences (Theorem 6.1.7), a locally integrable sigma martingale is shown to be a local martingale, a continuous sigma martingale is shown to be a local martingale, and the classical Ansel–Stricker Lemma is reproved using Fatou's lemma and basic properties of stochastic integrals. Each of these results admits a proof via semimartingale characteristics; the alternative routes given here avoid that apparatus and isolate which hypothesis is doing the work in each statement.

Jacod–Yor representation for sigma martingales. Theorem 6.2.6 establishes, for $X \in \mathcal{M}_\sigma^d$, the equivalence of:

(i) the predictable representation property for sigma martingales — every $Y \in \mathcal{M}_{\sigma,0}^1$ is of the form $H \bullet X$ for some $H \in L(X)$;

(ii) $X^\perp$ contains only constant processes;

(iii) every absolutely continuous sigma martingale measure for X coincides with $\mathbf{P}$;

(iv) $\mathbf{P}$ is an extremal point of $\mathfrak{P}_\sigma(X)$.

Previous versions of this theorem were available for square-integrable martingales [96] and for local martingales; for sigma martingales, the analogous statement does not appear to have been proved before. As a consequence, the Second Fundamental Theorem of Asset Pricing (Theorem 7.2.6) holds under the same market-model hypotheses as the First Fundamental Theorem (Theorem 7.2.5). This answers Item 3.

Implications. The mathematical implication is specific rather than sweeping: the Émery topology, originally introduced as a technical convenience, suffices as the defining structure for the general stochastic vector integral, and a number of arguments routinely phrased in terms of decompositions can be phrased in terms of continuity on a complete metric space instead.

The financial implication is concrete. In general semimartingale models, the equivalent measure supplied by the First Fundamental Theorem is in general only a sigma martingale measure, not a local martingale measure. Theorems formulated at the local-martingale level therefore do not apply in the most general arbitrage-free models considered in this thesis. The Second Fundamental Theorem of Section 7.2.1 and the pricing bounds of Section 7.2.2 are formulated at the sigma-martingale level and apply directly to such models.

Where the approach is and is not the natural choice. The principal setting in which the topological construction earns its keep is the general stochastic vector integral with

unbounded predictable integrands against vector-valued semimartingale integrators. Here, the classical and the functional analytic approaches require additional decomposition theory including long and technical proof strategies. Outside that setting the simpler constructions remain preferable. For càglàd integrands, the functional analytic construction of Protter [166] yields a cleaner definition under ucp convergence with shorter proofs of the basic properties. For continuous integrators, the classical Itô construction and its extensions to continuous local martingales suffice for many standard applications and avoid the abstraction inherent in a non-locally-convex topological vector space.

Two limitations of the topological approach deserve explicit acknowledgement. First, the dominated convergence theorem requires a self-contained technical proof (Theorem 4.4.1) with auxiliary seminorms and a monotone-class step; in the classical and functional analytic settings the same result follows almost immediately from the basic properties of the integral, because those basic properties are proved by different means. In the topological framework the basic properties themselves rely on bounded convergence, so dominated convergence must be established before, not after, the standard machinery. Second, the metric $\|\cdot\|_{\mathcal{S}}$ is not a norm, the space $\mathcal{S}_e^d$ is not locally convex, and continuity of scalar multiplication has to be imposed as a separate axiom (Remark 4.1.10); for readers meeting stochastic integration for the first time, the explicit Itô isometry of the classical approach is likely easier to motivate than the abstract continuous extension used here. The construction is therefore most appropriate for general or research-level treatments; introductory expositions are better served by the classical route.

Further directions. On the theoretical side, the topological construction of Chapter 4 could be extended to infinite-dimensional integrals, with Hilbert- or Banach-space-valued integrators, broadening the theory beyond the finite-dimensional setting treated here. On the applied side, it would be worthwhile to investigate concrete sigma martingale models arising in real markets, in particular highly volatile markets that admit negative prices, such as energy markets, where the classical local-martingale framework is too restrictive and the generality of sigma martingales becomes essential.

Bibliography

[1] Beatrice Acciaio, Julio Backhoff-Veraguas, and Gudmund Pammer. "Quantitative Fundamental Theorem of Asset Pricing". In: *Mathematical Finance* 35.3 (2025), pp. 636–660. DOI: 10.1111/mafi.12457.

[2] Anna Aksamit and Monique Jeanblanc. *Enlargement of Filtration with Finance in View*. SpringerBriefs in Quantitative Finance. Springer International Publishing, 2017. ISBN: 978-3-319-41255-9. DOI: 10.1007/978-3-319-41255-9.

[3] Jean-Pascal Ansel and Christophe Stricker. "Couverture des actifs contingents et prix maximum". In: *Annales de l'institut Henri Poincaré (B) Probabilités et Statistiques*. Vol. 30. Gauthier-Villars. 1994, pp. 303–315. URL: http://www.numdam.org/item/AIHPB_1994__30_2_303_0/.

[4] David Applebaum. *Lévy Processes and Stochastic Calculus*. 2nd ed. Vol. 116. Cambridge Studies in Advanced Mathematics. Cambridge: Cambridge University Press, 2009, p. 490. ISBN: 9780521738651. DOI: 10.1017/CBO9780511809781.

[5] Robert B. Ash and Catherine A. Doléans-Dade. *Probability and Measure Theory*. 2nd ed. San Diego: Academic Press, 2000. ISBN: 9780120652020. URL: https://www.elsevier.com/books/probability-and-measure-theory/ash/978-0-12-065202-0.

[6] Richard F. Bass. "The Doob–Meyer Decomposition Revisited". In: *Canadian Mathematical Bulletin* 39.2 (1996), pp. 138–150. DOI: 10.4153/CMB-1996-018-8.

[7] Mathias Beiglböck, Walter Schachermayer, and Bezirgen Veliyev. "A direct proof of the Bichteler–Dellacherie theorem and connections to arbitrage". In: *The Annals of Probability* 39.6 (2011), pp. 2424–2440. ISSN: 0091-1798. DOI: 10.1214/10-AOP602.

[8] Mathias Beiglböck, Walter Schachermayer, and Bezirgen Veliyev. "A short proof of the Doob–Meyer theorem". In: *Stochastic Processes and their Applications* 122.4 (2012), pp. 1204–1209. ISSN: 0304-4149. DOI: 10.1016/j.spa.2011.12.001.

[9] Mathias Beiglböck and Pietro Siorpaes. “Riemann-integration and a new proof of the Bichteler–Dellacherie theorem”. In: *Stochastic Processes and their Applications* 124.3 (2014), pp. 1226–1235. ISSN: 0304-4149. DOI: 10.1016/j.spa.2013.10.001.

[10] Francesca Biagini et al. *Stochastic Calculus for Fractional Brownian Motion and Applications*. Probability and Its Applications. Springer London, 2008. ISBN: 9781846287978. DOI: 10.1007/978-1-84628-797-8.

[11] Klaus Bichteler. “Stochastic integrators”. In: *Bulletin of the American Mathematical Society* 1.5 (1979), pp. 761–765. DOI: 10.1090/s0273-0979-1979-14655-x.

[12] Klaus Bichteler. “Stochastic Integration and L^p-Theory of Semimartingales”. In: *The Annals of Probability* (1981), pp. 49–89. ISSN: 0091-1798. DOI: 10.1214/aop/1176994308.

[13] Klaus Bichteler. *Stochastic Integration with Jumps*. Vol. 89. Encyclopedia of Mathematics and its Applications. Cambridge: Cambridge University Press, 2002. ISBN: 9780521811293. DOI: 10.1017/CBO9780511549878.

