City Research Online

City, University of London Institutional Repository

Citation: Černý, A. & Ruf, J. (2022). Simplified stochastic calculus via semimartingale representations. Electronic Journal of Probability, 27, pp. 1-32. doi: 10.1214/21-EJP729

This is the published version of the paper.

This version of the publication may differ from the final published version.

Permanent repository link: https://openaccess.city.ac.uk/id/eprint/27246/

Link to published version: https://doi.org/10.1214/21-EJP729

Copyright: City Research Online aims to make research outputs of City, University of London available to a wider audience. Copyright and Moral Rights remain with the author(s) and/or copyright holders. URLs from City Research Online may be freely distributed and linked to.

Reuse: Copies of full items can be used for personal research or study, educational, or not-for-profit purposes without prior permission or charge. Provided that the authors, title and full bibliographic details are credited, a hyperlink and/or URL is given for the original metadata page and the content is not changed in any way.

City Research Online: http://openaccess.city.ac.uk/ publications@city.ac.uk

Electronic Journal of Probability

Electron. J. Probab. **27** (2022), article no. 3, 1–32.
ISSN: 1083-6489 https://doi.org/10.1214/21-EJP729

Simplified stochastic calculus via semimartingale representations

Aleš Černý* Johannes Ruf†

Abstract

We develop a stochastic calculus that makes it easy to capture a variety of predictable transformations of semimartingales such as changes of variables, stochastic integrals, and their compositions. The framework offers a unified treatment of real-valued and complex-valued semimartingales. The proposed calculus is a blueprint for the derivation of new relationships among stochastic processes with specific examples provided below.

Keywords: semimartingale representation; complex-valued process; generalized Yor formula; Émery formula; Itô formula.
MSC2020 subject classifications: 60G07; 60G44; 60G48; 60H05; 60H05.
Submitted to EJP on October 9, 2020, final version accepted on December 7, 2021.
Supersedes arXiv:2006.11914.

1 Introduction

"Because in mathematics we pile inferences upon inferences, it is a good thing whenever we can subsume as many of them as possible under one symbol. For once we have understood the true significance of an operation, just the sensible apprehension of its symbol will suffice to obviate the whole reasoning process that earlier we had to engage anew each time the operation was encountered."

— Carl Jacobi (1804–1851) [31, p. 67]

We study the following concept. A semimartingale Y is said to be represented by a semimartingale X if, roughly speaking, there is a predictable function ξ acting on the increments of X such that the increments of Y satisfy $\mathrm{d}Y_t = \xi_t(\mathrm{d}X_t)$, where $\xi_t(\mathrm{d}X_t)$ is given some "natural" meaning. Such representation of Y in terms of X, if it exists, is measure-invariant. One hopes that common operations on Y yield processes that are again X–representable, for example,

*Bayes Business School, City, University of London. E-mail: ales.cerny.1@city.ac.uk
†Department of Mathematics, London School of Economics and Political Science. E-mail: j.ruf@lse.ac.uk

(i) a stochastic integral $\zeta_t \mathrm{d}Y_t$ "ought to" yield

$$\zeta_t \mathrm{d}Y_t = \zeta_t \xi_t(\mathrm{d}X_t); \tag{1.1}$$

(ii) for a change of variables by means of some smooth function f it should be true that

$$\mathrm{d}f(Y_t) = f(Y_{t-} + \xi_t(\mathrm{d}X_t)) - f(Y_{t-}); \tag{1.2}$$

(iii) for a new process Z such that $\mathrm{d}Z_t = \psi_t(\mathrm{d}Y_t)$ one would like to obtain the composition rule

$$\mathrm{d}Z_t = \psi_t(\xi_t(\mathrm{d}X_t)). \tag{1.3}$$

In integral form we shall write, e.g.,

$$Y_t = Y_0 + \xi \circ X_t = Y_0 + \int_0^t \xi_s(\mathrm{d}X_s).$$

The purpose of the calculus (1.1)–(1.3) is to reduce the computational burden in a generic modelling situation where one starts from a (multivariate) process X whose predictable P–characteristics relative to some truncation function are given as the primitive input to the problem. The process X, which is trivially representable with respect to itself, is transformed by several applications of Propositions 3.14 and 3.15 and Theorem 3.18, i.e., by the rules (1.1)–(1.3) above, to another process Y which is also X–representable. In many situations the required end product is the P–drift of Y, here denoted B^Y; e.g., when computing exponential compensators; see Duffie, Filipović, and Schachermayer [11, Proposition 11.2]. A generic formula for the drift of a represented process, $B^{\xi \circ X}$, is given in Proposition 5.6.

We begin with a brief description of some historical background of this paper. The first seeds of measure-invariant stochastic calculus were planted by McKean [25], who would write the classical Itô formula in the form

$$\mathrm{d}f(X_t) = f'(X_t)\mathrm{d}X_t + \frac{1}{2} f''(X_t)(\mathrm{d}X_t)^2, \tag{1.4}$$

and only afterwards substitute the canonical decomposition of X in the first term and the quadratic variation of X in the second term on the right-hand side of (1.4). With the development of general semimartingale integration it soon became clear that (1.4) is fully rigorous, as written, for any continuous semimartingale X and any sufficiently smooth function f on an appropriate domain; see Doléans-Dade and Meyer [10, Théorème 8].

From here it is not a big conceptual leap to want to study general transformations of the increments $\mathrm{d}X_t$ by means of some predictable function ξ. Precisely this was suggested by Émery [13] together with the notation $\mathrm{d}Y_t = \xi(\mathrm{d}X_t)$ and a specific measure-invariant formula for $\xi(\mathrm{d}X_t)$, for time-constant deterministic ξ and matrix(!)-valued X (because Émery's goal at the time was to study the natural exponential of a matrix with stochastic coefficients). As far as we know, nobody has up until now attempted to build a coherent calculus based on Émery's formula. See also Remark 3.10 for other connections to the literature.

Next, let us offer a flavour of the simplifications the calculus can achieve in conjunction with drift calculations. For example, for real-valued X and $\alpha \in \mathbb{C}$ the calculus permits one to write

$$\mathrm{e}^{\alpha(X - X_0)} = \mathscr{E}((\mathrm{e}^{\alpha\,\mathrm{id}} - 1) \circ X), \tag{1.5}$$

where $\mathscr{E}$ is the Doléans-Dade stochastic exponential (see [9]). If X has independent increments and the expectation is finite, one then obtains (see Černý and Ruf [3, Theorem 4.1])

$$\mathsf{E}\big[\mathrm{e}^{\alpha(X_t - X_0)}\big] = \mathscr{E}\big(B^{(\mathrm{e}^{\alpha\mathrm{id}} - 1)\circ X}\big)_t. \tag{1.6}$$

Table 1: Useful identities involving the power function, stochastic exponential $\mathscr{E}$, and stochastic logarithm $\mathcal{L}$.

Name	Assumptions	Conclusions
P4.1	$\alpha \in \mathbb{Z}_+$; or $\alpha \in \mathbb{Z}$; $\Delta X \neq -1$; or $\alpha \in \mathbb{C} \setminus \mathbb{Z}$; $\operatorname{Re}\mathscr{E}(X) > 0$	$\mathscr{E}(X)^\alpha = \mathscr{E}\left(((1+\mathrm{id})^\alpha - 1) \circ X\right)$
P4.1	analogous	$\mathscr{E}\big(X^{(1)}\big)^\alpha \mathscr{E}\big(X^{(2)}\big)^\beta = \mathscr{E}\left(\big((1+\mathrm{id}_1)^\alpha (1+\mathrm{id}_2)^\beta - 1\big) \circ X\right)$
P4.1	$\alpha, \beta \in \mathbb{Z}$; $X, X_- \neq 0$; or $\alpha, \beta \in \mathbb{C} \setminus \mathbb{Z}$; ...	$\mathcal{L}\left(\big(X^{(1)}\big)^\alpha \big(X^{(2)}\big)^\beta\right)$ $= \big((1+\mathrm{id}_1)^\alpha (1+\mathrm{id}_2)^\beta - 1\big) \circ \big(\mathcal{L}\big(X^{(1)}\big), \mathcal{L}\big(X^{(2)}\big)\big)$
P4.2	none	(i) $\mathcal{L}\big(\mathrm{e}^X\big) = (\mathrm{e}^{\mathrm{id}} - 1) \circ X$; (ii) $\lvert\mathscr{E}(X)\rvert = \mathscr{E}\left((\lvert 1+\mathrm{id}\rvert - 1) \circ X\right)$
P4.2	$\Delta X \neq -1$	(i) $\mathscr{E}(X) = \mathrm{e}^{\log(1+\mathrm{id}) \circ X}$; (ii) $\log\lvert\mathscr{E}(X)\rvert = \log\lvert 1+\mathrm{id}\rvert \circ X$
P4.2	$\operatorname{Re}\mathscr{E}(X) > 0$	$\log \mathscr{E}(X) = \log(1+\mathrm{id}) \circ X$
P4.3	$\alpha \in (0, \infty)$; or $\alpha \in \mathbb{C} \setminus \mathbb{R}_+$; $\Delta X \neq -1$	$\lvert\mathscr{E}(X)\rvert^\alpha = \mathscr{E}\left((\lvert 1+\mathrm{id}\rvert^\alpha - 1) \circ X\right)$
P4.3	analogous	$\big\lvert\mathscr{E}\big(X^{(1)}\big)\big\rvert^\alpha \big\lvert\mathscr{E}\big(X^{(2)}\big)\big\rvert^\beta = \mathscr{E}\left(\Big(\lvert 1+\mathrm{id}_1\rvert^\alpha \lvert 1+\mathrm{id}_2\rvert^\beta - 1\Big) \circ X\right)$
P4.3	$X, X_- \neq 0$	$\mathcal{L}\left(\big\lvert X^{(1)}\big\rvert^\alpha \big\lvert X^{(2)}\big\rvert^\beta\right)$ $= \big(\lvert 1+\mathrm{id}_1\rvert^\alpha \lvert 1+\mathrm{id}_2\rvert^\beta - 1\big) \circ \big(\mathcal{L}\big(X^{(1)}\big), \mathcal{L}\big(X^{(2)}\big)\big)$

When X is a Lévy process and α is purely imaginary, the right-hand side of (1.6) is just the Lévy-Khintchin formula; see [3, Corollary 4.2]. Useful formulae akin to (1.5) are collected in Table 1.

The calculus can do more. Staying with X that has independent increments, suppose that instead of the natural exponential $\mathrm{e}^{X - X_0}$ in (1.5), the starting object is the stochastic exponential $\mathscr{E}(X) > 0$. The calculus now provides a formula for the Mellin transform

$$\mathsf{E}\big[\mathscr{E}(X)_t^\alpha\big] = \mathsf{E}\big[\mathscr{E}(((1+\mathrm{id})^\alpha - 1) \circ X)_t\big] = \mathscr{E}\big(B^{((1+\mathrm{id})^\alpha - 1) \circ X}\big)_t.$$

When $\mathscr{E}(X)$ is signed, one can evaluate $\lvert\mathscr{E}(X)\rvert^\alpha$ and $\operatorname{sgn}(\mathscr{E}(X))\lvert\mathscr{E}(X)\rvert^\alpha$ separately to obtain

$$\mathsf{E}\big[\lvert\mathscr{E}(X)_t\rvert^\alpha\big] = \mathscr{E}\big(B^{(\lvert 1+\mathrm{id}\rvert^\alpha - 1) \circ X}\big)_t; \tag{1.7}$$

$$\mathsf{E}\big[\operatorname{sgn}(\mathscr{E}(X))\lvert\mathscr{E}(X)_t\rvert^\alpha\big] = \mathscr{E}\big(B^{(\operatorname{sgn}(1+\mathrm{id})\lvert 1+\mathrm{id}\rvert^\alpha - 1) \circ X}\big)_t; \tag{1.8}$$

see [3, Examples 4.4 and 4.5]. We refer the reader also to the introductory paper Černý and Ruf [5], where other concrete illustrations of the calculus are given.[1]

On the theoretical side, the paper introduces the class $\mathfrak{U}$ of universal representing functions that are well-behaved with respect to operations (1.1)–(1.3); if one uses only locally bounded integration, change of variables, and composition, one is guaranteed never to leave $\mathfrak{U}$, which makes the calculus completely straightforward. For example, the representing functions in (1.5)–(1.8) and also those in Table 1 are all in $\mathfrak{U}$. The most important results pertaining to the class $\mathfrak{U}$ are highlighted in Table 2.

Furthermore, we develop a coherent theory for a wider class $\mathfrak{I}(X)$ of representing functions specific to X, in which $\mathfrak{U}$ appears as a special case. Here the "natural" composition rules (1.1) and (1.3) sometimes fail. We study sufficient conditions for their

[1] This paper is conceptually different from [5] in two important respects. First, we provide a unified treatment of real-valued and complex-valued representations where [5] only considers two ad-hoc non-interacting subsets of representing functions that must be applied separately to real-valued and complex-valued processes, respectively. For example, [5] cannot handle the formulae (1.7) and (1.8). Second, [5] operates strictly inside a special class $\mathfrak{U}$, introduced here.

Table 2: Summary of statements for $\mathfrak{U}$ — the class of universal representing functions. Here $\mathfrak{I}(X)$ denotes the class of predictable functions for which $\xi \circ X$ is well-defined.

Name	Assumptions	Conclusions
P3.6	$\xi \in \mathfrak{U}$; $\xi(\Delta X)$ finite	$\xi \in \mathfrak{I}(X)$
P3.13	none	$\mathrm{id}, \mathrm{id}^2 \in \mathfrak{U}$; $X = X_0 + \mathrm{id} \circ X$; $[X, X] = \mathrm{id}^2 \circ X$
P3.14	ζ locally bounded predictable	$\zeta\mathrm{id} \in \mathfrak{U}$; $\zeta \cdot X = \zeta\mathrm{id} \circ X$
P3.15	$f : \mathbb{R} \to \mathbb{R}$ smooth	$f(X_- + \mathrm{id}) - f(X_-) \in \mathfrak{U}$; $f(X) = f(X_0) + (f(X_- + \mathrm{id}) - f(X_-)) \circ X$
T3.17	$\xi, \psi \in \mathfrak{U}$; $\psi(\xi(\Delta X))$ finite	$\psi(\xi) \in \mathfrak{U}$; $\xi, \psi(\xi) \in \mathfrak{I}(X)$; $\psi \in \mathfrak{I}(\xi \circ X)$; $\psi \circ (\xi \circ X) = \psi(\xi) \circ X$

Table 3: Summary of statements for $\mathfrak{I}(X)$ — the class of predictable functions for which $\xi \circ X$ is well-defined.

Name	Assumptions	Conclusions
P3.13	$\xi \in \mathfrak{I}(X)$	$\xi \in \mathfrak{I}(X - X_0)$; $\xi \circ X = \xi \circ (X - X_0)$; $\Delta(\xi \circ X) = \xi(\Delta X)$
P3.14	$\zeta \in L(X)$	$\zeta\mathrm{id} \in \mathfrak{I}(X)$; $\zeta \cdot X = \zeta\mathrm{id} \circ X$
R3.19	$\zeta \in L(X)$; $\psi \in \mathfrak{I}(\zeta \cdot X)$	$\psi(\zeta\mathrm{id}) \in \mathfrak{I}(X)$; $\psi \circ (\zeta\mathrm{id} \circ X) = \psi(\zeta\mathrm{id}) \circ X$
C3.20	$\xi \in \mathfrak{I}(X)$; $\psi \in \mathfrak{I}(\xi \circ X)$; $\psi'(0)$ locally bounded	$\psi(\xi) \in \mathfrak{I}(X)$; $\psi \circ (\xi \circ X) = \psi(\xi) \circ X$

validity and offer counterexamples when such conditions are not met. The proposed framework does deliver closedness under composition for general stochastic integrals without further assumptions; this and other important properties of the class $\mathfrak{I}(X)$ are collected in Table 3.

We shall say more on the benefits of the calculus in the concluding Section 6 once all notation has been introduced. The basic message is encouraging: with appropriate care one can hide much of the required stochastic analysis (stochastic integrals, jump-measure integrals) under the hood and treat common operations on stochastic processes *algebraically*, as compositions of functions (indeed, $\xi \circ X$ can be interpreted in some cases as the ξ–variation of X; see Remark 3.10). The benefits of doing so are significant, especially in the context of measure changes.

One might expect the operations (1.1)–(1.3) to always work when the representing process X is a pure-jump process of finite variation. Using only standard techniques, this intuition is false, however, because an integral of a finite variation semimartingale need not itself be of finite variation. We do obtain universal validity of rules (1.1)–(1.3) for pure-jump processes after suitably extending the standard integrals with respect to random measures. This universality then applies to all representing processes X that belong *sigma-locally* to the class of finite-variation pure-jump semimartingales; see Subsection 2.5.

The paper is organized as follows. Section 2 introduces notation and reviews important concepts such as integration with respect to a complex-valued semimartingale. Section 3 defines representation of a semimartingale and derives important properties thereof, such as (1.1)–(1.3). Here one gets to see an explicit formula for the object $\xi \circ X$, which is formulated in terms of real derivatives of the complex function ξ. This formula looks quite natural in the special case when both ξ and X are real-valued; such simplicity is also preserved when ξ is analytic at 0 but this is much harder to see in the original definition. Subsection 3.3 provides an alternative form of the most general representation formula in terms of so-called Wirtinger derivatives, where the simplification in the analytic case is plainly visible. Section 4 lists and proves a number of useful representations, among them generalizations of the Yor formula, thereby illustrating the strength of the proposed calculus. This section also provides counterexamples that document tightness of the results obtained in Section 3. Section 5 summarizes the computation of predictable characteristics of a represented semimartingale. Finally, Section 6 discusses additional benefits of the proposed calculus and directions for future research.

2 Setup and notation

This section provides background on complex numbers and the probabilistic setup. It furthermore reviews stochastic integration for complex-valued semimartingales, the notion of predictable functions, and sigma-localized integrals with respect to random measures.

2.1 The lift from $\mathbb{C}$ to $\mathbb{R}^2$

Below, we explicitly allow quantities to be complex-valued in order to allow for a consistent treatment of complex integrals, exponentials, etc., and in particular characteristic functions. The reader interested only in real-valued calculus can easily skip this subsection and always replace the general '$\mathbb{C}$–valued' by the special case '$\mathbb{R}$–valued' in their mind. Throughout this section, let $m \in \mathbb{N}$ denote an integer. To simplify notation later on, we write $\overline{\mathbb{C}}^m = \mathbb{C}^m \cup \{\mathrm{NaN}\}$ for some 'non-number' $\mathrm{NaN} \notin \bigcup_{k\in\mathbb{N}} \mathbb{C}^k$. We introduce the function $\mathrm{id} : \overline{\mathbb{C}}^m \to \overline{\mathbb{C}}^m$ by $\mathrm{id}(v) = v$.

