Electronic Journal of Probability

Electron. J. Probab. **28** (2023), article no. 89, 1–41.
ISSN: 1083-6489 https://doi.org/10.1214/23-EJP986

Itô–Föllmer calculus in Banach spaces I: the Itô formula*

Yuki Hirai†

Abstract

We prove Föllmer's pathwise Itô formula for a Banach space-valued càdlàg path. We also relax the assumption on the sequence of partitions along which we treat the quadratic variation of a path.

Keywords: Itô–Föllmer integral, quadratic variation, Itô formula.
MSC2020 subject classifications: Primary 60H99, Secondary 60H05.
Submitted to EJP on June 7, 2022, final version accepted on June 27, 2023.

Contents

*This work was partially supported by JSPS KAKENHI Grant Number JP19H00643 (Masanori Hino).

†Department of Creative Engineering, National Institute of Technology, Tsuruoka College, Japan. E-mail: hirai@tsuruoka-nct.ac.jp

1 Introduction

In his seminal paper [22], Föllmer presented a new perspective on Itô's stochastic calculus. The main theorem of Föllmer [22] states that a deterministic càdlàg path satisfies the Itô formula provided it has quadratic variation along a given sequence of partitions. This theorem enables us to construct the Itô integral $\int_0^t f(X_{s-})\mathrm{d}X_s$ for a sufficiently nice function f and a path X with a quadratic variation. This suggests the possibility of developing an analogue of the Itô calculus in completely analytic, probability-free situations. We call this framework Föllmer's pathwise Itô calculus or, more simply, the Itô–Föllmer calculus. It can be regarded as a deterministic counterpart of the classical Itô calculus.

Recently, the Itô–Föllmer calculus has been receiving increasing attention from the viewpoint of its financial applications. It is regarded as a useful tool to study financial theory under probability-free settings and has been used to construct financial strategies in a strictly pathwise manner (see, e.g., Föllmer and Schied [24], Schied [67], Davis, Obłój and Raval [13], and Schied, Speiser, and Voloshchenko [69]). We expect that the Itô–Föllmer calculus will have a growing presence in financial applications.

The Itô–Föllmer calculus can be applied to a stochastic process having quadratic variation. A standard example of such a process is a semimartingale. However, it is known that the class of processes possessing quadratic variation is strictly larger than that of semimartingales (see, e.g., Föllmer [22, 23]). In this sense, the Itô–Föllmer calculus enables us to extend stochastic integration theory beyond semimartingales.

There are several approaches to the pathwise construction of stochastic integration. First, we mention classic studies by Bichteler [3], Karandikar [38], and Willinger and Taqqu [78, 79], but see also Nutz [53] and Łochowski [40, 43]. The theory of Vovk's outer measure and typical paths was pioneered by Vovk [73, 74, 75, 76, 77] and further developed by several authors, including Perkowski and Prömel [56, 55], Łochowski, Perkowski, and Prömel [42], and Bartl, Kupper, and Neufeld [2]. Russo and Vallois [58, 59, 60, 61, 62, 63, 64] developed a theory called stochastic calculus via regularization. The rough path theory, pioneered by Lyons [44], and its generalization have become important in stochastic calculus and its applications. In addition, we refer to Gubinelli [29], Gubinelli and Tindel [30], Friz and Shekhar [26], and Friz and Zhang [27]. Some studies have investigated the relation between the Itô–Föllmer calculus and rough path theory (see, e.g., Perkowski and Prömel [56], and Friz and Hairer [25]).

Among the various pathwise methods, we consider Föllmer's approach to be the simplest and the most intuitively clear. It needs only elementary arguments to establish calculation rules such as Itô's formula within this framework. Moreover, the Itô–Föllmer calculus requires only a minimal assumption that the integrator has quadratic variation. We believe that these are advantages of Föllmer's theory, and also that careful observation of this theory helps us to understand the path properties of processes better when we consider semimartingales and their stochastic integration.

Increasingly many works related to Itô–Föllmer calculus have been appearing recently. First, we refer to Sondermann [70], Schied [67], Hirai [35], and Cont and Perkowski [11]. Schied [68], Mishura and Schied [52], and Cont and Das [7] construct deterministic continuous paths with nontrivial quadratic variations. See also Chiu and Cont [6]. Functional extensions of the Itô–Föllmer calculus have been developed by Dupire [21], Cont and Fournié [8, 9, 10], and Ananova and Cont [1], for example. Extension of the Itô–Föllmer calculus in terms of local times has been investigated in Davis, Obłój, and Raval [13], Davis, Obłój, and Siorpaes [14], Łochowski et al. [41], and Hirai [34, 36].

To our knowledge, however, the Itô–Föllmer calculus in an infinite dimensional setting has not yet been sufficiently studied. Stochastic integration in infinite dimensions naturally appears when we treat stochastic partial differential equations (see, e.g. Da Prato and Zabczyk [12]). These have played an important role in modelling term structures of interest rates or forward variances in mathematical finance, and also in models of statistical mechanics and quantum field theories. Then we aim to extend Föllmer's theory to Banach space-valued paths. In this paper, we prove the Itô formula for a path in a Banach space with a suitably defined quadratic variation. We will study relations between various quadratic variations and prove some transformation formulae for quadratic variations in our second paper in this series [37]. We not only generalize the state space of paths but also relax the assumption on the sequence of partitions along which we consider the quadratic variation. In the context of the Itô–Föllmer calculus, two types of assumptions about a sequence of partitions are frequently used. One is $|\pi_n| \to 0$, as used in Föllmer [22], and the other is the condition $O_t^-(X;\pi_n) \to 0$, which is used in many papers handling continuous paths and some dealing with discontinuous paths such as Vovk [76]. In this paper, we introduce new conditions to a sequence of partitions and a càdlàg path (Definition 3.4), which gives a unified approach.

There have been many attempts to extend classical stochastic calculus to Banach or Hilbert space-valued processes. Examples include Kunita [39], Metivier [45], Pellaumail [54], Yor [81], Gravereaux and Pellaumail [28], Metivier and Pistone [48], Meyer [49], Metivier and Pellaumail [47], Gyöngy and Krylov [32, 33], Gyöngy [31], Metivier [46], Pratelli [57], Brooks and Dinculeanu [5], Mikulevicius and Rozovskii [50, 51], Dinculeanu [20], De Donno and Pratelli [15], van Neerven, Veraar, and Weis [71], Veraar and Yaroslavtsev [72], and Yaroslavtsev [80]. Note that Di Girolami, Fabbri, and Russo [17] treat quadratic covariation of Banach space-valued processes within the framework of stochastic calculus via regularization, with Föllmer's calculus in mind.

Our method can be interpreted as a deterministic counterpart of these stochastic integration theories in Banach spaces. Some of the works listed above, such as Metivier and Pellaumail [47] and Dinculeanu [20], give proofs of Itô's formula in a similar manner to Föllmer's calculus. One of the advantages of our approach appears in the statement of the Itô formula. For a function f to satisfy the Itô formula, we require f to be just C^2 class, while a stochastic approach needs some additional assumptions about the boundedness of f and its derivatives.

Before explaining our contribution, we begin by summarizing the main result of Föllmer [22]. Let $\Pi = (\pi_n)_{n\in\mathbb{N}}$ be a sequence of partitions of $\mathbb{R}_{\geq 0}$ such that $|\pi_n| \coloneqq \sup_{]r,s]\in\pi_n}|s-r|$ tends to 0 as $n\to\infty$. We say that a càdlàg path $X\colon \mathbb{R}_{\geq 0}\to\mathbb{R}$ has quadratic variation along Π if there exists a càdlàg increasing function $[X,X]$ such that for all $t\in\mathbb{R}_{\geq 0}$

(i) $\sum_{]r,s]\in\pi_n}(X_{s\wedge t}-X_{r\wedge t})^2$ converges to $[X,X]_t$ as $n\to\infty$, and

(ii) $\Delta[X,X]_t = (\Delta X_t)^2$.

An $\mathbb{R}^d$-valued càdlàg path $X=(X^1,\dots,X^d)$ has quadratic variation along Π if the real-valued path X^i+X^j has quadratic variation along the same sequence for each i and j. Föllmer [22] proved that if X has quadratic variation, then for any $f\in C^2(\mathbb{R}^d)$ the path

$t \mapsto f(X_t)$ satisfies Itô's formula. That is,

$$f(X_t) = f(X_0) + \int_0^t \langle Df(X_{s-}), \mathrm{d}X_s \rangle + \frac{1}{2} \sum_{1 \leq i,j \leq d} \int_0^t \frac{\partial^2 f}{\partial x_i \partial x_j}(X_{s-}) \mathrm{d}[X^i, X^j]^{\mathrm{c}}_s$$
$$+ \sum_{0 < s \leq t} \{\Delta f(X_s) - \langle Df(X_{s-}), \Delta X_s \rangle\} \tag{1.1}$$

holds for all $t \in \mathbb{R}_{\geq 0}$. The first term on the right-hand side of (1.1) is defined as the limit

$$\int_0^t \langle Df(X_{s-}), \mathrm{d}X_s \rangle = \lim_{n \to \infty} \sum_{]r,s] \in \pi_n} \langle Df(X_r), X_{s \wedge t} - X_{r \wedge t} \rangle,$$

where $\langle \ , \ \rangle$ denotes the usual Euclidean inner product. We call this limit the Itô–Föllmer integral along Π. Föllmer's theorem claims that if X has quadratic variation along Π, then the Itô–Föllmer integral above exists and it satisfies equation (1.1).

As stated above, we aim to extend Föllmer's pathwise Itô formula to Banach space-valued paths. Let us describe a simplified version of our main result. The precise statement will be given as Theorem 3.6 and Corollary 3.7. Let E be a Banach space and $E \widehat{\otimes}_\alpha E$ be the tensor product Banach space with respect to a reasonable crossnorm α. We say that an E-valued càdlàg path X has strong/weak α-tensor quadratic variation along $\Pi = (\pi_n)$ if there is a càdlàg path ${}^\alpha[X,X] \colon \mathbb{R}_{\geq 0} \to E \widehat{\otimes}_\alpha E$ of finite variation such that, for all $t \geq 0$,

(i) the sequence $\sum_{]r,s] \in \pi_n} (X_{s \wedge t} - X_{r \wedge t})^{\otimes 2}$ converges to ${}^\alpha[X,X]_t$ in the norm/weak topology of $E \widehat{\otimes}_\alpha E$, and
(ii) the equation $\Delta {}^\alpha[X,Y]_t = \Delta X_t^{\otimes 2}$ holds.

The path X has finite 2-variation along Π if

$$V^2(X;\Pi)_t \coloneqq \sup_{n \in \mathbb{N}} \sum_{]r,s] \in \pi_n} \|X_{s \wedge t} - X_{r \wedge t}\|^2 < \infty$$

for all $t \geq 0$. We say that a sequence of partitions (π_n) satisfies Condition (C) for a path $X \colon [0, \infty[\to E$ in a Banach space if it satisfies Conditions (C1)–(C3) of Definition 3.4. Roughly speaking, Conditions (C1) and (C2) state that (π_n) reconstructs the information of the jumps of X. Condition (C3) means that (π_n) controls the oscillation of X in some sense. Under these settings, our main result, the Itô formula, is stated as follows. Let $X \colon \mathbb{R}_{\geq 0} \to E$ be a càdlàg path that has strong/weak α-tensor quadratic variation and finite 2-variation along $(\pi_n)_{n \in \mathbb{N}}$, and let $A \colon \mathbb{R}_{\geq 0} \to F$ be a càdlàg path of finite variation in a Banach space. Suppose that (π_n) satisfies Condition (C) for (A, X) and the left-side discretization of (A, X) along (π_n) approximates (A_-, X_-) pointwise (see Definition 3.1 for the exact definition). If $f \colon F \times E \to G$ is a $C^{1,2}$ function such that the second derivative induces a continuous map $D_x^2 f \colon F \times E \to \mathcal{L}(E \widehat{\otimes}_\alpha E, G)$, then the composite function $f(A, X)$ satisfies

$$\begin{aligned} f(A_t, X_t) &- f(A_0, X_0) \\ &= \int_0^t D_a f(A_{s-}, X_{s-}) \mathrm{d}A^{\mathrm{c}}_s + \int_0^t D_x f(A_{s-}, X_{s-}) \mathrm{d}X_s \\ &\quad + \frac{1}{2} \int_0^t D_x^2 f(A_{s-}, X_{s-}) \mathrm{d}^\alpha[X,X]^{\mathrm{c}}_s + \sum_{0 < s \leq t} \{\Delta f(A_s, X_s) - D_x f(A_{s-}, X_{s-}) \Delta X_s\}. \end{aligned}$$

The second integral on the right-hand side is defined respectively as the strong/weak limit of left-side Riemannian sums along (π_n).

To conclude this section, we give an outline of the remainder of this paper. Section 2 is a preliminary part of this article. We introduce basic notation and terminology in the first subsection. The next subsection is devoted to a review of càdlàg paths and Stieltjes integrals in Banach spaces. In Section 3, we set up basic notions in Itô–Föllmer calculus in Banach spaces and state the main results of the paper (Theorem 3.6 and Corollary 3.7). In Section 4, we study conditions on the sequence of partitions and the relation between them and càdlàg paths. Fundamental properties of quadratic variations are studied in Section 5. The purpose of Section 6 is to show Lemma 3.10, which is essentially used in the proof of the main theorem. In the last section of the main part, Section 7, we prove the Itô formula for a Banach space-valued path having quadratic variation. In appendices, we present some auxiliary results related to differential calculus and integration in Banach spaces.

2 Preliminaries

2.1 Notations and terminologies

In this section, we introduce basic notation and terminology used throughout this article.

The symbol $\mathbb{N}$ denotes the set of natural numbers $\{0, 1, 2, \dots\}$ and $\mathbb{R}$ denotes that of real numbers. If A is a subset of $\mathbb{R}$ and $a \in \mathbb{R}$, we define $A_{\geq a} = \{x \in A \mid x \geq a\}$.

If E and F are two real Banach spaces, $\mathcal{L}(E, F)$ denotes the space of bounded linear maps from E to F. In addition, given another Banach space G, we define $\mathcal{L}^{(2)}(E, F; G)$ as the space of bounded bilinear maps from $E \times F$ to G. Recall that $\mathcal{L}(E, F)$ and $\mathcal{L}^{(2)}(E, F; G)$ are Banach spaces with norms

$$\|L\| = \sup_{\|x\| \neq 0} \frac{\|Lx\|}{\|x\|}, \qquad \|B\| = \sup_{\|x\|, \|y\| \neq 0} \frac{\|B(x, y)\|}{\|x\| \|y\|},$$

respectively.

Now we introduce another topology on the space $\mathcal{L}(E, G)$. Let $\mathcal{K}_E$ be the family of all compact subsets of E. For each $K \in \mathcal{K}_E$, define a seminorm ρ_K by the formula

$$\rho_K(L) = \inf\{C > 0 \mid \forall x \in K,\ \|Lx\|_F \leq C\|x\|_E\} \tag{2.1}$$

for each $L \in \mathcal{L}(E, F)$. Then the family $(\rho_K)_{K \in \mathcal{K}_E}$ induces a locally convex Hausdorff topology on $\mathcal{L}(E, F)$. We use the symbol $\mathcal{L}_{\mathrm{c}}(E, F)$ for this topological vector space.

Let $[0, \infty[\, = \mathbb{R}_{\geq 0} = \{r \in \mathbb{R} \mid r \geq 0\}$ and let E be a Hausdorff topological vector space. A *càdlàg* path in E is a function $X \colon \mathbb{R}_{\geq 0} \to E$ that is right continuous at every $t \geq 0$ and has a left limit at every $t > 0$. The terms *RCLL* and *right-regular* are also used to stand for the same property. Similarly, a *càglàd* (also called *LCRL* or *left-regular*) path in E is a function $X \colon \mathbb{R}_{\geq 0} \to E$ that is left continuous on $]0, \infty[$ and has right limits on $[0, \infty[$. The symbols $D([0, \infty[, E)$ and $D(\mathbb{R}_{\geq 0}, E)$ denote the set of all càdlàg paths in E. If X is an element of $D(\mathbb{R}_{\geq 0}, E)$, we define

$$X(t-) = \lim_{s \uparrow\uparrow t} X(s) = \lim_{s \to t, s < t} X_s, \qquad \Delta X(t) = X(t) - X(t-).$$

We also use X_t, X_{t-}, and ΔX_t to indicate the values $X(t)$, $X(t-)$, and $\Delta X(t)$, respectively. Next, set

$$\begin{aligned} D(X) &= \{t \in \mathbb{R}_{\geq 0} \mid \|\Delta X_t\| \neq 0\}, \\ D_\varepsilon(X) &= \{t \in \mathbb{R}_{\geq 0} \mid \|\Delta X_t\| \geq \varepsilon\}, \\ D^\varepsilon(X) = D(X) \setminus D_\varepsilon(X) &= \{t \in \mathbb{R}_{\geq 0} \mid 0 < \|\Delta X_t\| < \varepsilon\}. \end{aligned}$$

We simply write D, D_ε, and D^ε if there is no ambiguity. Given a discrete set $D \subset [0,\infty[$ and a càdlàg path X, we define

$$J_D(X)_t = J(D;X)_t = \sum_{0<s\leq t} \Delta X_s 1_D(s).$$

Recall that a subset D is discrete if it is a discrete topological subspace of $[0,\infty[$; i.e. every element of D is an isolated point with respect to the subspace topology. By assumption, the set $D \cap [0,t]$ is finite for all t, and therefore the summation above is well-defined. Then $J_D(X)$ is a càdlàg path of finite variation. For abbreviation, we often write $J_\varepsilon(X)$ instead of $J(D_\varepsilon(X);X)$.

Throughout this paper, the term *partition of* $\mathbb{R}_{\geq 0}$ always means the set of intervals of the form $\pi = \{]t_i, t_{i+1}]; i \in \mathbb{N}\}$ which satisfies $0 = t_0 < t_1 < \cdots \to \infty$. The set of all partitions of $\mathbb{R}_{\geq 0}$ is denoted by $\mathrm{Par}(\mathbb{R}_{\geq 0})$ or $\mathrm{Par}([0,\infty[)$. Similarly, given a compact interval $[a,b] \subset \mathbb{R}_{\geq 0}$, let $\mathrm{Par}([a,b])$ be the set of all finite partitions of the form $\pi = \{]t_i, t_{i+1}]; 0 \leq i \leq n-1\}$ with $a = t_0 < t_1 < \cdots < t_n = b$. For a partition π of $\mathbb{R}_{\geq 0}$ or a compact interval, we define $\pi^{\mathrm{p}} \subset \mathbb{R}_{\geq 0}$ as the set of all endpoints of elements of π. If $\pi = \{]t_i, t_{i+1}]; i \in I\}$, then $\pi^{\mathrm{p}} = \{t_i, t_{i+1}; i \in I\}$.

2.2 Remarks on càdlàg paths and Stieltjes integration

In this subsection, we review some basic properties of càdlàg paths that will be referred to later.

Let E be a Banach space. E-valued right continuous and left continuous step functions are functions of the form

$$\sum_{i\in\mathbb{N}} 1_{[t_i,t_{i+1}[} a_i, \qquad 1_{\{0\}} b_0 + \sum_{i\in\mathbb{N}} 1_{]t_i,t_{i+1}]} b_{i+1},$$

respectively, where $0 = t_0 < t_1 < \cdots < t_n < \cdots \to \infty$ and a_i, $b_i \in E$ for all $i \in \mathbb{N}$. Right continuous step functions are càdlàg and left continuous step functions are càglàd. Every right continuous step function $f = \sum_{i\in\mathbb{N}} 1_{[t_i,t_{i+1}[} a_i$ is strongly $\mathcal{B}(\mathbb{R}_{\geq 0})/\mathcal{B}(E)$ measurable because it is the pointwise limit of the sequence (f_n) defined by $f_n = \sum_{0\leq i\leq n} 1_{[t_i,t_{i+1}[} a_i$.

A càdlàg path in a Banach space satisfies the following properties.