[14] Nicholas Hugh Bingham and Rüdiger Kiesel. *Risk-Neutral Valuation: Pricing and Hedging of Financial Derivatives*. 2nd ed. Springer Finance. London: Springer, 2013. ISBN: 9781447138563. DOI: 10.1007/978-1-4471-3856-3.

[15] Gergely Bodó and Markus Riedle. “Stochastic integration with respect to cylindrical Lévy processes in Hilbert spaces”. In: *Journal of the London Mathematical Society* (2025). DOI: 10.1112/jlms.70298.

[16] James K Brooks and David Neal. “The Optional Stochastic Integral”. In: *Seminar on Stochastic Processes, 1988*. Springer. 1989, pp. 45–54. DOI: 10.1007/978-1-4612-3698-6_2.

[17] Bernard Bru and Marc Yor. “Comments on the life and mathematical legacy of Wolfgang Doeblin”. In: *Finance and Stochastics* 6.1 (2002), pp. 3–47. ISSN: 0949-2984. DOI: 10.1007/s780-002-8399-0.

[18] Zdzisław Brzeźniak, Erika Hausenblas, and Jianliang Zhu. “2D Stochastic Navier–Stokes Equations Driven by Jump Noise”. In: *Stochastic Processes and their Applications* 118 (2008), pp. 2066–2092. DOI: 10.1016/j.spa.2008.09.009.

[19] Donald L Burkholder. “Martingale Transforms”. In: *The Annals of Mathematical Statistics* 37.6 (1966), pp. 1494–1504. DOI: 10.1214/aoms/1177699141.

[20] 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 II: Probability Theory*. Ed. by Lucien M. Le Cam, Jerzy Neyman, and Elizabeth L. Scott. University of California Press, 1972, pp. 223–240. DOI: 10.1525/9780520423671-018.

[21] Aleš Černý and Johannes Ruf. "Pure-jump semimartingales". In: *Bernoulli* 27.4 (2021), pp. 2624–2648. DOI: 10.3150/21-BEJ1325.

[22] Aleš Černý and Johannes Ruf. "Simplified stochastic calculus via semimartingale representations". In: *Electron. J. Probab.* 27 (2022), Paper No. 3, 32. DOI: 10.1214/21-ejp729.

[23] Aleš Černý and Johannes Ruf. "Simplified calculus for semimartingales: Multiplicative compensators and changes of measure". In: *Stochastic Processes and their Applications* 161 (2023), pp. 572–602. DOI: 10.1016/j.spa.2023.04.020.

[24] Umut Çetin, Robert A. Jarrow, and Philip Protter. "Liquidity risk and arbitrage pricing theory". In: *Finance and Stochastics* 8.3 (2004), pp. 311–341. DOI: 10.1007/s00780-004-0123-x.

[25] Geon Ho Choe. *Stochastic Analysis for Finance with Simulations*. Singapore: Springer, 2016. ISBN: 9783319255880. DOI: 10.1007/978-3-319-25589-7.

[26] Ching-Sung Chou. "Caractérisation d'une classe de semimartingales". In: *Séminaire de Probabilités XIII*. Springer, 1979, pp. 250–252. URL: https://www.numdam.org/item/SPS_1979__13__250_0/.

[27] Ching-Sung Chou, Paul-André Meyer, and Christophe Stricker. "Sur les intégrales stochastiques de processus prévisibles non bornés". In: *Séminaire de Probabilités XIV, 1978/79*. Vol. 784. Lecture Notes in Mathematics. Berlin: Springer, 1980, pp. 128–139. DOI: 10.1007/BFb0089478.

[28] Kai Lai Chung and Joseph L. Doob. "Fields, optionality and measurability". In: *Amer. J. Math.* 87 (1965), pp. 397–424. ISSN: 0002-9327. DOI: 10.2307/2373011.

[29] Kai Lai Chung and R. J. Williams. *Introduction to Stochastic Integration*. 2nd ed. Probability and its Applications. Boston, MA: Birkhäuser, 1990, pp. xvi+276. ISBN: 9781461288374. DOI: 10.1007/978-1-4612-4480-6.

[30] E. Çinlar et al. "Semimartingales and Markov processes". In: *Probability Theory and Related Fields* 54.2 (1980), pp. 161–219. DOI: 10.1007/BF00531446.

[31] Samuel N. Cohen and Robert J. Elliott. *Stochastic Calculus and Applications*. Probability and Its Applications. New York: Springer, 2015. ISBN: 9781493928675. DOI: 10.1007/978-1-4939-2867-5.

[32] Rama Cont and Peter Tankov. "Non-parametric calibration of jump-diffusion option pricing models". In: *Journal of Computational Finance* 7.3 (2004), pp. 1–49. DOI: 10.21314/JCF.2004.123.

[33] Philippe Courrège. "Intégrales stochastiques et martingales de carré intégrable". In: *Séminaire Brelot-Choquet-Deny. Théorie du potentiel* 7 (1963), pp. 1–20. URL: https://www.numdam.org/item/SBCD_1962-1963__7_2_A1_0/.

[34] Philippe Courrège. "Décomposition des martingales de carré intégrable". In: *Séminaire Brelot–Choquet–Deny. Théorie du potentiel* 8.1 (1964). Exposé No. 6, pp. 1–14. URL: https://www.numdam.org/item/SBCD_1963-1964__8_1_A2_0/.

[35] Christa Cuchiero and Josef Teichmann. "A Simplified Proof of the Fundamental Theorem of Asset Pricing under No Unbounded Profit with Bounded Risk". In: *Séminaire de Probabilités XLVII*. Vol. 2137. Lecture Notes in Mathematics. Springer, 2015, pp. 109–121. DOI: 10.1007/978-3-319-18585-4_6.

[36] Mark HA Davis and Pravin Varaiya. "The multiplicity of an increasing family of σ-fields". In: *The Annals of Probability* (1974), pp. 958–963. ISSN: 0091-1798. DOI: 10.1214/aop/1176996562.

[37] Marzia De Donno, Paolo Guasoni, and Maurizio Pratelli. "Super-replication and utility maximization in large financial markets". In: *Stochastic Processes and their Applications* 115.12 (2005), pp. 2006–2022. DOI: 10.1016/j.spa.2005.06.010.

[38] Marzia De Donno and Maurizio Pratelli. "On a lemma by Ansel and Stricker". In: *Séminaire de Probabilités XL*. Springer, 2007, pp. 411–414. DOI: 10.1007/978-3-540-71189-6_22.

[39] Freddy Delbaen and Walter Schachermayer. "A general version of the fundamental theorem of asset pricing". In: *Mathematische Annalen* 300.1 (1994), pp. 463–520. ISSN: 0025-5831. DOI: 10.1007/BF01450498.

[40] Freddy Delbaen and Walter Schachermayer. "The Fundamental Theorem of Asset Pricing for Unbounded Stochastic Processes". In: *Mathematische Annalen* 312.2 (1998), pp. 215–250. ISSN: 0025-5831. DOI: 10.1007/s002080050220.

[41] Freddy Delbaen and Walter Schachermayer. *The Mathematics of Arbitrage*. Springer Finance. Springer, 2006. ISBN: 9783540219927. DOI: 10.1007/978-3-540-31299-4.

[42] Claude Dellacherie. “Un survol de la théorie de l’intégrale stochastique”. In: *Stochastic Processes and their Applications* 10.2 (1980), pp. 115–144. ISSN: 0304-4149. DOI: 10.1016/0304-4149(80)90017-4.