The definitions below now hinge on the identification map $\hat{\mathrm{id}} : \overline{\mathbb{C}} \to \mathbb{R}^2 \cup \{\mathrm{NaN}\}$ given by

$$\hat{\mathrm{id}}(v) = \begin{bmatrix} \mathrm{Re}\, v \\ \mathrm{Im}\, v \end{bmatrix}, \quad v \in \mathbb{C}; \qquad \hat{\mathrm{id}}(\mathrm{NaN}) = \mathrm{NaN},$$

and its appropriate multidimensional extension, again denoted by $\hat{\mathrm{id}} : \overline{\mathbb{C}}^m \to \mathbb{R}^{2m} \cup \{\mathrm{NaN}\}$ given by

$$\hat{\mathrm{id}}(v) = (\mathrm{Re}\, v_1, \mathrm{Im}\, v_1, \ldots, \mathrm{Re}\, v_m, \mathrm{Im}\, v_m)^\top, \quad v \in \mathbb{C}^m; \qquad \hat{\mathrm{id}}(\mathrm{NaN}) = \mathrm{NaN}.$$

Observe that $\hat{\mathrm{id}}(v) \in \mathbb{R}^{2m}$ for $v \in \mathbb{C}^m$ contains the values of $\mathrm{Re}\, v$ and $\mathrm{Im}\, v$, interlaced. At times we silently use matrix-valued versions of these canonical maps, which are taken to double the row dimension but which we do not introduce formally to avoid excessive notation.

So as not to obscure the main ideas with notation, we will highlight the key properties of the lift $\hat{\mathrm{id}}$ for $m = 1$. To this end, the inverse map to $\hat{\mathrm{id}}$ is $\hat{\mathrm{id}}^{-1} : \mathbb{R}^2 \cup \{\mathrm{NaN}\} \to \overline{\mathbb{C}}$ given by

$$\hat{\mathrm{id}}^{-1}\left([x\ y]^\top\right) = x + iy, \quad [x\ y]^\top \in \mathbb{R}^2; \qquad \hat{\mathrm{id}}^{-1}(\mathrm{NaN}) = \mathrm{NaN}.$$

The following two properties of $\hat{\mathrm{id}}$ are of importance:

- $\hat{\mathrm{id}}$ and $\hat{\mathrm{id}}^{-1}$ are linear, when restricted to $\mathbb{C}$ and $\mathbb{R}^2$;
- for $u, v \in \mathbb{C}$ one obtains

$$\hat{\mathrm{id}}(uv) = \left[\hat{\mathrm{id}}(u)\ \hat{\mathrm{id}}(iu)\right] \hat{\mathrm{id}}(v) = \begin{bmatrix} \operatorname{Re} u & -\operatorname{Im} u \\ \operatorname{Im} u & \operatorname{Re} u \end{bmatrix} \begin{bmatrix} \operatorname{Re} v \\ \operatorname{Im} v \end{bmatrix}. \tag{2.1}$$

2.2 Probabilistic quantities

We fix a probability space $(\Omega, \mathscr{F}, \mathsf{P})$ with a right-continuous filtration $\mathfrak{F}$. We shall assume, without loss of generality, that all semimartingales are right-continuous, and have left limits almost surely. For a brief review of standard results without the assumption that the filtration is augmented by null sets, see Perkowski and Ruf [29, Appendix A]. We follow mostly the notation of Jacod and Shiryaev [18].

For a $\mathbb{C}^m$–valued stochastic process V we shall write $\hat{V} = \hat{\mathrm{id}}(V)$ for the corresponding $\mathbb{R}^{2m}$–valued process.

Definition 2.1 (Complex-valued process properties)**.** *A $\mathbb{C}^m$–valued stochastic process V is said to have a certain property, for example to be a semimartingale (respectively, martingale; local martingale; special semimartingale; process of finite variation; process with independent increments; predictable; locally bounded; etc.) if the $\mathbb{R}^{2m}$–valued process $\hat{V} = \hat{\mathrm{id}}(V)$ has that same property, i.e., if $\hat{V}$ is a semimartingale (respectively, martingale, etc.).* □

We denote the left-limit process of a (complex-valued) semimartingale V by V_- and use the convention $V_{0-} = V_0$. We also set $\Delta V = V - V_-$; in particular we have $\Delta V_0 = 0$. For complex-valued processes, the quadratic variation process is defined to be bilinear.[2] That is, for $\mathbb{C}$–valued semimartingales V and U we set

$$[V, U] = [\operatorname{Re} V, \operatorname{Re} U] - [\operatorname{Im} V, \operatorname{Im} U] + i\left([\operatorname{Re} V, \operatorname{Im} U] + [\operatorname{Im} V, \operatorname{Re} U]\right).$$

We have again $[V, U]_0 = 0$. If V is $\mathbb{C}^m$–valued, then $[V, V]$ denotes the corresponding $\mathbb{C}^{m\times m}$–valued quadratic variation, formally given by

$$[V, V] = (I_m \otimes [1\ i]) \left[\hat{V}, \hat{V}\right] \left(I_m \otimes \begin{bmatrix} 1 \\ i \end{bmatrix}\right),$$

where I_m denotes the $m \times m$ identity matrix and $\otimes$ the Kronecker product. Observe that for a $\mathbb{C}^{n\times m}$–valued matrix R we have $[RV, RV] = R[V, V]R^\top$. Furthermore, we write $[V, V]^c$ for the continuous part of the quadratic variation $[V, V]$ (the latter being of finite variation).

Remark 2.2 (Alternative characterisations of $[V, V]^c$)**.** We might call $V^c(\mathsf{P})$ the continuous local martingale part of a semimartingale V; see [18, I.4.27]. Note that $V^c(\mathsf{P})$ depends on the underlying measure P. To wit, for two equivalent measures $\mathsf{Q} \sim \mathsf{P}$, we usually have $V^c(\mathsf{Q}) \neq V^c(\mathsf{P})$ if $\mathsf{Q} \neq \mathsf{P}$. Nevertheless, we always have

$$[V, V]^c = [V^c(\mathsf{P}), V^c(\mathsf{P})] = [V^c(\mathsf{Q}), V^c(\mathsf{Q})];$$

see also Dellacherie and Meyer [7, Theorem VIII.27] and Protter [30, p. 70]. □

Let μ^V denote the jump measure of a semimartingale V and ν^V its predictable compensator (under a fixed probability measure P). Then for a $\mathbb{C}$–valued bounded predictable function ξ (a precise definition is provided in Subsection 2.4) with $\xi(0) = 0$ we have

$$\xi * \mu^V = \xi\big(\hat{\mathrm{id}}^{-1}\big) * \mu^{\hat{V}} = \sum_{t \le \cdot} \xi_t(\Delta V_t),$$

[2] The bilinear definition is more prevalent. It is used, for example, in Doléans-Dade [9], Émery [14], Revuz and Yor [32], and Protter [30]. The sesquilinear alternative appears in Getoor and Sharpe [15].

provided $|\xi| * \mu^V < \infty$. Then ν^V is a predictable random measure such that $\xi * \mu^V - \xi * \nu^V$ is a local martingale. Observe furthermore that for an m–dimensional semimartingale V we have

$$[V,V]^c = [V,V] - \mathrm{id}\,\mathrm{id}^\top * \mu^V.$$

If V is special, we let the triplet $(B^V, [\hat{V},\hat{V}]^c, \nu^V)$ denote the corresponding semimartingale characteristics of V under a fixed probability measure P.[3] In particular, the drift B^V, i.e., the predictable finite-variation part of the Doob–Meyer decomposition of V, is always assumed to start in zero, i.e., $B^V_0 = 0$. For a general m–dimensional semimartingale V, we write $V[1] = V - \mathrm{id}\mathbf{1}_{|\mathrm{id}|>1} * \mu^V$. We can then define the 'clock' (or 'activity') process

$$A^V = \sum_{i=1}^{2m} \mathrm{TV}\left(B_i^{\hat{V}[1]}\right) + \mathrm{trace}\left[\hat{V},\hat{V}\right]^c + (|\mathrm{id}|^2 \wedge 1) * \nu^V,$$

where TV denotes total variation. Then A^V is non-decreasing and locally bounded. Thanks to [18, II.2.9], there exists an appropriate transition kernel F^V such that

$$\nu^V(\mathrm{d}t, \mathrm{d}v) = F^V(\mathrm{d}v)\mathrm{d}A^V_t.$$

2.3 Stochastic integration

In this subsection we discuss stochastic integrals of predictable processes with respect to complex-valued semimartingales. To begin, consider a $\overline{\mathbb{C}}^{1\times m}$–valued process ζ and a $\mathbb{C}^m$–valued semimartingale V. Here ζ is explicitly allowed to take the value NaN, but needs to be $\mathbb{C}^{1\times m}$–valued, $(\mathsf{P}\times A^V)$–a.e., for the integral to be defined. If V is real-valued, then we write $\zeta \in L(V)$ if both $\mathrm{Re}\,\zeta$ and $\mathrm{Im}\,\zeta$ are integrable with respect to V (in the standard sense). We then set $\zeta \cdot V = (\mathrm{Re}\,\zeta)\cdot V + i(\mathrm{Im}\,\zeta)\cdot V$.

If V is complex-valued, then we say $\zeta \in L(V)$ if $(\zeta \otimes [1\ i]) \in L(\hat{V})$, where $\otimes$ represents the Kronecker product; recall also (2.1). We then write

$$\zeta \cdot V = (\zeta \otimes [1\ i]) \cdot \hat{V} \tag{2.2}$$

for the stochastic integral of ζ with respect to V. For real-valued V the class $L(V)$ is defined twice but it is clear that the two definitions are consistent and $\zeta \cdot V$ is well defined. For $m = 1$ one has $\zeta \in L(V)$ if and only if $[\zeta\ \ i\zeta] \in L(\hat{V})$. It is clear how to extend this definition to a $\overline{\mathbb{C}}^{n\times m}$–valued process ζ, where $n \in \mathbb{N}$.

Remark 2.3 (Caveat of complex-valued integration)**.** Complex-valued stochastic integrals appear in the literature in a very limited context such as stochastic differential equations (e.g., [18, I.4.60]) or the Itô formula (e.g., [32, Proposition V.2.3]). In those circumstances the integrands are locally bounded, meaning that vector-valued integration is not required and integrability itself is not an issue. Our definition coincides with these special cases when ζ is locally bounded but in general the (real) stochastic integrals on the right-hand side of (2.2) cannot be computed component-wise. □

Finally, for a $\mathbb{C}^{n\times m\times m}$–valued process ζ and a $\mathbb{C}^{m\times m}$–valued semimartingale V (usually a quadratic variation process), let $\mathrm{vec}^{\mathrm{r}}(\zeta)$ and $\mathrm{vec}^{\mathrm{c}}(V)$ denote the row-wise and column-wise flattening of ζ and V, respectively. Then $\mathrm{vec}^{\mathrm{r}}(\zeta)$ is $(n \times m^2)$–dimensional and $\mathrm{vec}^{\mathrm{c}}(V)$ is m^2–dimensional. We then write $\zeta \in L(V)$ if $\mathrm{vec}^{\mathrm{r}}(\zeta) \in L(\mathrm{vec}^{\mathrm{c}}(V))$ and $\zeta \cdot V = \mathrm{vec}^{\mathrm{r}}(\zeta)\cdot \mathrm{vec}^{\mathrm{c}}(V)$.

[3] We use the real-valued lift of V to describe the continuous part of the quadratic variation in the characteristic triplet. This is necessary to capture the full dynamics of V. For example, let V and W denote two independent $\mathbb{R}$–valued Brownian motions and set $Z = \sqrt{2}V + iW$. Then $[V,V]^c = [Z,Z]^c$ but indeed $[\hat{V},\hat{V}]^c \neq [\hat{Z},\hat{Z}]^c$.

2.4 Predictable functions

For this subsection, let $m, n \in \mathbb{N}$. As in [18, II.1.4], we consider the notion of a predictable function on $\overline{\Omega}^m = \Omega \times [0,\infty) \times \overline{\mathbb{C}}^m$. For two predictable functions $\xi : \overline{\Omega}^m \to \overline{\mathbb{C}}^n$ and $\psi : \overline{\Omega}^n \to \overline{\mathbb{C}}$ we shall write $\psi(\xi)$ to denote the function $(\omega, t, x) \mapsto \psi(\omega, t, \xi(\omega, t, x))$ with the convention $\psi(\omega, t, \mathrm{NaN}) = \mathrm{NaN}$. If ψ and ξ are predictable, then so is $\psi(\xi)$.

For a predictable function $\xi : \overline{\Omega}^m \to \overline{\mathbb{C}}^n$ we shall write $\hat{\xi} = \hat{\mathrm{id}}(\xi)$ and $\xi^{(k)}$ for the k–th component of ξ, where $k \in \{1, \cdots, n\}$. We also write $\hat{D}\xi$ and $\hat{D}^2\xi$ for the real derivatives of ξ, i.e., $\hat{D}_i\xi^{(k)}$ is the composition of the i–th element of the gradient of $\xi^{(k)}\big(\hat{\mathrm{id}}^{-1}\big)$ and the lift $\hat{\mathrm{id}}$ and $\hat{D}^2_{i,j}\xi^{(k)}$ is the composition of the (i,j)–th element of the Hessian of $\xi^{(k)}\big(\hat{\mathrm{id}}^{-1}\big)$ and the lift $\hat{\mathrm{id}}$, for $i, j \in \{1, \cdots, 2m\}$. Note that $\hat{D}\xi$ has dimension $n \times (2m)$, $\hat{D}^2\xi$ has dimension $n \times (2m) \times (2m)$, and the domains of $\hat{D}\xi$, $\hat{D}^2\xi$ equal $\overline{\Omega}^m$, i.e., they coincide with the domain of ξ.

We want to allow for predictable functions such as $\xi = \log(1 + \mathrm{id})$ whose effective domain is not the entire $\mathbb{C}$. For this reason, we define, for a given predictable function $\xi : \overline{\Omega}^m \to \overline{\mathbb{C}}^n$, the set of semimartingales whose jumps are compatible with ξ, i.e.,

$$\mathsf{Dom}(\xi) = \{\text{semimartingale } V : \xi(\Delta V) \text{ is } \mathbb{C}^n\text{–valued}, \mathsf{P}\text{–almost surely}\}.$$

If for another predictable function $\psi : \overline{\Omega}^n \to \overline{\mathbb{C}}^m$ we have $\psi(\xi(\Delta V)) = \Delta V$ for all $V \in \mathsf{Dom}(\xi)$, we say ξ allows for a left inverse. If $\xi(\psi(\Delta V)) = \Delta V$ for all $V \in \mathsf{Dom}(\psi)$ we say that ξ allows for a right inverse. If ψ represents both left and right inverse we shall use the notation $\xi^{-1} = \psi$.

2.5 Sigma-localized integrals with respect to random measures

We next recall from Černý and Ruf [4] relevant results about the sigma-localized version of the $*$ integral of a predictable function with respect to ν^V and μ^V for a semimartingale V, which we fix from now on to the end of this section. The following is adapted from [4, Definition 3.1].

Definition 2.4 (Extended integral with respect to random measure)**.** *Denote by $L(\mu^V)$ the set of predictable functions that are absolutely integrable with respect to μ^V. We say that a predictable function ξ belongs to $L_\sigma(\mu^V)$, the sigma–localized class of $L(\mu^V)$, if there is a sequence $(C_k)_{k\in\mathbb{N}}$ of predictable sets increasing to $\Omega \times [0,\infty)$ and a semimartingale Y such that $\mathbf{1}_{C_k}\xi \in L(\mu^V)$ for each $k \in \mathbb{N}$ and*

$$(\mathbf{1}_{C_k}\xi) * \mu^V = \mathbf{1}_{C_k} \cdot Y, \qquad k \in \mathbb{N}.$$

In such case the semimartingale Y is denoted by $\xi \star \mu^V$.

Similarly, we define $L_\sigma(\nu^V)$ and $\xi \star \nu^V$. □

In the following, we recall useful characterizations for $L_\sigma(\nu^V)$ and $L_\sigma(\mu^V)$.

Proposition 2.5 (Kallsen [21], Definition 4.1, Lemma 4.1)**.** *For a predictable function ξ the following statements are equivalent.*

(i) $\xi \in L_\sigma(\nu^V)$.

(ii) *The following two conditions hold:*

  (a) $\int |\xi_t(v)| F^V_t(\mathrm{d}v) < \infty \quad (\mathsf{P} \times A^V)$*–a.e.*

  (b) $\int_0^\cdot \left|\int \xi_t(v) F^V_t(\mathrm{d}v)\right| \mathrm{d}A^V_t < \infty$.

Moreover, for $\xi \in L_\sigma(\nu^V)$ one has

$$\xi \star \nu^V = \int_0^\cdot \left(\int \xi_t(v) F^V_t(\mathrm{d}v)\right) \mathrm{d}A^V_t.$$

Proposition 2.6 ([4], Proposition 3.4)**.** *For a predictable function ξ the following statements are equivalent.*

(i) $\xi \in L_\sigma(\mu^V)$.

(ii) *The following two conditions hold.*

(a) $|\xi|^2 * \mu^V < \infty$.

(b) $\xi \mathbf{1}_{\{|\xi|\le 1\}} \in L_\sigma(\nu^V)$.

Furthermore, for $\xi \in L_\sigma(\mu^V)$ one has

$$\xi \star \mu^V = \xi \mathbf{1}_{\{|\xi|>1\}} * \mu^V + \xi \mathbf{1}_{\{|\xi|\le 1\}} * (\mu^V - \nu^V) + \xi \mathbf{1}_{\{|\xi|\le 1\}} \star \nu^V, \tag{2.3}$$

where the integral with respect to $\mu^V - \nu^V$ is defined in [18, II.1.27(b)].

Remark 2.7 ([4], Remarks 3.2 and 3.5)**.** Let Q denote a probability measure absolutely continuous with respect to P. With the obvious notation, we then have $L^{\mathsf{P}}_\sigma(\mu^V) \subset L^{\mathsf{Q}}_\sigma(\mu^V)$. For $L^{\mathsf{P}}_\sigma(\nu^V(\mathsf{P}))$ and $L^{\mathsf{Q}}_\sigma(\nu^V(\mathsf{Q}))$, no such inclusions hold in general. However, for ξ with $\xi^2 * \mu^V < \infty$ Proposition 2.6 yields that if $\xi \mathbf{1}_{\{|\xi|\le 1\}} \in L_\sigma(\nu^V(\mathsf{P}))$ then also $\xi \mathbf{1}_{\{|\xi|\le 1\}} \in L_\sigma(\nu^V(\mathsf{Q}))$. □

Next we recall a composition property for stochastic integrals. Such result does not hold if the $\star$ integral were to be replaced by the $*$ integral.

Proposition 2.8 ([4], Proposition 3.9)**.** *For $\xi \in L_\sigma(\mu^V)$ taking values in $\overline{\mathbb{C}}^n$ for some $n \in \mathbb{N}$ and a $\overline{\mathbb{C}}^{1,n}$–valued predictable process ζ the following statements are equivalent.*

(i) $\zeta \in L(\xi \star \mu^V)$.

(ii) $\zeta\xi \in L_\sigma(\mu^V)$.

Furthermore, if either condition holds then $\zeta \cdot (\xi \star \mu^V) = (\zeta\xi) \star \mu^V$.

The previous three propositions and the remaining ones of this section are proved in the corresponding references for the case when V is $\mathbb{R}$–valued. The arguments for the general case are straightforward; see also [4, Remark 2.1].