Lemma 2.1. *Let f be a càdlàg path in a Banach space E.*

- *(i) For every $C > 0$, there are only finitely many s satisfying $\|\Delta f_s\|_E > C$ in each compact interval of $[0,\infty[$.*
- *(ii) The image $f(I)$ of any compact interval $I \subset [0,\infty[$ is relatively compact in E.*
- *(iii) Suppose that every jump of f is smaller than $C > 0$ on a compact interval $I \subset [0,\infty[$. Then for all $\varepsilon > 0$, there exists a $\delta > 0$ such that $\|f(s) - f(u)\|_E < C + \varepsilon$ holds for any $s, u \in I$ satisfying $|s - u| < \delta$.*
- *(iv) The path f is the uniform limit of some sequence of right continuous step functions on each compact interval.*
- *(v) For every $t > 0$ and $\varepsilon > 0$, there is a partition $\pi \in \mathrm{Par}([0,t])$ that satisfies*

$$\omega(f,[r,s[) := \sup_{u,v\in[r,s[} \|f(u) - f(v)\| < \varepsilon$$

for all $]r,s] \in \pi$.

One can find an analogue of Lemma 2.1 for càdlàg paths in an arbitrary separable complete metric space in Billingsley [4, 122].

Next, recall that a function $f\colon [0,\infty[\to E$ is of bounded variation on a compact interval $I \subset [0,\infty[$ if

$$V(f;I) \coloneqq \sup_{\pi\in\operatorname{Par} I} \sum_{]r,s]\in\pi} \|f(s)-f(r)\|_E < \infty.$$

For convenience, set $V(f;\emptyset) = 0$ and $V(f;[a,a]) = 0$ for every $a \in [0,\infty[$. The function f has finite variation if it has bounded variation on every compact subinterval of $[0,\infty[$. The set of all càdlàg paths of finite variation in E is denoted by $FV(\mathbb{R}_{\geq 0}, E)$ or $FV([0,\infty[, E)$. We define the total variation path $V(f)$ of a function f of $FV(\mathbb{R}_{\geq 0}, E)$ by $V(f)_t = V(f;[0,t])$. Then $V(f)$ is increasing and satisfies $V(f)_0 = 0$.

We list several basic properties of a path of finite variation below. See Dinculeanu [20, §18] for a proof.

Lemma 2.2. *Let $f\colon [0,\infty[\to E$ be a càdlàg path of finite variation in a Banach space.*

(i) If $a,b,c \in [0,\infty[$ satisfy $a \leq b \leq c$, then

$$V(f,[a,c]) = V(f,[a,b]) + V(f,[b,c]).$$

(ii) The total variation path $V(f)$ is càdlàg.

(iii) The jump of $V(f)$ at $t \geq 0$ is given by $\Delta V(f)(t) = \|\Delta f(t)\|_E$.

(iv) The family $(\Delta f(s))_{s\in[0,t]}$ is absolutely summable for all $t \geq 0$; i.e.

$$\sup\left\{ \sum_{s\in F} \|\Delta f(s)\|_E \;\middle|\; F\text{: a finite subset of } [0,t] \right\} < \infty.$$

(v) The function f^{d} defined by

$$f^{\mathrm{d}}(t) = \sum_{0<s\leq t} \Delta f(s).$$

is again a càdlàg path of finite variation.

Note that the summation in (v) of Lemma 2.2 is defined in the following manner. Let D be the set of all finite subsets of $]0,t]$. We regard D as a directed set with the order defined by inclusion. Then the net $(\sum_{s\in d} \Delta f(s))_{d\in D}$ converges in E by Condition (iv) of Lemma 2.2, and hence we can define

$$\sum_{0<s\leq t} \Delta f(s) = \lim_{d} \sum_{s\in d} \Delta f(s).$$

The function f^{d} defined in Proposition 2.2 is called the discontinuous part of f. We also define $f^{\mathrm{c}} = f - f^{\mathrm{d}}$ and call it the continuous part of f.

Let $\mathcal{I}$ the semiring of subsets of $\mathbb{R}_{\geq 0}$ consisting of all bounded intervals of the form $]a,b]$ and the singleton $\{0\}$. Given an $f \in D(\mathbb{R}_{\geq 0}, E)$, define

$$\mu_f(]a,b]) = f(b) - f(a), \quad \mu_f(\{0\}) = f(0)$$

for any two real numbers satisfying $0 \leq a \leq b$. If f has finite variation, the function $\mu_f\colon \mathcal{I} \to E$ can be uniquely extended to a σ-additive measure defined on the δ-ring $\mathcal{D}$ generated by $\mathcal{I}$. Refer to Theorem 18.19 of Dinculeanu [20, 208] for a proof. Notice that this correspondence between a function and a measure satisfies $(\mu_f)^{\mathrm{d}} = \mu_{f^{\mathrm{d}}}$ and $(\mu_f)^{\mathrm{c}} = \mu_{f^{\mathrm{c}}}$. See Appendix B.1 for the definition of the measures $(\mu_f)^{\mathrm{d}}$ and $(\mu_f)^{\mathrm{c}}$.

Because there is a measure μ_f associated with f, we can consider the Stieltjes integral with respect to f. Let $B\colon F \times E \to G$ be a continuous bilinear map between

Banach spaces. Set $L^1_{\mathrm{loc}}(\mu_f;F) = L^1_{\mathrm{loc}}(|\mu_f|;F)$, where $|\mu_f|$ denotes the variation measure of μ_f introduced in Appendix B.1. Then for each $g \in L^1_{\mathrm{loc}}(\mu_f;F)$ and a bounded interval $I =]a,b]$, the Stieltjes integral is defined by the formula

$$\int_a^b B(g(s), \mathrm{d}f(s)) = \int_I B(g(s), \mathrm{d}f(s)) := \int_I B(g(s), \mu_f(\mathrm{d}s)).$$

See Appendix B.1 for the definition of $\int_I B(g(s), \mu_f(\mathrm{d}s))$. If $g\colon \mathbb{R}_{\geq 0} \to \mathcal{L}_{\mathrm{c}}(E,G)$ is a càglàd path, we can also define the Stieltjes integral as

$$\int_I g(s)\,\mathrm{d}f(s) = \int_I g(s)\,\mu_f(\mathrm{d}s),$$

where the integral on the right-hand side is constructed in Appendix B.2. Finally, note that the decomposition $f = f^{\mathrm{c}} + f^{\mathrm{d}}$ gives a decomposition of the integral

$$\int_I B(g(s), \mathrm{d}f(s)) = \int_I B(g(s), \mathrm{d}f^{\mathrm{c}}(s)) + \int_I B(g(s), \mathrm{d}f^{\mathrm{d}}(s)).$$

3 Settings and the main result

In this section, we introduce the main results of this paper, namely, the Itô formula within the framework of the Itô–Föllmer calculus in Banach spaces. The statement is to be found in Theorem 3.6 and Corollary 3.7.

First, we introduce some notations about difference operators. Given a function $X\colon \mathbb{R}_{\geq 0} \to E$ and $t \geq 0$, define functions δX and δX_t on $\mathcal{I}$ into E by the formulae

$$\delta X(I) = \delta X(]r,s]) = X_s - X_r, \qquad \delta X_t(I) = \delta X(I \cap [0,t]) = X_{s\wedge t} - X_{r\wedge t}$$

for each $I =]r,s] \in \mathcal{I}$. We also define $\delta X(\{0\}) = \delta X_t(\{0\}) = X_0$. Moreover, consider a bilinear map $B\colon F \times E \to G$ in Banach spaces and another function $Y\colon \mathbb{R}_{\geq 0} \to F$. Then we define functions $B(Y, \delta X)$ and $B(Y, \delta X_t)$ from $\mathcal{I}$ to G by the formulae

$$B(Y,\delta X)(]r,s]) = B(Y_r, \delta X(]r,s])), \qquad B(Y,\delta X_t)(]r,s]) = B(Y_r, \delta X(]r,s] \cap [0,t]))$$

for $I =]r,s] \in \mathcal{I}$ and

$$B(Y,\delta X)(\{0\}) = B(Y,\delta X_t)(\{0\}) = B(Y_0, X_0).$$

By using this notation, the left-side Riemannian sum along a partition π is expressed as

$$\sum_{]r,s]\in\pi} B(Y_r, X_{t\wedge s} - X_{t\wedge r}) = \sum_{]r,s]\in\pi} B(Y_{t\wedge r}, X_{t\wedge s} - X_{t\wedge r}) = \sum_{I\in\pi} B(Y, \delta X_t)(I).$$

We also define $B(\delta Y, X)$ and $B(\delta Y_t, X)$ in a similar manner. The function $B(\delta X_t, \delta Y_t)\colon \mathcal{I} \to G$ is defined as the composition of B and $(\delta X_t, \delta Y_t)\colon \mathcal{I} \to F \times E$.

Definition 3.1. *Let E, F, and G be Banach spaces and let $B\colon E \times F \to G$ be a bounded bilinear map. Suppose that $\Pi = (\pi_n)_{n\in\mathbb{N}}$ is a sequence of partitions of $\mathbb{R}_{\geq 0}$.*

(i) A path $(X,Y) \in D(\mathbb{R}_{\geq 0}, E \times F)$ has B-quadratic covariation *along Π if there exists a G-valued càdlàg path $Q_B(X,Y)$ of finite variation satisfying the following conditions:*

(a) for all $t \in \mathbb{R}_{\geq 0}$

$$\lim_{n\to\infty} \sum_{I\in\pi_n} B(\delta X_t, \delta Y_t)(I) = Q_B(X,Y)_t$$

in the norm topology of G;

(b) for all $t \in \mathbb{R}_{\geq 0}$, the jump of $Q_B(X,Y)$ is given by

$$\Delta Q_B(X,Y)_t = B(\Delta X_t, \Delta Y_t).$$

Then the path $Q_B(X,Y)$ is called the B-quadratic covariation of X and Y. If $E = F$ and $X = Y$, we call $Q_B(X,X)$ the B-quadratic variation of X.

(ii) If the convergence of (i)-(a) holds in the weak topology of G, we say that (X,Y) has the weak B-quadratic covariation $Q_B(X,Y)$.

We often call a B-quadratic covariation a *strong* B-quadratic covariation if we stress that the convergence holds in the norm topology. The quadratic covariation $Q_B(X,Y)$ depends on the sequence of partitions Π. Given a partition π, we often write

$$Q_B^{\pi}(X,Y)_t = \sum_{I \in \pi} B(\delta X_t, \delta Y_t)(I).$$

Because X and Y are càdlàg, the map $t \mapsto Q_B^{\pi}(X,Y)_t$ is also càdlàg. It is not, however, of finite variation unless X and Y are of finite variation. With this notation, we can say that the strong/weak B-quadratic covariation $Q_B(X,Y)$ is the pointwise limit of $(Q_B^{\pi_n}(X,Y))_{n \in \mathbb{N}}$ respectively in the norm/weak topology satisfying the jump condition (i)-(b) of Definition 3.1.

An important class of a bounded bilinear map B is the canonical bilinear map into a tensor product of Banach spaces. Let α be a norm on the algebraic tensor product $E \otimes F$ of two Banach spaces. The norm α is called a reasonable crossnorm if it satisfies the following two conditions:

(i) the inequality $\alpha(x \otimes y) \leq \|x\|\|y\|$ holds for all $x \in E$ and $y \in F$;

(ii) the inequality $\|x^* \otimes y^*\|_{(E \otimes F, \alpha)^*} \leq \|x^*\|\|y^*\|$ holds for all $x^* \in E^*$ and $y^* \in F^*$, where $\| \ \|_{(E \otimes F, \alpha)^*}$ denotes the usual operator norm on the normed space $(E \otimes F, \alpha)$.

The completion of the normed space $(E \otimes F, \alpha)$, which is generally incomplete, is denoted by $E \widehat{\otimes}_{\alpha} F$. See Diestel and Uhl [18] and Ryan [65] for basic facts about tensor products of Banach spaces. The quadratic covariation of (X,Y) with respect to the canonical bilinear map $\otimes \colon E \times F \to E \widehat{\otimes}_{\alpha} F$ is denoted by ${}^{\alpha}[X,Y]$, and it is called the *α-tensor quadratic covariation*. We also write $[X,Y]^{\pi} = Q_{\otimes}^{\pi}(X,Y)$ and call it the discrete tensor quadratic covariation of (X,Y) along π. If $E = F$, we can consider α-tensor quadratic covariations ${}^{\alpha}[X,Y]$ and ${}^{\alpha}[Y,X]$. Although one of them exists if and only if the other does, they are not equal in general. The path ${}^{\alpha}[X,X]$ is called the *α-tensor quadratic variation*.

There are various important reasonable crossnorms in Banach space theory. The greatest crossnorm γ, also called the projective tensor norm [1], is defined by the following formula

$$\gamma(z) = \inf \left\{ \sum_{i \in I} \|x_i\|_E \|y_i\|_F \;\middle|\; I \text{ is finite, } x_i \in E \text{ and } y_i \in F \text{ for all } i \in I, \text{ and } z = \sum_{i \in I} x_i \otimes y_i \right\}.$$

We simply write $E \widehat{\otimes} F = E \widehat{\otimes}_{\gamma} F$ and call it the projective tensor product of E and F. γ-tensor quadratic variations are also called *projective tensor quadratic variations*. An advantage of the projective tensor product is that there is a canonical isometric isomorphism $\mathcal{L}(E, \mathcal{L}(E,G)) \simeq \mathcal{L}(E \widehat{\otimes} E, G) \simeq \mathcal{L}^{(2)}(E,E;G)$. We use this identification without mention. If E and F are Hilbert spaces, Hilbert–Schmidt crossnorm is also

[1] The projective tensor norm is often denoted by π. We use γ, following Diestel and Uhl [18], because the symbol π is used to indicate a partition in this article.

a natural object. For $L^p(\mu_1)$ and $L^p(\mu_2)$ defined on some measure spaces, there is a crossnorm $\triangle_p$ such that $L^p(\mu_1)\widehat{\otimes}_{\triangle_p} L^p(\mu_2) \cong L^p(\mu_1 \otimes \mu_2)$. See Defant and Floret [16, Section 7].

The tensor quadratic variation of an $\mathbb{R}^d$-valued path $X = (X^1, \dots, X^d)$ has the matrix representation

$$[X, X]_t = \begin{pmatrix} [X^1, X^1]_t & \cdots & [X^1, X^d]_t \\ \vdots & \ddots & \vdots \\ [X^d, X^1]_t & \cdots & [X^d, X^d]_t \end{pmatrix}.$$

A càdlàg path $X : \mathbb{R}_{\geq 0} \to \mathbb{R}^d$ has tensor quadratic variation if and only if it has quadratic variation in the sense of Definition 2.3 of Hirai [35].

Remark 3.2. (i) Our definition of strong tensor quadratic variations can be regarded as a pathwise analogue of tensor quadratic variations in classical stochastic integration theory in infinite dimensions such as Metivier and Pellaumail [47] and Metivier [46]. See also Dinculeanu [20]. Although one can define tensor quadratic variations in general Banach spaces, classical existence results only deal with Hilbert spaces.

(ii) Another important approach is developing recently in the context of the martingale theory in UMD Banach spaces. Yaroslavtsev [80] shows that a local martingale in a UMD Banach space has the covariation bilinear form $[\![M]\!]$. In our terminology, the covariation bilinear form corresponds to the cylindrical quadratic variation in the sense of Corollary 3.5 in Hirai [37]. By this corollary, we see that a càdlàg path with weak tensor quadratic variation has cylindrical quadratic variation.

Now we introduce a different type of quadratic variation, namely, scalar quadratic variation. Again we assume that $\Pi = (\pi_n)$ is a sequence of partitions of $\mathbb{R}_{\geq 0}$.

Definition 3.3. *Let E be a Banach space and X be an E-valued càdlàg path.*

(i) The 2-variation of X on $[0, t]$ along Π *is defined by*

$$V^2(X; \Pi)_t = \sup_{n \in \mathbb{N}} \sum_{I \in \pi_n} \|\delta X_t(I)\|^2.$$

We say the X has finite 2-variation along Π if $V^2(X; \Pi)_t < \infty$ for all $t \geq 0$.

(ii) A càdlàg path $X : \mathbb{R}_{\geq 0} \to E$ has scalar quadratic variation along Π *if there exists a real-valued càdlàg increasing path $Q(X)$ such that*

(a) for all $t \in \mathbb{R}_{\geq 0}$,

$$\sum_{I \in \pi_n} \|\delta X_t(I)\|^2 \xrightarrow[n \to \infty]{} Q(X)_t,$$

(b) for all $t \in \mathbb{R}_{\geq 0}$, the jump of $Q(X)$ at t is given by $\Delta Q(X)_t = \|\Delta X_t\|_E^2$.

We call the increasing path $Q(X)$ the scalar quadratic variation of X along (π_n).

If X has scalar quadratic variation along Π, then it has finite 2-variation along Π.

The scalar quadratic variation of a Hilbert space valued path X coincides with the bilinear quadratic variation $Q_{\langle\ ,\ \rangle}(X, X)$, where $\langle\ ,\ \rangle$ is the inner product of the state space. If a càdlàg path $X = (X^1, \dots, X^d) : \mathbb{R}_{\geq 0} \to \mathbb{R}^d$ has tensor quadratic variation along (π_n), then it has scalar quadratic variation given by

$$Q(X)_t = \mathrm{Trace}[X, X]_t = \sum_{1 \leq i \leq d} [X^i, X^i]_t.$$

This trace representation is still valid for Hilbert space-valued càdlàg paths. This result is proved in Hirai [37].

Next, we introduce some conditions on a sequence of partitions and a càdlàg path. Let $\pi \in \mathrm{Par}([0,\infty[)$ and $t \in]0,\infty[$. The symbol $\pi(t)$ denotes the element of π that contains t. By definition, there exists only one such interval. In addition, set $\overline{\pi}(t) = \sup \pi(t)$ and $\underline{\pi}(t) = \inf \pi(t)$. Then we have $\pi(t) =]\underline{\pi}(t), \overline{\pi}(t)]$.

Let $f\colon S \to E$ be a function into a Banach space and A be a subset of S. Define the oscillation of f on A by

$$\omega(f;A) = \sup_{x,y\in A} \|f(x) - f(y)\|_E.$$

Using this notation, we introduce two kinds of oscillation of a path $X \in D(\mathbb{R}_{\geq 0}, E)$ along a partition $\pi \in \mathrm{Par}(\mathbb{R}_{\geq 0})$ as follows.

$$O_t^+(X;\pi) = \sup_{]r,s]\in\pi} \omega(X;]r,s] \cap [0,t]),$$
$$O_t^-(X;\pi) = \sup_{]r,s]\in\pi} \omega(X;]r,s[\cap [0,t]) = \sup_{]r,s]\in\pi} \omega(X; [r,s[\cap [0,t]).$$

The second equality in the definition of $O_t^-(X;\pi)$ is valid because X is supposed to be right continuous. These oscillations satisfy the relation $O_t^-(X;\pi) \leq O_t^+(X;\pi)$ for all $t \geq 0$. If X is continuous, these two quantities coincide.

Definition 3.4. *Let E be a Banach space, $X \in D(\mathbb{R}_{\geq 0}, E)$, and $(\pi_n)_{n\in\mathbb{N}}$ be a sequence of partitions of $\mathbb{R}_{\geq 0}$.*

- *(i) The sequence (π_n)* satisfies Condition (C) for X *if it satisfies (C1)–(C3) below.*
  - *(C1) Let $t \in [0,\infty[$ and $\varepsilon > 0$. Then there exists an $N \in \mathbb{N}$ such that for all $n \geq N$ and for all $I \in \pi_n$, the set $I \cap [0,t] \cap D_\varepsilon(X)$ has at most one element.*
  - *(C2) Let $s \in D(X)$ and $t \in [s,\infty[$. Then*
    $$\lim_{n\to\infty} \delta X_t(\pi_n(s)) = \lim_{n\to\infty} \left\{ X(\overline{\pi_n}(s) \wedge t) - X(\underline{\pi_n}(s) \wedge t) \right\} = \Delta X_s.$$
  - *(C3) For all $t \in \mathbb{R}_{\geq 0}$,*
    $$\varlimsup_{\varepsilon\downarrow\downarrow 0} \varlimsup_{n\to\infty} O_t^+(X - J_\varepsilon(X);\pi_n) = 0.$$
- *(ii) The sequence (π_n)* approximates $X\colon \mathbb{R}_{\geq 0} \to E$ from the left *if $\lim_{n\to\infty} X(\underline{\pi_n}(t)) = X(t-)$ holds for all $t > 0$. Then we call (π_n)* a left approximation sequence *for X.*

Remark 3.5.