[43] Claude Dellacherie. “Mesurabilité des débuts et théorème de section : le lot à la portée de toutes les bourses”. In: *Séminaire de Probabilités XV 1979/80*. Springer, 1981, pp. 351–370. URL: https://www.numdam.org/item/SPS_1981__15__351_0/.

[44] Claude Dellacherie and Paul-André Meyer. *Probabilities and Potential*. North-Holland Mathematics Studies. Amsterdam: Hermann, 1978. ISBN: 9780720407013. URL: https://archive.org/details/probabilitespote0000dell.

[45] Claude Dellacherie and Paul-André Meyer. *Probabilities and Potential B: Theory of Martingales*. Vol. 72. North-Holland Mathematics Studies. Amsterdam: North-Holland, 1982. ISBN: 9780444865267. URL: https://archive.org/details/probabilitiespot0000della.

[46] Catherine Doléans. “Processus croissants naturels et processus croissants très-bien-mesurables”. In: *C. R. Acad. Sci. Paris Sér. A-B* 264 (1967), A874–A876. ISSN: 0151-0509. URL: https://gallica.bnf.fr/ark:/12148/cb343492882/date.

[47] Catherine Doléans. “Existence du processus croissant naturel associé à un potentiel de la classe (D)”. In: *Z. Wahrscheinlichkeitstheorie und Verw. Gebiete* 9 (1968), pp. 309–314. DOI: 10.1007/BF00531754.

[48] Catherine Doléans-Dade. “On the existence and unicity of solutions of stochastic integral equations”. In: *Z. Wahrscheinlichkeitstheorie und Verw. Gebiete* 36.2 (1976), pp. 93–101. DOI: 10.1007/BF00533992.

[49] Catherine Doléans-Dade and Paul-André Meyer. “Intégrales stochastiques par rapport aux martingales locales”. In: *Séminaire de Probabilités IV, Université de Strasbourg*. Vol. 124. Lecture Notes in Mathematics. Berlin: Springer, 1970, pp. 77–107. DOI: 10.1007/BFb0059337.

[50] Marzia De Donno and Maurizio Pratelli. “Stochastic Integration with Respect to a Sequence of Semimartingales”. In: *In Memoriam Paul-André Meyer: Séminaire de Probabilités XXXIX*. Vol. 1874. Lecture Notes in Mathematics. Berlin: Springer, 2006, pp. 119–135. DOI: 10.1007/978-3-540-35513-7_10.

[51] Joseph L. Doob. *Stochastic Processes*. Vol. 7. Wiley Classics Library. New York: John Wiley & Sons, 1953. URL: https://archive.org/details/stochasticproces0000doob.

[52] Richard Durrett. *Stochastic Calculus*. Probability and Stochastics Series. A practical introduction. Boca Raton, FL: CRC Press, 1996, pp. x+341. ISBN: 0-8493-8071-5. DOI: 10.1201/9780203738283.

[53] Eugene Borisovich Dynkin. *Markov Processes. Vols. I, II*. Vol. 121–122. Die Grundlehren der mathematischen Wissenschaften. Translated with the authorization and assistance of the author by J. Fabius, V. Greenberg, A. Maitra, and G. Majone. Berlin; New York: Springer; Academic Press, Inc., 1965, Vol. I: xii+365, Vol. II: viii+274. DOI: 10.1007/978-3-662-00031-1.

[54] Stefan Ebenfeld. *Grundlagen der Finanzmathematik: Mathematische Methoden, Modellierung von Finanzmärkten und Finanzprodukten*. Stuttgart: Schäffer-Poeschel, 2007. ISBN: 9783791026343. URL: https://www.schaeffer-poeschel.de/shop/grundlagen-der-finanzmathematik.html.

[55] Robert J. Elliott and Peter E. Kopp. *Mathematics of Financial Markets*. 2nd ed. Springer Finance. New York: Springer, 2005, pp. xii+352. ISBN: 0-387-21292-2. DOI: 10.1007/b97681.

[56] Paul Embrechts and Makoto Maejima. *Selfsimilar processes*. Princeton Series in Applied Mathematics. Princeton University Press, Princeton, NJ, 2002, pp. xii+111. ISBN: 0-691-09627-9. DOI: 10.1515/9781400825103.

[57] Michel Émery. "Une topologie sur l'espace des semimartingales". In: *Séminaire de Probabilités XIII*. Vol. 721. Lecture Notes in Mathematics. Berlin: Springer, 1979, pp. 260–280. DOI: 10.1007/BFb0066206.

[58] Michel Émery. "Compensation de processus à variation finie non localement intégrables". In: *Séminaire de Probabilités XIV (1978/79)*. Ed. by Jacques Azéma and Marc Yor. Vol. 784. Lecture Notes in Mathematics. Springer, 1980, pp. 152–160. DOI: 10.1007/BFb0089481.

[59] Michel Émery. "Métrisabilité de quelques espaces de processus aléatoires". In: *Séminaire de probabilités de Strasbourg* 14 (1980), pp. 140–147. URL: https://www.numdam.org/item/SPS_1980__14__140_0/.

[60] Michel Émery. *Stochastic Calculus in Manifolds*. Universitext. With an appendix by P.-A. Meyer. Berlin: Springer, 1989, pp. x+151. ISBN: 3-540-51664-6. DOI: 10.1007/978-3-642-75051-9.

[61] Hans Föllmer and Alexander Schied. *Stochastic Finance: An Introduction in Discrete Time*. 2nd. Vol. 27. De Gruyter Studies in Mathematics. Updated and expanded edition, introducing discrete-time methods for stochastic finance. Berlin: Walter de Gruyter, 2011. ISBN: 9783110218053. DOI: 10.1515/9783110218053.

[62] C. A. Fonseca-Mora. "Convergence uniform on compacts in probability with applications to stochastic analysis in duals of nuclear spaces". In: *Stochastic Analysis and Applications* 42.4 (2024), pp. 797–827. DOI: 10.1080/07362994.2024.2383678.

[63] Christian A. Fonseca-Mora. "Stochastic integration with respect to cylindrical semimartingales". In: *Electronic Journal of Probability* 26 (2021), paper no. 147, 1–38. DOI: 10.1214/21-EJP715.

[64] Christian A. Fonseca-Mora. *Vector-Valued Stochastic Integration with Respect to Semimartingales in the Dual of Nuclear Spaces*. 2025. arXiv: 2503.03647 [math.PR]. URL: https://arxiv.org/abs/2503.03647.

[65] Claudio Fontana. "Weak and strong no-arbitrage conditions for continuous financial markets". In: *International Journal of Theoretical and Applied Finance* 18.01 (2015), p. 1550005. ISSN: 0219-0249. DOI: 10.1142/s0219024915500053.

[66] Avner Friedman. *Stochastic Differential Equations and Applications. Vol. 1*. Vol. 28. Probability and Mathematical Statistics. New York-London: Academic Press, 1975, pp. xiii+231. URL: https://archive.org/details/stochasticdiffer0001frie.

[67] Adriano M. Garsia. *Martingale Inequalities: Seminar Notes on Recent Progress*. Mathematics Lecture Note Series. Reading, Mass.–London–Amsterdam: W. A. Benjamin, Inc., 1973, pp. viii+184. URL: https://search.worldcat.org/title/848215.

[68] Leszek Gawarecki and Vidyadhar Mandrekar. *Stochastic Differential Equations in Infinite Dimensions: With Applications to Stochastic Partial Differential Equations*. Probability and Its Applications. Berlin Heidelberg: Springer, 2010. ISBN: 9783642161940. DOI: 10.1007/978-3-642-14403-5.