We next denote by $\mathscr{V}$ the set of semimartingales with finite variation on compact time intervals and by $\mathscr{V}^{\mathrm{d}}$ the subset of finite variation pure-jump processes, i.e., those semimartingales $V \in \mathscr{V}$ that satisfy $V = V_0 + \mathrm{id} * \mu^V$. The statements in this subsection can also be expressed in terms of a special class of semimartingales $\mathscr{V}^{\mathrm{d}}_\sigma$, i.e., the σ–localized class of finite variation pure-jump processes. The key connection is the following.

Proposition 2.9 ([4], Proposition 3.12)**.** *If $\xi \in L_\sigma(\mu^V)$ then $\xi \star \mu^V$ is an element of $\mathscr{V}^{\mathrm{d}}_\sigma$. Conversely, if $V \in \mathscr{V}^{\mathrm{d}}_\sigma$ then $\mathrm{id} \in L_\sigma(\mu^V)$ and $V = V_0 + \mathrm{id} \star \mu^V$.*

We conclude this section with a natural decomposition of V into jumps at predictable times and a quasi-left-continuous process.

Proposition 2.10 ([4], Proposition 3.15)**.** *Every semimartingale V has the unique decomposition*

$$V = V_0 + V^{\mathrm{qc}} + V^{\mathrm{dp}},$$

where $V^{\mathrm{qc}}_0 = V^{\mathrm{dp}}_0 = 0$, V^{qc} is a quasi-left-continuous semimartingale, V^{dp} jumps only at predictable times, and $V^{\mathrm{dp}} \in \mathscr{V}^{\mathrm{d}}_\sigma$. We then have $[V^{\mathrm{qc}}, V^{\mathrm{dp}}] = 0$.

If we define the predictable set $\mathcal{H}_V = \{\nu^V(\{\cdot\}) = 0\}$, then indeed $V^{\mathrm{qc}} = \mathbf{1}_{\mathcal{H}_V} \cdot V$ and $V^{\mathrm{dp}} = \mathbf{1}_{\mathcal{H}^c_V} \cdot V$. Hence V is special if and only if both V^{qc} and V^{dp} are special.

Let $\mathcal{T}_V$ denote a countable family of stopping times that exhausts the jumps of V^{dp}.[4] For each V there may be many ways to choose $\mathcal{T}_V$. The following statement holds for any such $\mathcal{T}_V$.

[4] Note that $\mathsf{P}[\Delta V_\tau = 0] > 0$ is possible for $\tau \in \mathcal{T}_V$.

Proposition 2.11 (Drift of a pure-jump process jumping only at predictable times)**.** *Assume that V^{dp} is special. Then we have*

$$B^{V^{\mathrm{dp}}} = \sum_{\tau \in \mathcal{T}_V} \mathsf{E}_{\tau-}[\Delta V_\tau] \mathbf{1}_{[\![\tau,\infty[\![}.$$

Proof. Thanks to (2.3), we have $B^{V^{\mathrm{dp}}} = \mathrm{id} \star \nu^{V^{\mathrm{dp}}}$. Moreover, $B^{V^{\mathrm{dp}}}$ is of sigma-finite variation and $\mathcal{T}_V$ exhausts its jumps. Proposition 4.6 in [4] applied to $B^{V^{\mathrm{dp}}}$ then yields the result. □

3 Semimartingale representation

3.1 Definition and basic properties

From now on we shall fix some $d, n \in \mathbb{N}$ and consider a $\mathbb{C}^d$–valued semimartingale X. We shall then study a variety of predictable transformations of X. Of course, an $\mathbb{R}^d$–valued semimartingale can always be considered a special case.

Example 3.1 (A motivational example)**.** Let X denote an $\mathbb{R}$–valued semimartingale and let $f : \mathbb{R} \to \mathbb{R}$ denote a twice continuously differentiable function. Then it is well known that also the process $Y = f(X)$ is a semimartingale. More precisely, the Itô–Meyer change of variables formula, [18, I.4.57], provides the representation

$$Y = f(X_0) + f'(X_-) \cdot X + \frac{1}{2} f''(X_-) \cdot [X, X]^c + (f(X_- + \mathrm{id}) - f(X_-) - f'(X_-)\,\mathrm{id}) * \mu^X. \quad (3.1)$$

Let us now introduce the predictable function $\xi^{f,X} : \Omega \times [0, \infty) \times \mathbb{R} \to \mathbb{R}$ by

$$\xi^{f,X}(\omega, t, x) = f(X_{t-}(\omega) + x) - f(X_{t-}(\omega)).$$

Note that the derivatives $D\xi^{f,X}$ and $D^2\xi^{f,X}$ exist. The representation in (3.1) then can be written in the more compact form

$$Y = Y_0 + D\xi^{f,X}(0) \cdot X + \frac{1}{2} D^2\xi^{f,X}(0) \cdot [X, X]^c + \big(\xi^{f,X} - D\xi^{f,X}(0)\,\mathrm{id}\big) * \mu^X. \quad (3.2)$$

Observe that $\Delta Y = \xi(\Delta X)$ and that Y is fully determined by X and the predictable function $\xi^{f,X}$. □

The connection between (3.1) and (3.2) motivates the key concept of this paper, Definition 3.8 below. Recall from Subsection 2.5 the predictable set $\mathcal{H}_X$, on which X^{dp} has no 'activity.'

Definition 3.2 (Representing functions for a given semimartingale X)**.** *Let $\mathfrak{I}^n(X)$ denote the set of all predictable functions $\xi : \overline{\Omega}^d \to \overline{\mathbb{C}}^n$ such that the following properties hold.*

(1) $X \in \mathsf{Dom}(\xi)$, *viewed as a property of ξ for fixed X.*

(2) $\xi(0) = 0$, $(\mathsf{P} \times A^X)$*–a.e.*

(3) $x \mapsto \mathbf{1}_{\mathcal{H}_X} \xi(x)$ *is twice real-differentiable at zero,* $(\mathsf{P} \times A^X)$*–a.e.*

(4) $\mathbf{1}_{\mathcal{H}_X} \hat{D}\xi(0) \in L(\hat{X})$.

(5) $\hat{D}^2\xi(0) \in L([\hat{X}, \hat{X}]^c)$.

(6) $(\xi - \mathbf{1}_{\mathcal{H}_X} \hat{D}\xi(0)\,\hat{\mathrm{id}}) \in L_\sigma(\mu^X)$.

We write $\mathfrak{I}(X) = \bigcup_{k \in \mathbb{N}} \mathfrak{I}^k(X)$. □

Remark 3.3 (The role of the predictable set $\mathcal{H}_X$)**.** If a predictable function ξ satisfies the conditions of Definition 3.2 with $\mathcal{H}_X$ replaced by a larger predictable set $\mathcal{H} \supset \mathcal{H}_X$ (e.g., $\mathcal{H} = \Omega \times [0, \infty)$, corresponding to no indicators at all), then the conclusion $\xi \in \mathfrak{I}(X)$ still holds. To see this, we only need to argue (6). This follows from observing that we have $\mathbf{1}_{\mathcal{H}\setminus\mathcal{H}_X} \hat{D}\xi(0) \in L(\hat{X})$, yielding $\mathbf{1}_{\mathcal{H}\setminus\mathcal{H}_X} \hat{D}\xi(0)\,\hat{\mathrm{id}} \in L_\sigma(\mu^{\hat{X}})$ by Proposition 2.8.

Example 4.6 below provides an instance where $X = X^{\mathrm{dp}}$, $\xi \in \mathfrak{I}(X)$, ξ is twice differentiable at zero, but $D\xi(0) \notin L(X)$. Thus, allowing for the existence of an appropriate predictable set $\mathcal{H}_X$ such that only $\mathbf{1}_{\mathcal{H}_X} D\xi(0) \in L(X)$ is required, indeed allows for a bigger class $\mathfrak{I}(X)$. □

As Propositions 3.6 and 3.15 and Theorem 3.17 below argue, the following class $\mathfrak{U}$ enjoys closedness with respect to common operations and universality in the sense that a representing function $\xi \in \mathfrak{U}$ satisfies $\xi \in \mathfrak{I}(X)$ for *any* semimartingale X provided that $\xi(\Delta X)$ is finite.

Definition 3.4 (Universal representing functions)**.** *Let $\mathfrak{U}^n$ denote the set of all predictable functions $\xi : \overline{\Omega}^d \to \overline{\mathbb{C}}^n$ such that the following properties hold,* P*–almost surely.*

(1) $\xi_t(0) = 0$*, for all* $t \geq 0$*.*

(2) $x \mapsto \xi_t(x)$ *is twice real-differentiable at zero, for all* $t \geq 0$*.*

(3) $\hat{D}\xi(0)$ *and* $\hat{D}^2\xi(0)$ *are locally bounded.*

(4) *There is a predictable locally bounded process* $K > 0$ *such that*

$$\sup_{|x| \leq 1/K} \frac{\left|\xi(x) - \hat{D}\xi(0)\hat{\mathrm{id}}(x)\right|}{|x|^2} \mathbf{1}_{x \neq 0} \text{ is locally bounded.}$$

We write $\mathfrak{U} = \bigcup_{n \in \mathbb{N}} \mathfrak{U}^n$*.* □

Remark 3.5 (A special case: real-valued semimartingales)**.** If X is real-valued then we may consider ξ as a predictable function with real domain. In this case, it can be easily checked that in Definitions 3.2 and 3.4 we may omit the hats on top of D, id, and X, with D and D^2 being the standard gradient and Hessian, respectively. □

Proposition 3.6 (Universality of $\mathfrak{U}$)**.** *Fix some* $\xi \in \mathfrak{U}$ *such that* $X \in \mathsf{Dom}(\xi)$*. We then have* $\xi \in \mathfrak{I}^n(X)$*,* $(\xi - \hat{D}\xi(0)\hat{\mathrm{id}}) \in L(\mu^X)$*, and*

$$\left(\xi - \hat{D}\xi(0)\hat{\mathrm{id}}\right) \star \mu^X = \left(\xi - \hat{D}\xi(0)\hat{\mathrm{id}}\right) * \mu^X.$$

Proof. The first claim follows from Remark 3.3. For the second claim it suffices to observe that $|\xi - \hat{D}\xi(0)\hat{\mathrm{id}}| * \mu^X < \infty$ by localization. □

Proposition 3.7 (Properties of $\mathfrak{I}(X)$)**.** *The following statements hold.*

(1) *If* $\xi, \psi \in \mathfrak{I}^n(X)$ *for some* $n \in \mathbb{N}$ *and* $\lambda \in \mathbb{C}$ *then* $\xi + \lambda\psi \in \mathfrak{I}^n(X)$*.*

(2) *If* $X \in \mathscr{V}_\sigma^{\mathrm{d}}$ *then* $\mathfrak{I}(X) \subset L_\sigma(\mu^X)$*. Moreover, if* $X = X^{\mathrm{dp}}$ *then* $\mathfrak{I}(X) = L_\sigma(\mu^X)$*.*

(3) *Let* $\mathcal{H}$ *denote a predictable set. Then* $\mathfrak{I}(X) = \mathfrak{I}(\mathbf{1}_{\mathcal{H}} \cdot X) \cap \mathfrak{I}(\mathbf{1}_{\mathcal{H}^c} \cdot X)$*; in particular,* $\mathfrak{I}(X) = \mathfrak{I}(X^{\mathrm{qc}}) \cap \mathfrak{I}(X^{\mathrm{dp}})$*.*

(4) *Let* Y *denote another semimartingale and let* $\psi \in \mathfrak{I}(X, Y)$*. If* ψ *is constant in the* y*–argument then* $\xi : x \mapsto \psi(x, 0)$ *is in* $\mathfrak{I}(X)$*.*

Proof. Parts (1) and (4) follow directly from Definition 3.2. Parts (2) and (3) rely on an application of Proposition 2.8. □

Definition 3.8 (Semimartingale representation)**.** *For a predictable function $\xi \in \mathfrak{I}(X)$ we use the notation*

$$\xi \circ X = \mathbf{1}_{\mathcal{H}_X} \hat{D}\xi(0) \cdot \hat{X} + \frac{1}{2}\hat{D}^2\xi(0) \cdot \left[\hat{X}, \hat{X}\right]^c + \left(\xi - \mathbf{1}_{\mathcal{H}_X} \hat{D}\xi(0)\, \hat{\mathrm{id}}\right) \star \mu^X. \tag{3.3}$$

If there exists $\xi \in \mathfrak{I}(X)$ such that

$$Y = Y_0 + \xi \circ X, \tag{3.4}$$

we say that the semimartingale Y is represented in terms of the semimartingale X. □

Remark 3.9 (Émery formula)**.** The right-hand side of equation (3.3) appears almost verbatim in Émery [13, eq. (13)] in the special case where X is real-valued and ξ is a real-valued twice continuously differentiable time-constant and deterministic function; see also Proposition 3.6. In this case the $\star$ integral can be replaced by the standard $*$ integral. □

Remark 3.10 (Interpretation of $\xi \circ X$ as ξ–variation)**.** The object $\xi \circ X$ with time-constant deterministic ξ, most often a power function, resurfaces several times in the literature under the name ξ–variation, see Doléans [8], Monroe [27, 28], Lépingle [24], Jacod [17], and Carr and Lee [2]. The terminology and Émery's [13] notation $\int_0^{\cdot} \xi(\mathrm{d}X_s)$ originate from the fact that, for suitably regular time-constant deterministic ξ, the partial sums $\sum_{n\in\mathbb{N}} \xi_{t_{n-1}} \left(X^{t_n} - X^{t_{n-1}}\right)$ converge uniformly on compact time intervals in probability to $\xi \circ X$ as the time partition $(t_n)_{n\in\mathbb{N}}$ becomes finer; see [13, Théorème 2a] and [17, Theorem 2.2] for a related statement. For a precise statement of such convergence for predictable ξ, see [6]. □

Remark 3.11 (Generalizations of Émery formula)**.** The conditions in Definition 3.2 ensure that all terms in (3.4) are defined. One could extend the class $\mathfrak{I}(X)$ further. The idea of such generalisation would be to focus on the activity of the individual components of X. For example, one could abstain from requesting that $x \mapsto \mathbf{1}_{\mathcal{H}_X}\xi(x)$ is real-differentiable in the i–th component for times when $\mathrm{d}A^{X^{(i)}} = 0$. Moreover, one could assume that the second real derivative of $x \mapsto \mathbf{1}_{\mathcal{H}_X}\xi(x)$ only needs to exist ($\mathsf{P} \times \mathrm{trace}[\hat{X}, \hat{X}]^c$)–a.e. However, such generalisations would come with more complicated notation and would obscure the main results, hence we do not pursue them here. □

Remark 3.12 (Measure invariance of representations)**.** Note that $\mathfrak{I}(X)$ is invariant under equivalent changes of measures. More precisely, if Q is a probability measure absolutely continuous with respect to P and if $\xi \in \mathfrak{I}(X)$ under P, then also $\xi \in \mathfrak{I}(X)$ under Q (recall Remark 2.7 to see this). Moreover, if we define $Y = \xi \circ X$ under P, then we also have $Y = \xi \circ X$ under Q. Hence, $\xi \circ X$ is measure-invariant in the sense that (3.3) only depends on the null sets. A similar statement holds for $\mathfrak{U}$. This is in contrast to the common (and frequently also very useful) representation of Y in terms of predictable characteristics. □

We now list some immediate consequences of the definition of representability.

Proposition 3.13 (Properties of representation)**.** *The following statements hold.*

(1) *Let $\xi \in \mathfrak{I}(X)$. Then*

$$\Delta(\xi \circ X) = \xi(\Delta X).$$

(2) *$\mathfrak{I}(X) = \mathfrak{I}(X - X_0)$ and for any $\xi \in \mathfrak{I}(X)$ one has*

$$\xi \circ X = \xi \circ (X - X_0).$$

(3) *We have* $\mathrm{id}_i, \mathrm{id}_i\mathrm{id}_j \in \mathfrak{U}^1$*, for all* $i, j \in \{1, \cdots, d\}$*, with*

$$X^{(i)} = X_0^{(i)} + \mathrm{id}_i \circ X;$$
$$\left[X^{(i)}, X^{(j)}\right] = (\mathrm{id}_i\mathrm{id}_j) \circ X.$$

(4) *If* $\xi \in \mathfrak{I}(X)$ *then* $\xi^* \in \mathfrak{I}(X)$ *and* $(\xi \circ X)^* = \xi^* \circ X$*, where the superscript* $*$ *denotes the complex conjugate.*

(5) *If* $\xi \in \mathfrak{I}(X)$ *then* $\xi \circ X^{\mathrm{qc}} = (\xi \circ X)^{\mathrm{qc}}$ *and* $\xi \circ X^{\mathrm{dp}} = (\xi \circ X)^{\mathrm{dp}}$*. (Recall also Proposition 3.7*(3)*).*

(6) *Let* Y *be a predictable semimartingale of finite variation and* $\xi \in \mathfrak{I}(X, Y)$ *such that* $\xi(\cdot, \Delta Y) \in \mathfrak{I}(X)$ *and* $\xi(0, \cdot) \in \mathfrak{I}(Y)$*. Then we have*

$$\xi \circ (X, Y) = \xi(\cdot, \Delta Y) \circ X + \xi(0, \cdot) \circ Y. \tag{3.5}$$

Proof. Parts (1), (2), (4), and (5) follow directly from Definitions 3.2 and 3.8. Part (3) follows directly from Proposition 3.6 and Definition 3.8.

For (6), note that $\xi(\cdot, \Delta Y) \in \mathfrak{I}(X)$ yields that $\mathbf{1}_{\mathcal{H}_{X,Y}} \hat{D}_x \xi(0, 0) \in L(\hat{X})$ and $\hat{D}^2_{xx}\xi(0, 0) \in L([X, X]^c)$ with

$$\mathbf{1}_{\mathcal{H}_X} \hat{D}_x \xi(0, \Delta Y) \cdot \hat{X} = \mathbf{1}_{\mathcal{H}_{X,Y}} \hat{D}_x \xi(0, 0) \cdot \hat{X};$$
$$\hat{D}^2_{xx}\xi(0, \Delta Y) \cdot \left[\hat{X}, \hat{X}\right]^c = \hat{D}^2_{xx}\xi(0, 0) \cdot \left[\hat{X}, \hat{X}\right]^c.$$

Similarly $\xi(0, \cdot) \in \mathfrak{I}(Y)$ yields that $\mathbf{1}_{\mathcal{H}_{X,Y}} \hat{D}_y \xi(0, 0) \in L(\hat{Y})$ with

$$\mathbf{1}_{\mathcal{H}_Y} \hat{D}_y \xi(0, 0) \cdot \hat{Y} = \mathbf{1}_{\mathcal{H}_{X,Y}} \hat{D}_y \xi(0, 0) \cdot \hat{Y}.$$

Now, the result follows by comparing the jumps on the left and right hand side of (3.5), for example by using (1). □

Proposition 3.14 (Representation of stochastic integrals)**.** *Let* ζ *be a* $\mathbb{C}^{1\times d}$*–valued predictable process in* $L(X)$*. Then* $\zeta\mathrm{id} \in \mathfrak{I}^1(X)$ *and*

$$\zeta \cdot X = \zeta\mathrm{id} \circ X.$$

Proof. Let I_d be a $d \times d$ identity matrix. Observe that $\xi = \zeta\mathrm{id}$ verifies

$$\hat{D}\xi = \zeta \otimes [1\ i]; \qquad \hat{D}^2\xi = 0; \qquad \xi - \hat{D}\xi(0)\,\hat{\mathrm{id}} = \zeta\left(\mathrm{id} - (I_d \otimes [1\ i])\,\hat{\mathrm{id}}\right) = 0.$$

Hence, ξ belongs to $\mathfrak{I}(X)$ as per Definition 3.2, and (3.3) together with (2.2) yield the claim. □

Proposition 3.15 (Representation of a change of variables)**.** *Let* $\mathcal{U} \subset \mathbb{C}^d$ *be an open set such that* $X_-, X \in \mathcal{U}$ *and let* $f : \mathcal{U} \to \mathbb{C}^n$ *be twice continuously real-differentiable. Then the predictable function* $\xi^{f,X} : \overline{\Omega}^d \to \overline{\mathbb{C}}^n$ *defined by*

$$\xi^{f,X}(x) = \begin{cases} f(X_- + x) - f(X_-), & X_- + x \in \mathcal{U} \\ \mathrm{NaN}, & X_- + x \notin \mathcal{U} \end{cases}, \qquad x \in \mathbb{C}^d,$$

belongs to $\mathfrak{U}^n$*,* $X \in \mathsf{Dom}(\xi^{f,X})$*, and*

$$f(X) = f(X_0) + \xi^{f,X} \circ X.$$ □

Proof. Denote by $R > 0$ the distance from X_- to the boundary of $\mathcal{U}$, by R^* its running infimum, and by $\tau > 0$ the first time R^* hits zero. The left-continuity of R now yields $\tau = \infty$ and $R^* > 0$. Therefore, $(\tau_n)_{n\in\mathbb{N}}$ given by $\tau_n = \inf\{t : R^*_t \leq 1/n\}$, is a localizing sequence of stopping times that makes both $K = {}^2/_R$ and $\sup_{|x|\leq 1/K}\big|\hat{D}^2\xi(x)\big|$ locally bounded, yielding $\xi^{f,X} \in \mathfrak{U}$.