- (i) Let (π_n) be a sequence of partitions along which we consider a quadratic variation of a path. In this paper, we often require that (π_n) satisfies Condition (C) defined above. Therefore, we can say that (C) is a condition for integrators of Itô–Föllmer integration.
- (ii) In contrast to (i), Condition (ii) of Definition 3.4 needs to be satisfied by integrands of Itô–Föllmer or Stieltjes integrals. This will be mainly used in Theorem 3.6 and Lemma 3.10.

Under these assumptions, we have the following $C^{1,2}$ type Itô formula for Banach space-valued paths.

Theorem 3.6 (Itô formula)**.** *Let E, E_1, F, and G be Banach spaces, $B\colon E \times E \to E_1$ be a bounded bilinear map, $(A,X) \in D(\mathbb{R}_{\geq 0}, F \times E)$, and (π_n) be a sequence of partitions that satisfies Condition (C) for (A,X) and approximates (A,X) from the left. Suppose that X has weak B-quadratic variation and finite 2-variation along (π_n) and suppose also that A has finite variation.*

Moreover, let $f\colon F \times E \to G$ be a function satisfying the following conditions:

- *(i) the map $F \ni a \mapsto f(a,x) \in G$ is Gâteaux differentiable for all $x \in E$ and $D_a f\colon F \times E \to \mathcal{L}_{\mathrm{c}}(F,G)$ is continuous;*

(ii) *the map* $E \ni x \mapsto f(a,x) \in G$ *is twice Gâteaux differentiable for all* $a \in F$ *and* $D_x f\colon F \times E \to \mathcal{L}_{\mathrm{c}}(E,G)$ *is continuous;*

(iii) *there is a continuous function* $D_B^2 f\colon F \times E \to \mathcal{L}_{\mathrm{c}}(E_1,G)$ *that commutes the diagram*

$$\begin{array}{ccc} F \times E & \xrightarrow{D_x^2 f} & \mathcal{L}_{\mathrm{c}}(E\widehat{\otimes}E,G) \\ & \searrow^{D_B^2 f} & \uparrow B^* \\ & & \mathcal{L}_{\mathrm{c}}(E_1,G) \end{array}$$

where $B^*\colon \mathcal{L}_{\mathrm{c}}(E_1,G) \to \mathcal{L}_{\mathrm{c}}(E\widehat{\otimes}E,G)$ *is defined by* $B^*(T)(x\otimes y) = T \circ B(x,y)$*.*

Then the Itô–Föllmer integral $\int_0^t D_x f(A_{s-},X_{s-})\mathrm{d}X_s$ *exists in the weak topology and it satisfies*

$$\begin{aligned} f(A_t,X_t) - f(A_0,X_0) &= \int_0^t D_a f(A_{s-},X_{s-})\mathrm{d}A_s^{\mathrm{c}} + \int_0^t D_x f(A_{s-},X_{s-})\mathrm{d}X_s \\ &+ \frac{1}{2}\int_0^t D_B^2 f(A_{s-},X_{s-})\mathrm{d}Q_B(X,X)_s^{\mathrm{c}} + \sum_{0<s\le t}\{\Delta f(A_s,X_s) - D_x f(A_{s-},X_{s-})\Delta X_s\}. \end{aligned} \tag{3.1}$$

Moreover, if the quadratic variation Q_B *exits in the strong sense, the convergence of the Itô–Föllmer integral holds in the norm topology of* G*.*

Here, note that (3.1) is an equation in the Banach space G. The Itô–Föllmer integral in Theorem 3.6 is defined by

$$\int_0^t D_x f(A_{s-},X_{s-})\,\mathrm{d}X_s = \lim_{n\to\infty}\sum_{]r,s]\in\pi_n} D_x f(A_r,X_r)(X_{t\wedge s} - X_{t\wedge r})$$

with suitable topology (see Definition 7.1.)

As a direct consequence of Theorem 3.6, we can derive the Itô formula related to tensor quadratic variations.

Corollary 3.7. *Let* E *and* F *be Banach spaces and let* α *be a reasonable cross norm on* $E \otimes E$*. Suppose that* $X \in D(\mathbb{R}_{\ge 0},E)$ *has strong or weak* α*-tensor quadratic variation and finite 2-variation along* (π_n) *and suppose also that* $A \in FV(\mathbb{R}_{\ge 0},F)$*. Moreover, let* $f\colon F \times E \to G$ *be a function satisfying the following conditions:*

(i) *the map* $F \ni a \mapsto f(a,x) \in G$ *is Gâteaux differentiable for all* $x \in E$ *and* $D_a f\colon F \times E \to \mathcal{L}_{\mathrm{c}}(F,G)$ *is continuous;*

(ii) *the map* $E \ni x \mapsto f(a,x) \in G$ *is twice Gâteaux differentiable for all* $a \in F$ *and* $D_x f\colon F \times E \to \mathcal{L}_{\mathrm{c}}(E,G)$ *is continuous;*

(iii) *the second derivative of* f *induces a continuous function* $D_x^2 f\colon F \times E \to \mathcal{L}_{\mathrm{c}}(E\widehat{\otimes}_\alpha E,G)$*.*

Then $f(A,X)$ *admits the following Itô formula:*

$$\begin{aligned} f(A_t,X_t) - f(A_0,X_0) &= \int_0^t D_a f(A_{s-},X_{s-})\mathrm{d}A_s^{\mathrm{c}} + \int_0^t D_x f(A_{s-},X_{s-})\mathrm{d}X_s \\ &\quad + \frac{1}{2}\int_0^t D_x^2 f(A_{s-},X_{s-})\mathrm{d}^\alpha[X,X]_s^{\mathrm{c}} + \sum_{0<s\le t}\{\Delta f(A_s,X_s) - D_x f(A_{s-},X_{s-})\Delta X_s\}. \end{aligned}$$

The convergence of the Itô–Föllmer integral holds in the norm or weak topology, respectively.

Remark 3.8. (i) Let $f\colon F \times E \to G$ be a $C^{1,2}$ function in the sense of Fréchet differentiation, i.e.

(a) the function $a \mapsto f(a, x)$ is Fréchet differentiable for all $x \in E$ and $D_a f\colon F \times E \to \mathcal{L}(F, G)$ is continuous;

(b) the function $x \mapsto f(a, x)$ is twice Fréchet differentiable for all $a \in F$, and $Df\colon F \times E \to \mathcal{L}(E, G)$ and $D^2 f\colon F \times E \to \mathcal{L}(E \widehat{\otimes} E, G)$ are continuous.

Then f satisfies the assumptions in Corollary 3.7 for the projective tensor norm γ.

(ii) Define a class $FC_{\mathrm{b}}^2(E)$ to be the set of all real functions of the form $G = g \circ P$ with $P\colon E \to V$ being the projection to a finite-dimensional subspace and $g \in C_{\mathrm{b}}^2(V)$. Let $\mathcal{C}$ be the completion of FC_{b}^2 with respect to the norm

$$\|G\| \coloneqq \sup_{x \in E} |G(x)| + \sup_{x \in E} \|DG(x)\|_{E^*} + \sup_{x \in E} \|D^2 G(x)\|_{(E \widehat{\otimes}_\alpha E)^*}.$$

Then every element of $\mathcal{C}$ satisfies the assumptions in Corollary 3.7.

Remark 3.9. A possible direction of development is a functional extension of Theorem 3.6 and Corollary 3.7, that is, infinite-dimensional functional Itô calculus. For such an extension, we should introduce Dupire's derivative of a functional of infinite dimensional paths. This seems an important future work. See, as referred to in Section 1, Dupire [21], Cont and Fournié [8, 9, 10], and Ananova and Cont [1], for functional Itô calculus in finite-dimensional state space.

The following lemma is essentially used to prove Theorem 3.6.

Lemma 3.10. *Let $B\colon E \times E \to E_1$ be a bounded bilinear map between Banach spaces and (π_n) be a sequence of partitions of $\mathbb{R}_{\geq 0}$. Suppose that a path $X \in D(\mathbb{R}_{\geq 0}, E)$ has weak B-quadratic variation and finite 2-variation along (π_n). Moreover, assume that (π_n) satisfies Condition (C) for X and approximates a $\xi \in D(\mathbb{R}_{\geq 0}, \mathcal{L}_{\mathrm{c}}(E_1, G))$ from the left. Then for all $t \in [0, \infty[$*

$$\sum_{I \in \pi_n} \xi B(\delta X_t, \delta X_t)(I) \xrightarrow[n\to\infty]{} \int_{]0,t]} \xi_{u-} \mathrm{d}Q_B(X, X)_u \tag{3.2}$$

holds in the weak topology of G.

If, in addition, $Q_B(X, X)$ is the strong B-quadratic variation, then (3.2) holds in the norm topology.

4 Auxiliary results regarding sequences of partitions

In this section, we investigate conditions on a sequence of partitions along which we deal with quadratic variations and Itô–Föllmer integrals. Recall that basic notions were defined in Definition 3.4.

Definition 4.1. *Let E be a Banach space, $X \in D(\mathbb{R}_{\geq 0}, E)$, and $(\pi_n)_{n \in \mathbb{N}}$ be a sequence of partitions of $\mathbb{R}_{\geq 0}$.*

(i) The sequence (π_n) controls the oscillation of X *if $\lim_{n\to\infty} O_t^-(X; \pi_n) = 0$ holds for all t.*

(ii) The sequence (π_n) exhausts the jumps of X *if $D(X) \subset \bigcup_{n \in \mathbb{N}} \bigcap_{k \geq n} \pi_k^{\mathrm{p}}$.*

Example 4.2. (i) Let r be an irrational number and $X = 1_{[r,\infty[}$. For each $n \in \mathbb{N}$, we set $\pi_n = \{]k2^{-n}, (k+1)2^{-n}]; k \in \mathbb{N}\}$. Then the sequence (π_n) satisfies $|\pi_n| \to 0$ as $n \to \infty$. This sequence, however, does not control the oscillation of X.

(ii) Let $X = 1_{[1,\infty[}$ and $\pi_n = \{]k, k+1]; k \in \mathbb{N}\}$. Though the sequence (π_n) controls the oscillation of X, it does not satisfy $|\pi_n| \to 0$.

(iii) Let $X = 1_{[1/2,\infty[}$ and $\pi_n = \{]k, k+1]; k \in \mathbb{N}\}$. The sequence (π_n) neither controls the oscillation of X nor satisfies $|\pi_n| \to 0$. However, it satisfies Condition (C) for X in the sense of Definition 3.4.

Condition (i) of Definition 4.1 is characterized as follows.

Lemma 4.3. *Let X be a càdlàg path in E and (π_n) be a sequence of partitions of $\mathbb{R}_{\geq 0}$. Then the following conditions are equivalent.*

(i) The sequence (π_n) controls the oscillation of X.

(ii) The sequence (π_n) satisfies the following conditions:

(a) the sequence (π_n) exhausts the jumps of X;

(b) if X is not constant on $]s,t[\subset \mathbb{R}_{\geq 0}$, then $]s,t[$ contains at least one element of π_n^{p} for sufficiently large n.

Lemma 4.3 is a generalization of Lemma 1 of Vovk [76, 272]. Note that the condition '$]s,t[$ *contains at least one point of* π_n^{p}' is equivalent to '*there is no $I \in \pi_n$ including* $]s,t[$.'

Proof. Step 1.1: (i) $\Longrightarrow$ (ii)-(a). Let $s \in D(X)$ and set $\varepsilon = \|\Delta X(s)\|_E$. Moreover, fix $T > s$ arbitrarily. Then, by assumption, we can choose an $N \in \mathbb{N}$ such that $O_T^-(X, \pi_n) < \varepsilon/2$ holds for all $n \geq N$. We will show that $s = \overline{\pi_n}(s)$ holds for all $n \geq N$. Take an s' from the interval $]\underline{\pi_n}(s), s[$ such that

$$\|X_{s'} - X_{s-}\|_E \leq O_T^-(X, \pi_n) < \frac{\varepsilon}{2}.$$

Then, we have

$$\|X_s - X_{s'}\|_E \geq \|\Delta X_s\|_E - \|X_{s-} - X_{s'}\|_E > \varepsilon - \frac{\varepsilon}{2} = \frac{\varepsilon}{2}.$$

This shows that $s \notin]\underline{\pi_n}(s), \overline{\pi_n}(s)[$, and therefore $s = \overline{\pi_n}(s)$. Hence (π_n) exhausts the jumps of X.

Step 1.2: (i) $\Longrightarrow$ (ii)-(b). Assume that $\varepsilon \coloneqq \omega(X;]s,t[) > 0$. Choose an $N \in \mathbb{N}$ that satisfies $O_t^-(X, \pi_n) < \varepsilon$ for all $n \geq N$. For an arbitrarily fixed $n \geq N$, choose a unique i satisfying $s \in [t_i^n, t_{i+1}^n[$. It remains to show that $t_{i+1}^n \in]s,t[$. Seeking a contradiction, suppose $t_{i+1}^n \geq t$. Then $[s,t[\subset [t_i^n, t_{i+1}^n[$ and hence

$$O_t^-(X, \pi_n) \geq \sup_{u,v \in [t_i^n, t_{i+1}^n[\cap [0,t]} \|X_u - X_v\| \geq \sup_{u,v \in [s,t[} \|X_u - X_v\| = \varepsilon.$$

This contradicts the condition that $O_t^-(X, \pi_n) < \varepsilon$.

Step 2: (ii) $\Longrightarrow$ (i). Suppose that (π_n) satisfies the conditions (ii)–(a) and (b).

Fix an $\varepsilon > 0$ and a $t > 0$ arbitrarily. Because X is càdlàg, we can take a sequence $0 = s_0 < s_1 < \cdots < s_N = t$ such that $\omega(X;]s_i, s_{i+1}[) < \varepsilon/2$ for all i (see Lemma 2.1). By assumption, we can choose an $N \in \mathbb{N}$ satisfying the following conditions:

1. If $n \geq N$, there are no $I \in \pi_n$ and $i \in \{0, \dots, N\}$ satisfying $]s_i, s_{i+1}[\subset I$ and $\omega(X;]s_i, s_{i+1}[) > 0$.
2. $\{s_0, \dots, s_N\} \cap D(X) \subset \bigcap_{n \geq N} \pi_n^{\mathrm{p}}$.

Let $n \geq N$ and $]u,v] \in \pi_n$. First, assume that $\omega(X;]u,v[) > 0$. By Condition 1, we see that there are only two cases for the relationship between $]u,v]$ and $(s_i)_{0 \leq i \leq N}$ as follows.

A. There is a unique i such that $]u,v] \subset]s_i, s_{i+1}[$.

B. There is a unique i such that $s_i \in]u,v[$.

In Case A, the oscillation of X on $]u,v[$ is estimated as

$$\omega(X;]u,v[\cap [0,t]) \leq \omega(X;]s_i, s_{i+1}[) < \frac{\varepsilon}{2}.$$

On the other hand, in Case B, X is continuous at $s_i \in]u,v[$ by Condition 2. Therefore,

$$\omega(X;]u,v[\cap[0,t]) \leq \omega(X;]s_{i-1},s_i[) + \omega(X;]s_i,s_{i+1}[) < \varepsilon.$$

If $\omega(X;]u,v[) = 0$, we clearly have the same estimate. By the discussion above, we find that $\omega(X;]u,v[\cap[0,t]) < \varepsilon$ holds for all $]u,v] \in \pi_n$, and consequently

$$O_t^-(X;\pi_n) = \sup_{]r,s]\in\pi_n} \omega(X;]r,s[\cap[0,t]) \leq \varepsilon$$

for every $n \geq N$. This completes the proof. □

We next consider the condition that (π_n) approximates a path from the left. Given a partition π, define the left discretization of a path $\xi\colon \mathbb{R}_{\geq 0} \to E$ along π by

$${}^{\pi}\xi = \sum_{]r,s]\in\pi} \xi(r)1_{]r,s]}.$$

If ξ is càdlàg, then the sequence (π_n) approximates ξ from the left in the sense of Definition 3.4 if and only if $({}^{\pi_n}\xi)_{n\in\mathbb{N}}$ converges to ξ_- pointwise. Consider the following example.

Example 4.4. Let $X = 1_{[1/2,\infty[}$ and $\pi_n = \{]k,k+1];k\in\mathbb{N}\}$. As we saw in Example 4.2, the sequence (π_n) satisfies Condition (C) for X. For each $n\in\mathbb{N}$, the left discretization of X is given by ${}^{\pi_n}X = 1_{]1,\infty[}$. The sequence $({}^{\pi_n}X)$ does not converge to X pointwise, and hence (π_n) does not approximate X from the left.

As we mentioned in Section 1, two types of assumptions about a sequence of partitions are frequently used in the context of the Itô–Föllmer calculus. One is that $|\pi_n| \to 0$ and the other is that (π_n) controls the oscillation of X. In the next proposition, we show that both conditions imply that (π_n) satisfies Condition (C) for X.

Proposition 4.5. *Let (π_n) be a sequence of partitions of $\mathbb{R}_{\geq 0}$ and let E be a Banach space.*

(i) Suppose that (π_n) satisfies $|\pi_n| \to 0$. Then it satisfies Condition (C) for every càdlàg path in E. Moreover, it approximates every càdlàg path in E from the left.

(ii) Suppose that (π_n) controls the oscillation of $X \in D(\mathbb{R}_{\geq 0},E)$. Then it satisfies Condition (C) for X and approximates X from the left.

Proof. (i) If $|\pi_n| \to 0$, then $\overline{\pi_n}(t) \to t$ and $\underline{\pi_n}(t) \to t$ hold for every $t \geq 0$. This directly implies that (π_n) approximates X from the left. Moreover, we have $\delta X_u(\pi_n(t)) \to X_t$ for every $t,u > 0$ with $t \leq u$. Hence, (π_n) satisfies Condition (C2). Condition (C3) follows from (iii) of Lemma 2.1. Condition (C1) remains to be shown. Given an $\varepsilon > 0$, define

$$r := \inf\{|u-v| \mid u,v \in D_\varepsilon(X)\cap[0,t], u\neq v\} > 0.$$

If $D_\varepsilon(X)\cap[0,t]$ has only one element, there is nothing to do. Otherwise, r is not zero, because $D_\varepsilon(X)\cap[0,t]$ has at most finitely many elements (see (i) of Lemma 2.1). Now we take an N satisfying $|\pi_n| < r$ for all $n \geq N$. Then for each $n \geq N$ and $]u,v] \in \pi_n$, the set $]u,v]\cap[0,t]$ contains at most one element of D_ε.

(ii) Assume that (π_n) controls the oscillation of X. Then we see that (π_n) approximates X from the left by the following estimate.

$$\|X_{\underline{\pi_n}(t)} - X_{t-}\|_E \leq \omega(X;[\underline{\pi_n}(t),\overline{\pi_n}(t)[\cap[0,t]) \leq O_t^-(X,\pi_n).$$

Now, let us show that (π_n) satisfies (C) for X. To obtain (C2), take a $t \in D(X)$. Because (π_n) exhausts the jumps of X (Lemma 4.3), we have $\overline{\pi_n}(t) = t$ for sufficiently large n. This combined with the fact that (π_n) is a left-approximation sequence implies (C2). Next, consider (C1). Let $\varepsilon > 0$ and fix an $N_\varepsilon \in \mathbb{N}$ so that $O_t^-(X, \pi_n) < \varepsilon$ holds for any $n \geq N_\varepsilon$. Then, for every $n \geq N_\varepsilon$ and $]r,s] \in \pi_n$, the interval $]r,s[\cap [0,t]$ does not contain any jump of X that is greater than ε. Therefore, $]r,s]$ possesses at most one element of $D_\varepsilon(X)$. This means that (π_n) satisfies (C1). All that is left is to check Condition (C3). Choose an M_ε that satisfies $O_t^-(X;\pi_n) < \varepsilon/2$ for all $I \in \pi_n$ and $n \geq M$. As we just have shown, $J_{\varepsilon/2}(X)$ is zero on the interior of each $I \in \pi_n$ and $n \geq M_\varepsilon$. Hence,

$$\begin{aligned}\omega(X - J_{\varepsilon/2}(X);]r,s] \cap [0,t]) &\leq \omega(X - J_{\varepsilon/2}(X);]r,s[\cap [0,t]) + \|\Delta(X - J_{\varepsilon/2}(X))_s\|_E \\ &\leq O_t^-(X;\pi_n) + \|\Delta(X - J_{\varepsilon/2}(X))_s\|_E < \varepsilon\end{aligned}$$

holds for all $]r,s] \in \pi_n$ and $n \geq M_\varepsilon$, which implies (C3). □

In the last part of this section, we give an additional lemma about a sequence of partitions.