[69] Ī. Ī. Gihman and A. V. Skorohod. *Stochastic Differential Equations*. Vol. 72. Ergebnisse der Mathematik und ihrer Grenzgebiete [Results in Mathematics and Related Areas]. New York–Heidelberg: Springer, 1972, pp. viii+354. DOI: 10.1007/978-3-642-88264-7.

[70] Thomas Goll, Jan Kallsen, et al. "A complete explicit solution to the log-optimal portfolio problem". In: *The Annals of Applied Probability* 13.2 (2003), pp. 774–799. DOI: 10.1214/aoap/1050689603.

[71] Bronius Grigelionis and Remigijus Mikulevičius. "Stochastic Evolution Equations and Densities of the Conditional Distributions". In: *Theory and Application of Random Fields*. Vol. 49. Lecture Notes in Control and Information Sciences. Berlin: Springer, 1983, pp. 49–88. DOI: 10.1007/BFb0006320.

[72] Mircea Grigoriu. *Stochastic Calculus*. Applications in Science and Engineering. Boston, MA: Birkhäuser Boston, Inc., 2002, pp. xiv+774. ISBN: 0-8176-4242-0. DOI: 10.1007/978-0-8176-8228-6.

[73] Paolo Guasoni. "No arbitrage under transaction costs, with fractional Brownian motion and beyond". In: *Mathematical Finance* 16.3 (2006), pp. 569–582. ISSN: 0960-1627. DOI: 10.1111/j.1467-9965.2006.00283.x.

[74] Alexander A. Gushchin. *Stochastic Calculus for Quantitative Finance*. London: ISTE Press - Elsevier, 2015. ISBN: 978-1-78548-034-8. URL: https://www.elsevier.com/books/stochastic-calculus-for-quantitative-finance/gushchin/978-1-78548-034-8.

[75] István Gyöngy and N. V. Krylov. "On stochastic equations with respect to semimartingales II. Itô formula in Banach spaces". In: *Stochastics* 6.3-4 (1982), pp. 153–173. DOI: 10.1080/17442508208833208.

[76] John Michael Harrison and David M Kreps. "Martingales and arbitrage in multiperiod securities markets". In: *Journal of economic theory* 20.3 (1979), pp. 381–408. DOI: 10.1016/0022-0531(79)90043-7.

[77] John Michael Harrison and Stanley R Pliska. "Martingales and stochastic integrals in the theory of continuous trading". In: *Stochastic Processes and their Applications* 11.3 (1981), pp. 215–260. ISSN: 0304-4149. DOI: 10.1016/0304-4149(81)90026-0.

[78] John Michael Harrison and Stanley R Pliska. "A stochastic calculus model of continuous trading: Complete markets". In: *Stochastic Processes and their Applications* 15.3 (1983), pp. 313–316. ISSN: 0304-4149. DOI: 10.1016/0304-4149(83)90038-8.

[79] Sheng-Wu He, Jia-Gang Wang, and Jia-An Yan. *Semimartingale Theory and Stochastic Calculus*. Boca Raton, FL: Kexue Chubanshe (Science Press), Beijing; CRC Press, 1992, pp. xiv+546. ISBN: 7-03-003066-4. DOI: 10.1201/9780203738306.

[80] Gilbert A. Hunt. "Markoff processes and potentials I". In: *Illinois Journal of Mathematics* 1.1 (1957), pp. 44–93. DOI: 10.1215/ijm/1255380682.

[81] Gilbert A. Hunt. "Markoff processes and potentials II". In: *Illinois Journal of Mathematics* 1.3 (1957), pp. 316–369. DOI: 10.1215/ijm/1255380388.

[82] Gilbert A. Hunt. "Markoff processes and potentials III". In: *Illinois Journal of Mathematics* 2.2 (1958), pp. 151–213. DOI: 10.1215/ijm/1255381342.

[83] Albrecht Irle. *Finanzmathematik*. 3rd ed. Studienbücher Wirtschaftsmathematik. Die Bewertung von Derivaten. [The evaluation of derivatives]. Wiesbaden: Springer Spektrum, 2012, p. 343. ISBN: 978-3-8348-1574-3. DOI: 10.1007/978-3-8348-8314-8.

[84] Kiyosi Itô. “Stochastic integral”. In: *Proc. Imp. Acad.* 20.8 (1944), pp. 519–524. DOI: 10.3792/pia/1195572786.

[85] Kiyosi Itô. “On a stochastic integral equation”. In: *Proc. Japan Acad.* 22.2 (1946), pp. 32–35. DOI: 10.3792/pja/1195572371.

[86] Kiyosi Itô. “Stochastic differential equations in a differentiable manifold”. In: *Nagoya Math. J.* 1 (1950), pp. 35–47. DOI: 10.1017/S0027763000022819.

[87] Kiyosi Itô. “Multiple Wiener integral”. In: *Journal of the Mathematical Society of Japan* 3.1 (1951), pp. 157–169. DOI: 10.2969/jmsj/00310157.

[88] Kiyosi Itô. “On a formula concerning stochastic differentials”. In: *Nagoya Math. J.* 3 (1951), pp. 55–65. DOI: 10.1017/S0027763000012216.

[89] Kiyosi Itô. *On Stochastic Differential Equations*. Vol. 4. Memoirs of the American Mathematical Society. Providence, RI: American Mathematical Society, 1951. DOI: 10.1090/memo/0004.

[90] Kiyosi Itô. *Lectures on Stochastic Processes*. Vol. 24. Bombay: Tata Institute of Fundamental Research, 1961. URL: https://www.math.tifr.res.in/~publ/ln/tifr24.pdf.

[91] Kiyosi Itô. “Transformation of Markov processes by multiplicative functionals”. In: *Ann. Inst. Fourier* 15.1 (1965), pp. 15–30. URL: http://www.numdam.org/item/AIF_1965__15_1_13_0/.

[92] Jean Jacod. “Sur la construction des intégrales stochastiques et les sous-espaces stables de martingales”. In: *Séminaire de Probabilités XI*. Vol. 581. Lecture Notes in Mathematics. Berlin: Springer, 1977, pp. 390–410. URL: https://www.numdam.org/item/SPS_1977__11__390_0/.

[93] Jean Jacod. *Calcul Stochastique et Problèmes de Martingales*. Vol. 714. Lecture Notes in Mathematics. Berlin–Heidelberg: Springer, 1979, pp. x+280. ISBN: 978-3-540-09587-3. DOI: 10.1007/BFb0064907.

[94] Jean Jacod. “Intégrales stochastiques par rapport à une semimartingale vectorielle et changements de filtration”. In: *Séminaire de Probabilités XIV, 1978/79*. Vol. 784. Lecture Notes in Mathematics. Berlin: Springer, 1980, pp. 161–172. URL: https://www.numdam.org/item/SPS_1980__14__161_0/.

[95] Jean Jacod and Albert N. Shiryaev. *Limit Theorems for Stochastic Processes*. 2nd ed. Vol. 288. Grundlehren der Mathematischen Wissenschaften. Berlin: Springer, 2003, pp. xx+661. ISBN: 3-540-43932-3. DOI: 10.1007/978-3-662-05265-5.

[96] Jean Jacod and Marc Yor. "Étude des solutions extrémales et représentation intégrale des solutions pour certains problèmes de martingales". In: *Probability Theory and Related Fields* 38.2 (1977), pp. 83–125. DOI: 10.1007/BF00533669.

[97] Adam Jakubowski. "An Almost Sure Approximation for the Predictable Process in the Doob–Meyer Decomposition Theorem". In: *Séminaire de Probabilités XXXVIII*. Vol. 1857. Lecture Notes in Mathematics. Berlin: Springer, 2005, pp. 158–164. DOI: 10.1007/978-3-540-31449-3_11.