Since $X \in \mathsf{Dom}(\xi^{f,X})$, Proposition 3.6 now yields that

$$f\big(\hat{\mathrm{id}}^{-1}\big(\hat{X}\big)\big) = f\big(\hat{\mathrm{id}}^{-1}\big(\hat{X}_0\big)\big) + \xi^{f,X} \circ X$$

is the Itô–Meyer change of variables formula for the real-valued function $\hat{\mathrm{id}}\big(f\big(\hat{\mathrm{id}}^{-1}\big)\big)$ applied to the real-valued process $\hat{X}$; see [18, I.4.57]. In view of $f\big(\hat{\mathrm{id}}^{-1}\big(\hat{X}\big)\big) = f(X)$ the proof is complete. □

Remark 3.16 (Itô's formula requires smoothness)**.** It is possible to exhibit a function $f : \mathbb{R} \to \mathbb{R}$ and a semimartingale X such that $\xi^{f,X} \in \mathfrak{I}(X)$, in the notation of Proposition 3.15, and such that $f(X)$ is a semimartingale, but

$$f(X) \neq f(X_0) + \xi^{f,X} \circ X.$$

For example, choose X equal to Brownian motion started at 0 and $f = |\mathrm{id}|$. Here f is not twice differentiable but $\xi^{f,X} \in \mathfrak{I}(X)$ anyway as it is Lebesgue-a.e. twice differentiable. Then $\xi^{f,X} \circ X = \mathrm{sgn}(X) \cdot X$ is another Brownian motion while $f(X) - f(X_0)$ is the absolute value of X. □

3.2 Composition of representations

We now describe the composition of representations. It is this result along with its consequences that makes the calculus simple.

Theorem 3.17 (Composition of universal representing functions)**.** *The class $\mathfrak{U}$ is closed under (dimensionally correct) composition, i.e., if $\xi \in \mathfrak{U}^n$ and $\psi : \overline{\Omega}^n \to \overline{\mathbb{C}}$ is another predictable function with $\psi \in \mathfrak{U}$, then $\psi(\xi) \in \mathfrak{U}$. Furthermore, if $\psi(\xi(\Delta X))$ is finite-valued, then one has $\psi, \psi(\xi) \in \mathfrak{I}(X)$, $\psi \in \mathfrak{I}(\xi \circ X)$, and*

$$\psi \circ (\xi \circ X) = \psi(\xi) \circ X.$$

Proof. Properties (1)–(3) of Definition 3.4 follow easily by direct calculation; see also (3.8) and (3.9) below. To show property (4), by localization we may assume that $\hat{D}\xi(0)$ is bounded and that there exists a constant $K_\xi > 0$ such that

$$\sup_{0<|x|\leq 1/K_\xi} \frac{|\xi(x) - \hat{D}\xi(0)\hat{\mathrm{id}}(x)|}{|x|^2} < \infty.$$

An analogous statement applies to ψ, with some constant $K_\psi > 0$. By possibly making K_ξ larger we may also assume that $\sup_{|x|\leq 1/K_\xi}|\xi(x)| \leq {}^1/_{K_\psi}$.

For $\eta = \psi(\xi)$ we then have $\hat{D}\eta(0) = \hat{D}\psi(0)\hat{D}\hat{\xi}(0)$ and by the triangular inequality

$$\begin{aligned}\frac{|\eta(x) - \hat{D}\eta(0)\hat{\mathrm{id}}(x)|}{|x|^2}\mathbf{1}_{x\neq 0} \leq{}& \frac{|\psi(\xi(x)) - \hat{D}\psi(0)\hat{\mathrm{id}}(\xi(x))|}{|\xi(x)|^2}\mathbf{1}_{\xi(x)\neq 0}\frac{|\xi(x)|^2}{|x|^2}\mathbf{1}_{x\neq 0}\\ &+ \frac{|\hat{D}\psi(0)(\hat{\xi}(x) - \hat{D}\hat{\xi}(0)\hat{\mathrm{id}}(x))|}{|x|^2}\mathbf{1}_{x\neq 0}.\end{aligned}$$

In view of the boundedness of $\hat{D}\psi(0)$ and $\hat{D}\xi(0)$, the supremum on the right-hand side over $|x| < {}^1/_{K_\xi}$ is finite. The statement $\psi(\xi) \in \mathfrak{U}$ follows. By Proposition 3.6 we have $\xi \in \mathfrak{I}(X)$ and $\psi \in \mathfrak{I}(\xi \circ X)$. As $\hat{D}\psi(0)$ is locally bounded, the rest of the statement follows from Theorem 3.18 below. □

When ξ and ψ are deterministic and time-constant functions, Theorem 3.17 reduces to the tower property in Carr and Lee [2, Proposition 2.4].

Theorem 3.18 (Composition of semimartingale representations)**.** *Let $\xi \in \mathfrak{I}(X)$. Moreover, fix $\psi \in \mathfrak{I}(\xi \circ X)$ such that $\psi(0) = 0$ and $\mathbf{1}_{\mathcal{H}_X}\psi$ is twice real-differentiable at zero, $(\mathsf{P} \times A^X)$–a.e., and*

$$\mathbf{1}_{\mathcal{H}_X}\hat{D}\psi(0) \in L\left(\hat{D}^2\hat{\xi}(0)\cdot\left[\hat{X},\hat{X}\right]^c\right) \cap L\left(\left(\hat{\xi} - \mathbf{1}_{\mathcal{H}_X}\hat{D}\hat{\xi}(0)\,\hat{\mathrm{id}}\right)\star\mu^X\right). \tag{3.6}$$

Then $\psi(\xi) \in \mathfrak{I}(X)$ and we have

$$\psi \circ (\xi \circ X) = \psi(\xi) \circ X. \tag{3.7}$$

Proof. Let $Y = \xi \circ X$. Without loss of generality, we may assume $\psi \in \mathfrak{I}^1(Y)$. We need to check the six properties of Definition 3.2 for $\eta = \psi(\xi)$. We clearly have (1), (2), and (3). In analogy to the chain rule for real derivatives, on $\mathcal{H}_X$ we also have

$$\hat{D}\eta(0) = \sum_{k=1}^{2n}\hat{D}_k\psi(0)\hat{D}\hat{\xi}^{(k)}(0) = \hat{D}\psi(0)\hat{D}\hat{\xi}(0); \tag{3.8}$$

$$\hat{D}^2\eta(0) = \sum_{k,l=1}^{2n}\hat{D}^2_{k,l}\psi\,(0)\,\hat{D}\hat{\xi}^{(k)}(0)^\top\hat{D}\hat{\xi}^{(l)}(0) + \sum_{k=1}^{2n}\hat{D}_k\psi\,(0)\,\hat{D}^2\hat{\xi}^{(k)}(0). \tag{3.9}$$

By assumption, we have $\mathbf{1}_{\mathcal{H}_Y}\hat{D}\psi(0) \in L(\hat{Y})$. Since $\mathcal{H}_X \subset \mathcal{H}_Y$, this also yields $\mathbf{1}_{\mathcal{H}_X}\hat{D}\psi(0) \in L(\hat{Y})$. Together with (3.6), we obtain

$$\mathbf{1}_{\mathcal{H}_X}\hat{D}\psi(0) \in L\left(\mathbf{1}_{\mathcal{H}_X}\hat{D}\hat{\xi}(0)\cdot\hat{X}\right),$$

hence also Definition 3.2(4) with ξ replaced by $\eta = \psi(\xi)$. Similarly, we also get Definition 3.2(5).

Next, observe in view of identity (3.8) that

$$\begin{aligned}\eta - \mathbf{1}_{\mathcal{H}_X}\hat{D}\eta(0)\,\hat{\mathrm{id}} &= \left(\psi(\xi) - \mathbf{1}_{\mathcal{H}_Y}\hat{D}\psi(0)\hat{\xi}\right) + \mathbf{1}_{\mathcal{H}_X}\hat{D}\psi(0)\left(\hat{\xi} - \hat{D}\hat{\xi}(0)\,\hat{\mathrm{id}}\right) + \mathbf{1}_{\mathcal{H}_Y\setminus\mathcal{H}_X}\hat{D}\psi(0)\hat{\xi}\\ &\in L_\sigma(\mu^X)\end{aligned}$$

by Proposition 2.8, the assumptions, and $\mathbf{1}_{\mathcal{H}_Y\setminus\mathcal{H}_X}|\xi| * \mu^X = 0$. This yields Definition 3.2(6) with ξ replaced by η.

Finally, (3.7) follows from (3.6), (3.8), and (3.9) together with the computations

$$\begin{aligned}\psi \circ (\xi \circ X) &= \mathbf{1}_{\mathcal{H}_Y}\hat{D}\psi(0)\cdot\hat{Y} + \frac{1}{2}\hat{D}^2\psi(0)\cdot\left[\hat{Y},\hat{Y}\right]^c + \left(\psi - \mathbf{1}_{\mathcal{H}_Y}\hat{D}\psi(0)\,\hat{\mathrm{id}}\right)\star\mu^Y\\ &= \mathbf{1}_{\mathcal{H}_Y}\hat{D}\psi(0)\cdot\left(\mathbf{1}_{\mathcal{H}_X}\hat{D}\hat{\xi}(0)\cdot\hat{X} + \frac{1}{2}\hat{D}^2\hat{\xi}(0)\cdot\left[\hat{X},\hat{X}\right]^c + \left(\hat{\xi} - \mathbf{1}_{\mathcal{H}_X}\hat{D}\hat{\xi}(0)\hat{\mathrm{id}}\right)\star\mu^X\right)\\ &\quad + \frac{1}{2}\hat{D}^2\psi(0)\cdot\left(\sum_{i,j=1}^{d}\hat{D}_i\hat{\xi}(0)\hat{D}_j\hat{\xi}(0)\cdot\left[\hat{X}^{(i)},\hat{X}^{(j)}\right]^c\right) + \left(\psi(\xi) - \mathbf{1}_{\mathcal{H}_Y}\hat{D}\psi(0)\hat{\xi}\right)\star\mu^X\\ &= \mathbf{1}_{\mathcal{H}_X}\hat{D}\eta(0)\cdot\hat{X} + \frac{1}{2}\hat{D}^2\eta(0)\cdot\left[\hat{X},\hat{X}\right]^c + \left(\eta - \mathbf{1}_{\mathcal{H}_X}\hat{D}\eta(0)\hat{\mathrm{id}}\right)\star\mu^X\\ &= \psi(\xi)\circ X.\end{aligned}$$

Here, we have used the associativity of the stochastic integrals with respect to $\hat{X}$ and $[\hat{X},\hat{X}]^c$ as well as the associativity of the $\star$ jump-measure integral. □

Example 4.11 below shows that without the assumption that ψ is twice real-differentiable at zero, $(\mathsf{P} \times A^X)$–a.e., the conclusion of Theorem 3.18 does not necessarily hold.

Remark 3.19 (The linear case)**.** If $\xi \in \mathfrak{I}(X)$ is linear, i.e., of the form $\xi = \zeta \mathrm{id}$ for some predictable process ζ, then (3.6) is automatically satisfied. □

Corollary 3.20 (Sufficient condition for composition of representations)**.** *Let $\xi \in \mathfrak{I}(X)$. Moreover, let $\psi \in \mathfrak{I}(\xi \circ X)$ such that $\psi(0) = 0$ and $\mathbf{1}_{\mathcal{H}_{\xi\circ X}}\psi$ is twice real-differentiable at zero, $(\mathsf{P} \times A^X)$–a.e., and such that $\mathbf{1}_{\mathcal{H}_{\xi\circ X}}\hat{D}\psi(0)$ is locally bounded (e.g., if $\psi \in \mathfrak{U}$ and $\xi \circ X \in \mathsf{Dom}(\psi)$). Then $\psi(\xi) \in \mathfrak{I}(X)$ and* (3.7) *holds.*

Remark 3.21 (Algebra of X–representable processes)**.** Thanks to Proposition 3.7(1), the space of $\mathbb{C}$–valued X–representable processes is a vector space. Proposition 3.15 and Corollary 3.20 yield that this space is also an algebra, namely closed under multiplication. Indeed, for $U = U_0 + \xi^U \circ X$ and $V = V_0 + \xi^V \circ X$ we have

$$UV = U_0V_0 + \left(\left(U_- + \xi^U\right)\left(V_- + \xi^V\right) - U_-V_-\right) \circ X. \qquad \square$$

Remark 3.22 (Converse of the composition theorem)**.** A reverse direction of Theorem 3.18 holds, too. To wit, let $\xi \in \mathfrak{I}^n(X)$. Moreover, fix some predictable function $\psi : \overline{\Omega}^n \to \overline{\mathbb{C}}$ such that $\psi(\xi) \in \mathfrak{I}^1(X)$, $\mathbf{1}_{\mathcal{H}_X}\psi$ is twice real-differentiable at zero, $(\mathsf{P} \times A^X)$–a.e., and (3.6) holds. Then $\psi \in \mathfrak{I}(\xi \circ X)$.

To see this, first note that $\psi(0) = \psi(\xi(0)) = 0$, $(\mathsf{P} \times A^X)$–a.e. We next follow the arguments of Theorem 3.18, using (3.8) and (3.9) with $\eta = \psi(\xi) \in \mathfrak{I}(X)$. We then directly obtain that Definition 3.2(1), (2), (3), (4), and (5) hold with ξ replaced by ψ. Here we used again $\xi(\Delta X) = 0$ on $\mathcal{H}_X \setminus \mathcal{H}_{\xi\circ X}$. Next, observe that

$$\psi(\xi) - \mathbf{1}_{\mathcal{H}_{\xi\circ X}}\hat{D}\psi(0)\hat{\xi} = \left(\eta - \mathbf{1}_{\mathcal{H}_X}\hat{D}\eta(0)\,\hat{\mathrm{id}}\right) + \mathbf{1}_{\mathcal{H}_X}D\hat{\psi}(0)\left(\hat{D}\hat{\xi}(0)\,\hat{\mathrm{id}} - \hat{\xi}\right) \in L_\sigma(\mu^X),$$

yielding the claim.

Examples 4.9 and 4.10 illustrate again how essential (3.6) is for the remark to hold. □

Corollary 3.23 (Inverse functions)**.** *Let $\xi \in \mathfrak{I}^d(X)$ and $Y = Y_0 + \xi \circ X$. Moreover, assume that the smallest singular value of $\hat{D}\hat{\xi}(0)$ is locally bounded away from zero and that ξ allows for a predictable left inverse ψ (see Subsection 2.4). Then $\psi \in \mathfrak{I}^d(Y)$ and $X = X_0 + \psi \circ Y$.*

Proof. Since $\hat{\xi}\big(\hat{\mathrm{id}}^{-1}\big)$ is continuously differentiable at zero on $\mathcal{H}_X$, $\hat{\psi}\big(\hat{\mathrm{id}}^{-1}\big)$ is actually an inverse of $\hat{\xi}\big(\hat{\mathrm{id}}^{-1}\big)$ in a neighbourhood of zero on $\mathcal{H}_X$. Thus $\mathbf{1}_{\mathcal{H}_X}\psi$ is twice real-differentiable at zero with $\hat{D}\hat{\psi}(0) = (\hat{D}\hat{\xi}(0))^{-1}$ on $\mathcal{H}_X$. If now the smallest singular value of $\hat{D}\hat{\xi}(0)$ is locally bounded away from zero, then the largest singular value of $\hat{D}\hat{\psi}(0)$ is locally bounded and by equivalence of the Schatten and maximum matrix norms each element of $\hat{D}\psi(0)$ is locally bounded. The assertion follows from Remark 3.22. □

If the assumption that the smallest singular value of $\hat{D}\hat{\xi}(0)$ is locally bounded away from zero is replaced by the weaker assumption that is is merely positive, then Corollary 3.23 is wrong as Examples 4.9 and 4.10 below illustrate, even if ψ is an inverse of ξ and $d = 1$.

Remark 3.24 (Advantages of the proposed calculus)**.** Results like Theorem 3.18 make this stochastic calculus simple and powerful. Consider the situation when one has to perform a change of variables $f(Y)$ on an X–representable $\mathbb{R}$–valued process Y. A direct

application of the Itô–Meyer formula (3.1) to the representation of Y in (3.2) yields

$$f(Y) = f(Y_0) + Df(Y_-) \cdot \left(D\xi(0) \cdot X + \frac{1}{2} D^2\xi(0) \cdot [X,X]^c + (\xi - D\xi(0)\,\mathrm{id}) \star \mu^X \right) \\ + \frac{1}{2} D^2 f(Y_-) \cdot [Y,Y]^c + (f(Y_- + \mathrm{id}) - f(Y_-) - Df(Y_-)\,\mathrm{id}) * \mu^Y.$$

One then has to collect all terms manually in order to simplify this expression and eventually recast it in terms of μ^X.