Lemma 4.6. *(i) Let X be a càdlàg path in a Banach space E. If (π_n) approximates ξ from the left, then (π_n) also approximates $f \circ \xi$ from the left for every continuous function $f\colon E \to E'$ to an arbitrary Banach space.*

(ii) Let X and Y be càdlàg paths in Banach spaces E and F, respectively. If (π_n) satisfies Condition (C) for the path (X,Y) in $E \times F$, then (π_n) satisfies (C) for each of X and Y. Here, we regard $E \times F$ as a Banach space endowed with the direct sum norm $\| \ \|_E + \| \ \|_F$.

Proof. (i) immediately follows from the continuity of f.

To show (ii), suppose that (π_n) satisfies (C) for (X,Y). It suffices to show that (π_n) satisfies (C) for X. First, fix $t \in \mathbb{R}_{\geq 0}$ and $\varepsilon > 0$ arbitrarily, and then choose an N so that $D_\varepsilon(X,Y) \cap I \cap [0,t]$ has at most one element for all $n \geq N$ and $I \in \pi_n$. The inclusion $D_\varepsilon(X) \cap I \cap [0,t] \subset D_\varepsilon(X,Y) \cap I \cap [0,t]$ implies that the cardinality of $I \cap [0,t] \cap D_\varepsilon(X)$ is no greater than 1. Condition (C2) follows directly from the definition of product topology. Condition (C3) remains to be shown. For an arbitrary $\delta > 0$, choose an $\varepsilon_0 > 0$ satisfying

$$\sup_{\varepsilon \leq \varepsilon_0} \overline{\lim_{n\to\infty}} O_t^+((X,Y) - J_\varepsilon(X,Y); \pi_n) < \frac{\delta}{2}.$$

Set $\varepsilon_1 = \varepsilon_0 \wedge (\delta/2)$. Given $\varepsilon \leq \varepsilon_1$, we can take $M_\varepsilon > 0$ such that

(a) $I \cap [0,t] \cap D_\varepsilon(X,Y)$ has at most one element for all $I \in \pi_n$ and $n \geq M_\varepsilon$.
(b) $\sup_{n \geq M_\varepsilon} O_t^+((X,Y) - J_\varepsilon(X,Y); \pi_n) < \delta/2$.

If $n \geq M_\varepsilon$ and $I \in \pi_n$, then for any $u, v \in I$, we have

$$\begin{aligned}&\|(X - J_\varepsilon(X))_u - (X - J_\varepsilon(X))_v\|_E \\ &\quad\leq \|(X - J(D_\varepsilon(X,Y),X))_u - (X - J(D_\varepsilon(X,Y),X))_v\|_E \\ &\qquad + \|(J(D_\varepsilon(X,Y),X) - J(D_\varepsilon(X),X))_u - (J(D_\varepsilon(X,Y),X) - J(D_\varepsilon(X),X))_v\|_E \\ &\quad\leq \sup_{n \geq M_\varepsilon} O_t^+((X,Y) - J_\varepsilon(X,Y); \pi_n) + \varepsilon \\ &\quad\leq \frac{\delta}{2} + \frac{\delta}{2} = \delta.\end{aligned}$$

Here, note that the second inequality holds by Condition (a) above. Thus, we get

$$\overline{\lim_{n\to\infty}} O_t^+(X - J_\varepsilon(X); \pi_n) \leq \sup_{n \geq M_\varepsilon} O_t^+(X - J_\varepsilon(X); \pi_n) \leq \delta.$$

for arbitrary $\varepsilon \leq \varepsilon_1$. This implies (C3) for X. □

5 Properties of quadratic variations

This section is devoted to studying some basic properties of quadratic variations introduced in Section 3. Throughout this section, suppose that we are given a sequence $\Pi = (\pi_n)_{n\in\mathbb{N}}$ of partitions of $\mathbb{R}_{\geq 0}$.

First, we give some examples of quadratic variations.

Example 5.1. Let A be a path of finite variation in a Banach space E. If (π_n) satisfies $|\pi_n| \to 0$, then A has projective tensor and scalar quadratic variations given by

$$^{\gamma}[A,A]_t = \sum_{0<s\leq t} (\Delta A_s)^{\otimes 2}, \qquad Q(A)_t = \sum_{0<s\leq t} \|\Delta A_s\|^2.$$

This result will be proved later in this section.

Example 5.2. Let $x\colon [0,\infty[\to \mathbb{R}$ be a càdlàg path and (π_n) a sequence of partitions such that $|\pi_n| \to 0$. Given arbitrary C^1-function $f\colon \mathbb{R} \to E$ into a Banach space E, let us set $X(t) = f(x(t))$. If x has quadratic variation along (π_n), then X has projective tensor quadratic variation given by

$$^{\gamma}[X,X]_t = \int_0^t Df(x_{s-})^{\otimes 2} \mathrm{d}[x,x]^{\mathrm{c}}_s + \sum_{0<s\leq t} \Delta f(x_s)^{\otimes 2}.$$

This is one of the main results of the second article in this series [37].

For the construction of a real continuous path of nontrivial quadratic variation, refer to Schied [68], Mishura and Schied [52], and Cont and Das [7].

The next examples are from the theory of stochastic processes.

Example 5.3. (i) Let $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, P)$ be a filtered probability space satisfying the usual conditions. Consider a semimartingale $X = (X_t)_{t\geq 0}$ in a separable Hilbert space H. Moreover, let $\pi = (\tau^n_k)_{n\in\mathbb{N}}$ be an increasing sequence of bounded stopping times such that $\tau^n_k \to \infty$ as $k \to \infty$ and $\sup_k(\tau^{n+1}_k - \tau^n_k) \to 0$ as $n \to \infty$ almost surely. Then the process $[X,X]^{\pi}$ converges to the quadratic variation process $[X,X]$ uniformly on compacts in probability (ucp). By passing to an appropriate subsequence, we see that almost all paths have quadratic variation along the subsequence. See Gravereaux and Pellaumail [28] or Metivier and Pellaumail [47] for details.

(ii) In addition to the assumptions of (i), let $f\colon H \to E$ be a C^1 function into a Banach space E. Then, along a suitable subsequence of (π_n), almost all paths of $f(X)$ have quadratic variation. If, moreover, f is of C^2 class, the Itô–Föllmer integral $\int_0^{\cdot} Df(X_{s-})\mathrm{d}X_s$ exists and its paths have quadratic variation along the same subsequence. Now, since the Banach space E is chosen arbitrarily, it may fail to satisfy some useful properties required by the martingale theory, e.g. UMD property or martingale type 2 property. The path $f(X)$, however, behaves well enough from the viewpoint of the Itô–Föllmer calculus.

Now we consider the transpose of quadratic covariation. Let $B \in \mathcal{L}^{(2)}(E,F;G)$ and define the transpose ${}^tB\colon F\times E \to G$ of B by ${}^tB(y,x) = B(x,y)$. Then $(X,Y) \in D(\mathbb{R}_{\geq 0}, E\times F)$ has strong/weak B-quadratic covariation if and only if (Y,X) does with respect to the transpose ${}^tB$.

Recall that a d-dimensional càdlàg path $X = (X_1,\dots,X_d)$ has tensor quadratic variation if and only if X_i and X_j have quadratic covariation for each i and j. This characterization is generalized to bilinear quadratic covariations in Banach spaces.

Proposition 5.4. *Let E_i, F_j, and G_{ij} be Banach spaces and let $B_{ij}\colon E_i \times F_j \to G_{ij}$ be a bounded bilinear map for $i,j \in \{1,2\}$. Define a continuous bilinear map $\mathbf{B}\colon (E_1\times E_2)\times$*

$(F_1 \times F_2) \to \prod_{i,j} G_{ij}$ *by*

$$\mathbf{B}((x_1, x_2), (y_1, y_2)) = (B_{ij}(x_i, y_i))_{i,j \in \{1,2\}}.$$

If $\mathbf{X} = (X_1, X_2) \in D(\mathbb{R}_{\geq 0}, E_1 \times E_2)$ *and* $\mathbf{Y} = (Y_1, Y_2) \in D(\mathbb{R}_{\geq 0}, F_1 \times F_2)$*, then* $(\mathbf{X}, \mathbf{Y})$ *has strong or weak* $\mathbf{B}$*-quadratic covariation if and only if* (X_i, Y_j) *has respectively strong or weak* B_{ij}*-quadratic covariation for all* $i, j \in \{1, 2\}$*. In this case, these quadratic covariations satisfy*

$$Q_{\mathbf{B}}((X_1, X_2), (Y_1, Y_2)) = (Q_{B_{ij}}(X_i, X_j))_{i,j \in \{1,2\}}. \tag{5.1}$$

Using the matrix notation, we can also express equation (5.1) as

$$Q_{\mathbf{B}}((X_1, X_2), (Y_1, Y_2)) = \begin{pmatrix} Q_{B_{11}}(X_1, Y_1) & Q_{B_{12}}(X_1, Y_2) \\ Q_{B_{21}}(X_2, Y_1) & Q_{B_{22}}(X_2, Y_2) \end{pmatrix}.$$

Proof. By the definition of $\mathbf{B}$, we can easily check that

$$\sum_{I \in \pi_n} \mathbf{B}\Big(\big((\delta X_1)_t, (\delta X_2)_t\big), \big((\delta Y_1)_t, (\delta Y_2)_t\big)\Big)(I) = \left(\sum_{I \in \pi_n} B_{ij}\big((\delta X_i)_t, (\delta Y_i)_t\big)(I)\right)_{i,j \in \{1,2\}},$$
$$\mathbf{B}\left(\Delta(X_1, X_2)_t, \Delta(Y_1, Y_2)_t\right) = \left(B_{ij}(\Delta X_i(t), \Delta Y_j(t))\right)_{i,j \in \{1,2\}},$$

hold for all $t \geq 0$. These immediately prove the assertion. □

Applying Proposition 5.4 to the canonical bilinear map $\otimes \colon E_i \times E_j \to E_i \widehat{\otimes}_\alpha E_j$, we obtain the following corollary.

Corollary 5.5. *Let* $(X_1, X_2) \in D(\mathbb{R}_{\geq 0}, E_1 \times E_2)$ *and let* α *be a uniform crossnorm. Then* (X_1, X_2) *has strong or weak* α*-tensor quadratic variation if and only if* (X_i, X_j) *has strong or weak* α*-tensor quadratic covariation, respectively, for every* $i, j \in \{1, 2\}$*.*

Like the quadratic covariation $[X, Y]$ of scalar paths, quadratic covariation Q_B is bilinear in an appropriate sense.

Proposition 5.6. *Let* $X_1, X_2 \in D(\mathbb{R}_{\geq 0}, E)$ *and* $Y_1, Y_2 \in D(\mathbb{R}_{\geq 0}, F)$*. Suppose that* (X_i, Y_j) *has strong or weak quadratic covariation with respect to* $B \in \mathcal{L}^{(2)}(E, F; G)$ *for each* $i, j \in \{1, 2\}$*. Then,* $(X_1 + X_2, Y_1 + Y_2)$ *has respectively strong or weak* B*-quadratic covariation given by*

$$Q_B(X_1 + X_2, Y_1 + Y_2) = \sum_{i,j \in \{1,2\}} Q_B(X_i, Y_j).$$

Proof. By the bilinearity of B, we see that

$$\sum_{I \in \pi_n} B(\delta(X_1 + X_2)_t, \delta(Y_1 + Y_2)_t)(I) = \sum_{i,j \in \{1,2\}} \sum_{I \in \pi_n} B(\delta(X_i)_t, \delta(Y_j)_t)(I)$$

for every $t \geq 0$. Therefore, by assumption, the left-hand side converges to $\sum_{ij} Q_B(X_i, Y_j)$ in the corresponding topology. Again by bilinearity,

$$\Delta\left(\sum_{i,j \in \{1,2\}} Q_B(X_i, Y_j)\right)(t) = \sum_{i,j \in \{1,2\}} B(\Delta(X_i)_t, \Delta(Y_j)_t) = B\left(\Delta(X_1 + X_2)_t, \Delta(Y_1 + Y_2)_t\right).$$

Hence, $\sum_{ij} Q_B(X_i, Y_j)$ is the B-quadratic covariation of $X_1 + X_2$ and $Y_1 + Y_2$. □

Corollary 5.7. *Let E and F be Banach spaces and α be a reasonable crossnorm on $E \otimes F$. Suppose that $(X_i, Y_j) \in D(\mathbb{R}_{\geq 0}, E \times F)$ has strong or weak α-tensor quadratic covariation for every $i, j \in \{1, 2\}$. Then, $(X_1 + X_2, Y_1 + Y_2)$ has strong or weak α-tensor quadratic covariation, respectively, and it satisfies*

$$^{\alpha}[X_1 + X_2, Y_1 + Y_2] = {}^{\alpha}[X_1, Y_1] + {}^{\alpha}[X_1, Y_2] + {}^{\alpha}[X_2, Y_1] + {}^{\alpha}[X_2, Y_2].$$

Next, we investigate the quadratic variation of a path of finite variation. For convenience, we introduce the following notation. Let $D \subset \mathbb{R}_{\geq 0}$ and define functions e^1_D and e^2_D from $\mathcal{P}(\mathbb{R}_{\geq 0})$ to $\{0, 1\}$ by

$$e^1_D(A) = \begin{cases} 1 & \text{if} \quad A \cap D \neq \emptyset \\ 0 & \text{if} \quad A \cap D = \emptyset \end{cases}, \qquad e^2_D = 1 - e^1_D.$$

The symbol $\mathcal{P}(\mathbb{R}_{\geq 0})$ above denotes the power set of $\mathbb{R}_{\geq 0}$.

Proposition 5.8. *Let E, F, and G be Banach spaces and let $B \in \mathcal{L}^{(2)}(E, F; G)$. Assume that $A \in FV(\mathbb{R}_{\geq 0}, E)$, $X \in D(\mathbb{R}_{\geq 0}, F)$, and (π_n) satisfies Condition (C) for (A, X). Then (A, X) has the strong B-quadratic variation given by*

$$Q_B(A, X)_t = \sum_{0 < s \leq t} B(\Delta A_s, \Delta X_s).$$

Proof. Fix $t \in \mathbb{R}_{\geq 0}$ and take an arbitrary $\varepsilon > 0$. For convenience, set $D = D(A, X)$, $D_\varepsilon = D_\varepsilon(A, X)$, and $D^\varepsilon = D^\varepsilon(A, X)$. Then,

$$\begin{aligned} &\left\| \sum_{I \in \pi_n} B(\delta A_t, \delta X_t)(I) - \sum_{0 < u \leq t} B(\Delta A_u, \Delta X_u) \right\| \\ &\quad \leq \left\| \sum_{I \in \pi_n} B(\delta A_t, \delta X_t)(I) e^1_{D_\varepsilon}(I) - \sum_{0 < u \leq t} B\left(\Delta A_u, \Delta X_u\right) \right\| + \left\| \sum_{I \in \pi_n} B(\delta A_t, \delta X_t)(I) e^2_{D_\varepsilon}(I) \right\| \end{aligned} \tag{5.2}$$

for any $t \in \mathbb{R}_{\geq 0}$. We will observe the behaviour of each term on the right-hand side of (5.2).

Because (π_n) satisfies Condition (C) for X, there exists an N_1 such that $D_\varepsilon \cap [0, t] \cap I$ contains at most one point for all $n \geq N_1$ and $I \in \pi_n$. If $n \geq N_1$, we have

$$\sum_{I \in \pi_n} B(\delta A_t, \delta X_t)(I) e^1_{D_\varepsilon}(I) = \sum_{u \in D_\varepsilon} B\left(\delta A_t, \delta X_t\right)(\pi_n(u)).$$

Therefore, by Condition (C2),

$$\lim_{n \to \infty} \sum_{I \in \pi_n} B(\delta A_t, \delta X_t)(I) e^1_{D_\varepsilon}(I) = \sum_{u \in D_\varepsilon \cap [0, t]} B(\Delta A_u, \Delta X_u).$$

This implies that

$$\begin{aligned} &\lim_{n \to \infty} \left\| \sum_{I \in \pi_n} B(\delta A_t, \delta X_t)(I) e^1_{D_\varepsilon}(I) - \sum_{0 < u \leq t} B(\Delta A_u, \Delta X_u) \right\| \\ &\quad = \left\| \sum_{u \in D^\varepsilon \cap [0, t]} B(\Delta A_u, \Delta X_u) \right\| \leq \|B\| \sup_{u \in [0, t]} \|\Delta X_u\| \sum_{u \in D^\varepsilon \cap [0, t]} \|\Delta A_u\|. \end{aligned}$$

Next, we consider the second term on the right-hand side of (5.2). Since X has no jumps greater than ε on I whenever $e^2_{D_\varepsilon}(I) = 1$, we have the estimate

$$\|\delta X_t(I)\|_F \, e^2_{D_\varepsilon}(I) = \|\delta(X - J_{D_\varepsilon}(X))_t(I)\| e^2_{D_\varepsilon}(I) \leq O^+_t(X - J_{D_\varepsilon}(X); \pi_n) e^2_{D_\varepsilon}(I).$$

Hence

$$\begin{aligned} \left\| \sum_{I \in \pi_n} B(\delta A_t, \delta X_t)(I) e^2_{D_\varepsilon}(I) \right\|_G &\leq \|B\| \sum_{I \in \pi_n} \|\delta A_t(I)\|_E \, \|\delta X_t(I)\|_F \, e^2_{D_\varepsilon}(I) \\ &\leq \|B\| O^+_t(X - J_{D_\varepsilon}(X); \pi_n) V(A)_t. \end{aligned}$$

From the discussion above, we can deduce that

$$\begin{aligned} &\varlimsup_{n\to\infty} \left\| \sum_{I \in \pi_n} B(\delta A_t, \delta X_t)(I) - \sum_{0<u\leq t} B\left(\Delta A_u, \Delta X_u\right) \right\|_G \\ &\quad \leq \|B\| \sup_{u \in [0,t]} \|\Delta X_u\|_F \sum_{u \in D^\varepsilon \cap [0,t]} \|\Delta A_u\|_E + \|B\| V(A)_t \varlimsup_{n\to\infty} O^+_t(X - J_{D_\varepsilon}(X); \pi_n). \end{aligned}$$

Consequently,

$$\varlimsup_{n\to\infty} \left\| \sum_{I \in \pi_n} B(\delta A_t, \delta X_t)(I) - \sum_{0<u\leq t} B\left(\Delta A_u, \Delta X_u\right) \right\|_G = 0,$$

which is the desired conclusion. □

Applying Proposition 5.8 to the canonical bilinear maps $\otimes \colon E \times F \to E \widehat{\otimes}_\alpha F$ and $\otimes \colon F \times E \to F \widehat{\otimes}_{\alpha'} E$, we get the following corollary.

Corollary 5.9. *Let $A \in FV(\mathbb{R}_{\geq 0}, E)$ and $X \in D(\mathbb{R}_{\geq 0}, F)$. If (π_n) satisfies Condition (C) for (A, X), then it has tensor quadratic covariations ${}^\alpha[A, X]$ and ${}^{\alpha'}[X, A]$ given by*

$${}^\alpha[A, X]_t = \sum_{0<s\leq t} \Delta A_s \otimes \Delta X_s, \qquad {}^{\alpha'}[X, A]_t = \sum_{0<s\leq t} \Delta X_s \otimes \Delta A_s$$

for every reasonable crossnorms α and α' on $E \otimes F$ and $F \otimes E$, respectively.