[98] Adam Jakubowski and Markus Riedle. "Stochastic integration with respect to cylindrical Lévy processes". In: *Ann. Probab.* 45.6B (2017), pp. 4273–4306. ISSN: 0091-1798. DOI: 10.1214/16-AOP1164.

[99] Klaus Jänich. *Topology*. Undergraduate Texts in Mathematics. Springer New York Inc., 1984. ISBN: 9781461270188. DOI: 10.1007/978-1-4757-1793-8.

[100] Robert Jarrow and Philip Protter. "A short history of stochastic integration and mathematical finance: the early years, 1880–1970". In: IMS Lecture Notes Monogr. Ser. 45 (2004), pp. 75–91. DOI: 10.1214/lnms/1196285381.

[101] Guy Johnson and LL Helms. "Class D supermartingales". In: *Bulletin of the American Mathematical Society* 69.1 (1963), pp. 59–62. DOI: 10.1090/S0002-9904-1963-10892-9.

[102] Yuri M. Kabanov. "Hedging and Liquidation under Transaction Costs in Currency Markets". In: *Finance and Stochastics* 3.2 (1999), pp. 237–248. DOI: 10.1007/s007800050062.

[103] Svenja Kaden and Jürgen Potthoff. "Progressive stochastic processes and an application to the Itô integral". In: *Stochastic Anal. Appl.* 22.4 (2004), pp. 843–865. ISSN: 0736-2994. DOI: 10.1081/SAP-120037621.

[104] Olav Kallenberg. *Foundations of Modern Probability*. 2nd ed. Probability and Its Applications. New York: Springer, 2002, pp. xvi+650. ISBN: 978-0-387-95313-7. DOI: 10.1007/978-1-4757-4015-8.

[105] Gopinath Kallianpur. *Stochastic Filtering Theory*. Vol. 13. Applications of Mathematics. New York–Berlin: Springer, 1980, pp. xvi+316. ISBN: 0-387-90445-X. DOI: 10.1007/978-1-4757-6592-2.

[106] Jan Kallsen. "σ-localization and σ-martingales". In: *Theory of Probability & Its Applications* 48.1 (2004), pp. 152–163. DOI: 10.1137/S0040585X97980827.

[107] Dhandapani Kannan and Vangipuram Lakshmikantham. *Handbook of Stochastic Analysis and Applications*. Boca Raton, FL: CRC Press, 2001, p. 808. ISBN: 9780367578732. DOI: 10.1201/9780203908440.

[108] Rajeeva L. Karandikar and B. V. Rao. "On the second fundamental theorem of asset pricing". In: *Communications on Stochastic Analysis* 10.1 (2016), Article 5, 57–81. DOI: 10.31390/cosa.10.1.05.

[109] Rajeeva L. Karandikar and B. V. Rao. *Introduction to Stochastic Calculus*. Indian Statistical Institute Series. Singapore: Springer, 2018, pp. xiii+441. ISBN: 978-981-10-8317-4. DOI: 10.1007/978-981-10-8318-1.

[110] Ioannis Karatzas and Steven E. Shreve. *Brownian Motion and Stochastic Calculus*. 2nd ed. Vol. 113. Graduate Texts in Mathematics. New York: Springer, 1991, pp. xvi+470. ISBN: 978-0-387-97655-6. DOI: 10.1007/978-1-4612-0949-2.

[111] Ioannis Karatzas and Steven E. Shreve. *Methods of Mathematical Finance*. Applications of mathematics. Springer, 1998. ISBN: 9780387948393. DOI: 10.1007/978-1-4939-6845-9.

[112] Constantinos Kardaras. "Stochastic integration with respect to arbitrary collections of continuous semimartingales and applications to Mathematical Finance". In: *Annals of Applied Probability* 34.3 (2024), pp. 2566–2599. DOI: 10.1214/23-AAP1942.

[113] Fima C. Klebaner. *Introduction to Stochastic Calculus with Applications*. 2nd ed. London: Imperial College Press, 2005, pp. xvi+412. ISBN: 9781848168329. DOI: 10.1142/p386.

[114] Achim Klenke. *Probability Theory: A Comprehensive Course*. Universitext. London: Springer London, 2008. ISBN: 978-1-84800-048-3. DOI: 10.1007/978-1-84800-048-3.

[115] Martin Koller. *Life Insurance Risk Management Essentials*. EAA Series. Berlin: Springer, 2011, pp. xii+339. ISBN: 978-3-642-20720-4. DOI: 10.1007/978-3-642-20721-1.

[116] János Komlós. "A generalisation of a problem of Steinhaus". In: *Acta Mathematica Academiae Scientiarum Hungaricae* 18.1-2 (1967), pp. 217–229. ISSN: 0001-5954. DOI: 10.1007/BF02020976.

[117] Hayri Korezlioglu and Claude Martaias. "Stochastic integration for operator valued processes on Hilbert spaces and on nuclear spaces". In: *Stochastics: An International Journal of Probability and Stochastic Processes* 24.3 (1988), pp. 171–219. DOI: 10.1080/17442508808833515.

[118] Tomasz Kosmala and Markus Riedle. "Stochastic integration with respect to cylindrical Lévy processes by p-summing operators". In: *J. Theoret. Probab.* 34.1 (2021), pp. 477–497. ISSN: 0894-9840. DOI: 10.1007/s10959-019-00978-x.

[119] Hiroshi Kunita. "Stochastic integrals based on martingales taking values in Hilbert space". In: *Nagoya Mathematical Journal* 38 (1970), pp. 41–52. DOI: 10.1017/s0027763000013507.

[120] Hiroshi Kunita. *Stochastic Flows and Stochastic Differential Equations*. Vol. 24. Cambridge Studies in Advanced Mathematics. Cambridge: Cambridge University Press, 1997, pp. xvi+346. ISBN: 9780521631297. DOI: 10.1017/CBO9780511609572.

[121] Hiroshi Kunita and Shinzo Watanabe. "On square integrable martingales". In: *Nagoya Mathematical Journal* 30 (1967), pp. 209–245. DOI: 10.1017/s0027763000012484.

[122] Hui-Hsiung Kuo. *Introduction to Stochastic Integration*. 1st ed. Universitext. New York: Springer, 2006, pp. x+278. ISBN: 978-0-387-28720-1. DOI: 10.1007/0-387-31057-6.

[123] Hui-Hsiung Kuo. "The Itô calculus and white noise theory: a brief survey toward general stochastic integration". In: *Communications on Stochastic Analysis* 8.1 (2014), p. 8. URL: https://digitalcommons.lsu.edu/cosa/vol8/iss1/7/.

[124] Alfred U. Kussmaul. *Stochastic Integration and Generalized Martingales*. Vol. 11. Research Notes in Mathematics. London–San Francisco, Calif.–Melbourne: Pitman Publishing, 1977, pp. v+163. URL: https://search.worldcat.org/title/3036050.

[125] Ronald Larsen. *Functional analysis: an introduction*. Pure and applied mathematics. M. Dekker, 1973. ISBN: 9780824760427. URL: https://search.worldcat.org/title/685053.

[126] Jean-François Le Gall. *Brownian Motion, Martingales, and Stochastic Calculus*. Vol. 274. Graduate Texts in Mathematics. Cham: Springer, 2016, pp. xvii+188. ISBN: 9783319310886. DOI: 10.1007/978-3-319-31089-9.

[127] Érik Lenglart. "Semi-martingales et intégrales stochastiques en temps continu". In: *Revue du CETHEDEC* 20.75 (1983), pp. 91–160. URL: https://search.worldcat.org/issn/0397-0892.