In contrast, the notation of (3.4) gives $f(Y) = f(Y_0) + \xi^{f,Y}(\xi) \circ X$. Only the function $\xi^{f,Y}(\xi)$ needs to be computed and then the corresponding representation applies. This is pedagogically pleasing because $\xi^{f,Y}(\xi)$ describes the jumps of $f(Y)$ in terms of the jumps of X, i.e.,

$$\Delta f(Y) = f(Y_- + \Delta Y) - f(Y_-) = f(Y_- + \xi(\Delta X)) - f(Y_-) = \xi^{f,Y}(\xi(\Delta X)). \qquad \square$$

3.3 Alternative Émery formula

In the non-analytic case, which too is of practical importance, it can be helpful to rephrase the Émery formula (3.3) in terms of the $\mathbb{C}^{2d}$–valued process (X, X^*). Here X^* denotes the complex conjugate of X. This allows the use of Wirtinger partial derivatives (see [35]), given by

$$\frac{\partial}{\partial x} = \frac{1}{2}\left(\frac{\partial}{\partial \operatorname{Re} x} - i \frac{\partial}{\partial \operatorname{Im} x} \right) \quad \text{and} \quad \frac{\partial}{\partial x^*} = \frac{1}{2}\left(\frac{\partial}{\partial \operatorname{Re} x} + i \frac{\partial}{\partial \operatorname{Im} x} \right). \tag{3.10}$$

This turns out to be convenient in some applications; see Proposition 3.26 and Example 3.28. Observe, however, that the proposed calculus allows one to write simply $\xi \circ X$ and operate on the level of ξ, where the specific physical implementation of $\xi \circ X$ is immaterial.

To arrive at the alternative Émery formula, we introduce the function $\check{\mathrm{id}} : \overline{\mathbb{C}}^d \to \overline{\mathbb{C}}^{2d}$ by

$$\check{\mathrm{id}} = \left(I_d \otimes \begin{bmatrix} 1 & i \\ 1 & -i \end{bmatrix} \right) \hat{\mathrm{id}}; \qquad \check{\mathrm{id}}(\mathrm{NaN}) = \mathrm{NaN},$$

where $\otimes$ again denotes the Kronecker product. This allows us to introduce the process

$$\check{X} = \check{\mathrm{id}}(X). \tag{3.11}$$

Observe that $\hat{X}$ is the $\mathbb{R}^{2d}$–valued process containing the values of $\operatorname{Re} X$ and $\operatorname{Im} X$, interlaced, while $\check{X}$ is the $\mathbb{C}^{2d}$–valued process containing X and its conjugate X^*, interlaced.

Next, we denote by $\check{D}\xi$ the row vector of Wirtinger derivatives, given by

$$\check{D}\xi = \frac{1}{2} \hat{D}\xi \left(I_d \otimes \begin{bmatrix} 1 & 1 \\ -i & i \end{bmatrix} \right), \tag{3.12}$$

and by $\check{D}^2\xi$ the corresponding 'Wirtinger Hessian,' given by

$$\check{D}^2\xi^{(k)} = \check{D}(\check{D}\xi^{(k)})^\top = \left(I_d \otimes \begin{bmatrix} 1 & -i \\ 1 & i \end{bmatrix} \right) \frac{1}{4} \hat{D}^2\xi^{(k)} \left(I_d \otimes \begin{bmatrix} 1 & 1 \\ -i & i \end{bmatrix} \right), \qquad k = 1, \dots, n.$$

The following technical observation will be very useful in the subsequent proposition.

Lemma 3.25 (Invertible linear transformations in a stochastic integral)**.** *Fix $m \in \mathbb{N}$. Let Λ_1, Λ_2 be arbitrary invertible matrices in $\mathbb{C}^{m\times m}$. Let ζ denote a $\mathbb{C}^{m\times m}$–valued predictable process and let V denote a $\mathbb{C}^{m\times m}$–valued semimartingale. Then the following are equivalent.*

(i) $\zeta \in L(V)$.

(ii) $\Lambda_1 \zeta \Lambda_2 \in L\left(\Lambda_2^{-1} V \Lambda_1^{-1}\right)$.

If one (hence both) of these conditions holds then

$$\zeta \cdot V = \Lambda_1 \zeta \Lambda_2 \cdot \left(\Lambda_2^{-1} V \Lambda_1^{-1}\right). \tag{3.13}$$

Proof. Note that it suffices to argue the implication from (i) to (ii) and to show (3.13). Moreover, since $\zeta \in L(V) \iff \zeta^\top \in L(V^\top)$ and $\zeta \cdot V = \zeta^\top \cdot V^\top$, it is enough to prove the statement with Λ_1 being the identity matrix. To this end, assume (i) holds. Let $\mathrm{vec}^{\mathrm{r}}(\zeta)$ (respectively, $\mathrm{vec}^{\mathrm{c}}(V)$) denote the row-wise (column-wise) flattening of ζ (respectively, V), that produces a $(1 \times m^2)$–dimensional row (m^2–dimensional column) vector. Then (i) is equivalent to $\mathrm{vec}^{\mathrm{r}}(\zeta) \in L(\mathrm{vec}^{\mathrm{c}}(V))$ and one has $\zeta \cdot V = \mathrm{vec}^{\mathrm{r}}(\zeta) \cdot \mathrm{vec}^{\mathrm{c}}(V)$. Thanks to Proposition 3.14 and Remark 3.22, this then yields $(\mathrm{vec}^{\mathrm{r}}(\zeta) R) \in L(R^{-1}\,\mathrm{vec}^{\mathrm{c}}(V))$ for any invertible $m^2 \times m^2$ matrix R, along with

$$\zeta \cdot V = (\mathrm{vec}^{\mathrm{r}}(\zeta) R) \cdot (R^{-1}\,\mathrm{vec}^{\mathrm{c}}(V)).$$

Choosing $R = I_m \otimes \Lambda_2$ yields $R^{-1} = I_m \otimes \Lambda_2^{-1}$, $\mathrm{vec}^{\mathrm{r}}(\zeta) R = \mathrm{vec}^{\mathrm{r}}(\zeta \Lambda_2)$, and hence the desired statement. □

Proposition 3.26 (Émery formula in terms of Wirtinger derivatives)**.** *For $\xi \in \mathfrak{I}(X)$, the following terms are well defined and we have*

$$\xi \circ X = \mathbf{1}_{\mathcal{H}_X} \check{D}\xi(0) \cdot \check{X} + \frac{1}{2} \check{D}^2 \xi(0) \cdot \left[\check{X}, \check{X}\right]^c + \left(\xi - \mathbf{1}_{\mathcal{H}_X} \check{D}\xi(0)\,\check{\mathrm{id}}\right) \star \mu^X. \tag{3.14}$$

Furthermore, if $\mathbf{1}_{\mathcal{H}_X} \xi$ is analytic at 0, $(\mathsf{P} \times A^X)$–a.e., the following terms are well defined and we have

$$\xi \circ X = \mathbf{1}_{\mathcal{H}_X} D\xi(0) \cdot X + \frac{1}{2} D^2 \xi(0) \cdot [X, X]^c + (\xi - \mathbf{1}_{\mathcal{H}_X} D\xi(0)\,\mathrm{id}) \star \mu^X. \tag{3.15}$$

Here $D\xi(0)$ and $D^2\xi(0)$ stand for complex derivatives. □

Proof. Let us first prove (3.14). To this end, we introduce the matrix

$$\Sigma = \frac{1}{2} I_d \otimes \begin{bmatrix} 1 & 1 \\ -i & i \end{bmatrix},$$

satisfying $\Sigma^{-1} = 2(\Sigma^*)^\top$. Then $\check{D}\xi = \hat{D}\xi\Sigma$ and $\check{X} = \Sigma^{-1}\hat{X}$. Thanks to Proposition 3.14 and Remark 3.22, we then have $\mathbf{1}_{\mathcal{H}_X} \check{D}\xi(0) \in L(\check{X})$ and $\mathbf{1}_{\mathcal{H}_X} \check{D}\xi(0) \cdot \check{X} = \mathbf{1}_{\mathcal{H}_X} \hat{D}\xi(0) \cdot \hat{X}$. By the same token, we also have $\check{D}\xi(0)\,\check{\mathrm{id}} = \hat{D}\xi(0)\,\hat{\mathrm{id}}$. Next, note that

$$\check{D}^2\xi(0) = \Sigma^\top \hat{D}^2\xi(0)\Sigma; \qquad \left[\check{X}, \check{X}\right] = \Sigma^{-1}\left[\hat{X}, \hat{X}\right](\Sigma^{-1})^\top.$$

An application of Lemma 3.25 now concludes the proof of (3.14).

The simplifications in the analytic case follow from the standard properties of Wirtinger derivatives, see for example Remmert [31, Section I.4]. □

We now provide two examples of complex-valued representations where the representing functions are not assumed analytic at 0. Recall the notation for Wirtinger derivatives in (3.10).

Example 3.27 (Quadratic covariation of represented semimartingales)**.** Let X be a $\mathbb{C}$–valued semimartingale. Then by Proposition 3.13(3) and Corollary 3.20, for $\xi, \psi \in \mathfrak{I}^1(X)$, we have $[\xi \circ X, \psi \circ X] = \xi\psi \circ X$. In the explicit form (3.14), this is written as

$$\begin{aligned}[\xi \circ X, \psi \circ X] &= \partial_x \xi(0)\partial_x \psi(0) \cdot [X, X]^c + \partial_{x^*}\xi(0)\partial_{x^*}\psi(0) \cdot [X^*, X^*]^c \\ &\quad + (\partial_x \xi(0)\partial_{x^*}\psi(0) + \partial_{x^*}\xi(0)\partial_x\psi(0)) \cdot [X, X^*]^c + \xi\psi \star \mu^X.\end{aligned}$$

This formula seems very intuitive. The first three terms capture the continuous covariation of $\xi \circ X$ and $\psi \circ X$. The last term is the pure-jump component which multiplies together the jumps in $\xi \circ X$ and $\psi \circ X$. □

Example 3.28 (Explicit complex-valued expression for $(|1+\mathrm{id}|^\alpha - 1) \circ X$, $\alpha \in \mathbb{C}$, $\Delta X \neq -1$)**.** Consider the predictable function $\xi = |1 + \mathrm{id}|^\alpha - 1$, which on a sufficiently small neighbourhood of zero satisfies

$$|1 + \mathrm{id}|^\alpha - 1 = (1 + \mathrm{id})^{\frac{\alpha}{2}}(1 + \mathrm{id}^*)^{\frac{\alpha}{2}} - 1.$$

On this neighbourhood, apply formal Wirtinger calculus to obtain

$$\begin{aligned}\partial_x \xi &= \frac{\alpha}{2}(1 + \mathrm{id})^{\frac{\alpha}{2}-1}(1 + \mathrm{id}^*)^{\frac{\alpha}{2}}; & \partial_{x^*}\xi &= \frac{\alpha}{2}(1 + \mathrm{id})^{\frac{\alpha}{2}}(1 + \mathrm{id}^*)^{\frac{\alpha}{2}-1}; \\ \partial^2_{xx}\xi &= \frac{\alpha}{2}\left(\frac{\alpha}{2} - 1\right)(1 + \mathrm{id})^{\frac{\alpha}{2}-2}(1 + \mathrm{id}^*)^{\frac{\alpha}{2}}; & \partial^2_{x^*x^*}\xi &= \frac{\alpha}{2}\left(\frac{\alpha}{2} - 1\right)(1 + \mathrm{id})^{\frac{\alpha}{2}}(1 + \mathrm{id}^*)^{\frac{\alpha}{2}-2};\end{aligned}$$

$$\partial^2_{xx^*}\xi = \frac{\alpha^2}{4}|1 + \mathrm{id}|^{\alpha-1}.$$

Next, $\xi \in \mathfrak{U}$, hence $\xi \in \mathfrak{I}(X)$ for *any* $\mathbb{C}$–valued semimartingale X with $X \in \mathsf{Dom}(\xi)$, in particular for any X with $\Delta X \neq -1$. Formula (3.14) now yields

$$\begin{aligned}(|1 + \mathrm{id}|^\alpha - 1) \circ X &= \frac{\alpha}{2} \cdot (X + X^*) + \frac{\alpha}{4}\left(\frac{\alpha}{2} - 1\right)([X, X]^c + [X^*, X^*]^c) + \frac{\alpha^2}{4}[X, X^*] \\ &\quad + \left(|1 + \mathrm{id}|^\alpha - 1 - \frac{\alpha}{2}(\mathrm{id} + \mathrm{id}^*)\right) * \mu^X \\ &= \alpha \cdot \mathrm{Re}\, X + \frac{\alpha}{2}(\alpha - 1)[\mathrm{Re}\, X, \mathrm{Re}\, X]^c + \frac{\alpha}{2}[\mathrm{Im}\, X, \mathrm{Im}\, X]^c \\ &\quad + (|1 + \mathrm{id}|^\alpha - 1 - \alpha\, \mathrm{Re}\, \mathrm{id})) * \mu^X.\end{aligned}$$

We continue discussing this setup in Example 5.10 below and apply it in [3, Examples 4.3 and 4.4]. There one obtains the Mellin transform of the positive and negative parts of a signed stochastic exponential of a process with independent increments. □

4 Specific examples of the semimartingale representation

4.1 Generic applications

If X is a $\mathbb{C}$–valued semimartingale, then by [9, Théorème 1] (see also [18, I.4.60]) the stochastic exponential $\mathscr{E}(X)$ of X is the unique solution to the stochastic differential equation

$$\mathscr{E}(X) = 1 + \mathscr{E}(X)_- \cdot X. \tag{4.1}$$

The stochastic logarithm $\mathcal{L}(X)$ of a semimartingale X that can hit zero only by a jump (but not continuously) and is absorbed in zero is given by

$$\mathcal{L}(X) = \frac{1}{X_-}\mathbf{1}_{\{X_- \neq 0\}} \cdot X,$$

where $\mathbf{1}_{\{X_{t-} \neq 0\}}/X_{t-}$ is defined to be zero on the set $\{X_{t-} = 0\}$, for all $t \geq 0$.

All representing functions shown in this subsection belong to the universal class $\mathfrak{U}$ and can therefore be applied to any semimartingale whose jumps are compatible with the

given function (Proposition 3.6). The simplified stochastic calculus yields many identities by straightforward computations. Using only the Itô–Meyer change of variables formula, those identities would involve convoluted arguments. As an example, we now establish a generalization of Yor's formula and its converse (see [18, II.8.19–20]).

Proposition 4.1 (Generalized Yor formula and its converse)**.** *Consider a $\mathbb{C}^2$–valued semimartingale X and α, $\beta \in \mathbb{C}$.*

(1) *Assume that the following conditions hold.*

- *If $\alpha \in \{0, -1, -2, \dots\}$, then $\Delta X^{(1)} \neq -1$.*
- *If $\alpha \in \mathbb{C} \setminus \mathbb{Z}$, then $\operatorname{Re} \mathscr{E}\big(X^{(1)}\big) > 0$.*

Assume that these two conditions also hold with α and $X^{(1)}$ replaced by β and $X^{(2)}$, respectively. We then have

$$\mathscr{E}\big(X^{(1)}\big)^{\alpha} \mathscr{E}\big(X^{(2)}\big)^{\beta} = \mathscr{E}\left(\big((1+\mathrm{id}_1)^{\alpha}(1+\mathrm{id}_2)^{\beta} - 1\big) \circ X\right), \tag{4.2}$$

where complex powers with exponent in $\mathbb{C} \setminus \mathbb{Z}$ are defined via the principal value logarithm. In particular, with $\alpha = \beta = 1$ we have

$$\mathscr{E}\big(X^{(1)}\big)\mathscr{E}\big(X^{(2)}\big) = \mathscr{E}\big(X^{(1)} + X^{(2)} + \big[X^{(1)}, X^{(2)}\big]\big). \tag{4.3}$$

(2) *Assume next the following conditions.*

- *If $\alpha \in \mathbb{N}$, then $X^{(1)}$ does not reach zero continuously and is absorbed in zero.*
- *If $\alpha \in \{0, -1, -2, \dots\}$, then $X^{(1)} \neq 0$ and $X^{(1)}_- \neq 0$.*
- *If $\alpha \in \mathbb{C} \setminus \mathbb{Z}$, then $\operatorname{Re} X^{(1)} > 0$ and $\operatorname{Re} X^{(1)}_- > 0$.*

Assume that these three conditions also hold with α and $X^{(1)}$ replaced by β and $X^{(2)}$, respectively. Finally, denote by τ the first time $X^{(1)}$ or $X^{(2)}$ hit zero. We then have

$$\mathcal{L}\big(\big(X^{(1)}\big)^{\alpha}\big(X^{(2)}\big)^{\beta}\big) = \big((1+\mathrm{id}_1)^{\alpha}(1+\mathrm{id}_2)^{\beta} - 1\big) \circ \big(\mathcal{L}\big(X^{(1)}\big)^{\tau}, \mathcal{L}\big(X^{(2)}\big)^{\tau}\big). \tag{4.4}$$

In particular, with $\alpha = \beta = 1$ and $X^{(1)}, X^{(2)}$ not hitting zero we have

$$\mathcal{L}\big(X^{(1)}X^{(2)}\big) = \mathcal{L}\big(X^{(1)}\big) + \mathcal{L}\big(X^{(2)}\big) + \big[\mathcal{L}\big(X^{(1)}\big), \mathcal{L}\big(X^{(2)}\big)\big].$$

Proof. To start, from (4.1) for example, recall that by Proposition 3.14

$$\mathscr{E}\big(X^{(k)}\big) = 1 + \mathscr{E}\big(X^{(k)}\big)_- \mathrm{id}_k \circ X, \qquad k \in \{1, 2\}. \tag{4.5}$$

Now, the change of variables formula in Proposition 3.15 applied to the function $f = \mathrm{id}_1^{\alpha}\mathrm{id}_2^{\beta}$ over an appropriate domain $\mathcal{U}$ (obtained as the Cartesian product of appropriate one-dimensional domains, i.e., $\mathbb{C}$ for $\alpha \in \mathbb{N}$; $\mathbb{C} \setminus \{0\}$ for $\alpha \in \{0, -1, \dots\}$; $\{z \in \mathbb{C} : \operatorname{Re} z > 0\}$ for $\alpha \in \mathbb{C} \setminus \mathbb{Z}$; and likewise with β in place of α) combined with Proposition 3.14 and Theorem 3.17 yield

$$\mathscr{E}\big(X^{(1)}\big)^{\alpha}\mathscr{E}\big(X^{(2)}\big)^{\beta} = 1 + \Big(\mathscr{E}\big(X^{(1)}\big)^{\alpha}_- \mathscr{E}\big(X^{(2)}\big)^{\beta}_-\Big) \cdot \Big(\big(1+\mathrm{id}_1\big)^{\alpha}\big(1+\mathrm{id}_2\big)^{\beta} - 1\Big) \circ X.$$

The uniqueness of strong solutions to the stochastic differential equation (4.1) then yields (4.2).

Next, define $Y^{(k)} = \mathcal{L}(X^{(k)})$ for $k \in \{1, 2\}$. Then from (4.2) we obtain

$$\big(X^{(1)}\big)^{\alpha}\big(X^{(2)}\big)^{\beta} = \mathscr{E}\big(Y^{(1)}\big)^{\alpha}\mathscr{E}\big(Y^{(2)}\big)^{\beta} = \mathscr{E}\left(\big((1+\mathrm{id}_1)^{\alpha}(1+\mathrm{id}_2)^{\beta} - 1\big) \circ Y\right).$$

Taking stochastic logarithms on both sides yields (4.4). □

With $\alpha = 1$ and $\beta = -1$, identity (4.2) yields a $\mathbb{C}$–valued extension of Equation (1-5) in Mémin [26]. If we additionally assume $X^{(1)} = 0$ and $Y = X^{(2)}$, we get

$$\mathcal{L}\left(\frac{1}{\mathscr{E}(Y)}\right) = \left(\frac{1}{1+\mathrm{id}} - 1\right) \circ Y = Y_0 - Y + [Y,Y]^c + \frac{\mathrm{id}^2}{1+\mathrm{id}} * \mu^Y.$$

Here we used that the function $(1+\mathrm{id})^{-1} - 1 \in \mathfrak{U}$ is analytic at 0 in conjunction with Propositions 3.6 and 3.26. See also Larsson and Ruf [23, Theorem 4.1] for an $\mathbb{R}$–valued version.