Using Corollaries 5.7 and 5.9, we obtain the following.

Corollary 5.10. *Let $(X, A) \in D(\mathbb{R}_{\geq 0}, E) \times FV(\mathbb{R}_{\geq 0}, E)$ and suppose that (π_n) satisfies Condition (C) for (X, A). If $X \colon \mathbb{R}_{\geq 0} \to E$ has α-tensor quadratic variation along (π_n), then $X + A$ has α-tensor quadratic variation given by*

$${}^\alpha[X + A, X + A] = {}^\alpha[X, X] + {}^\alpha[X, A] + {}^\alpha[A, X] + {}^\alpha[A, A]$$

for every reasonable crossnorm α on $E \otimes E$.

By a discussion similar to the proof of Proposition 5.8, we see that a path of finite variation has scalar quadratic variation.

Proposition 5.11. *Let A be a càdlàg path of finite variation in a Banach space E. If (π_n) satisfies Condition (C) for A, then A has the scalar quadratic covariation given by*

$$Q(A)_t = \sum_{0<s\leq t} \|\Delta A_s\|^2.$$

In the preceding part of this paper, we have used the summation

$$\sum_{]r,s]\in\pi_n} B(X_{s\wedge t}-X_{r\wedge t}, Y_{s\wedge t}-Y_{r\wedge t})$$

to define the quadratic covariation. We can also consider a different summation

$$\sum_{]r,s]\in\pi_n} 1_{[0,t[}(r)B(X_s-X_r, Y_s-Y_r),$$

which is a slight modification of that used in the original paper by Föllmer [22]. Let us investigate the relation between these two summations.

Proposition 5.12. *Let $(X,Y)\in D(\mathbb{R}_{\geq 0}, E\times F)$ and $B\in\mathcal{L}^{(2)}(E,F;G)$. Suppose that $(X,Y)_{\overline{\pi_n}(t)}\to(X,Y)_t$ holds for all $t\in\mathbb{R}_{\geq 0}$. Then the following two conditions are equivalent, respectively.*

(i) The path (X,Y) has strong/weak B-quadratic covariation along (π_n).

(ii) There exists a càdlàg path $V\in FV(\mathbb{R}_{\geq 0},G)$ such that

(a) for all $t\in\mathbb{R}_{\geq 0}$

$$\sum_{]r,s]\in\pi_n} 1_{[0,t[}(r)B(\delta X,\delta Y)(]r,s]) \xrightarrow[n\to\infty]{} V_t$$

in the norm/weak topology of G,

(b) for all $t\in\mathbb{R}_{\geq 0}$

$$\Delta V_t = B(\Delta X_t, \Delta Y_t).$$

If these conditions are satisfied, then V coincides with the quadratic covariation $Q_B(X,Y)$.

Proof. We show the assertion about strong convergence. For each $t>0$

$$\begin{aligned}
&\left\|\sum_{]r,s]\in\pi_n} B(\delta X_t,\delta Y_t)(]r,s]) - \sum_{]r,s]\in\pi_n} 1_{[0,t[}(r)B(\delta X,\delta Y)(]r,s])\right\| \\
&\quad = \|B(\delta X_t,\delta Y_t)(\pi_n(t)) - B(\delta X,\delta Y)(\pi_n(t))\| \\
&\quad \leq \|B(\delta X_t-\delta X,\delta Y_t)(\pi_n(t))\| + \|B(\delta X,\delta Y_t-\delta Y)(\pi_n(t))\| \\
&\quad \leq \|B\|\|X_{\overline{\pi_n}(t)}-X_t\|\|\delta Y_t(\pi_n(t))\| + \|B\|\|\delta X(\pi_n(t))\|\|Y_{\overline{\pi_n}(t)}-Y_t\|.
\end{aligned}$$

Hence, by assumption, the convergences of these two sequences are equivalent and their limits coincide.

In the weak case, we have a similar estimate for the pairing $\langle z^*,\ \rangle$, which shows the assertion about the weak quadratic covariation. □

According to Proposition 5.12, we see that the two definitions of quadratic covariation are equivalent provided that (π_n) satisfies the assumption in the proposition. The first definition, which is given in Definition 3.1, is more intuitive. The second one has some technical advantages because the path $t\mapsto\sum_{I\in\pi_n}1_{[0,t[}B(\delta X,\delta Y)(I)$ is of finite variation [2].

Remark 5.13. Following a discussion similar to that in Proposition 5.12, we can obtain an equivalent definition of scalar quadratic variation using the summation $\sum_{I\in\pi_n}\|1_{[0,t[}\delta X(I)\|^2$ if (π_n) satisfies the same condition as Proposition 5.12.

[2]Note that the path $t\mapsto\sum_{\pi_n}1_{[0,t[}B(\delta X,\delta Y)(I)$ is *càglàd* but not càdlàg.

6 Proof of Lemma 3.10

In this section, we prove Lemma 3.10, which is essentially used to show the main theorems of this paper. To prove it, we prepare some additional lemmas. Though Lemma 3.10 includes both weak and strong convergence results, we mainly focus on the proof of the weak case [3].

Throughout this section, let the symbols E, E_1, and G denote Banach spaces and $B\colon E\times E\to E_1$ denote a bounded bilinear map.

Lemma 6.1. *Let $s>0$ and let $(\pi_n)_{n\in\mathbb{N}}$ be a sequence of partitions. Then (π_n) approximates $1_{[0,s[}$ from the left if and only if $\overline{\pi_n}(s)\to s$ as $n\to\infty$.*

Proof. First, assume that (π_n) approximates $1_{[0,s[}$ from the left and take an $\varepsilon>0$ arbitrarily. By definition, $1_{[0,s[}(s+\varepsilon-)=0$. Since (π_n) approximates $1_{[0,s[}$ from the left and $1_{[0,s[}$ takes only two values 0 and 1, one sees that $1_{[0,s[}(\underline{\pi_n}(s+\varepsilon))=0$ holds for a large enough n. Therefore, $\underline{\pi_n}(s+\varepsilon)\geq s$ for large enough n. This leads to $\overline{\pi_n}(s)\leq\underline{\pi_n}(s+\varepsilon)<s+\varepsilon$ for a sufficiently large n. Hence $\overline{\pi_n}(s)\to s$ as $n\to\infty$.

Conversely, assume that $(\overline{\pi_n}(s))_{n\in\mathbb{N}}$ converges to s as $n\to\infty$. If $0<t\leq s$, then the convergence $1_{[0,s[}(\underline{\pi_n}(t))\to 1_{[0,s[}(t-)=1$ is obvious. Let $t>s$. We see from the assumption that $\overline{\pi_n}(s)<t$ for large enough n. For such an n, we have $s\leq\overline{\pi_n}(s)\leq\underline{\pi_n}(t)<t$, and hence $1_{[0,s[}(\underline{\pi_n}(t))=0=1_{[0,s[}(t-)$. This shows that $\lim_{n\to\infty}1_{[0,s[}(\underline{\pi_n}(t))=1_{[0,s[}(t-)$. Hence (π_n) is a left approximation sequence for $1_{[0,s[}$. ☐

Lemma 6.2. *Suppose that $X\in D(\mathbb{R}_{\geq 0},E)$ has weak B-quadratic variation along a sequence (π_n) satisfying Condition (C) for X. Let r,s be two real numbers satisfying $0\leq r<s$. If (π_n) approximates $1_{[r,s[}$ from the left, then*

$$\lim_{n\to\infty}\sum_{I\in\pi_n}1_{[r,s[}B(\delta X_t,\delta X_t)(I)=Q_B(X,X)_{s\wedge t}-Q_B(X,X)_{r\wedge t}\tag{6.1}$$

holds for all $t\in\mathbb{R}_{\geq 0}$ in the weak topology.

If X has strong B-quadratic variation, then the convergence of (6.1) *holds in the norm topology.*

Proof. We show the case of weak convergence. By considering the decomposition

$$\sum_{I\in\pi_n}1_{[r,s[}B(\delta X_t,\delta X_t)(I)=\sum_{I\in\pi_n}1_{[0,s[}B(\delta X_t,\delta X_t)(I)-\sum_{I\in\pi_n}1_{[0,r[}B(\delta X_t,\delta X_t)(I),$$

we can assume that $r=0$ without loss of generality.

If $t\leq s$, the equation

$$\sum_{I\in\pi_n}1_{[0,s[}B(\delta X_t,\delta X_t)(I)=\sum_{I\in\pi_n}B(\delta X_t,\delta X_t)(I)$$

holds for all $n\in\mathbb{N}$. Therefore,

$$\lim_{n\to\infty}\left\langle z^*,\sum_{I\in\pi_n}1_{[0,s[}B(\delta X_t,\delta X_t)(I)\right\rangle=\langle z^*,Q_B(X,X)_t\rangle=\langle z^*,Q_B(X,X)_{t\wedge s}\rangle$$

for all $z^*\in E_1^*$.

[3] For the strong case, the reader can refer to an earlier version of this article at arXiv: 2104.08138v2

Next, assume $s < t$. Then, by Lemma 6.1, $\overline{\pi_n}(s) \to s$ as $n \to \infty$. Hence,

$$\begin{aligned}\lim_{n\to\infty}\left\langle z^*, \sum_{I\in\pi_n} 1_{[0,s[}B(\delta X_t,\delta X_t)(I)\right\rangle &= \lim_{n\to\infty}\left\langle z^*, \sum_{I\in\pi_n} 1_{[0,s[}B(\delta X,\delta X)(I)\right\rangle \\ &= \langle z^*, Q_B(X,X)_s\rangle \\ &= \langle z^*, Q_B(X,X)_{s\wedge t}\rangle.\end{aligned}$$

Note that the second equality follows from the same argument as the proof of Proposition 5.12.

In both cases, we have the desired convergence.

If X has strong B-quadratic variation, we can directly show the norm convergence of the sequence without taking the pairing $\langle z^*, \ \rangle$. □

Lemma 6.3. *Let X be a càdlàg path in E with weak B-quadratic variation along a sequence (π_n) satisfying Condition (C) for X. Suppose that $\xi \in D(\mathbb{R}_{\geq 0}, \mathcal{L}(E_1, G))$ has the representation*

$$\xi = \sum_{i\geq 1} 1_{[\tau_{i-1},\tau_i[}a_i, \tag{6.2}$$

where $0 = \tau_0 < \tau_1 < \cdots < \tau_i < \tau_{i+1} < \cdots \to \infty$ and each a_i is an element of $\mathcal{L}(E_1, G)$. If (π_n) approximates ξ from the left, then the Stieltjes integral of ξ_- with respect to $Q_B(X,X)$ is approximated as

$$\lim_{n\to\infty}\sum_{I\in\pi_n}\xi B(\delta X_t,\delta X_t)(I) = \int_{]0,t]}\xi_{s-}\mathrm{d}Q_B(X,X)_s. \tag{6.3}$$

in the weak topology.

If X has strong B-quadratic variation, then (6.3) *holds in the norm topology of G.*

Proof. We show the weak convergence case. First, note that the Stieltjes integral on the right-hand side of (6.3) has the representation

$$\int_{]0,t]}\xi_{s-}\mathrm{d}Q_B(X,X)_s = \sum_{i\geq 1} a_i\{Q_B(X,X)_{\tau_i\wedge t} - Q_B(X,X)_{\tau_{i-1}\wedge t}\}.$$

On the other hand, the summation on the left-hand side of (6.3) is calculated as

$$\sum_{I\in\pi_n}\xi B(\delta X_t,\delta X_t)(I) = \sum_{i\geq 1} a_i \sum_{I\in\pi_n} 1_{[\tau_{i-1},\tau_i[}B(\delta X_t,\delta X_t)(I).$$

Therefore, it suffices to show that

$$\lim_{n\to\infty}\left\langle z^*a_i, \sum_{I\in\pi_n} 1_{[\tau_{i-1},\tau_i[}B(\delta X_t,\delta X_t)(I)\right\rangle = \langle z^*a_i, Q_B(X,X)_{\tau_i\wedge t} - Q_B(X,X)_{\tau_{i-1}\wedge t}\rangle$$

for all $z^* \in G^*$ and $i \geq 1$. This follows directly from Lemma 6.2.

If $Q_B(X,X)$ is the strong B-quadratic variation, the sequence of discrete sums converges in the norm topology by the strong version of Lemma 6.2. □

Lemma 6.4. *Let V be a locally convex Hausdorff topological vector space of which topology is generated by the family of seminorms $(\rho_i)_{i\in I}$.*

(i) Let $f\colon \mathbb{R}_{\geq 0} \to V$ be a càdlàg path. Then for every $i \in I$ and $\varepsilon > 0$, there is a right continuous step function h such that $\rho_i(f(t) - h(t)) \leq \varepsilon$ for all $t \geq 0$.

(ii) If a sequence of partitions (π_n) approximates f from the left, then the step function h in (i) can be chosen so that (π_n) still approximates h from the left.

Proof. Fix $\varepsilon > 0$ and $i \in I$. We define the oscillation of f on $S \subset \mathbb{R}_{\geq 0}$ by

$$\omega_i(f; S) = \sup\{\rho_i(f(t) - f(s)) \mid s, t \in S\}.$$

(i) Let us construct a partition of $\mathbb{R}_{\geq 0}$ recursively. First, let $t_0 = 0$. Next, assume that there is a sequence $0 = t_0 < \cdots < t_n$ such that $\omega_i(f; [t_k, t_{k+1}[) \leq \varepsilon$ for $k \in \{0, \ldots, n-1\}$.

Case A. If $\omega_i(f; [t_n, \infty[) = 0$, then we set $t_{n+1} = t_n + 1$.

Case B. If $\omega_i(f; [t_n, \infty[) > \varepsilon$, first let

$$t'_{n+1} = \inf\{t > t_n \,|\, \omega_i(f; [t_n, t[) > \varepsilon\},$$

and then define

$$t_{n+1} = \begin{cases} \sup\{t > t'_{n+1} \mid \omega_i(f; [t'_{n+1}, t[) = 0\} & \text{if } \omega_i(f; [t'_{n+1}, \infty[) > 0 \\ t'_{n+1} & \text{otherwise.} \end{cases} \tag{6.4}$$

Note that f is not constant on any interval of the form $[t_{n+1}, t[$ $(t > t_{n+1})$ when $\omega_i(f; [t'_{n+1}, \infty[) > 0$.

Case C. If $0 < \omega_i(f; [t_n, \infty[) \leq \varepsilon$, then let $t'_{n+1} = t_n + 1$ and define t_{n+1} by the formula (6.4).

The sequence defined above satisfies $t_n \to \infty$. Indeed, if $\omega_i(f; [t_n, \infty[) \leq \varepsilon$ for some $n \in \mathbb{N}$, we have $t_{n+k} \geq t_n + k$ for all $k \in \mathbb{N}$, by definition. Now assume that $\omega_i(f; [t_n, \infty[) > \varepsilon$ for all $n \in \mathbb{N}$ and $t_n \uparrow t^* < \infty$. In this case, $\omega_i(f; [t^* - \delta, t^*[) \geq \varepsilon$ holds for arbitrary small δ. This contradicts the existence of the left limit at t^*.

Let $P \coloneqq (t_n)_{n \in \mathbb{N}}$ and define a function f_P by the formula

$$f_P = \sum_{n \geq 0} 1_{[t_n, t_{n+1}[} f(t_{n+1}-).$$

For $t \in [t_n, t_{n+1}[$, we see that

$$\rho_i(f(t) - f_P(t)) = \rho_i(f(t) - f(t_{n+1}-)) \leq \omega_i(f; [t_n, t_{n+1}[) \leq \varepsilon.$$

Hence f_P is a right continuous step function satisfying the desired condition.

(ii) We shall show that the function f_P defined above is approximated from left by the left-approximation sequence (π_k) for f. Note that, by the definition of f_P,

$$f_P(s-) = \sum_{n \geq 0} 1_{]t_n, t_{n+1}]}(s) f(t_{n+1}-)$$

holds for all $s > 0$.

Now fix $s > 0$ arbitrarily and choose a unique $n_0 \in \mathbb{N}$ such that $s \in]t_{n_0}, t_{n_0+1}]$. If f is discontinuous at t_{n_0}, then we see that $\overline{\pi_k}(t_{n_0}) \to t_{n_0}$ by the same argument as the proof of Lemma 6.1. In this case, $t_{n_0} \leq \overline{\pi_k}(t_{n_0}) \leq \underline{\pi_k}(s) < s$ for large enough k and therefore we have

$$\lim_{k \to \infty} f_P(\underline{\pi_k}(s)) = f(t_{n_0+1}-) = f_P(s-).$$

Next, assume that f is continuous at t_{n_0} and f is not constant on any interval of the form $[t_{n_0}, t[$ for $t > t_{n_0}$. Take t' such that $t_{n_0} < t' < s$ and $f(t'-) \neq f(t_{n_0}) = f(t_{n_0}-)$. If $\underline{\pi_k}(t') \leq t_{n_0}$ for infinitely many k, then we can take a subsequence such that $\underline{\pi_{k_l}}(t') \leq t_{n_0}$ for all l. By definition, $\underline{\pi_{k_l}}(t') = \underline{\pi_{k_l}}(t_{n_0})$ and hence

$$f(t'-) = \lim_{l \to \infty} f(\underline{\pi_{k_l}}(t')) = \lim_{l \to \infty} f(\underline{\pi_{k_l}}(t_{n_0})) = f(t_{n_0}-),$$

which contradicts the assumption on t'. This shows that $t_{n_0} < \underline{\pi_k}(t') \leq \underline{\pi_k}(s) < s$ holds for large enough k and consequently

$$\lim_{k\to\infty} f_P(\underline{\pi_k}(s)) = f(t_{n_0+1}-) = f_P(s-).$$

Finally, assume that f is continuous at t_{n_0} and f is constant on $[t_{n_0}, t[$ for some $t > t_{n_0}$. In this case, we have $\omega_i(f; [t_{n_0}, \infty[) = 0$ by the definition of $P = (t_n)$. Also note that $t_{n_0} \leq \underline{\pi_k}(s) < s$ or $\underline{\pi_k}(t_{n_0}) = \underline{\pi_k}(s)$ holds for each k. In both cases,

$$\lim_{k\to\infty} f_P(\underline{\pi_k}(s)) = f(t_{n_0}-) = f(t_{n_0+1}-) = f_P(s-).$$

The second equality follows from the continuity of f at t_{n_0} and the property that $\omega_i(f; [t_{n_0}, \infty[) = 0$. This completes the proof. □

Finally, we start dealing with the proof of Lemma 3.10.