[128] Giorgio Letta. *Martingales et intégration stochastique*. Publications of the Scuola Normale Superiore. Pisa, Italy: Edizioni della Normale, Scuola Normale Superiore, 1984, p. 190. ISBN: 978-88-7642-290-4. URL: https://books.google.de/books?id=ORfvAAAAMAAJ.

[129] Paul Lévy. *Théorie de l'addition des variables aléatoires*. French. Paris: Gauthier-Villars, 1937. URL: https://gallica.bnf.fr/ark:/12148/bpt6k3221520.

[130] Shu Lin. "Stochastic analysis of fractional Brownian motions". In: *Stochastics: An International Journal of Probability and Stochastic Processes* 55.1-2 (1995), pp. 121–140. DOI: 10.1080/17442509508834002.

[131] Robert Liptser and A.N. Shiryaev. *Theory of Martingales*. Vol. 49. Mathematics and its Applications. Dordrecht: Kluwer Academic Publishers, 1989. ISBN: 978-0-7923-0395-4. DOI: 10.1007/978-94-009-2438-3.

[132] Henry P. McKean. *Stochastic Integrals*. New York: Academic Press, 1969. URL: https://archive.org/details/stochasticintegr0000mcke.

[133] Peter Medvegyev. *Stochastic Integration Theory*. Vol. 14. Oxford Graduate Texts in Mathematics. Oxford: Oxford University Press, 2007, pp. xii+338. ISBN: 978-0-19-921525-6. DOI: 10.1093/oso/9780199215256.001.0001.

[134] Dieter Meintrup and Stefan Schäffler. *Stochastik: Theorie und Anwendungen. Statistik und ihre Anwendungen*. German. 1st ed. Berlin: Springer, 2005, pp. xii+624. ISBN: 978-3-540-21676-6. DOI: 10.1007/b137972.

[135] Jean Mémin. "Espaces de semi martingales et changement de probabilité". In: *Probability Theory and Related Fields* 52.1 (Jan. 1980), pp. 9–39. ISSN: 1432-2064. DOI: 10.1007/BF00534184.

[136] Michel Métivier. *Semimartingales: A Course on Stochastic Processes*. Vol. 2. De Gruyter studies in mathematics. Berlin/New York: Walter de Gruyter, 1982, pp. xi+287. ISBN: 9783110086744. DOI: 10.1515/9783110844603.

[137] Michel Métivier and Jean Pellaumail. *Stochastic Integration*. Translated from the French. New York–London: Academic Press, 1980, pp. xiv+344. ISBN: 978-0-12-491450-6. URL: https://archive.org/details/stochasticintegr0000meti.

[138] Michel Métivier and G. Pistone. "Une formule d'isométrie pour l'intégrale stochastique hilbertienne et équations d'évolution linéaires stochastiques". In: *Zeitschrift für Wahrscheinlichkeitstheorie und verwandte Gebiete* 33.1 (1975), pp. 1–18. DOI: 10.1007/BF00531862.

[139] Paul-André Meyer. "A decomposition theorem for supermartingales". In: *Illinois Journal of Mathematics* 6.2 (1962), pp. 193–205. DOI: 10.1215/ijm/1255632318.

[140] Paul-André Meyer. "Decomposition of supermartingales: The uniqueness theorem". In: *Illinois Journal of Mathematics* 7.1 (1963), pp. 1–17. DOI: 10.1215/ijm/1255637477.

[141] Paul-André Meyer. *Probabilités et Potentiel*. Paris: Hermann, 1966. URL: https://search.worldcat.org/title/491108.

[142] Paul-André Meyer. *Probabilités et Potentiel*. Vol. 14. Publications de l'Institut de Mathématique de l'Université de Strasbourg. Paris: Hermann, 1966. URL: https://search.worldcat.org/title/491108.

[143] Paul-André Meyer. "Intégrales stochastiques I". In: *Séminaire de probabilités de Strasbourg* 1 (1967), pp. 72–94. URL: https://www.numdam.org/item/SPS_1967__1__72_0/.

[144] Paul-André Meyer. "Intégrales stochastiques II". In: *Séminaire de probabilités de Strasbourg* 1 (1967), pp. 95–117. URL: https://www.numdam.org/item/SPS_1967__1__95_0/.

[145] Paul-André Meyer. "Intégrales stochastiques III". In: *Séminaire de probabilités de Strasbourg* 1 (1967), pp. 118–141. URL: https://www.numdam.org/item/SPS_1967__1__118_0/.

[146] Paul-André Meyer. "Intégrales stochastiques IV". In: *Séminaire de Probabilités X (1974/75)*. Ed. by Claude Dellacherie and Paul-André Meyer. Vol. 511. Lecture Notes in Mathematics. Springer, 1976, pp. 252–298. DOI: 10.1007/BFb0083875.

[147] Paul-André Meyer. "Le théorème fondamental sur les martingales locales". In: *Séminaire de Probabilités XI* 581 (1977), pp. 463–464. DOI: 10.1007/BFb0087213.

[148] Paul-André Meyer. "Notes sur les intégrales stochastiques. IV. Caractérisation de BMO par un opérateur maximal". In: Lecture Notes in Math., Vol. 581 (1977), pp. 446–481. DOI: 10.1007/bf01932733.

[149] Paul-André Meyer. "Caractérisation des semimartingales, d'après Dellacherie". In: *Séminaire de Probabilités XIII*. Ed. by Claude Dellacherie, Paul-André Meyer, and Marc Weil. Vol. 721. Lecture Notes in Mathematics. Berlin: Springer Berlin Heidelberg, 1979, pp. 620–623. ISBN: 978-3-540-09505-7. DOI: 10.1007/BFb0070899.

[150] Paul-André Meyer. "Les processus stochastiques de 1950 à nos jours". In: *Development of Mathematics 1950–2000*. Basel: Birkhäuser, 2000, pp. 813–848. DOI: 10.1007/978-3-0348-8470-9_31.

[151] Paul-André Meyer. "Un Cours sur les Intégrales Stochastiques". French. In: *Séminaire de Probabilités 1967–1980*. Ed. by M. Émery and M. Yor. Vol. 1771. Lecture Notes in Mathematics. Berlin: Springer Berlin Heidelberg, 2002, pp. 174–329. ISBN: 978-3-540-42813-8. DOI: 10.1007/978-3-540-45530-1_11.

[152] R. Mikulevičius and B. L. Rozovskii. "Martingale problems for stochastic PDEs". In: *Stochastic partial differential equations: six perspectives* 64 (1999), pp. 243–326. DOI: 10.1090/pspum/064/1742115.

[153] Remigijus Mikulevičius and Boris L. Rozovskii. "Normalized Stochastic Integrals in Topological Vector Spaces". In: *Séminaire de Probabilités XXXII*. Vol. 1686. Lecture Notes in Mathematics. Berlin: Springer, 1998, pp. 137–165. DOI: 10.1007/BFb0101756.

[154] Minoru Motoo and Shinzo Watanabe. "On a class of additive functionals of Markov processes". In: *Journal of Mathematics of Kyoto University* 4.3 (1965), pp. 429–469. DOI: 10.1215/kjm/1250524698.

[155] Michael Mürmann. *Wahrscheinlichkeitstheorie und stochastische Prozesse*. German. Wiesbaden: Springer Vieweg, 2014. ISBN: 978-3-642-38159-1. DOI: 10.1007/978-3-658-04076-0.

[156] Marek Musiela and Marek Rutkowski. *Martingale Methods in Financial Modelling*. 2nd ed. Vol. 36. Stochastic Modelling and Applied Probability. Berlin: Springer, 2005, pp. xx+638. ISBN: 9783540209669. DOI: 10.1007/b137866.