Parts of the following proposition, restricted to real-valued semimartingales, appear in Jacod and Shiryaev [18, II.8.8–12].

Proposition 4.2 (Identities involving natural / stochastic exponential / logarithm)**.** *Let X denote a $\mathbb{C}$–valued semimartingale. Then*

$$\mathcal{L}\big(\mathrm{e}^X\big) = (\mathrm{e}^{\mathrm{id}} - 1) \circ X; \tag{4.6}$$

$$|\mathscr{E}(X)| = \mathscr{E}\left((|1+\mathrm{id}| - 1) \circ X\right). \tag{4.7}$$

If $\Delta X \neq -1$, then

$$\mathscr{E}(X) = \mathrm{e}^{\log(1+\mathrm{id})\circ X}; \tag{4.8}$$

$$\log|\mathscr{E}(X)| = \log|1+\mathrm{id}| \circ X,$$

where $\log$ denotes again the principal value logarithm. Moreover, if $\operatorname{Re}\mathscr{E}(X) > 0$ then

$$\log \mathscr{E}(X) = \log(1+\mathrm{id}) \circ X.$$

Proof. Apply the change of variables formula in Proposition 3.15 to the function e^{id} to obtain

$$\mathrm{e}^X = \mathrm{e}^{X_0} + \big(\mathrm{e}^{X_- + \mathrm{id}} - \mathrm{e}^{X_-}\big) \circ X.$$

As $1/\mathrm{e}^{X_-}$ is locally bounded, Proposition 3.14 in conjunction with Theorem 3.17 yield

$$\mathcal{L}\big(\mathrm{e}^X\big) = \mathrm{e}^{-X_-} \cdot \mathrm{e}^X = \mathrm{e}^{-X_-}\big(\mathrm{e}^{X_- + \mathrm{id}} - \mathrm{e}^{X_-}\big) \circ X,$$

which on simplification gives (4.6).

Next we will prove (4.7) under the additional assumption $\Delta X \neq -1$. Since the function $|\mathrm{id}|$ is twice continuously real-differentiable on $\mathcal{U} = \mathbb{C} \setminus \{0\}$ and $\mathscr{E}(X)_-, \mathscr{E}(X)$ take values in $\mathcal{U}$, Propositions 3.14 and 3.15, representation (4.5), and Theorem 3.17 yield

$$\begin{aligned}\big|\mathscr{E}(X)\big| &= (|\mathscr{E}(X)_- + \mathrm{id}| - |\mathscr{E}(X)_-|) \circ \mathscr{E}(X) = (|\mathscr{E}(X)_- + \mathscr{E}(X)_-\mathrm{id}| - |\mathscr{E}(X)_-|) \circ X \\ &= (|\mathscr{E}(X)_-|\,(|1+\mathrm{id}| - 1)) \circ X = |\mathscr{E}(X)_-| \cdot ((|1+\mathrm{id}| - 1) \circ X)\,,\end{aligned}$$

therefore (4.7) holds in this special case.

Define next

$$Y^{(1)} = \mathrm{id}\mathbf{1}_{\mathrm{id}\neq -1} \circ X \qquad \text{and} \qquad Y^{(2)} = -\mathbf{1}_{\mathrm{id}=-1} \circ X. \tag{4.9}$$

Observe that $X = Y^{(1)} + Y^{(2)}$, $[Y^{(1)}, Y^{(2)}] = 0$, and $|\mathscr{E}(Y^{(2)})| = \mathscr{E}(Y^{(2)})$ as the latter only takes on values 0 and 1. The Yor formula in (4.3) now yields

$$|\mathscr{E}(X)| = \big|\mathscr{E}\big(Y^{(1)}\big)\mathscr{E}\big(Y^{(2)}\big)\big| = \big|\mathscr{E}\big(Y^{(1)}\big)\big|\mathscr{E}\big(Y^{(2)}\big). \tag{4.10}$$

Moreover, note that $\Delta Y^{(1)} \neq -1$ hence by the special case of (4.7) shown earlier we have

$$|\mathscr{E}(Y^{(1)})| = \mathscr{E}\left((|1+\mathrm{id}| - 1) \circ Y^{(1)}\right) = \mathscr{E}\big((|1+\mathrm{id}\mathbf{1}_{\mathrm{id}\neq -1}| - 1) \circ X\big), \tag{4.11}$$

where the second equality follows from (4.9) and Theorem 3.17. Equations (4.9)–(4.11) now yield

$$|\mathscr{E}(X)| = \mathscr{E}\big((|1+\mathrm{id}\mathbf{1}_{\mathrm{id}\neq -1}| - 1) \circ X\big)\mathscr{E}\left(-\mathbf{1}_{\mathrm{id}=-1} \circ X\right)$$

and a second application of the Yor formula (4.3) concludes the proof of (4.7) in full generality.

Assume from now on that $\Delta X \neq -1$. Observe that $\log(1+\mathrm{id})$ is the right-inverse of the function $\mathrm{e}^{\mathrm{id}} - 1$ over the domain $\mathbb{C} \setminus \{-1\}$ and that $\log(1+\mathrm{id}) \in \mathfrak{U} \cap \mathfrak{I}(X)$. We may therefore define $Y = \log(1+\mathrm{id}) \circ X$. From (4.6) and the composition Theorem 3.17 one obtains

$$\mathcal{L}\left(\mathrm{e}^{\log(1+\mathrm{id})\circ X}\right) = \mathcal{L}\big(\mathrm{e}^{Y}\big) = \big(\mathrm{e}^{\mathrm{id}} - 1\big) \circ Y = \big(\mathrm{e}^{\log(1+\mathrm{id})} - 1\big) \circ X = X - X_0,$$

which yields (4.8).

Finally, for a semimartingale Y satisfying $Y > 0$ and $Y_- > 0$ one obtains by Proposition 3.15, the identity $Y = Y_-\mathrm{id} \circ \mathcal{L}(Y)$, and Theorem 3.17 that

$$\begin{aligned}\log Y &= (\log(Y_- + \mathrm{id}) - \log(Y_-)) \circ Y \\ &= (\log(Y_- + Y_-\mathrm{id}) - \log(Y_-)) \circ \mathcal{L}(Y) = \log(1+\mathrm{id}) \circ \mathcal{L}(Y),\end{aligned}$$

hence

$$\log|\mathscr{E}(X)| = \log \mathscr{E}\left((|1+\mathrm{id}| - 1) \circ X\right) = \log(1+\mathrm{id}) \circ \left((|1+\mathrm{id}| - 1) \circ X\right) = \log|1+\mathrm{id}| \circ X,$$

again by composition.

Consider now X such that $\mathrm{Re}\,\mathscr{E}(X) > 0$, hence $\mathrm{Re}\,\mathscr{E}(X)_- \geq 0$ and $\mathscr{E}(X)_- \neq 0$. As in the previous step, by Proposition 3.15 and Theorem 3.17 one obtains

$$\begin{aligned}\log \mathscr{E}(X) &= (\log(\mathscr{E}(X)_- + \mathrm{id}) - \log \mathscr{E}(X)_-) \circ \mathscr{E}(X) \\ &= (\log(\mathscr{E}(X)_-(1+\mathrm{id})) - \log \mathscr{E}(X)_-) \circ X = \log(1+\mathrm{id}) \circ X,\end{aligned}$$

where the last equality follows by comparing the respective Émery formulae. □

Proposition 4.3 (Generalized Yor formula involving absolute values)**.** *Consider a $\mathbb{C}^2$–valued semimartingale X and α, $\beta \in \mathbb{C}$.*

(1) *Assume that the following condition holds.*

- *If $\alpha \in \mathbb{C} \setminus (0,\infty)$ then $\Delta X^{(1)} \neq -1$.*

Assume that this condition also holds with α and $X^{(1)}$ replaced by β and $X^{(2)}$, respectively. We then have

$$\big|\mathscr{E}\big(X^{(1)}\big)\big|^{\alpha}\big|\mathscr{E}\big(X^{(2)}\big)\big|^{\beta} = \mathscr{E}\left(\left(|1+\mathrm{id}_1|^{\alpha}|1+\mathrm{id}_2|^{\beta} - 1\right) \circ X\right). \tag{4.12}$$

(2) *Assume next the following conditions.*

- *If $\alpha \in (0,\infty)$ then $X^{(1)}$ does not reach zero continuously and is absorbed in zero.*
- *If $\alpha \in \mathbb{C} \setminus (0,\infty)$ then $X^{(1)} \neq 0$ and $X^{(1)}_- \neq 0$.*

Assume that these two conditions also hold with α and $X^{(1)}$ replaced by β and $X^{(2)}$, respectively. Finally, denote by τ the first time $X^{(1)}$ or $X^{(2)}$ hit zero. We then have

$$\mathcal{L}\left(\big|X^{(1)}\big|^{\alpha}\big|X^{(2)}\big|^{\beta}\right) = \left(|1+\mathrm{id}_1|^{\alpha}|1+\mathrm{id}_2|^{\beta} - 1\right) \circ \big(\mathcal{L}\big(X^{(1)}\big)^{\tau}, \mathcal{L}\big(X^{(2)}\big)^{\tau}\big). \tag{4.13}$$

Proof. First, for $\alpha \in \mathbb{C} \setminus (0,\infty)$ and $k \in \{1,2\}$ we have thanks to (4.7)

$$\big|\mathscr{E}\big(X^{(k)}\big)\big| = \mathscr{E}\left((|1+\mathrm{id}| - 1) \circ X^{(k)}\right) > 0. \tag{4.14}$$

Formula (4.12) now follows via an application of Proposition 4.1(1) in conjunction with Theorem 3.17. In the case $\alpha \in (0,\infty)$, the function $|\mathrm{id}|^{\alpha}$ is well defined on $\mathbb{C}$ rather than just $\mathbb{C} \setminus \{0\}$, hence condition (4.14) is not needed. For $\alpha \in \mathbb{N}$ we may appeal again to

Proposition 4.1(1) to obtain (4.12) and the result easily carries over to all $\alpha \in (0, \infty)$. Item (1) is proved.

Representation (4.13) is now obtained by writing $X^{(k)} = \mathscr{E}\big(\mathcal{L}\big(X^{(k)}\big)\big)$ for each $k \in \{1, 2\}$, applying formula (4.12) to the stochastic exponentials, and finally taking stochastic logarithms on both sides of (4.12). □

Example 4.4 (Iterated composition)**.** Let us now consider the following construction for a $\mathbb{C}$–valued semimartingale X and for a constant $\alpha \in \mathbb{C}$. Define inductively the processes $Y^0 = X$;

$$Y^k = \mathcal{L}\big(\exp\big(\alpha Y^{k-1}\big)\big), \qquad k \in \mathbb{N}.$$

Then an induction argument, (4.6), and Theorem 3.17 yield that $Y^k = \xi^k \circ X$ for all $k \in \mathbb{N} \cup \{0\}$, with $\xi^0 = \mathrm{id}$;

$$\xi^k = \exp\big(\alpha \xi^{k-1}\big) - 1, \qquad k \in \mathbb{N}.$$

Explicitly, for each $k \in \mathbb{N}$, ξ^k is a nested function of the form

$$\xi^k = \underbrace{\exp\left(\alpha\left(\ldots\left(\exp\left(\alpha\left(\exp\left(\alpha\,\mathrm{id}\right) - 1\right)\right) - 1\right)\ldots\right)\right) - 1}_{k \text{ times}}.$$

Using the chain rule, one infers that ξ^k is analytic at zero for each $k \in \mathbb{N}$ with

$$\begin{aligned} D\xi^k(0) &= \alpha D\xi^{k-1}(0); \\ D^2\xi^k(0) &= \alpha D^2\xi^{k-1}(0) + \alpha^2\big(D\xi^{k-1}(0)\big)^2, \end{aligned}$$

which implies

$$D\xi^k(0) = \alpha^k; \qquad D^2\xi^k(0) = \alpha D^2\xi^{k-1}(0) + \alpha^{2k} = \alpha^{k+1}\frac{\alpha^k - 1}{\alpha - 1}, \qquad k \in \mathbb{N},$$

where for $\alpha = 1$ we interpret $(\alpha^k - 1)/(\alpha - 1)$ as k. We conclude that, for each $k \in \mathbb{N}$,

$$Y^k = \alpha^k(X - X_0) + \frac{1}{2}\alpha^{k+1}\frac{\alpha^k - 1}{\alpha - 1}[X, X]^c + \big(\xi^k - \alpha^k\,\mathrm{id}\big) * \mu^X. \tag{4.15}$$

Note that this representation of Y^k is the same for *any* starting process X, for each $k \in \mathbb{N}$. For example, let $X_t = \mu t + \sigma W_t$ for all $t \geq 0$, where W is Brownian motion with $W_0 = 0$. Here $\mu \in \mathbb{R}$ denotes the drift rate and $\sigma \in \mathbb{R}$ the volatility. Then (4.15) yields

$$Y_t^k = \alpha^k \sigma W_t + \left(\alpha^k \mu + \frac{1}{2}\alpha^{k+1}\frac{\alpha^k - 1}{\alpha - 1}\sigma^2\right) t, \qquad t \geq 0,$$

for all $k \in \mathbb{N}$. Classical calculus would yield the same result, of course. For each $k \in \mathbb{N}$, one would repeatedly compute

$$Y^k = \exp\Big(-\alpha Y_-^{k-1}\Big) \cdot \exp\Big(\alpha Y^{k-1}\Big).$$

This is not too complicated but can easily become quite cumbersome, even in the case of drifted Brownian motion. □

Example 4.5 (Itô–Wentzell formula)**.** The semimartingale representation proposed in this paper naturally leads to a parsimonious generalization of the Itô–Wentzell formula; see Jeanblanc, Yor, and Chesney [19, Theorem 1.5.3.2] and also Bank and Baum [1, Proposition 1.3]. To this end, consider an $\mathbb{R}^n$–valued semimartingale V and a predictable function ψ such that $\psi(x, \cdot) \in \mathfrak{I}(V)$ for each $x \in \mathbb{R}^d$. Define next a family of semimartingales $(F(x))_{x \in \mathbb{R}^d}$ by setting

$$F(x) = \psi\left(x, \mathrm{id}\right) \circ V.$$

One can now randomize the family F by allowing x to switch values stochastically in line with the $\mathbb{R}^d$–valued semimartingale X. Assuming F is sufficiently smooth, the

randomized process $F(X)$ defined by

$$F(X)_t = F_t(x)|_{x=X_t}, \qquad t \geq 0$$

will again be a semimartingale. The observation

$$\Delta F(X) = (F(X) - F_-(X)) + (F_-(X) - F_-(X_-))$$

then yields, under suitable technical conditions, that

$$F(X) = \xi \circ (X, V) \tag{4.16}$$

with

$$\xi(x, v) = \psi(X_- + x, v) + (\psi(\theta + x, \mathrm{id}) - \psi(\theta, \mathrm{id})) \circ V_-|_{\theta = X_-}, \qquad x \in \mathbb{R}^d,\ v \in \mathbb{R}^n.$$

We leave the technical details to future work. For the moment, we only note that for $\mathbb{R}$–valued continuous processes X and V and for $\psi(x, v) = f(x)v$, where $x, v \in \mathbb{R}$ and $f : \mathbb{R} \to \mathbb{R}$ is twice continuously differentiable, one formally obtains

$$D_1\xi(0,0) = f'(X)\,\mathrm{id} \circ V = f'(X) \cdot V; \qquad D_2\xi(0,0) = f(X);$$
$$D^2_{1,1}\xi(0,0) = f''(X)\,\mathrm{id} \circ V = f''(X) \cdot V; \qquad D^2_{1,2}\xi(0,0) = f'(X); \qquad D^2_{2,2}\xi(0,0) = 0.$$

If one can now show that $f'(X) \cdot V = F'(X)$ and $f''(X) \cdot V = F''(X)$, then (4.16) yields the statement of Jeanblanc et al. [19, Theorem 1.5.3.2]. □

As the examples illustrate, the stochastic calculus introduced above is powerful and simple. Stochastic integration, Itô's formula, and the composition rule of Theorem 3.18 allow for a wide range of applications. Within the confines of their assumptions they show that it is enough to study jump transformations; i.e., to represent Y in terms of X it suffices to trace how the jump ΔX_t is transformed into the jump ΔY_t at time $t \geq 0$.

4.2 Counterexamples

This subsection illustrates the tightness of the results in Section 3 by providing several counterexamples.

Example 4.6 ($\xi \in \mathfrak{I}(X)$, but $\xi'(0) \notin L(X)$)**.** Here, we construct a process $X \in \mathscr{V}^{\mathrm{d}}_\sigma$ and a predictable function $\xi \in \mathfrak{I}(X)$, twice continuously differentiable at zero, such that $\xi'(0) \notin L(X)$. This illustrates the role of the predictable set $\mathcal{H}_X$ in Definition 3.2.

Let $U \in \mathscr{V}^{\mathrm{d}}_\sigma$ denote a piecewise constant martingale that jumps at times $2 - 1/n$ by $\pm 1/n^2$. Let $(\Theta_n)_{n\in\mathbb{N}}$ denote an independent sequence of independent $\{0, 1\}$–valued random variables with $\mathsf{P}[\Theta_n = 1] = 1/n^4$. Let $(\Psi_n)_{n\in\mathbb{N}}$ denote a sequence, independent of U and $(\Theta_n)_{n\in\mathbb{N}}$, of independent standard normally distributed random variables. Let now $V \in \mathscr{V}^{\mathrm{d}}_\sigma$ denote a piecewise constant martingale that jumps at times $2 - 1/n$ by Ψ_n if $\Theta_n = 1$ and does not jump if $\Theta_n = 0$.

Next, set $X = U + V \in \mathscr{V}^{\mathrm{d}}_\sigma$ and assume the filtration is the natural filtration of X. An application of Borel–Cantelli then yields that V only has finitely many jumps, hence $\Delta X = \Delta U$ except finitely many times. Consider next the deterministic predictable function ξ given by

$$\xi_t = \mathbf{1}_{t<2}\left(\mathrm{id}\mathbf{1}_{|\mathrm{id}|\geq(2-t)^2} + \frac{1}{(2-t)^2}\mathrm{id}\mathbf{1}_{|\mathrm{id}|<(2-t)^2}\right), \qquad t \geq 0.$$

Note that

$$\xi'_t(0) = \mathbf{1}_{t<2}(2-t)^{-2}, \qquad t \geq 0;$$

hence $\xi'_{2-1/n}(0) = n^2$ for all $n \in \mathbb{N}$ and $|\xi'(0)\mathrm{id}|^2 * \mu^X_2 = \infty$. Then thanks to Propositions 2.6, 2.8, 2.9, and 3.7(2), $\xi'(0) \notin L(X)$ but $\xi \in L_\sigma(\mu^X) = \mathfrak{I}(X)$.