Proof of Lemma 3.10. First, fix $t > 0$ and $\varepsilon > 0$ arbitrarily and choose a compact set $K \subset E_1$ satisfying

$$B(\delta X_t, \delta X_t)(\mathcal{I} \cap [0,t]) \cup \delta Q_B(X,X)_t(\mathcal{I} \cap [0,t]) \subset K.$$

By Lemma 6.4, we can find an $\mathcal{L}_c(E_1, G)$-valued right continuous step function

$$h = \sum_{i\geq 1} 1_{[\tau_{i-1}, \tau_i[} a_i$$

so that $\rho_K(h(s) - \xi(s)) \leq \varepsilon$ holds for all $s \in [0,t]$ and (π_n) approximates h from the left. Then, for all $z^* \in G^*$,

$$\begin{aligned}
&\left| \left\langle z^*, \sum_{I\in\pi_n} \xi B(\delta X_t, \delta X_t)(I) \right\rangle - \left\langle z^*, \int_{]0,t]} \xi_{u-} \mathrm{d}Q_B(X,X)_u \right\rangle \right| \\
&\quad \leq \left| \sum_{I\in\pi_n} \langle z^*\xi, B(\delta X_t, \delta X_t)\rangle (I) - \sum_{I\in\pi_n} \langle z^*h, B(\delta X_t, \delta X_t)\rangle (I) \right| \\
&\qquad + \left| \left\langle z^*, \sum_{]r,s]\in\pi_n} hB(\delta X_t, \delta X_t)(I) \right\rangle - \left\langle z^*, \int_{]0,t]} h(u-) \mathrm{d}Q_B(X,X)_u \right\rangle \right| \\
&\qquad + \left| \int_{]0,t]} \langle z^*h(u-), \mathrm{d}Q_B(X,X)_u \rangle - \int_{]0,t]} \langle z^*\xi(u-), \mathrm{d}Q_B(X,X)_u \rangle \right|. \qquad (6.5)
\end{aligned}$$

We will observe the behaviour of each part of the right-hand side. We can deduce from Lemma 6.3 that the second term converges to 0 as $n \to \infty$. By the choice of h, we find that

$$\left| \sum_{I\in\pi_n} \langle z^*\xi, B(\delta X_t, \delta X_t)\rangle (I) - \sum_{I\in\pi_n} \langle z^*h, B(\delta X_t, \delta X_t)\rangle (I) \right| \leq \varepsilon \|z^*\| \|B\| \sum_{I\in\pi_n} \|\delta X_t(I)\|_E^2 .$$

Therefore,

$$\varlimsup_{n\to\infty} \left| \sum_{I\in\pi_n} \langle z^*\xi, B(\delta X_t, \delta X_t)\rangle (I) - \sum_{I\in\pi_n} \langle z^*h, B(\delta X_t, \delta X_t)\rangle (I) \right| \leq \varepsilon \|z^*\| \|B\| V^2(X;\Pi)_t.$$

On the other hand, we have

$$\left| \int_{]0,t]} \langle z^*(h(u-) - \xi(u-)), \mathrm{d}Q_B(X,X)_u \rangle \right| \leq \|z^*\| \int_{]0,t]} \rho_K(h(u-) - \xi(u-)) \mathrm{d}V(Q_B(X,X))_t$$
$$\leq \varepsilon \|z^*\| V(Q_B(X,X))_t$$

by Proposition B.7. Consequently,

$$\varlimsup_{n\to\infty} \left| \left\langle z^*, \sum_{I\in\pi_n} \xi B(\delta X_t, \delta X_t)(I) - \int_{]0,t]} \xi_{u-} \mathrm{d}Q_B(X,X)_u \right\rangle \right|$$
$$\leq \varepsilon \|z^*\| \{V^2(X;\Pi)_t + V(Q_B(X,X))_t\}.$$

Because ε is chosen arbitrarily, we get the desired conclusion.

If $Q_B(X,X)$ is the strong quadratic variation, we replace (6.5) with a similar norm inequality. In this case, we see that the corresponding second term converges to 0 by the strong version of Lemma 6.3. The remaining terms are estimated in almost the same way as above. □

7 The Itô formula

This section is devoted to showing the Itô formula within our framework of the Itô–Föllmer calculus in Banach spaces. Let us begin by defining Itô–Föllmer integrals.

Definition 7.1. *Let E be a locally convex space, and let F and G be Banach spaces. Consider càdlàg paths H and X in E and F, respectively, and a continuous bilinear map $B\colon E\times F\to G$. Suppose that a sequence of partitions (π_n) approximates H from the left. We call the limit*

$$\int_0^t B(H_{s-}, \mathrm{d}X_s) = \int_{]0,t]} B(H_{s-}, \mathrm{d}X_s) \coloneqq \lim_{n\to\infty} \sum_{]r,s]\in\pi_n} B\left(H_r, \delta X_t(]r,s])\right) \in G$$

the (strong) Itô–Föllmer integral of H with respect to X along (π_n) *if it exists. If this convergence holds in the weak topology, we call the limit the* weak Itô–Föllmer integral.

Similarly, the strong and the weak Itô–Föllmer integral for a $B' \in \mathcal{L}^{(2)}(F,E;G)$ are defined as the limit

$$\int_{]0,t]} B'(\mathrm{d}X_s, H_{s-}) = \lim_{n\to\infty} \sum_{]r,s]\in\pi_n} B'\left(\delta X_t(]r,s]), H_r\right) \in G$$

with the corresponding topology.

If B is the canonical bilinear map $\otimes\colon E\times F\to E\widehat{\otimes}_\alpha F$, we write

$$\int_{]0,t]} B(H_{s-}, \mathrm{d}X_s) = \int_{]0,t]} H_{s-} \otimes \mathrm{d}X_s.$$

Remark 7.2. The Itô–Föllmer integral of Definition 7.1 inherits the bilinear property from $B\in\mathcal{L}^{(2)}(E,F;G)$ in the following sense.

(i) Suppose that the following two Itô–Föllmer integrals exist:

$$\int_{]0,t]} B(H_{s-}, \mathrm{d}X_s), \qquad \int_{]0,t]} B(K_{s-}, \mathrm{d}X_s).$$

Then, for every α, $\beta\in\mathbb{R}$, the Itô–Föllmer integral of $\alpha H + \beta K$ with respect to X exists and satisfies

$$\int_{]0,t]} B(H_{s-} + K_{s-}, \mathrm{d}X_s) = \alpha \int_{]0,t]} B(H_{s-}, \mathrm{d}X_s) + \beta \int_{]0,t]} B(K_{s-}, \mathrm{d}X_s).$$

(ii) Suppose that the following two Itô–Föllmer integrals exist:

$$\int_{]0,t]} B(H_{s-}, \mathrm{d}X_s), \qquad \int_{]0,t]} B(H_{s-}, \mathrm{d}Y_s).$$

Then, for every α, $\beta \in \mathbb{R}$, the Itô–Föllmer integral of H with respect to $\alpha X + \beta Y$ exists and satisfies

$$\int_{]0,t]} B(H_{s-}, \mathrm{d}(\alpha X + \beta Y)_s) = \alpha \int_{]0,t]} B(H_{s-}, \mathrm{d}X_s) + \beta \int_{]0,t]} B(H_{s-}, \mathrm{d}Y_s).$$

First, we consider the case where the integrator is a path of finite variation. From the dominated convergence theorem, we can easily deduce the following proposition.

Proposition 7.3. *Let $H \in D(\mathbb{R}_{\geq 0}, E)$, $A \in FV(\mathbb{R}_{\geq 0}, F)$, and $B \in \mathcal{L}^{(2)}(E, F; G)$. If a sequence of partitions (π_n) approximates H from the left, we have*

$$(\mathrm{IF}) \int_{]0,t]} B(H_{s-}, \mathrm{d}A_s) = (\mathrm{S}) \int_{]0,t]} B(H_{s-}, \mathrm{d}A_s).$$

Here, the integral of the left-hand side is the Itô–Föllmer integral by Definition 7.1, and that of the right-hand side is the usual Stieltjes integral.

Now we start to prove our main theorem.

Proof of Theorem 3.6. We show the weak convergence of the Itô–Föllmer integral.

First, fix $t > 0$ arbitrarily and choose compact convex sets $K_1 \subset F$, $K_2 \subset E$, and $K_3 \subset G$ such that

$$A([0,t]) \subset K_1, \quad X([0,t]) \subset K_2, \quad B(\delta X, \delta X)(\mathcal{I} \cap [0,t]) \subset K_3.$$

Step 1: Convergence of the summation in the formula (3.3). In this step, we confirm that the summation of jump terms converges absolutely. This is proved by the following estimate, which follows from Taylor's formula (Proposition A.3):

$$\begin{aligned}
&\sum_{0<s\leq t} \|f(A_s, X_s) - f(A_{s-}, X_{s-}) - D_x f(A_{s-}, X_{s-})\Delta X_s\| \\
&\quad \leq \sum_{0<s\leq t} \|f(A_{s-}, X_s) - f(A_{s-}, X_{s-}) - D_x f(A_{s-}, X_{s-})\Delta X_s\| + \sum_{0<s\leq t} \|f(A_s, X_s) - f(A_{s-}, X_s)\| \\
&\quad \leq \sup_{(a,x)\in K_1\times K_2} \|D_B^2 f(a,x)\| \|B\| \sum_{0<s\leq t} \|\Delta X_s\|^2 + \sup_{(a,x)\in K_1\times K_2} \|D_a f(a,x)\| \sum_{0<s\leq t} \|\Delta A_s\| \\
&\quad < \infty.
\end{aligned}$$

Note that the uniform boundedness principle combined with the strong continuity shows that $\sup_{(a,x)} \left\|D_B^2 f(a,x)\right\|$ and $\sup_{(a,x)} \|D_a f(a,x)\|$ are finite.

Moreover, we see that

$$\begin{aligned}
\sum_{0<s\leq t} \|D_x^2 f(A_{s-}, X_{s-})\Delta X_s^{\otimes 2}\| &= \sum_{0<s\leq t} \|D_B^2 f(A_{s-}, X_{s-}) B(\Delta X_s^{\otimes 2})\| \\
&\leq \sup_{(a,x)\in K_1\times K_2} \left\|D_B^2 f(a,x)\right\| \|B\| \sum_{0<s\leq t} \|\Delta X_s\|^2 < \infty
\end{aligned}$$

and

$$\sum_{0<s\leq t} |D_a f(A_{s-}, X_{s-})\Delta A_s| \leq \sup_{(a,x)\in K_1\times K_2} \|D_a f(a,x)\| \sum_{0<s\leq t} \|\Delta A_s\| < \infty.$$

This shows that equation (3.1) is equivalent to the following equation:

$$\begin{aligned} f(A_t, X_t) - f(A_0, X_0) &- \int_{]0,t]} D_x f(A_{s-}, X_{s-})\,\mathrm{d}X_s \\ &= \int_{]0,t]} D_a f(A_{s-}, X_{s-})\,\mathrm{d}A_s + \frac{1}{2}\int_{]0,t]} D_B^2 f(A_{s-}, X_{s-})\,\mathrm{d}Q_B(X,X)_s \\ &\quad + \sum_{0<s\le t} \{\Delta f(A_s, X_s) - D_x f(A_{s-}, X_{s-})\Delta X_s\} \\ &\quad - \sum_{0<s\le t} D_a f(A_{s-}, X_{s-})\Delta A_s - \frac{1}{2}\sum_{0<s\le t} D_B^2 f(A_{s-}, X_{s-}) B(\Delta X_s, \Delta X_s). \end{aligned} \tag{7.1}$$

We will therefore prove (7.1) instead of (3.1).

Step 2: The Taylor expansion. Let $I =]u, v] \in \pi_n$ and consider the first-order Taylor expansion with respect to the variable a on $I \cap [0,t] =]u \wedge t, v \wedge t]$. Then we have

$$f(A_{v\wedge t}, X_{v\wedge t}) - f(A_{u\wedge t}, X_{v\wedge t}) = D_a f(A, X)\delta A_t(I) + r_t(I) \tag{7.2}$$

where

$$r_t(I) = \int_{[0,1]} \{D_a f\left(A_{u\wedge t} + \theta\delta A_t(I), X_{v\wedge t}\right) - D_a f\left(A_{u\wedge t}, X_{u\wedge t}\right)\}\,\delta A_t(I)\,\mathrm{d}\theta.$$

Here, recall that the notation '$D_a f(A,X)\delta A_t$' was introduced in the second paragraph of Section 3.

Next, we consider the second-order Taylor expansion

$$f(A_{u\wedge t}, X_{v\wedge t}) - f(A_{u\wedge t}, X_{u\wedge t}) = D_x f(A, X)\delta X_t(I) + \frac{1}{2}D_x^2 f(A, X)(\delta X_t)^{\otimes 2}(I) + R_t(I), \tag{7.3}$$

with $R_t(I)$ given by

$$\begin{aligned} R_t(I) &= \frac{1}{2}\int_{[0,1]} (1-\theta)\left\{D_x^2 f(A_{u\wedge t}, X_{u\wedge t} + \theta\delta X_t(I)) - D_x^2 f(A_{u\wedge t}, X_{u\wedge t})\right\}\delta X_t \otimes \delta X_t(I)\,\mathrm{d}\theta \\ &= \frac{1}{2}\int_{[0,1]} (1-\theta)\left\{D_B^2 f(A_{u\wedge t}, X_{u\wedge t} + \theta\delta X_t(I)) - D_B^2 f(A_{u\wedge t}, X_{u\wedge t})\right\} B(\delta X_t, \delta X_t)(I)\,\mathrm{d}\theta. \end{aligned}$$

By the definition of $D_B^2 f$, we can rewrite (7.3) as

$$f(A_{u\wedge t}, X_{v\wedge t}) - f(A_{u\wedge t}, X_{u\wedge t}) = D_x f(A, X)\delta X_t(I) + \frac{1}{2}D_B^2 f(A, X) B(\delta X_t^{\otimes 2})(I) + R_t(I), \tag{7.4}$$

Combining equations (7.3) and (7.4), we obtain

$$\delta f(A, X)_t = D_a f(A, X)\delta A_t + D_x f(A, X)\delta X_t + \frac{1}{2}D_B f(A, X) B(\delta X_t^{\otimes 2}) + r_t + R_t. \tag{7.5}$$

Now take $\varepsilon > 0$ arbitrarily. For simplicity, write $D = D(A, X)$, $D_\varepsilon = D_\varepsilon(A, X)$, and $D^\varepsilon = D^\varepsilon(A, X)$. Using the notations $e^1_{D_\varepsilon}$ and $e^2_{D_\varepsilon}$ introduced in Section 5, we can derive from (7.5) the equation

$$\begin{aligned} \delta f(A, X)_t - \delta f(A, X)_t e^1_{D_\varepsilon} &= D_x f(A, X)\delta X_t - D_x f(A, X)\delta X_t e^1_{D_\varepsilon} \\ &\quad + D_a f(A, X)\delta A_t - D_a f(A, X)\delta A_t e^1_{D_\varepsilon} + r_t e^2_{D_\varepsilon} \\ &\quad + \frac{1}{2}D_B^2 f(A, X) B(\delta X_t^{\otimes 2}) - \frac{1}{2}D_B^2 f(A, X) B(\delta X_t^{\otimes 2}) e^1_{D_\varepsilon} + R_t e^2_{D_\varepsilon}. \end{aligned}$$

Moreover, by summing up this equality along π_n, we see that

$$\begin{aligned}
&f(A_t, X_t) - f(A_0, X_0) - \sum_{I \in \pi_n} D_x f(A, X)\delta X_t(I) \\
&\quad = \sum_{I \in \pi_n} \delta f(A, X)_t(I) e^1_{D_\varepsilon}(I) - \sum_{I \in \pi_n} D_x f(A, X)\delta X_t(I) e^1_{D_\varepsilon}(I) \\
&\qquad + \sum_{I \in \pi_n} D_a f(A, X)\delta A_t(I) - \sum_{I \in \pi_n} D_a f(A, X)\delta A_t(I) e^1_{D_\varepsilon}(I) \\
&\qquad + \sum_{I \in \pi_n} \frac{1}{2} D_B^2 f(A, X) B(\delta X_t^{\otimes 2})(I) - \sum_{I \in \pi_n} \frac{1}{2} D_B^2 f(A, X) B(\delta X_t^{\otimes 2})(I) e^1_{D_\varepsilon}(I) \\
&\qquad + \sum_{I \in \pi_n} r_t(I) e^2_{D_\varepsilon}(I) + \sum_{I \in \pi_n} R_t(I) e^2_{D_\varepsilon}(I) \\
&\quad =: I_1^{(n)}(t) - I_2^{(n)}(t) + I_3^{(n)}(t) - I_4^{(n)}(t) + I_5^{(n)}(t) - I_6^{(n)}(t) + I_7^{(n)}(t) + I_8^{(n)}(t).
\end{aligned} \tag{7.6}$$

Step 3: Behaviour of $I_1^{(n)}(t), \dots, I_8^{(n)}(t)$ of (7.6). Since (π_n) satisfies Condition (C) for (A, X), we can easily verify that

$$\begin{aligned}
\lim_{n\to\infty} I_4^{(n)}(t) &= \sum_{s \in D_\varepsilon \cap [0,t]} D_a f(A_{s-}, X_{s-})\Delta A_s \\
\lim_{n\to\infty} I_2^{(n)}(t) &= \sum_{s \in D_\varepsilon \cap [0,t]} D_x f(A_{s-}, X_{s-})\Delta X_s \\
\lim_{n\to\infty} I_6^{(n)}(t) &= \frac{1}{2} \sum_{s \in D_\varepsilon \cap [0,t]} D_x^2 f(A_{s-}, X_{s-})(\Delta X_s)^{\otimes 2} = \frac{1}{2} \sum_{s \in D_\varepsilon \cap [0,t]} D_B^2 f(A_{s-}, X_{s-}) B(\Delta X_s^{\otimes 2}).
\end{aligned}$$

If $s \in D$ and $s \le t$, we can deduce from Proposition A.4 and the dominated convergence theorem that

$$\begin{aligned}
\delta f(A, X)_t(\pi_n(s)) &= \int_0^1 \{Df[(A, X)_{\underline{\pi_n}(s)} + \theta \delta(A, X)_t(\pi_n(s))]\}\{\delta(A, X)_t(\pi_n(s))\}\,\mathrm{d}\theta \\
&\xrightarrow[n\to\infty]{} \int_0^1 Df[(A, X)_{s-} + \theta \Delta(A, X)_s]\Delta(A, X)_s\,\mathrm{d}\theta = \Delta f(A, X)_s
\end{aligned}$$

Hence,

$$\lim_{n\to\infty} I_1^{(n)}(t) = \sum_{s \in D_\varepsilon \cap [0,t]} \Delta f(A_s, X_s).$$

By Lemma 3.10, we have

$$\lim_{n\to\infty} I_5^{(n)}(t) = \frac{1}{2} \int_{]0,t]} D_B^2 f(A_{s-}, X_{s-})\mathrm{d}Q_B(X, X)_s$$

in the weak topology.

The dominated convergence theorem (Proposition B.8) gives

$$\lim_{n\to\infty} I_3^{(n)}(t) = \int_{]0,t]} D_a f(A_{s-}, X_{s-})\mathrm{d}A_s.$$

It remains to estimate the residual terms. If $]u, v] \cap D_\varepsilon = \emptyset$, then

$$\begin{aligned}
\omega(X,]u, v] \cap [0, t]) &\le O_t^+(X - J(D_\varepsilon(X); X); \pi_n), \\
\omega(A,]u, v] \cap [0, t]) &\le O_t^+(A - J(D_\varepsilon(A); A); \pi_n).
\end{aligned}$$

Now we write, for convenience,

$$\alpha(\varepsilon,n) = O_t^+(X - J(D_\varepsilon(X);X);\pi_n),$$
$$\beta(\varepsilon,n) = O_t^+(A - J(D_\varepsilon(A);A);\pi_n).$$

By the assumption that (π_n) satisfies Condition (C) for (A,X) — and hence so does each of A and X — we see that

$$\varlimsup_{\varepsilon\downarrow\downarrow 0}\varlimsup_{n\to\infty}\alpha(\varepsilon,n) = 0, \qquad \varlimsup_{\varepsilon\downarrow\downarrow 0}\varlimsup_{n\to\infty}\beta(\varepsilon,n) = 0.$$

By definition,

$$\begin{aligned} I_7^{(n)}(t) &\le \sum_{]u,v]\in\pi_n} e^2_{D_\varepsilon}(]u,v]) \\ &\quad \times \int_{[0,1]} \|\{D_a f(A_{u\wedge t}+\theta\delta A_t(]u,v]), X_{v\wedge t}) - D_a f(A_{u\wedge t}, X_{u\wedge t})\}\,\delta A_t(]u,v])\|\,\mathrm{d}\theta \\ &\le V(A)_t \sup_{\substack{z,w\in K_1\times K_2 \\ |z-w|\le\alpha(\varepsilon,n)+\beta(\varepsilon,n)}} \rho_{K_1-K_1}(D_a f(z) - D_a f(w)). \end{aligned}$$

Similarly, we have

$$\begin{aligned} I_8^{(n)}(t) &\le \frac{1}{2}\sum_{]u,v]\in\pi_n} e^2_{D_\varepsilon}(]u,v])\int_{[0,1]}(1-\theta) \\ &\quad \cdot \left\|\left\{D_B^2 f(A_{u\wedge t}, X_{u\wedge t}+\theta\delta X_t(]u,v])) - D_B^2 f(A_{u\wedge t}, X_{u\wedge t})\right\}B(\delta X_t^{\otimes 2})(]u,v])\right\|\mathrm{d}\theta \\ &\le \sup_{\substack{z,w\in K_1\times K_2 \\ |z-w|\le\alpha(\varepsilon,n)}} \rho_{K_3}(D_B^2 f(z) - D_B^2 f(w))\|B\|\sum_{I\in\pi_n}\|\delta X_t(I)\|_E^2, \end{aligned}$$

Consequently, for every $z^*\in G^*$ with $\|z^*\|=1$,

$$\begin{aligned} &\varlimsup_{n\to\infty} |\langle z^*, (\text{RHS of (7.1)}) - (\text{RHS of (7.6)})\rangle| \\ &\le \left\|\sum_{s\in D^\varepsilon\cap[0,t]}\{\Delta f(A_s,X_s) - \langle D_x f(A_{s-},X_{s-}),\Delta X_s\rangle\}\right\| \\ &\quad + \left\|\sum_{s\in D^\varepsilon\cap[0,t]} D_a f(A_{s-},X_{s-})\Delta A_s\right\| + \left\|\sum_{s\in D^\varepsilon\cap[0,t]} D_x^2 f(A_{s-},X_{s-})(\Delta X_s)^{\otimes 2}\right\| \\ &\quad + \varlimsup_{n\to\infty}\sup_{\substack{z,w\in K_1 \\ |z-w|\le\alpha(\varepsilon,n)}} \rho_{K_3}(D_B^2 f(z) - D_B^2 f(w))V^2(X;\Pi)_t, \\ &\quad + \varlimsup_{n\to\infty}\sup_{\substack{z,w\in K_1\times K_2 \\ |z-w|\le\alpha(\varepsilon,n)+\beta(\varepsilon,n)}} \rho_{K_1-K_1}(D_a f(z) - D_a f(w))V(A)_t. \end{aligned} \tag{7.7}$$

Finally, by letting $\varepsilon\to 0$, we see that the right-hand side of (7.6) converges weakly to that of (7.1). This completes the proof for the weak case.