[157] Jacques Neveu. *Martingales à temps discret*. Paris: Masson et Cie, Éditeurs, 1972, pp. vii+218. URL: https://search.worldcat.org/title/798497.

[158] David Nualart and Wim Schoutens. "Chaotic and predictable representations for Lévy processes". In: *Stochastic Processes and their Applications* 90.1 (2000), pp. 109–122. ISSN: 0304-4149. DOI: 10.1016/S0304-4149(00)00038-8.

[159] Bernt Øksendal. *Stochastic Differential Equations: An Introduction with Applications*. 6th ed. Universitext. Berlin: Springer, 2010, p. 404. ISBN: 9783540047582. DOI: 10.1007/978-3-642-14394-6.

[160] Raymond Edward Alan Christopher Paley, Norbert Wiener, and Antoni Zygmund. "Notes on random functions". In: *Mathematische Zeitschrift* 37.1 (1933), pp. 647–668. DOI: 10.1007/bf01474606.

[161] Eckhard Platen and Nicola Bruti-Liberati. *Numerical Solution of Stochastic Differential Equations with Jumps in Finance*. Vol. 64. Springer Science & Business Media, 2010. DOI: 10.1007/978-3-642-13694-8.

[162] David Pollard, Richard D. Gill, and B. D. Ripley. *A User's Guide to Measure Theoretic Probability*. Cambridge Series in Statistical and Probabilistic Mathematics. Cambridge University Press, 2002. ISBN: 9780521002899. DOI: 10.1017/CBO9780511811555.

[163] Jean-Luc Prigent. *Weak Convergence of Financial Markets*. Springer Finance. Berlin: Springer, 2003, pp. xiv+422. ISBN: 3-540-42333-8. DOI: 10.1007/978-3-540-24831-6.

[164] Philip Protter. "Semimartingales and stochastic differential equations". In: *Ecole d'été de probabilités de Saint-Flour XIII–1983*. Ed. by Jacques Azéma and Marc Yor. Vol. 1215. Lecture Notes in Mathematics. Springer, 1986, pp. 85–175. DOI: 10.1007/BFb0074923.

[165] Philip Protter. "Stochastic integration without tears". In: *Stochastics: An International Journal of Probability and Stochastic Processes* 16.3-4 (1986), pp. 295–325. DOI: 10.1080/17442508608833378.

[166] Philip E. Protter. *Stochastic Integration and Differential Equations: Version 2.1*. 2nd ed. Vol. 21. Stochastic Modelling and Applied Probability. Berlin: Springer, 2005, pp. xviii+415. ISBN: 9783642055607. DOI: 10.1007/978-3-662-10061-5.

[167] K. M. Rao. "Quasi-martingales". In: *Mathematica Scandinavica* 24.1 (1969), pp. 79–92. DOI: 10.7146/math.scand.a-10987.

[168] K. Murali Rao. "On decomposition theorems of Meyer". In: *Math. Scand.* 24 (1969), pp. 66–78. ISSN: 0025-5521. DOI: 10.7146/math.scand.a-10920.

[169] Malempati M. Rao. *Stochastic Processes: General Theory*. Dordrecht: Springer, 1995, pp. xii+330. ISBN: 978-0-7923-3725-6. DOI: 10.1007/978-94-015-8475-3.

[170] Daniel Revuz and Marc Yor. *Continuous Martingales and Brownian Motion*. 3rd ed. Vol. 293. Grundlehren der mathematischen Wissenschaften A series of comprehensive studies in mathematics. Berlin, Heidelberg: Springer, 1999, p. 606. ISBN: 9783540643258. DOI: 10.1007/978-3-662-06400-9.

[171] Markus Riedle. "Stochastic integration with respect to cylindrical Lévy processes in Hilbert spaces: an L^2 approach". In: *Infin. Dimens. Anal. Quantum Probab. Relat. Top.* 17.1 (2014), pp. 1450008, 19. ISSN: 0219-0257. DOI: 10.1142/S0219025714500088.

[172] Leonard Christopher Gordon Rogers. "Arbitrage with fractional Brownian motion". In: *Mathematical Finance* 7.1 (1997), pp. 95–105. ISSN: 0960-1627. DOI: 10.1111/1467-9965.00025.

[173] Leonard Christopher Gordon Rogers and David Williams. *Diffusions, Markov Processes and Martingales: Volume 2, Itô Calculus*. Cambridge Mathematical Library. Cambridge: Cambridge University Press, 2000, pp. xvi+480. ISBN: 9780521775939. DOI: 10.1017/CBO9781107590120.

[174] Walter Rudin. *Real & Complex Analysis*. McGraw-Hill, 1987. ISBN: 9780070619876. URL: https://search.worldcat.org/title/13002293.

[175] Walter Rudin. *Functional Analysis*. 2nd ed. New York: McGraw-Hill, 1991. ISBN: 978-0-07-054236-5. URL: https://www.worldcat.org/isbn/9780070542365.

[176] Ludger Rüschendorf. *Stochastic Processes and Financial Mathematics*. Cham: Springer, 2023, pp. xii+428. ISBN: 978-3-662-64710-3. DOI: 10.1007/978-3-662-64711-0.

[177] Albert N. Shiryaev. *Essentials of Stochastic Finance. Facts, Models, Theory*. Vol. 3. Advanced Series on Statistical Science & Applied Probability. River Edge, NJ: World Scientific Publishing Co., Inc., 1999, pp. xvi+834. ISBN: 981-02-3605-0. DOI: 10.1142/9789812385192.

[178] Albert N. Shiryaev and Alexander Cherny. "Vector stochastic integrals and the fundamental theorems of asset pricing". In: *Proceedings of the Steklov Institute of Mathematics-Interperiodica Translation* 237 (2002), pp. 6–49. URL: https://www.mathnet.ru/eng/tm310.

[179] Moritz Sohns. "Utility Indifference Pricing in the Heston Model: Pricing, Hedging and Shortcomings". In: *Proceedings of the 9th Ročník Konference Doktorandů Na Vysoké škole Finanční A Správní*. Ed. by O. Roubai. 2022, pp. 132–153. ISBN: 978-80-7408-250-4. URL: https://is.vsfs.cz/lide/clovek.pl?id=22842.

[180] Moritz Sohns. "σ-Martingales: Foundations, Properties, and a New Proof of the Ansel–Stricker Lemma". In: *Mathematics* 13.4 (2025), p. 682. ISSN: 2227-7390. DOI: 10.3390/math13040682.

[181] Moritz Sohns. "General Stochastic Vector Integration - Three Approaches". In: *Preprints* (Jan. 2025). DOI: 10.20944/preprints202501.0645.v1.

[182] Moritz Sohns. "Minimal Entropy and Entropic Risk Measures: A Unified Framework via Relative Entropy". In: *Risks* 13.4 (2025). ISSN: 2227-9091. DOI: 10.3390/risks13040070.

[183] Moritz Sohns. "The General Semimartingale Market Model". In: *AppliedMath* 5.3 (2025). ISSN: 2673-9909. DOI: 10.3390/appliedmath5030097.

[184] Moritz Sohns. "The Jacod–Yor Theorem for Sigma Martingales and the Second Fundamental Theorem of Asset Pricing". In: *Journal of Stochastic Analysis* 6.2 (2025), Article 1. DOI: 10.31390/josa.6.2.01.