Moreover, if $Y = \xi \circ X$ also satisfies $Y = \eta \circ X$ for some $\eta \in \mathfrak{I}(X)$ and $\eta'(0)$ exists, then $\eta_{2-1/n} = \xi_{2-1/n}$ for all $n \in \mathbb{N}$, hence also $\eta'(0) \notin L(X)$. □

Example 4.7 ($\xi \in \mathfrak{I}(X)$ and $\psi \in \mathfrak{I}(\xi \circ X)$, but $\psi(\xi) \notin \mathfrak{I}(X)$; also: $\xi \in \mathfrak{I}(X)$, $\zeta \in L(\xi \circ X)$, but $\zeta\xi \notin \mathfrak{I}(X)$)**.** Consider a continuous semimartingale X given by

$$X_t = W_t - \int_0^t s^{-2/3}\mathrm{d}s = W_t - 3t^{\frac{1}{3}}, \qquad t \geq 0,$$

where W denotes a standard Brownian motion. Define the predictable, indeed, deterministic functions

$$\xi_t = \mathrm{id} + \mathrm{id}^2\, t^{-2/3}\mathbf{1}_{t>0}; \qquad \psi_t = \mathrm{id}\, t^{-1/3}\mathbf{1}_{t>0}, \qquad\qquad t \geq 0.$$

Thus, $\xi \in \mathfrak{I}^1(X)$; in particular, $Y = \xi \circ X$ satisfies $Y = W$. Hence, also $\psi \in \mathfrak{I}^1(Y)$ and $\psi \circ Y = \int_0^\cdot s^{-1/3}\mathrm{d}W_s$.

Now let $\eta = \psi(\xi)$. Thanks to (3.8) we have $\eta_t'(0) = t^{-1/3}$ for all $t > 0$; hence, $\eta \notin \mathfrak{I}(X)$ despite $\xi \in \mathfrak{I}(X)$ and $\psi \in \mathfrak{I}(\xi \circ X)$. In this example, $\psi'(0)$ is deterministic, but (3.6) does not hold. Hence, there is no contradiction to Theorem 3.18. □

Example 4.8 (Alternative construction: $\xi \in \mathfrak{I}(X)$ and $\psi \in \mathfrak{I}(\xi \circ X)$, but $\psi(\xi) \notin \mathfrak{I}(X)$; also: $\xi \in \mathfrak{I}(X)$, $\zeta \in L(\xi \circ X)$, but $\zeta\xi \notin \mathfrak{I}(X)$)**.** Let $(\tau_k)_{k\in\mathbb{N}}$ be a sequence of independent random variables with τ_k uniformly distributed on $(1/(k+1), 1/k)$. Let $(J_k)_{k\in\mathbb{N}}$ be an independent sequence of independent $\{2,4\}$–valued random variables with $\mathsf{P}[J_k = 2] = 1/2 = \mathsf{P}[J_k = 4]$. Set now

$$X_t = t^3 + \sum_{k=1}^{\infty} \tau_k^{J_k}\mathbf{1}_{[\![\tau_k,\infty[\![}(t), \qquad t \geq 0,$$

and assume that the filtration be the right-continuous modification of the one generated by the finite-variation process X.

Consider now deterministic ξ and ψ given by $\xi_t = \mathbf{1}_{\mathrm{id}\leq t^4}\mathrm{id}$ and $\psi_t = \mathrm{id}/t^2\mathbf{1}_{t>0}$ for all $t \geq 0$. Then $\xi \in \mathfrak{I}(X)$ with

$$Y_t = \xi \circ X_t = t^3 + \sum_{k=1}^{\infty} \tau_k^4\mathbf{1}_{\{J_k=4\}}\mathbf{1}_{[\![\tau_k,\infty[\![}(t), \qquad t \geq 0,$$

and $\psi \in \mathfrak{I}(Y)$ with

$$\psi \circ Y_t = 3t + \sum_{k=1}^{\infty} \tau_k^2\mathbf{1}_{\{J_k=4\}}\mathbf{1}_{[\![\tau_k,\infty[\![}(t), \qquad t \geq 0.$$

However, it is clear that $\psi(\xi) \notin \mathfrak{I}(X)$. An even stronger statement holds, namely that there exists no $\eta \in \mathfrak{I}(X)$ such that $\psi \circ Y = \eta \circ X$. □

Example 4.9 ($\xi^{-1} \notin \mathfrak{I}(\xi \circ X)$)**.** Assume that $X = W$ is standard Brownian motion. Let ξ denote some deterministic predictable function that satisfies $\xi_t(x) = tx + x^2/2$ for all $t > 0$ and x in a neighbourhood of zero (which may depend on t) and allows for an inverse. Then $\xi_t'(0) = t$ and $\xi_t''(0) = 1$ for all $t > 0$. Hence $\xi \in \mathfrak{I}^1(X)$ and we can define $Y = \xi \circ X$, satisfying

$$Y_t = \int_0^t s\mathrm{d}W_s + t, \qquad t > 0.$$

Moreover, with $\psi = \xi^{-1}$ we have $\psi(\xi) \in \mathfrak{I}^1(X)$. Observe, however, that $\psi'(0) \notin L(Y)$ and $\psi''(0) \notin L([Y,Y]^c)$ since $\psi_t'(0) = 1/t$ and

$$\psi_t''(0) = \frac{-\xi_t''(0)}{(\xi_t'(0))^3} = -t^{-3}, \qquad t \geq 0.$$

Thus $\psi \notin \mathfrak{I}(Y)$ but there is no contradiction to Remark 3.22 as (3.6) is not met. Because $\psi'(0) = 1/\xi'(0)$ is not locally bounded, this example does not contradict Corollary 3.23 either. Note that there exists no $\eta \in \mathfrak{I}(Y)$ such that $X = \eta \circ Y$. □

Example 4.10 (Alternative construction: $\xi^{-1} \notin \mathfrak{I}(\xi \circ X)$; additionally $\xi \circ X = X$)**.** Let

$(\tau_k)_{k\in\mathbb{N}}$ be a sequence of independent random variables with τ_k uniformly distributed on $(1/(k+1), 1/k)$. Let $(U_k)_{k\in\mathbb{N}}$ be an independent sequence of independent and identically distributed $\{-1,1\}$–valued random variables with $\mathsf{P}[U_1 = 1] = 1/2$. Set now

$$X = \sum_{k=1}^{\infty} U_k \tau_k \mathbf{1}_{[\![\tau_k,\infty[\![} \in \mathscr{V}_\sigma^{\mathrm{d}},$$

and assume that the filtration be the right-continuous modification of the one generated by X.

Let ξ denote some deterministic predictable function that allows for an inverse and satisfies, for all $t > 0$, $\xi_t(x) = x$ for all x with $|x| \geq 1/(t+1)$ and $\xi_t(x) = tx$ for all x in a neighbourhood of zero (which may depend on t). Then $\xi \in \mathfrak{I}(X)$ and $X = \xi \circ X$. However, since $\mathbf{1}_{\mathcal{H}_X} = 1$ and $D\xi^{-1}(0) \notin L(X)$, we have $\xi^{-1} \notin \mathfrak{I}(X)$, concluding the example. □

Example 4.11 ($\xi, \psi(\xi) \in \mathfrak{I}(X)$ and $\psi \in \mathfrak{I}(\xi \circ X)$ but $\psi(\xi) \circ X \neq \psi \circ (\xi \circ X)$)**.** Let X be a continuous semimartingale not equal to the zero process. Consider $\xi = \mathrm{id}^3 \in \mathfrak{U}$ and $\psi = \xi^{-1}$. Note that $\psi(\xi) = \mathrm{id} \in \mathfrak{U}$ and that $\xi \circ X = 0$, hence $\psi \in \mathfrak{I}(\xi \circ X)$. However, we have

$$\psi(\xi) \circ X = X \neq 0 = \psi \circ (\xi \circ X).$$

Theorem 3.18 is not contradicted because ψ is not differentiable at zero, $(\mathsf{P} \times A^X)$–a.e. □

5 Predictable characteristics

Up to this point we have relied on a 'pathwise' perspective in the sense that the representation of the process Y by means of $\xi \circ X$ depends on the probability measure only through the null sets; see also Remark 3.12. Now we will demonstrate the ability to convert an X–representation into predictable characteristics. In this section, we shall use generalized conditional expectation; see Shiryaev [33, pp. 475–476] and Jacod and Shiryaev [18, I.1.1].

5.1 Truncation functions

In [18, II.2.3–4], a bounded function $h : \mathbb{C}^d \to \mathbb{C}^d$ is called a truncation function for X if $h(x) = x$ in a neighbourhood of zero. For such h the process $X[h] = X - (\mathrm{id} - h) * \mu^X$ only has bounded jumps and is therefore special. Below, it will be useful to not only control the jumps of X, but also those of a stochastic integral with respect to $\widehat{X[h]} = \hat{\mathrm{id}}(X[h])$. This leads to the following generalization of the classical truncation function where the boundedness and integrability requirements are relaxed. Moreover, h is no longer restricted to be a time-constant deterministic predictable function.

Definition 5.1 (Truncation function for X and its compatibility with $\xi \in \mathfrak{I}(X)$)**.** *We call a predictable function $h : \overline{\Omega}^d \to \overline{\mathbb{C}}^d$ a truncation function for X if*

$$\mathrm{id} - h \in L_\sigma(\mu^X) \qquad \textit{and if} \qquad X[h] = X - (\mathrm{id} - h) \star \mu^X \textit{ is special.}$$

Moreover, for $\xi \in \mathfrak{I}(X)$, we say that a truncation function h for X is ξ–compatible if $\mathbf{1}_{\mathcal{H}_X} \hat{D}\xi(0) \in L(\widehat{X[h]})$ and $\mathbf{1}_{\mathcal{H}_X} \hat{D}\xi(0) \cdot \widehat{X[h]}$ is special. □

Remark 5.2 (Observations on truncation functions)**.** If h is a truncation function for X and $\xi \in \mathfrak{I}(X)$, implying $\mathbf{1}_{\mathcal{H}_X} \hat{D}\xi(0) \in L(\hat{X})$, it does not follow that $\mathbf{1}_{\mathcal{H}_X} \hat{D}\xi(0) \in L(\widehat{X[h]})$. Indeed, there exists an $\mathbb{R}$–valued quasi-left-continuous process $V \in \mathscr{V}_\sigma^{\mathrm{d}}$ with $V_0 = 0$ whose jumps are bounded and a predictable process $\zeta \in L(V - B^V)$ such that $\zeta \notin L(B^V)$ (see [4, Example 3.11]). Let now

$$X = V - B^V.$$

Then $h = 0$ is a truncation function for X with $X[0] = -B^V$ and $\mathcal{H}_X^{\mathrm{c}}$ is empty. The predictable function $\xi = \zeta \mathrm{id}$ satisfies $\zeta = \mathbf{1}_{\mathcal{H}_X} \xi'(0) \in L(X)$ but $\mathbf{1}_{\mathcal{H}_X} \xi'(0) \notin L(X[0])$.

Consider now the process $Y = \zeta \cdot X$. We claim that $h = 0$ is not a truncation function for Y. Hence, this provides an example of a process Y such that $Y[0]$ does *not* exist. Assume it did. Then $\mathrm{id} \in L_\sigma(\mu^Y)$, yielding $\zeta\mathrm{id} \in L_\sigma(\mu^X) = L_\sigma(\mu^V)$. In view of Proposition 2.8 and the fact that $\zeta \notin L(V)$ this yields a contradiction.

As a final observation for the moment, note that the process $\mathbf{1}_{\mathcal{H}_X}\hat{D}\xi(0) \cdot \widehat{X[h]}$ may not be special even if it is known that h is a truncation function for X, $\xi \in \mathfrak{I}(X)$, and $\mathbf{1}_{\mathcal{H}_X}\hat{D}\xi(0)$ integrates $\widehat{X[h]}$. For example, let $X = \mathbf{1}_{[\![U,\infty[\![}$, where U is uniformly distributed, let $\mathfrak{F}$ denote the smallest right-continuous filtration that makes X adapted, let $\xi_t = {}^{\mathrm{id}}/_t$ for all $t > 0$, and let $h = \mathrm{id}$. Then h is a truncation function for X with $X = X[h]$, $\xi \in \mathfrak{I}(X)$, and $\mathcal{H}_X^c$ empty, but $\mathbf{1}_{\mathcal{H}_X}\xi'(0) \cdot X[h] = {}^1/_U\mathbf{1}_{[\![U,\infty[\![}$ is not special. □

Lemma 5.3 (Compatible truncation)**.** *There exists a ξ–compatible truncation function h for X, for every $\xi \in \mathfrak{I}(X)$. Moreover, h can be chosen such that $h \in \mathfrak{I}(X)$ with*

$$X[h] = X_0 + h \circ X. \tag{5.1}$$

Furthermore, if $\hat{D}\xi(0)$ is locally bounded (in particular, if $\xi \in \mathfrak{U} \cap \mathfrak{I}(X)$), then any truncation function for X is ξ–compatible.

Proof. By assumption, $\varsigma = \mathbf{1}_{\mathcal{H}_X}\hat{D}\xi(0)$ is in $L(\hat{X})$. We claim that

$$h = \mathrm{id}\mathbf{1}_{|\mathrm{id}|\le 1 \text{ and } |\varsigma\,\hat{\mathrm{id}}|\le 1} \tag{5.2}$$

has the desired properties. Indeed, $\mathrm{id} - h \in L(\mu^X)$ because both X and $\varsigma \cdot \hat{X}$ have finitely many jumps larger than one in absolute value on any compact time interval. This also yields $\varsigma \in L\big(\hat{X} - \widehat{X[h]}\big)$ and consequently $\varsigma \in L(\widehat{X[h]})$. The jumps of $X[h]$ and $\varsigma \cdot \widehat{X[h]}$ are bounded by 1 in absolute value; therefore both processes are also special.

Observe that $h(\omega, t, x) = x$ on a (ω, t)–dependent neighbourhood of zero, $(\mathsf{P} \times A^X)$–a.e. This yields that h is analytic at 0, $Dh(0)$ is an identity matrix, and $D^2h(0) = 0$. The representation formula (3.15) now gives

$$X_0 + h \circ X = X + (h - \mathrm{id}) \star \mu^X = X[h].$$

The final claim follows by localization. □

Remark 5.4 (Truncation at zero)**.** The previous lemma shows that sufficiently many truncation functions can be applied via the natural formula (5.1). We elect *not* to make (5.1) the *only* way to truncate because (5.1) does not hold for $h = 0$. Truncation at zero is convenient when X has jumps of finite variation; more generally, it can be applied whenever $\mathrm{id} \in L_\sigma(\mu^X)$. □

The next proposition recognizes that the Émery formula (3.3) represents a whole spectrum of equivalent expressions where the jumps of X can be dialled down in the first term of (3.3) as long as they are equivalently modified in the last term of (3.3). In most applications, it is possible to choose as truncation one of the polar cases $h = 0$ or $h = \mathrm{id}$; less frequently one may have to opt for an intermediate truncation such as $h = \mathrm{id}\mathbf{1}_{|\mathrm{id}|\le 1}$; in full generality it may be necessary to use the compatible truncation (5.2).

Proposition 5.5 (Émery formula involving truncation)**.** *Fix $\xi \in \mathfrak{I}^n(X)$ and let g be a truncation function for $\xi \circ X$. Moreover, let h be a ξ–compatible truncation function for X. Then the following terms are well defined and we have*

$$(\xi \circ X)[g] = \mathbf{1}_{\mathcal{H}_X}\hat{D}\xi(0) \cdot \widehat{X[h]} + \frac{1}{2}\hat{D}^2\xi(0) \cdot \big[\hat{X}, \hat{X}\big]^c + \Big(g(\xi) - \mathbf{1}_{\mathcal{H}_X}\hat{D}\xi(0)\hat{h}\Big) \star \mu^X \tag{5.3}$$

$$= \mathbf{1}_{\mathcal{H}_X}\check{D}\xi(0) \cdot \widetilde{X[h]} + \frac{1}{2}\check{D}^2\xi(0) \cdot \big[\check{X}, \check{X}\big]^c + \big(g(\xi) - \mathbf{1}_{\mathcal{H}_X}\check{D}\xi(0)\check{h}\big) \star \mu^X. \tag{5.4}$$

If $\mathbf{1}_{\mathcal{H}_X}\xi$ is analytic at 0, $(\mathsf{P}\times A^X)$–a.e., the following terms are well defined and we have

$$(\xi\circ X)[g] = \mathbf{1}_{\mathcal{H}_X} D\xi(0)\cdot X[h] + \frac{1}{2}D^2\xi(0)\cdot[X,X]^c + (g(\xi) - \mathbf{1}_{\mathcal{H}_X} D\xi(0)h)\star\mu^X. \tag{5.5}$$

If g satisfies $g(w)=w$ on an (ω,t)–dependent neighbourhood of 0, $(\mathsf{P}\times A^X)$–a.e., then we also have $g(\xi)\in\mathfrak{I}(X)$ and

$$(\xi\circ X)[g] = g(\xi)\circ X. \qquad \square$$

Proof. Thanks to g being a truncation function for $\xi\circ X$ and h being ξ–compatible (in conjunction with Proposition 2.8) we have $(g(\xi) - \mathbf{1}_{\mathcal{H}_X}\hat{D}\xi(0)\hat{h}\in L_\sigma(\mu^X)$. It is now simple to establish (5.3). Next, (5.4) and (5.5) follow as in Proposition 3.26.

The additional hypothesis on g yields $\hat{D}\hat{g}(0) = I_{2n}$ and $\hat{D}^2 g(0) = 0$, hence $g\in\mathfrak{I}(\xi\circ X)$. The last statement follows from the definition of $\mathfrak{I}(X)$, the $\circ$–notation, and from Corollary 3.20. $\square$

5.2 Characteristics under the measure P

Proposition 5.5 yields the next observation, which is the key step towards computing predictable characteristics of represented semimartingales.