If $Q_B(X,X)$ is the strong B-quadratic variation, then $I_5^{(n)}$ converges in the norm topology by the strong version of Lemma 3.10. In this case, we obtain the norm convergence of the Riemannian sums by replacing (7.7) with a similar norm estimate. □

Combining Corollaries 3.7 and 5.9, we obtain the integration by parts formula. Note that the existence of the Itô–Föllmer integral $\int_0^t A_{s-}\otimes \mathrm{d}X_s$ follows from Corollary 3.7 and Proposition 7.3.

Corollary 7.4. *Let (π_n), X, and A satisfy the same assumptions as Corollary 3.7. Then,*

$$A_t\otimes X_t = \int_0^t \mathrm{d}A_s\otimes X_{s-} + \int_0^t A_{s-}\otimes\mathrm{d}X_s + {}^{\alpha}[A,X]_t.$$

A Differential calculus in Banach spaces

In this section, we review some basic results about differential calculus in Banach spaces.

Definition A.1. *Let E and F be Banach spaces and U an open subset of E.*

(i) A function $f\colon U \to F$ is Gâteaux differentiable at $x \in U$ if there exits an $L \in \mathcal{L}(E, F)$ such that

$$\lim_{t\to 0, t\neq 0} \frac{1}{t}\{f(x+th) - f(x)\} = Lh$$

for all $h \in E$. The function f is Gâteaux differentiable if it is Gâteaux differentiable at all points in U.

(ii) A function $f\colon U \to F$ is Fréchet differentiable at $x \in U$ if there exits an $L \in \mathcal{L}(E, F)$ such that

$$\|f(x+h) - f(x) - Lh\| = o(\|h\|).$$

The function f is Fréchet differentiable if it is Fréchet differentiable at all points in U.

The bounded operators L in (i) and (ii) are called Gâteaux and Fréchet derivatives, respectively, and denoted by the same symbol $Df(x)$. If f is Fréchet differentiable, then it is Gâteaux differentiable and both derivatives coincide. Hence this notation is consistent.

Higher-order Gâteaux and Fréchet derivatives are defined inductively by the formula $D^{n+1}f = DD^n f$. The n-th Gâteaux derivative $D^n f(x)$ at $x \in U$ is an element of $\mathcal{L}(E^{\widehat{\otimes} n}, F) \cong \mathcal{L}^{(n)}(E^n; F)$, where $\mathcal{L}^{(n)}(E^n; F)$ denotes the set of bounded n-linear maps. Moreover, if f is n-times Fréchet differentiable, the multilinear map $D^n f(x)\colon E^n \to F$ is symmetric.

Before introducing Taylor's theorem, we define a mild continuity condition for a function defined on a subset of a Banach space.

Definition A.2. *Let E be Banach spaces and U an open convex subset of E. A function $f\colon U \to T$ into a topological space is* continuous along line segments *if, for every x and y in U, its restriction to the line segment $\{\theta x + (1-\theta)y \mid \theta \in [0,1]\}$ is continuous.*

One can easily see that a Gâteaux differentiable function $f\colon U \to F$ is continuous along line segments.

The following proposition gives a version of Taylor's formula that we use in this article. It plays an essential role to prove the Itô formula in Section 7. Although that theorem seems classical, we give a proof for the reader's convenience.

Proposition A.3. *Let E and F be Banach spaces and U an open convex subset of E. Assume that $f\colon U \to F$ is n-times Gâteaux differentiable and $D^n f$ is strongly continuous along line segments. Then it admits Taylor's formula*

$$f(x+u) - f(x) = \sum_{1\leq k\leq n-1} \frac{1}{k!} D^k f(x)u + \int_0^1 \frac{(1-\theta)^{n-1}}{(n-1)!} D^n f(x+\theta u)u^{\otimes n}\, \mathrm{d}\theta$$

for all $x \in U$ and $u \in E$ with $x + u \in U$.

Proof. The proof is by induction on n.

Suppose that f is Gâteaux differentiable and Df is strongly continuous along line segments. Then the map $\theta \mapsto f(x+\theta u)$ has the continuous derivative

$$\frac{\mathrm{d}}{\mathrm{d}\theta} f(x+\theta u) = Df(x+\theta u)u.$$

Hence, by the fundamental theorem of calculus, we get

$$f(x+u)-f(x)=\int_0^1 Df(x+\theta u)u\,\mathrm{d}\theta.$$

Next, assume that the assertion holds for n. Let $f\colon U\to F$ be an $n+1$-times Gâteaux differentiable function whose derivative $D^{n+1}f$ is strongly continuous along line segments. By assumption, f satisfies

$$f(x+u)-f(x)=\sum_{1\le k\le n-1}\frac{1}{k!}D^k f(x)u^{\otimes k}+\int_0^1\frac{(1-\theta)^{n-1}}{(n-1)!}D^n f(x+\theta u)u^{\otimes n}\mathrm{d}\theta. \tag{A.1}$$

Computation by the Leibniz rule shows

$$\frac{\mathrm{d}}{\mathrm{d}\theta}\left(\frac{(1-\theta)^n}{n!}D^n f(x+\theta u)u^{\otimes n}\right)=-\frac{(1-\theta)^{n-1}}{(n-1)!}D^n f(x+\theta u)u^{\otimes n}+\frac{(1-\theta)^n}{n!}D^{n+1}f(x+\theta u)u^{\otimes n+1}.$$

Therefore,

$$-\frac{1}{n!}D^n f(x)u^{\otimes n}=-\int_0^1\frac{(1-\theta)^{n-1}}{(n-1)!}D^n f(x+\theta u)u^{\otimes n}\mathrm{d}\theta+\int_0^1\frac{(1-\theta)^n}{n!}D^{n+1}f(x+\theta u)u^{\otimes n+1}\mathrm{d}\theta. \tag{A.2}$$

Combining (A.1) and (A.2), we get the equation

$$f(x+u)-f(x)=\sum_{1\le k\le n}\frac{1}{k!}D^k f(x)u^{\otimes k}+\int_0^1\frac{(1-\theta)^n}{n!}D^{n+1}f(x+\theta u)u^{\otimes n+1}\mathrm{d}\theta.$$

This completes the proof. □

The following proposition is also used in the proof of the main theorem.

Proposition A.4. *Let E_1, E_2, and F be Banach spaces. Assume that $f\colon E_1\times E_2\to F$ satisfies the following conditions:*

(i) The map $x_1\mapsto f(x_1,x_2)$ is Gâteaux differentiable for all $x_2\in E_2$ and $D_{x_1}f\colon E_1\times E_2\to\mathcal{L}(E_1,F)$ is strongly continuous;
(ii) The map $x_2\mapsto f(x_1,x_2)$ is Gâteaux differentiable for all $x_1\in E_1$ and $D_{x_2}f\colon E_1\times E_2\to\mathcal{L}(E_2,F)$ is strongly continuous.

Then $f\colon E_1\times E_2\to F$ has the Gâteaux derivative given by $Df=(D_{x_1}f,D_{x_2}f)$ and $Df\colon E_1\times E_2\to\mathcal{L}(E_1\oplus E_2,F)$ is strongly continuous. Moreover, $Df|_K$ induces a continuous map $Df|_K\colon K\times(E_1\times E_2)\to F$ for every compact set $K\subset E_1\times E_2$.

Proof. Fix an $(x_1,x_2)\in E_1\times E_2$ and take an arbitrary directional vector $(h_1,h_2)\in E_1\times E_2$. Applying Taylor's formula to the first variable, we get

$$f(x_1+th_1,x_2+th_2)-f(x_1,x_2)=\int_0^1 D_{x_1}f(x_1+\theta th_1,x_2+th_2)th_1\,\mathrm{d}\theta+f(x_1,x_2+th_2)-f(x_1,x_2)$$

for all $t\neq 0$. Since $D_{x_1}f$ is strongly continuous, we see that

$$\lim_{t\to 0}D_{x_1}f(x_1+\theta th_1,x_2+th_2)h_1=D_{x_1}f(x_1,x_2)h_1,\qquad \sup_{\theta,t\in]0,1]}\|D_{x_1}f(x_1+\theta th_1,x_2+th_2)\|<\infty.$$

Note that the inequality above follows from the uniform boundedness principle. Therefore, by the dominated convergence theorem,

$$\frac{1}{t}\int_0^1 D_{x_1}f(x_1+\theta th_1,x_2+th_2)th_1\,\mathrm{d}\theta=\int_0^1 D_{x_1}f(x_1+\theta th_1,x_2+th_2)h_1\,\mathrm{d}\theta\xrightarrow[t\to 0]{} f(x_1,x_2)h_1.$$

This shows

$$\lim_{t\to 0, t\neq 0} \frac{1}{t}\{f(x_1+th_1, x_2+th_2) - f(x_1,x_2)\} = D_{x_1}f(x_1,x_2)h_1 + D_{x_2}f(x_1,x_2)h_2,$$

which means that $(D_{x_1}f, D_{x_2}f)$ is the Gâteaux derivative of f at (x_1,x_2). The strong continuity of $Df = (D_{x_1}f, D_{x_2}f)$ is obvious.

The last claim follows from Lemma A.5 below. □

Lemma A.5. *Let K be a compact topological space and let E and F be Banach spaces. Assume that a map $\varphi\colon K \to \mathcal{L}(E,F)$ is strongly continuous. Then, the map $\widetilde{\varphi}\colon K\times E \to F$ induced by φ is continuous.*

Proof. Fix $(t,x)\in K\times E$ and take an arbitrary net $(t_\lambda, x_\lambda)_{\lambda\in\Lambda}$ that converges to (t,x) in the product topology. Then

$$\|\varphi(t)x - \varphi(t_\lambda)x_\lambda\| \leq \|\varphi(t)x - \varphi(t_\lambda)x\| + \sup_{t\in K}\|\varphi(t)\|\|x - x_\lambda\|.$$

Because φ is strongly continuous, the first term on the right-hand side converges to 0. The strong continuity and the uniform boundedness principle imply that $\sup_{t\in K}\|\varphi(t)\| < \infty$. Hence the second term also converges to 0. As a consequence, we find that $\widetilde{\varphi}(t_\lambda, x_\lambda) \coloneqq \varphi(t_\lambda)x_\lambda$ converges to $\widetilde{\varphi}(t,x) \coloneqq \varphi(t)x$. □

B Vector integration

B.1 A brief review of vector integration in Banach spaces

In this subsection, we will review an integration theory for vector functions and vector measures of finite variation. Let E, F, and G be three Banach spaces and $B\colon F\times E \to G$ be a bounded bilinear map. We aim to introduce integrals of the form $\int B(f, \mathrm{d}\mu)$, where f is an F-valued 'nice' function and μ is an E-valued σ-additive measure defined on a δ-ring of subsets of $\mathbb{R}_{\geq 0}$. See Dinculeanu [20, Chapter 1 § 2] for details about the contents of this section. We also refer to Diestel and Uhl [18] and Dinculeanu [19], which are classical references in the theory of vector measures and integration.

Let $\mathcal{I}$ be the semiring of subsets of $\mathbb{R}_{\geq 0}$ consisting of all bounded intervals of the form $]a,b]$ and the singleton $\{0\}$. Moreover, let $\mathcal{D}$ be the δ-ring generated by $\mathcal{I}$. Note that both $\mathcal{I}$ and $\mathcal{D}$ generate the Borel σ-algebra. The variation of a σ-additive vector measure $\mu\colon \mathcal{D}\to E$ on a subset A of $\mathbb{R}_{\geq 0}$ is defined by the formula

$$|\mu|(A) \coloneqq \sup\left\{ \sum_{\lambda\in\Lambda} \|\mu(A_\lambda)\|_E \;\middle|\; (A_\lambda)_{\lambda\in\Lambda}\text{: a finite disjoint family of elements of } \mathcal{D}, \bigcup_{\lambda\in\Lambda} A_\lambda \subset A \right\}.$$

We say that μ has *finite variation* if $|\mu|(A) < \infty$ for all $A\in\mathcal{D}$. The measure μ has *bounded variation* whenever $|\mu|(\mathbb{R}_{\geq 0}) < \infty$. Since μ is assumed to be σ-additive, the variation measure $|\mu|\colon \mathcal{D}\to\mathbb{R}_{\geq 0}$ is also σ-additive. Then there is a unique σ-additive measure, denoted by the same symbol $|\mu|$, on the σ-algebra $\mathcal{B}(\mathbb{R}_{\geq 0})$ that extends $|\mu|\colon \mathcal{D}\to\mathbb{R}_{\geq 0}$.

An F-valued $\mathcal{D}$-simple function is a function $f\colon \mathbb{R}_{\geq 0}\to F$ of the form

$$f = \sum_{\lambda\in\Lambda} 1_{A_\lambda} a_\lambda,$$

where Λ is a finite set, $(a_\lambda)_{\lambda\in\Lambda}\in E^\Lambda$, and $(A_\lambda)_{\lambda\in\Lambda}\in\mathcal{D}^\Lambda$ is a disjoint family. We can immediately define the integral of f by

$$\int_{\mathbb{R}_{\geq 0}} B(f, \mathrm{d}\mu) = \sum_{\lambda\in\Lambda} B(a_\lambda, \mu(A_\lambda)).$$

Let $S_F(\mathcal{D})$ denote the set of all F-valued $\mathcal{D}$-simple functions. Then $S_F(\mathcal{D})$ is dense in the Lebesgue–Bochner space

$$L^1(|\mu|;F) := \left\{ f\colon \mathbb{R}_{\geq 0} \to F \,\middle|\, f \text{ is strongly } \mathcal{B}(\mathbb{R}_{\geq 0})\text{-measurable}, \int_{\mathbb{R}_{\geq 0}} \|f(s)\|_F \,|\mu|(\mathrm{d}s) < \infty \right\}.$$

Since the integration map $f \mapsto \int_{\mathbb{R}_{\geq 0}} B(f, \mathrm{d}\mu)$ is bounded and linear, it can be uniquely extended to a bounded linear operator $T_\mu\colon L^1(|\mu|;F) \to G$. If a strongly measurable function $f\colon \mathbb{R}_{\geq 0} \to F$ and an $A \in \mathcal{B}(\mathbb{R}_{\geq 0})$ satisfy $f1_A \in L^1(|\mu|;F)$, then we define the integral of f on A as

$$\int_A B(f(s), \mu(\mathrm{d}s)) = \int_A B(f, \mathrm{d}\mu) := T_\mu(f1_A).$$

In particular, every $f \in L^1_{\mathrm{loc}}(|\mu|;F)$ can be integrated on every bounded interval. We also use the notation

$$\int_A B(f, \mathrm{d}\mu) = \int_a^b B(f, \mathrm{d}\mu)$$

if A is a bounded interval of the form $A =]a, b]$. By a direct calculation, one can derive the inequality

$$\left\| \int_A B(f, \mathrm{d}\mu) \right\|_G \leq \|B\| \int_A \|f(s)\|_E |\mu|(\mathrm{d}s)$$

for all such A and f. This estimate guarantees that the dominated convergence theorem remains valid in this situation.

Finally, we introduce a decomposition of a vector measure on $\mathbb{R}_{\geq 0}$ into atomic and nonatomic parts. As before, let $\mu\colon \mathcal{D} \to E$ be a σ-additive measure of finite variation. Set

$$D = \{t \in \mathbb{R}_{\geq 0} \mid |\mu|(\{t\}) > 0\}.$$

Because μ has finite variation, we see that D is countable. Besides, since each singleton $\{t\}$ belongs to $\mathcal{D}$, we have the equality $\|\mu(\{t\})\| = |\mu|(\{t\})$ for all $t \in \mathbb{R}_{\geq 0}$. Now let

$$\mu^{\mathrm{d}} = \sum_{s \in D} \delta_s \mu(\{s\}),$$

where δ_s denotes the Dirac measure at $s \in \mathbb{R}_{\geq 0}$. Then μ^{d} is a σ-additive vector measure of finite variation that satisfies

$$|\mu^{\mathrm{d}}|(A) = \sum_{s \in D} \|\mu(\{s\})\|_E \, \delta_s(A) < \infty$$

for each $A \in \mathcal{D}$. If we define $\mu^{\mathrm{c}} = \mu - \mu^{\mathrm{d}}$, then μ^{c} and μ^{d} give a mutually singular decomposition of μ satisfying $|\mu| = |\mu^{\mathrm{c}}| + |\mu^{\mathrm{d}}|$. This decomposition of a measure gives a decomposition of an integral as

$$\int_{\mathbb{R}_{\geq 0}} B(f, \mathrm{d}\mu) = \int_{\mathbb{R}_{\geq 0}} B(f, \mathrm{d}\mu^{\mathrm{c}}) + \int_{\mathbb{R}_{\geq 0}} B(f, \mathrm{d}\mu^{\mathrm{d}})$$

for every $f \in L^1(|\mu|;F)$.

B.2 Extension of vector integration

In Section B.1, integrands are assumed to be strongly measurable with respect to the norm topology of F; i.e. integrands can necessarily be approximated pointwise by simple functions in the norm topology. A function $f\colon \mathbb{R}_{\geq 0} \to F$ that is càdlàg in a weaker topology may not satisfy this condition. We need, however, to integrate such functions for our Itô–Föllmer formula (Theorem 3.6). For this purpose, we extend the vector integration introduced in a previous section to a suitable setting.

As in Appendix B.1, let E and G be Banach spaces and let $\mu\colon \mathcal{D} \to E$ be a σ-additive vector measure of finite variation.

Lemma B.1. *The range $\mu(\mathcal{I} \cap [0,T])$ is relatively compact in E for all $T > 0$.*

Proof. First, define a function $F\colon \mathbb{R}_{\geq 0} \to E$ by the formula $F(t) = \mu([0,t])$. σ-additivity of μ implies that F is càdlàg in E. For a given $T > 0$, we can find a compact set K including the image $F([0,T])$. If $I =]a,b] \subset [0,T]$, then

$$\mu(]a,b]) = F(b) - F(a) \in K - K.$$

On the other hand, if $I = \{0\} \subset [0,T]$,

$$\mu(\{0\}) = F(0) \in K.$$

Thus $\mu(\mathcal{I}) \subset K \cup (K-K)$. Because K and $K-K$ are both compact, $\mu(\mathcal{I})$ is relatively compact. □

Lemma B.2. *Let $F = \mathcal{L}(E,G)$ and let $S_F(\mathcal{I})$ be the set of all F-valued $\mathcal{I}$-simple functions. Suppose that $T > 0$ and K is a compact set such that $\mu(\mathcal{I} \cap [0,T]) \subset K$. Then*

$$\left\| \int_{[0,T]} f \mathrm{d}\mu \right\|_G \leq \sup_{s \in [0,T]} \rho_K(f(s)) |\mu|([0,T]).$$

Here, recall that a seminorm ρ_K is defined by (2.1) in Section 2.1.