[185] Peter Tankov. *Financial Modeling with Lévy Processes*. Lecture notes, Special Lectures for PhD Students. Lecture delivered October 4–8, 2010 at IM PAN, Warsaw. Available online. Institute of Mathematics of the Polish Academy of Sciences (IM PAN) / Banach Center, 2010. URL: https://www.impan.pl/images/Notatki-z-wykadw/Skrypty/tankov2.pdf.

[186] John B. Walsh. “An Introduction to Stochastic Partial Differential Equations”. In: *École d’Été de Probabilités de Saint-Flour XIV–1984*. Vol. 1180. Lecture Notes in Mathematics. Berlin: Springer, 1986, pp. 265–439. DOI: 10.1007/BFb0074920.

[187] Albert T. Wang. “Quadratic variation of functionals of Brownian motion”. In: *The Annals of Probability* (1977), pp. 756–769. ISSN: 0091-1798. DOI: 10.1214/aop/1176995803.

[188] Shinzo Watanabe. “The Japanese contributions to martingales”. In: *J. Électron. Hist. Probab. Stat.* 5.1 (2009), p. 13. ISSN: 1773-0074. DOI: 10.5802/jep.1.

[189] Norbert Wiener. “Differential-space”. In: *Journal of Mathematics and Physics* 2.1-4 (1923), pp. 131–174. DOI: 10.1002/sapm192321131.

[190] Norbert Wiener. “The Dirichlet problem”. In: *Journal of Mathematics and Physics* 3.3 (1924), pp. 127–146. DOI: 10.1002/sapm192433127.

[191] David Williams. *Probability with Martingales*. Cambridge: Cambridge University Press, 1991. ISBN: 9780521406055. DOI: 10.1017/CBO9780511813658.

[192] Jia-An Yan. *Introduction to Stochastic Finance*. 1st ed. Universitext. Singapore: Springer, 2018. ISBN: 978-981-13-1657-9. DOI: 10.1007/978-981-13-1657-9.

[193] James J. Yeh. *Real Analysis: Theory Of Measure And Integration (3rd Edition)*. World Scientific Publishing Company, 2014. ISBN: 9789814578561. URL: https://books.google.de/books?id=R0o8DQAAQBAJ.

[194] Ch. Yoeurp. “Décompositions des martingales locales et formules exponentielles”. In: *Séminaire de Probabilités X, Université de Strasbourg*. Vol. 511. Lecture Notes in Mathematics. Berlin: Springer, 1976, pp. 432–480. DOI: 10.1007/BFb0080192.

[195] Marc Yor. “Un exemple de processus qui n’est pas une semi-martingale”. In: *Temps Locaux* 52 (1978), pp. 219–222. URL: https://www.numdam.org/item/SL_1978__52__219_0/.

[196] Kôsaku Yosida. *Functional Analysis*. 6th ed. Vol. 123. Grundlehren der mathematischen Wissenschaften. Berlin: Springer, 1980. ISBN: 978-3-540-58654-8. DOI: 10.1007/978-3-642-61859-8.

Index

$(\mathcal{H}^2, X)$-integrable, 58
$L(X)$, 60, 64, 97
$L^1(X)$, 106
L^2-martingale, 32
$L^2(M)$, 35
$L^p_{\mathrm{loc}}(M)$, 62
X-integrable, 47, 49
$\mathcal{BMO}$, 12
Λ, 53
Λ^d, 72
Λ^d_u, 72
Λ_{ucp}, 55
Σ-localizing sequence, 154
$\mathcal{H}^2$, 57
$\mathbb{D}_{\mathrm{ucp}}$, 55
$\mathbb{L}_{\mathrm{ucp}}$, 55
$\mathcal{E}$, 23
$\mathcal{E}_0$, 26
$\mathcal{E}_M$, 35
$\mathcal{E}_{\mathrm{Itô}}$, 25
$\mathcal{H}^2$-norm, 57
$\mathcal{M}^2$, 32
$\mathcal{S}$, 49
$\mathcal{S}^d$, 74
$\mathcal{S}^d_e$, 74
$\|\cdot\|_X$, 94
$\|\cdot\|_{\mathcal{BMO}}$, 12
$\|\cdot\|_{\mathcal{H}^p}$, 10
$\|\cdot\|_M$, 34
$\overline{\mathcal{E}}$, 26
σ-localizing sequence, 151
d_X, 58
$d_X(\cdot,\cdot)$, 94
$L^{\mathrm{var}}(A)$, 63
X-integrable, 51
$\mathscr{L}^2$, 26
$\mathcal{BMO}$-Norm, 12

absolutely continuous, 14
adapted, 7
adapted process
    with càdlàg paths, 55
    with càglàd paths, 55
admissible, 188

Bichteler-Dellacherie-Mokobodzki Theorem, 54, 92
Bichteler–Dellacherie–Mokobodzki Theorem, 54
BMO, 12
Bounded Convergence Theorem, 106
Brownian motion, 13

Class (D), 30
classical semimartingale, 43
close, 9
conditional quadratic covariation, 33
càdlàg, 7
càglàd, 7

Doléans-Dade exponential, 141
Dominated Convergence Theorem for stochastic integrals, 100
Doob class, *see* Class (D)
Doob-Meyer decomposition
  case without Class (D), 38
  general case, 30

elementary predictable process, 23
equivalent, 14
equivalent sigma martingale measure, 188
ESMM, 188
essential supremum, 7
evanescent, 7
exponential
  stochastic, 141
extremal, 175

finite variation process, 13
Fundamental Theorem of Local Martingales, 46

good integrator, 54

indicator function, 7
indistinguishable, 7
integrand, 72
integrator, 72
Itô norm, 25
Itô formula, 130
  multi-dimensional, 136
Itô Isometry, 27, 36, 59

jointly measurable, 15

Kunita-Watanabe Inequality, 36

local martingale, 9
local times, 138
locally, 8
locally bounded, 8
locally integrable, 8
Lévy's theorem, 143
  multi-dimensional, 145

martingale, 8
  L^2, 32
  square-integrable, 32
mean variation, 14
modifications, 7

natural process, 30
NFLVR, 189
No Free Lunch with Vanishing Risk, 189
numéraire, 185

orthogonal, 174

polarization identity, 121
predictable, 16, 17
predictable σ-algebra, 13
predictable process, 13
predictable quadratic variation, 33, 38
predictable representation property (PRP), 174
prelocally, 8
process
  elementary predictable, 23
  jointly measurable, 16
  natural, 30
  predictable, 16
  progressive, 16
  simple predictable, 53, 72
product measurable, 15, 17
progressively measurable, 15, 17

quadratic covariation, 41, 56, 121
quadratic variation, 41, 56, 121
quasimartingale, 14

Radon-Nikodym derivative, 14

random partition, 122

semimartingale
  $\mathcal{H}^2$, 57
  classical, 43
  special, 47
  topological, 74
semimartingale topology, 48, 73
sharp-bracket process, 33, 38
sigma martingale, 150, 151
simple predictable process, 53, 72
simple predictable processes, 23
space of semimartingales $\mathcal{H}^2$, 57
special semimartingale, 47
square-bracket process, 41
square-integrable martingale, 32
stable, 175
stochastic differential equation, 139
stochastic exponential, 141
stochastic integral, 24, 26, 34, 36, 39, 44, 47, 49–51, 55, 56, 59, 60, 62, 72, 97
stochastic intervals, 8
stochastic process, 7
stopped process, 8
stopping time, 8
submartingale, 8
supermartingale, 8

tensor quadratic variation, 121
topological semimartingale, 74
topology of ucp convergence, 73
trading strategy, 184

ucp, 55
uniformly integrable, 9
uniformly on compacts in probability, 55
unique, 141
usual conditions, 6

wealth process, 184

Émery norm, 48, 73
Émery topology, 48, 73