Proposition 5.6 (Drift of a truncated represented semimartingale)**.** *Fix $\xi\in\mathfrak{I}(X)$ and let g be a truncation function for $\xi\circ X$. Moreover, let h be a ξ–compatible truncation function for X. Then the following terms are well defined and the predictable compensator of $(\xi\circ X)[g]$ under P is given by*

$$B^{(\xi\circ X)[g]} = \mathbf{1}_{\mathcal{H}_X}\hat{D}\xi(0)\cdot B^{\widehat{X[h]}} + \frac{1}{2}\hat{D}^2\xi(0)\cdot\big[\hat{X},\hat{X}\big]^c + \Big(g(\xi) - \mathbf{1}_{\mathcal{H}_X}\hat{D}\xi(0)\hat{h}\Big)\star\nu^X \tag{5.6}$$

$$= \mathbf{1}_{\mathcal{H}_X}\check{D}\xi(0)\cdot B^{\widetilde{X[h]}} + \frac{1}{2}\check{D}^2\xi(0)\cdot\big[\check{X},\check{X}\big]^c + \big(g(\xi) - \mathbf{1}_{\mathcal{H}_X}\check{D}\xi(0)\check{h}\big)\star\nu^X. \tag{5.7}$$

If $\mathbf{1}_{\mathcal{H}_X}\xi$ is analytic at 0, $(\mathsf{P}\times A^X)$–a.e., the following terms are well defined and we have

$$B^{(\xi\circ X)[g]} = \mathbf{1}_{\mathcal{H}_X} D\xi(0)\cdot B^{X[h]} + \frac{1}{2}D^2\xi(0)\cdot[X,X]^c + (g(\xi) - \mathbf{1}_{\mathcal{H}_X} D\xi(0)h)\star\nu^X. \tag{5.8}$$

Proof. In (5.3), the last term is special since all the other terms are special. Hence (5.6) follows from Shiryaev and Cherny [34, Lemma 4.2]. Equations (5.7) and (5.8) follow as in Proposition 3.26. $\square$

Remark 5.7 (Discrete-time and continuous-time components of a drift)**.** Recall the unique decomposition in Proposition 2.10. Consider now a predictable function $\xi\in\mathfrak{I}(X)$. Proposition 3.7(3) asserts $(\xi\circ X)^{\mathrm{qc}} = \xi\circ X^{\mathrm{qc}}$ and $(\xi\circ X)^{\mathrm{dp}} = \xi\circ X^{\mathrm{dp}}$. Next, suppose $\xi\circ X$ is special. By Propositions 2.11 and 3.13(1), the drift at predictable jump times then takes a particularly simple form, namely,

$$B^{\xi\circ X^{\mathrm{dp}}} = \sum_{\tau\in\mathcal{T}_X}\mathsf{E}_{\tau-}[\xi_\tau(\Delta X_\tau)]\mathbf{1}_{[\![\tau,\infty[\![}. \tag{5.9}$$

Observe that this formula is simpler than Proposition 5.6 applied to X^{dp} in place of X. Therefore, in practice, Proposition 5.6 is used with $X = X^{\mathrm{qc}}$ to obtain $B^{\xi\circ X^{\mathrm{qc}}}$. One then has

$$B^{\xi\circ X} = B^{\xi\circ X^{\mathrm{qc}}} + B^{\xi\circ X^{\mathrm{dp}}}.$$

Finally, recall that X^{qc} is quasi-left-continuous, hence $B^{\xi\circ X^{\mathrm{qc}}}$ is continuous, yielding

$$\Delta B^{\xi\circ X} = \Delta B^{\xi\circ X^{\mathrm{dp}}}.$$

The literature employs the following weakening of (5.9), typically with $\xi = \mathrm{id}$,

$$\Delta B^{\xi\circ X} = \int_{\mathbb{C}^d}\xi(x)\nu^X(\{\cdot\},\mathrm{d}x);$$

see, for example, [18, II.2.14]. □

Corollary 5.8 (Characteristics of a represented semimartingale)**.** *Let $Y = Y_0 + \xi \circ X$ for some $\xi \in \mathfrak{I}^n(X)$. Then the semimartingale characteristics of Y with respect to the truncation function g for Y are given by*

$$B^{Y[g]} = B^{(\xi\circ X)[g]};$$

$$\big[\hat{Y}^{(k)}, \hat{Y}^{(l)}\big]^c = \Big(\hat{D}\hat{\xi}^{(k)}(0)^\top \hat{D}\hat{\xi}^{(l)}(0)\Big) \cdot \big[\hat{X}, \hat{X}\big]^c, \quad k, l \in \{1, \cdots, 2n\}; \tag{5.10}$$

$$\big[\check{Y}^{(k)}, \check{Y}^{(l)}\big]^c = \Big(\check{D}\check{\xi}^{(k)}(0)^\top \check{D}\check{\xi}^{(l)}(0)\Big) \cdot \big[\check{X}, \check{X}\big]^c, \quad k, l \in \{1, \cdots, 2n\}; \tag{5.11}$$

$$\begin{aligned}&\nu^Y \text{ is the push-forward measure of } \nu^X \text{ under } \xi, \text{ that is, } \psi * \nu^Y = \psi(\xi) * \nu^X \\ &\text{for all non-negative bounded predictable functions } \psi \text{ with } \psi(0) = 0.\end{aligned} \tag{5.12}$$

Proof. Definition 3.8 yields (5.10) and (5.11) then follows from (3.11) and (3.12) in view of the identity

$$\big(\check{D}\xi^{(k)}(0) \cdot \check{X}\big)^* = \big(\check{D}\xi^{(k)}(0)\big)^* \cdot \big(\check{X}\big)^* = \check{D}\xi^{*(k)}(0) \cdot \check{X}, \qquad k \in \{1, \cdots, n\},$$

where the superscript $*$ denotes again the complex conjugate. The statement in (5.12) follows from Proposition 5.6 on observing that $\nu^Y(G) = B^{\mathbf{1}_G \circ (\xi\circ X)}$, where $G = G_1 \times G_2$ with $G_1 \subset [0, \infty)$ predictable and G_2 a closed set in $\mathbb{C}^n$ not containing a neighbourhood of zero. □

When ξ is of the form $\xi = f(X_- + \mathrm{id}) - f(X_-)$ for a twice continuously differentiable real-valued function f and when X is real, then Corollary 5.8 reduces to the situation in Goll and Kallsen [16, Corollary A.6]. When $\xi = R\,\mathrm{id}$ for some $\mathbb{R}^{n\times d}$–valued matrix R and X is real-valued, Corollary 5.8 yields the statement of Eberlein, Papapantoleon, and Shiryaev [12, Proposition 2.4].

Example 5.9 (Generalized Yor formula continued)**.** We continue the discussion of Proposition 4.1. Consider $\alpha, \beta \in \mathbb{C}$ and a $\mathbb{C}^2$–valued semimartingale X satisfying the assumptions of Proposition 4.1(1) and additionally X is stopped when $\Delta X^{(1)} = -1$ or $\Delta X^{(2)} = -1$. We are interested in the drift of

$$Y = \mathcal{L}\left(\mathscr{E}\big(X^{(1)}\big)^\alpha \mathscr{E}\big(X^{(2)}\big)^\beta\right) = ((1 + \mathrm{id}_1)^\alpha (1 + \mathrm{id}_2)^\beta - 1) \circ X,$$

see (4.2). Here $(1 + \mathrm{id}_1)^\alpha (1 + \mathrm{id}_2)^\beta - 1$ belongs to $\mathfrak{U}$ and is analytic at zero. Let g and h denote truncation functions for Y and X, respectively. Thanks to Lemma 5.3 and (5.8) we now have

$$\begin{aligned}B^{Y[g]} = \alpha B^{X[h]^{(1)}} + \beta B^{X[h]^{(2)}} + \frac{1}{2}\alpha(\alpha-1)\big[X^{(1)}, X^{(1)}\big]^c + \frac{1}{2}\beta(\beta-1)\big[X^{(2)}, X^{(2)}\big]^c \\ + \alpha\beta\big[X^{(1)}, X^{(2)}\big]^c + \big(g\left((1 + \mathrm{id}_1)^\alpha (1 + \mathrm{id}_2)^\beta - 1\right) - [\alpha\ \beta]h\big) * \nu^X.\end{aligned} \tag{5.13}$$

Moreover, Corollary 5.8 yields

$$[Y, Y]^c = \alpha^2\big[X^{(1)}, X^{(1)}\big]^c + \beta^2\big[X^{(2)}, X^{(2)}\big]^c + 2\alpha\beta\big[X^{(1)}, X^{(2)}\big]^c.$$

For a direct derivation of (5.13) in the real-valued case when $\alpha = 1$ and $\beta = -1$, see for example Kallsen [20, Lemma 4.3]. □

Example 5.10 (Example 3.28 continued)**.** Consider for some $\alpha \in \mathbb{C}$ and $\mathbb{C}$–valued X with $\Delta X \neq -1$, the representation $Y = (|1 + \mathrm{id}|^\alpha - 1) \circ X$. Assume for simplicity that Y is special. The function $|1 + \mathrm{id}|^\alpha - 1$ is in $\mathfrak{U}$ but not analytic at 0. Example 3.28, Lemma 5.3, and (5.6) now yield

$$B^Y = \alpha \cdot B^{\mathrm{Re}\, X[h]} + \frac{\alpha}{2}(\alpha - 1)[\mathrm{Re}\, X, \mathrm{Re}\, X]^c + \frac{\alpha}{2}[\mathrm{Im}\, X, \mathrm{Im}\, X]^c + (|1 + \mathrm{id}|^\alpha - 1 - \alpha\, \mathrm{Re}\, h) * \nu^X$$

for any truncation function h for X. □

6 Concluding remarks

Let us review the benefits of the proposed 'calculus of predictable variations.' Some of the advantages, such as universality of representations in $\mathfrak{U}$ and the ease with which calculations can be performed in a very general class of complex-valued functions, have been showcased in the introduction and subsequently in the main body of the paper. Here we want to mention several other benefits that are of a more philosophical kind or whose detailed treatment is beyond the scope of this paper and will be pursued in other work.

The literature has a number of fragmented and specialized results that fit into the framework of semimartingale representations. On their own, these results are hard to generalize and do not suggest fruitful unification, hence are also difficult to recall and disseminate. The new calculus overcomes this barrier by providing a compact, systematic way of recording existing (and new) results. Let us mention two classical examples to illustrate these advantages.

- Recall that a $\mathbb{C}$–valued continuous local martingale is called conformal if $[X,X]^c = 0$. Hence by (3.15), an analytic representation with respect to a continuous conformal local martingale is again a conformal local martingale. This not only covers a change of variables by means of an analytic function, as in Getoor and Sharpe [15, Proposition 5.4], but includes *arbitrary* representation analytic at the origin. For example, the stochastic logarithm of a natural exponential preserves continuous conformal local martingales as its representing function $\mathrm{e}^{\mathrm{id}} - 1$ is analytic at 0.

- Consider now the explicit characterization of the complex stochastic exponential due to Doléans-Dade [9, Théorème 1]. This is captured by the representation (4.8),
$$\mathscr{E}(X) = \mathrm{e}^{\log(1+\mathrm{id})\circ X}, \qquad \text{provided } \Delta X \neq -1.$$
As $\log(1+\mathrm{id})$ is in $\mathfrak{U}$ and analytic at 0, the Émery formula (3.15) yields
$$\log(1+\mathrm{id})\circ X = X - X_0 - \frac{1}{2}[X,X]^c + (\log(1+\mathrm{id}) - \mathrm{id}) * \mu^X,$$
hence the jump integral converges pathwise, P–almost surely. After exponentiation this yields the aforementioned important formula
$$\mathscr{E}(X) = \mathrm{e}^{X - X_0 - \frac{1}{2}[X,X]^c} \prod_{t\leq\cdot} \mathrm{e}^{-\Delta X_t}(1+\Delta X_t),$$
this time in full generality, because the jump to zero may be treated separately.

Further advantages of the new calculus emerge when one is tasked with computing the drift of a represented process under some new probability measure Q whose density Z with respect to P is also represented, say by $\mathcal{L}(Z) = \psi\circ X$. It now suffices to observe that by Girsanov's theorem the Q–drift of X equals the P–drift of $X + [X, \mathcal{L}(Z)] = \mathrm{id}(1+\psi)\circ X$. We refer the reader to Černý and Ruf [3] for a detailed treatment of measure changes by means of non-negative, represented, multiplicatively compensated semimartingales and once again to [5] for specific applications.

The suggested calculus has one other benefit for applied stochastic modelling. In an applied setting it is impractical to work with the raw characteristics
$$\left(B^{X[h]}, [\hat{X},\hat{X}]^c, \nu^X\right).$$
This issue can be addressed by decomposing the process X uniquely into a 'discrete-time' component X^{dp} involving only jumps at predictable times and a 'continuous-time' part X^{qc}, see Proposition 2.10. When it comes to computing drifts, the jumps at predictable

times τ can be treated separately via the natural formula

$$\Delta B_\tau^{\xi\circ X} = \mathsf{E}_{\tau_-}[\xi_\tau(\Delta X_\tau)].$$

The remaining quasi-left-continuous part X^{qc} is usually an Itô semimartingale in applications, i.e., the characteristics of X^{qc} are assumed to be absolutely continuous with respect to time. One may then rephrase the drift computation for this component in terms of time rates, reverting to drift rates, quadratic variation rates (squared volatilities), and jump intensities (Lévy measures). Thus, the calculus naturally accommodates the two most common ways of specifying the underlying stochastic process X (discrete time vs. an Itô semimartingale) and even allows them to be combined in intricate ways, see [3, Example 4.5].

We shall close by mentioning possible directions for future research. As for extensions of the classes $\mathfrak{U}$ and $\mathfrak{I}(X)$, the most immediate generalization concerns the level of smoothness of the representing function at the origin. Lack of differentiability is associated with the need to consider local times in the Itô–Meyer formula; see Karatzas and Shreve [22, Theorem 3.6.22]. This suggests an appropriate modification of the Émery formula (3.3), for which the three key operations would have to be checked again. In Example 4.5, we have broached the subject of the Itô–Wentzell formula that we believe merits further investigation.

References

[1] P. Bank and D. Baum, *Hedging and portfolio optimization in financial markets with a large trader*, Math. Finance **14** (2004), no. 1, 1–18. MR2030833

[2] P. Carr and R. Lee, *Variation and share-weighted variation swaps on time-changed Lévy processes*, Finance Stoch. **17** (2013), no. 4, 685–716. MR3105930

[3] A. Černý and J. Ruf, *Simplified calculus for semimartingales: Multiplicative compensators and changes of measure*, arXiv:2006.12765, 2020.

[4] A. Černý and J. Ruf, *Pure-jump semimartingales*, Bernoulli **27** (2021), no. 4, 2624–2648. MR4303898

[5] A. Černý and J. Ruf, *Simplified stochastic calculus with applications in Economics and Finance*, European J. Oper. Res. **293** (2021), no. 2, 547–560. MR4241583

[6] A. Černý and J. Ruf, *On predictable variations of a semimartingale*, Working paper, 2021.

[7] C. Dellacherie and P.-A. Meyer, *Probabilities and Potential. B*, North-Holland Mathematics Studies, vol. 72, North-Holland, Amsterdam, 1982. MR0745449

[8] C. Doléans, *Variation quadratique des martingales continues à droite*, Ann. Math. Statist. **40** (1969), 284–289. MR0236982

[9] C. Doléans-Dade, *Quelques applications de la formule de changement de variables pour les semimartingales*, Z. Wahrscheinlichkeitstheorie und verw. Gebiete **16** (1970), 181–194. MR0283883

[10] C. Doléans-Dade and P.-A. Meyer, *Intégrales stochastiques par rapport aux martingales locales*, Séminaire de Probabilités IV, Strasbourg, Lecture Notes in Math., vol. 124, Springer, Berlin, 1970, pp. 77–107. MR0270425

[11] D. Duffie, D. Filipović, and W. Schachermayer, *Affine processes and applications in finance*, Ann. Appl. Probab. **13** (2003), no. 3, 984–1053. MR1994043

[12] E. Eberlein, A. Papapantoleon, and A. N. Shiryaev, *Esscher transform and the duality principle for multidimensional semimartingales*, Ann. Appl. Probab. **19** (2009), no. 5, 1944–1971. MR2569813

[13] M. Émery, *Stabilité des solutions des équations différentielles stochastiques application aux intégrales multiplicatives stochastiques*, Z. Wahrscheinlichkeitstheorie und verw. Gebiete **41** (1978), no. 3, 241–262. MR0464400

[14] M. Émery, *Stochastic Calculus in Manifolds*, Universitext, Springer-Verlag, Berlin, 1989, With an appendix by P.-A. Meyer. MR1030543

[15] R. K. Getoor and M. J. Sharpe, *Conformal martingales*, Invent. Math. **16** (1972), 271–308. MR0305473

[16] T. Goll and J. Kallsen, *Optimal portfolios for logarithmic utility*, Stochastic Process. Appl. **89** (2000), no. 1, 31–48. MR1775225

[17] J. Jacod, *Asymptotic properties of realized power variations and related functionals of semimartingales*, Stochastic Process. Appl. **118** (2008), no. 4, 517–559. MR2394762

[18] J. Jacod and A. N. Shiryaev, *Limit Theorems for Stochastic Processes*, 2nd ed., Comprehensive Studies in Mathematics, vol. 288, Springer, Berlin, 2003. MR1943877

[19] M. Jeanblanc, M. Yor, and M. Chesney, *Mathematical Methods for Financial Markets*, Springer Finance, Springer-Verlag London, Ltd., London, 2009. MR2568861

[20] J. Kallsen, *Optimal portfolios for exponential Lévy processes*, Math. Methods Oper. Res. **51** (2000), no. 3, 357–374. MR1778648

[21] J. Kallsen, *σ–localization and σ–martingales*, Theory Probab. Appl. **48** (2004), no. 1, 152–163. MR2013413

[22] I. Karatzas and S. E. Shreve, *Brownian Motion and Stochastic Calculus*, 2nd ed., Graduate Texts in Mathematics, vol. 113, Springer-Verlag, New York, 1991. MR1121940

[23] M. Larsson and J. Ruf, *Stochastic exponentials and logarithms on stochastic intervals—A survey*, J. Math. Anal. Appl. **476** (2019), no. 1, 2–12. MR3944415

[24] D. Lépingle, *La variation d'ordre p des semi-martingales*, Z. Wahrscheinlichkeitstheorie und verw. Gebiete **36** (1976), no. 4, 295–316. MR0420837

[25] H. P. McKean, Jr., *Stochastic Integrals*, Probability and Mathematical Statistics, No. 5, Academic Press, New York, 1969. MR0247684

[26] J. Mémin, *Décompositions multiplicatives de semimartingales exponentielles et applications*, Séminaire de Probabilités XII, Strasbourg, Lecture Notes in Math., vol. 649, Springer, Berlin, 1978, pp. 35–46. MR0519991

[27] I. Monroe, *On the γ–variation of processes with stationary independent increments*, Ann. Math. Statist. **43** (1972), 1213–1220. MR0312578

[28] I. Monroe, *Almost sure convergence of the quadratic variation of martingales: a counterexample*, Ann. Probab. **4** (1976), no. 1, 133–138. MR0400384

[29] N. Perkowski and J. Ruf, *Supermartingales as Radon-Nikodym densities and related measure extensions*, Ann. Probab. **43** (2015), no. 6, 3133–3176. MR3433578

[30] P. E. Protter, *Stochastic Integration and Differential Equations*, 2nd ed., Stochastic Modelling and Applied Probability, vol. 21, Springer, Berlin, 2005. MR2273672

[31] R. Remmert, *Theory of Complex Functions*, Graduate Texts in Mathematics, vol. 122, Springer-Verlag, New York, 1991, Translated from the second German edition by Robert B. Burckel, Readings in Mathematics. MR1084167

[32] D. Revuz and M. Yor, *Continuous Martingales and Brownian Motion*, Grundlehren der mathematischen Wissenschaften, vol. 293, Springer-Verlag, Berlin, 1991. MR1083357

[33] A. N. Shiryaev, *Probability*, 2nd ed., Graduate Texts in Mathematics, vol. 95, Springer, New York, 1996, Translated from the first (1980) Russian edition by R. P. Boas. MR1368405

[34] A. N. Shiryaev and A. S. Cherny, *A vector stochastic integral and the fundamental theorem of asset pricing*, Proc. Steklov Inst. Math. **237** (2002), 6–49. MR1975582

[35] W. Wirtinger, *Zur formalen Theorie der Funktionen von mehr komplexen Veränderlichen*, Math. Ann. **97** (1927), no. 1, 357–375. MR1512366

Acknowledgments. We thank Jan Kallsen, Christoph Kühn, Johannes Muhle-Karbe, Pietro Siorpaes, two anonymous referees, and an associate editor for helpful comments and suggestions.