Proof. Let $f = \sum_{I \in \Lambda} 1_I a_I$ be an $\mathcal{I}$-simple function with Λ being disjoint. Then, by the definition of the integral and the variation of μ, we see that

$$\left\| \int_{[0,T]} f \mathrm{d}\mu \right\|_G \leq \sum_{I \in \Lambda} \rho_K(a_I) \|\mu(I \cap [0,T])\|_E \leq \sup_{s \in [0,T]} \rho_K(f(s)) |\mu|([0,T]).$$

This is the desired inequality. □

In what follows, let $D^-(\mathbb{R}_{\geq 0}, \mathcal{L}_{\mathrm{c}}(E,G))$ stands for the locally convex Hausdorff topological vector space consisting of all càglàd functions endowed with the topology of uniform convergence on compacta.

Lemma B.3. *The space $S_F(\mathcal{I})$ is dense in $D^-(\mathbb{R}_{\geq 0}, \mathcal{L}_{\mathrm{c}}(E,G))$.*

Proof. First, let $\mathcal{K}$ the family of all compact subsets of E and define a family of seminorms $(\rho_{T,K}; T > 0, K \in \mathcal{K})$ by the formula

$$\rho_{T,K}(f) = \sup_{s \in [0,T]} \rho_K(f(s)).$$

Then the topology of $D^-(\mathbb{R}_{\geq 0}, \mathcal{L}_{\mathrm{c}}(E,G))$ is induced by $(\rho_{T,K})$.

Set $\Lambda = \mathbb{R}_{>0} \times \mathcal{K} \times \mathbb{R}_{>0}$ and define an order on Λ by letting $(T_1, K_1, \varepsilon_1) \leq (T_2, K_2, \varepsilon_2)$ whenever $T_1 \leq T_2$, $K_1 \subset K_2$, and $\varepsilon_1 \geq \varepsilon_2$. This order makes Λ a directed set because $(T_1 \vee T_2, K_1 \cup K_2, \varepsilon_1 \wedge \varepsilon_2) \in \Lambda$. For each $\lambda = (T, K, \varepsilon)$, we can find a $\mathcal{I}$-simple function s_λ satisfying $\rho_{T,K}(s_\lambda - f) < \varepsilon$ as in the proof of Lemma 6.4. The net $(s_\lambda)_{\lambda \in \Lambda}$ converges to f in the topology of $D^-(\mathbb{R}_{\geq 0}, \mathcal{L}_{\mathrm{c}}(E,G))$. Indeed, take $\varepsilon_0 > 0$ and (T_0, K_0) arbitrarily. If $\lambda = (T, K, \varepsilon) \geq (T_0, K_0, \varepsilon_0)$, then

$$\rho_{T_0,K_0}(s_\lambda - f) \leq \rho_{T,K}(s_\lambda - f) < \varepsilon \leq \varepsilon_0.$$

Hence $(s_\lambda)_{\lambda \in \Lambda}$ converges to f in ρ_{T_0,K_0}. Since the choice of (T_0, K_0) is arbitrary, we can conclude that $S_F(\mathcal{I})$ is dense in $D^-(\mathbb{R}_{\geq 0}, \mathcal{L}_{\mathrm{c}}(E,G))$. □

Theorem B.4. *Let $T > 0$. The integration map*

$$S_F(\mathcal{I}) \ni f \longmapsto \int_{[0,T]} f \mathrm{d}\mu \in G$$

can be uniquely extended to a continuous linear map on $D^-(\mathbb{R}_{\geq 0}, \mathcal{L}_{\mathrm{c}}(E,G))$.

Proof. Define the integral map $I_T \colon \mathcal{S}_F(\mathcal{I}) \to G$ by $I_T(f) = \int_{[0,T]} f \mathrm{d}\mu$. Because $\mathcal{S}_F(\mathcal{I})$ is dense in $D^-(\mathbb{R}_{\geq 0}, \mathcal{L}_{\mathrm{c}}(E,G))$ by Lemma B.3, it suffices to show that I_T is linear and continuous. Refer to Schaefer and Wolff [66, Chapter III Section 1] for the extension of a continuous linear map between topological vector spaces.

The linearity of I_T is obvious. The continuity follows from Lemma B.2 and a standard continuity criterion for linear maps in topological vector spaces. See, e.g. Schaefer and Wolff [66, Chapter III Section 1 (1.1)]. □

We will show a variant of the dominated convergence theorem for our vector integrals.

Lemma B.5. *For each $f \in D^-(\mathbb{R}_{\geq 0}, \mathcal{L}_{\mathrm{c}}(E,G))$ and compact subset K of E, the function $t \mapsto \rho_K(f(t))$ is Borel measurable.*

Proof. If $f \in D^-(\mathbb{R}_{\geq 0}, \mathcal{L}_{\mathrm{c}}(E,G))$, then the map $t \mapsto \rho_K(f(t))$ is càglàd. Hence it is Borel measurable. □

Lemma B.6. *Let $f = \sum_{I \in \Lambda} 1_I a_I$ be an $\mathcal{L}_{\mathrm{c}}(E,G)$-valued $\mathcal{I}$-simple function such that Λ is disjoint. Then,*

$$\rho_K(f(t)) = \sum_{I \in \Lambda} \rho_K(a_I) 1_I(t), \qquad t \geq 0.$$

Proposition B.7. *Let $T > 0$ and let K be a compact set satisfying $\mu(\mathcal{I} \cap [0,T]) \subset K$. Then*

$$\left\| \int_{[0,T]} f \mathrm{d}\mu \right\|_G \leq \int_{[0,T]} \rho_K(f(s)) |\mu|(\mathrm{d}s) \tag{B.1}$$

for all $f \in D^-(\mathbb{R}_{\geq 0}, \mathcal{L}_{\mathrm{c}}(E,G))$.

Proof. If a simple function f has a disjoint representation $f = \sum_{I \in \Lambda} 1_I a_I$, then

$$\begin{aligned}
\left\| \int_{[0,T]} f \mathrm{d}\mu \right\|_G &\leq \sum_{I \in \Lambda} \rho_K(a_I) \| \mu(I \cap [0,T]) \|_E \\
&\leq \sum_{I \in \Lambda} \rho_K(a_I) |\mu|(I \cap [0,T]) \\
&= \int_{[0,T]} \rho_K(f(s)) |\mu|(\mathrm{d}s).
\end{aligned}$$

Note that the last equality follows from Lemma B.6. Therefore (B.1) holds on $S_F(\mathcal{I})$.

The general case is proved by approximation. □

Proposition B.8 (Dominated convergence theorem)**.** *Let $T > 0$ and let K be a compact subset of E $\mu(\mathcal{I} \cap [0,T]) \subset K$. Suppose that a sequence (f_n) and an element f in $D^-(\mathbb{R}_{\geq 0}, \mathcal{L}_{\mathrm{c}}(E,G))$ satisfy the following conditions:*

(i) The sequence (f_n) converges to f pointwise on $[0,T]$ with respect to ρ_K;
(ii) There is a $g \in L^1([0,T], |\mu|)$ such that $\rho_K(f_n(t)) \leq g(t)$ almost everywhere on $[0,T]$.

Then,

$$\left\| \int_{[0,T]} f_n \mathrm{d}\mu - \int_{[0,T]} f \mathrm{d}\mu \right\|_G \xrightarrow[n \to \infty]{} 0.$$

Proof. Proposition B.7 implies the inequality

$$\left\| \int_{[0,T]} f_n \mathrm{d}\mu - \int_{[0,T]} f \mathrm{d}\mu \right\|_G \leq \int_{[0,T]} \rho_K(f_n(s) - f(s)) |\mu|(\mathrm{d}s).$$

By applying the dominated convergence theorem to the integral on the right-hand side, we obtain the desired convergence. □

Proposition B.9. *If $f \in D^-(\mathbb{R}_{\geq 0}, \mathcal{L}_c(E, G))$, then*

$$\left\langle z^*, \int_{]0,t]} f(s)\mu(\mathrm{d}s) \right\rangle = \int_{]0,t]} \langle z^* f(s), \mu(\mathrm{d}s) \rangle$$

for all $z^* \in G^*$ and $t \in \mathbb{R}_{\geq 0}$.

Proof. First, assume that f has the form $f = \sum_{I \in \Lambda} 1_I a_I$ with $\Lambda \subset \mathcal{I}$ disjoint. By a direct calculation, we see that

$$\left\langle z^*, \int_{]0,t]} f(s)\mu(\mathrm{d}s) \right\rangle = \sum_{I \in \Lambda} z^* a_I \mu(I \cap]0,t]) = \int_{]0,t]} \langle z^* f(s), \mu(\mathrm{d}s) \rangle .$$

Hence the formula holds for simple functions. It can be extended to general integrands by the density argument. □

References

[1] Anna Ananova and Rama Cont, *Pathwise integration with respect to paths of finite quadratic variation*, Journal de Mathématiques Pures et Appliquées **107** (2017), no. 6, 737–757. MR3650323

[2] Daniel Bartl, Michael Kupper, and Ariel Neufeld, *Stochastic integration and differential equations for typical paths*, Electronic Journal of Probability **24** (2019), no. 97, 1–21. MR4017115

[3] Klaus Bichteler, *Stochastic integration and L^p-theory of semimartingales*, The Annals of Probability **9** (1981), no. 1, 49–89. MR606798

[4] Patrick Billingsley, *Convergence of probability measures*, 2 ed., Wiley Series in Probability and Statistics, John Wiley & Sons, 1999. MR1700749

[5] J. K. Brooks and N. Dinculeanu, *Stochastic integration in Banach spaces*, Advances in Mathematics **81** (1990), no. 1, 99–104. MR1051225

[6] Henry Chiu and Rama Cont, *On pathwise quadratic variation for càdlàg functions*, Electronic Communications in Probability **23** (2018), no. 85, 1–12. MR3882226

[7] Rama Cont and Purba Das, *Quadratic variation along refining partitions: Constructions and examples*, Journal of Mathematical Analysis and Applications **512** (2022), no. 2, 126173. MR4401981

[8] Rama Cont and David-Antoine Fournié, *Change of variable formulas for non-anticipative functionals on path space*, Journal of Functional Analysis **259** (2010), no. 4, 1043–1072. MR2652181

[9] Rama Cont and David-Antoine Fournié, *A functional extension of the Ito formula*, Comptes Rendus Mathematique **348** (2010), no. 1, 57–61. MR2586744

[10] Rama Cont and David-Antoine Fournié, *Functional Itô calculus and stochastic integral representation of martingales*, The Annals of Probability **41** (2013), no. 1, 109–133. MR3059194

[11] Rama Cont and Nicolas Perkowski, *Pathwise integration and change of variable formulas for continuous paths with arbitrary regularity*, Transactions of the American Mathematical Society. Series B **6** (2019), 161–186. MR3937343

[12] Giuseppe Da Prato and Jerzy Zabczyk, *Stochastic equations in infinite dimensions*, 2 ed., Encyclopedia of Mathematics and its Applications, no. 152, Cambridge University Press, 2014. MR3236753

[13] Mark Davis, Jan Obłój, and Vimal Raval, *Arbitrage bounds for prices of weighted variance swaps*, Mathematical Finance. An International Journal of Mathematics, Statistics and Financial Economics **24** (2014), no. 4, 821–854. MR3274933

[14] Mark Davis, Jan Obłój, and Pietro Siorpaes, *Pathwise stochastic calculus with local times*, Annales de l'Institut Henri Poincaré Probabilités et Statistiques **54** (2018), no. 1, 1–21. MR3765878

[15] Marzia De Donno and Maurizio Pratelli, *Stochastic integration with respect to a sequence of semimartingales*, In Memoriam Paul-André Meyer: Séminaire de Probabilités XXXIX (Michel Émery and Marc Yor, eds.), Lecture Notes in Mathematics, no. 1874, Springer Berlin Heidelberg, Berlin, Heidelberg, 2006, pp. 119–135. MR2276892

[16] Andreas Defant and Klaus Floret, *Tensor norms and operator ideals*, North-Holland Mathematics Studies, no. 176, North-Holland Publishing Co., 1993. MR1209438

[17] Cristina Di Girolami, Giorgio Fabbri, and Francesco Russo, *The covariation for Banach space valued processes and applications*, Metrika. International Journal for Theoretical and Applied Statistics **77** (2014), no. 1, 51–104. MR3152020

[18] J. Diestel and J. J. Uhl, Jr., *Vector measures*, Mathematical Surveys and Monographs, no. 15, American Mathematical Society, 1977. MR0453964

[19] Nicolae Dinculeanu, *Vector measures*, International Series of Monographs in Pure and Applied Mathematics, no. 95, Pergamon Press, Oxford-New York-Toronto, Ont.; VEB Deutscher Verlag der Wissenschaften, Berlin, 1967. MR0206190

[20] Nicolae Dinculeanu, *Vector integration and stochastic integration in Banach spaces*, Pure and Applied Mathematics: A Wiley Series of Texts, Monographs, and Tracts, John Wiley & Sons, 2000. MR1782432

[21] Bruno Dupire, *Functional itô calculus*, Bloomberg Portfolio Research Paper No. 2009-04-FRONTIERS (2009).

[22] H. Föllmer, *Calcul d'Itô sans probabilités*, Lecture Notes in Mathematics, no. 850, pp. 143–150, Springer Berlin Heidelberg, 1981. MR622559

[23] H. Föllmer, *Dirichlet processes*, Stochastic Integrals: Proceedings of the LMS Durham Symposium, July 7 – 17, 1980 (Berlin, Heidelberg) (David Williams, ed.), Lecture Notes in Mathematics, no. 851, Springer Berlin Heidelberg, 1981, pp. 476–478. MR621001

[24] Hans Föllmer and Alexander Schied, *Probabilistic aspects of finance*, Bernoulli **19** (2013), no. 4, 1306–1326. MR3102553

[25] Peter K. Friz and Martin Hairer, *A course on rough paths*, Universitext, Springer International Publishing, 2014. MR3289027

[26] Peter K. Friz and Atul Shekhar, *General rough integration, Lévy rough paths and a Lévy-Kintchine-type formula*, The Annals of Probability **45** (2017), no. 4, 2707–2765. MR3693973

[27] Peter K. Friz and Huilin Zhang, *Differential equations driven by rough paths with jumps*, Journal of Differential Equations **264** (2018), no. 10, 6226–6301. MR3770049

[28] J. B. Gravereaux and Jean Pellaumail, *Formule de Ito pour des processus non continus à valeurs dans des espaces de Banach*, Annales de l'I.H.P. Probabilités et statistiques **10** (1974), no. 4, 399–422 (fr). MR373003

[29] M. Gubinelli, *Controlling rough paths*, Journal of Functional Analysis **216** (2004), no. 1, 86–140. MR2091358

[30] Massimiliano Gubinelli and Samy Tindel, *Rough evolution equations*, The Annals of Probability **38** (2010), no. 1, 1–75. MR2599193

[31] I. Gyöngy, *On stochastic equations with respect to semimartingales III*, Stochastics **7** (1982), no. 4, 231–254. MR674448

[32] I. Gyöngy and N. V. Krylov, *On stochastic equations with respect to semimartingales I*, Stochastics **4** (1980), no. 1, 1–21. MR587426

[33] I. Gyöngy and N. V. Krylov, *On stochastic equations with respect to semimartingales II. Itô formula in Banach spaces*, Stochastics **6** (1982), no. 3–4, 153–173. MR665398

[34] Yuki Hirai, *Itô-Föllmer integrals and their applications to finance*, Master's thesis, Osaka University, 2016.

[35] Yuki Hirai, *Remarks on Föllmer's pathwise Itô calculus*, Osaka Journal of Mathematics **56** (2019), no. 3, 631–660. MR3981304

[36] Yuki Hirai, *A study on Föllmer's Pathwise Itô calculus*, Ph.D. thesis, Osaka University, 2021.

[37] Yuki Hirai, *Itô–Föllmer calculus in Banach spaces II: Transformations of quadratic variations*, arXiv:2105.08262v2, 2022, Version 2.

[38] Rajeeva L. Karandikar, *On pathwise stochastic integration*, Stochastic Processes and their Applications **57** (1995), 11–18. MR1327950

[39] Hiroshi Kunita, *Stochastic integrals based on martingales taking values in Hilbert space*, Nagoya Mathematical Journal **38** (1970), 41–52. MR264754

[40] Rafał M. Łochowski, *On pathwise stochastic integration with respect to semimartingales*, Probability and Mathematical Statistics **34** (2014), no. 1, 23–43. MR3225999

[41] Rafał M. Łochowski, Jan Obłój, David J. Prömel, and Pietro Siorpaes, *Local times and Tanaka–Meyer formulae for càdlàg paths*, Electronic Journal of Probability **26** (2021), 1–29. MR4269207

[42] Rafał M. Łochowski, Nicolas Perkowski, and David J. Prömel, *A superhedging approach to stochastic integration*, Stochastic Processes and their Applications **128** (2018), no. 12, 4078–4103. MR3906979

[43] Rafał Marcin Łochowski, *Quadratic variation of a càdlàg semimartingale as a.s. limit of the normalized truncated variations*, Stochastics **91** (2019), 629–642. MR3939349

[44] Terry J. Lyons, *Differential equations driven by rough signals.*, Revista Matemática Iberoamericana **14** (1998), no. 2, 215–310 (eng). MR1302388

[45] Michel Métivier, *Intégrale stochastique par rapport à des processus à valeurs dans un espace de Banach réflexif*, Publications mathématiques et informatique de Rennes (1972), no. 2, 99–157 (fr). MR0365708

[46] Michel Métivier, *Semimartingales: a course on stochastic processes*, De Gruyter studies in mathematics, no. 2, Walter de Gruyter, 1982. MR688144

[47] Michel Métivier and Jean Pellaumail, *Stochastic integration*, Probability and Mathematical Statistics, Academic Press, 1980. MR578177

[48] Michel Métivier and Giovanni Pistone, *Une formule d'isométrie pour l'intégrale stochastique hilbertienne et équations d'évolution linéaires stochastiques*, Zeitschrift für Wahrscheinlichkeitstheorie und Verwandte Gebiete **33** (1975), no. 1, 1–18. MR383527

[49] P. A. Meyer, *Notes sur les integrales stochastiques. I integrales hilbertiennes*, Séminaire de Probabilités XI: Université de Strasbourg (C. Dellacherie, P. A. Meyer, and M. Weil, eds.), Lecture Notes in Mathematics, no. 581, Springer Berlin Heidelberg, 1977, pp. 446–462. MR0501333

[50] R. Mikulevicius and B. L. Rozovskii, *Normalized stochastic integrals in topological vector spaces*, Séminaire de Probabilités XXXII (Jacques Azéma, Marc Yor, Michel Émery, and Michel Ledoux, eds.), Springer Berlin Heidelberg, 1998, pp. 137–165. MR1655149

[51] R. Mikulevicius and B. L. Rozovskii, *Martingale problems for stochastic PDE's*, Stochastic partial differential equations: six perspectives, Mathematical Surveys and Monographs, no. 64, American Mathematical Society, 1999, pp. 243–325. MR1661767

[52] Yuliya Mishura and Alexander Schied, *Constructing functions with prescribed pathwise quadratic variation*, Journal of Mathematical Analysis and Applications **442** (2016), no. 1, 117–137. MR3498321

[53] Marcel Nutz, *Pathwise construction of stochastic integrals*, Electronic Communications in Probability **17** (2012), no. 24, 1–7. MR2950190

[54] Jean Pellaumail, *Sur l'intégrale stochastique et la décomposition de Doob-Meyer*, Astérisque, no. 9, Société mathématique de France, 1973 (fr). MR339333

[55] Nicolas Perkowski and David J. Prömel, *Local times for typical price paths and pathwise Tanaka formulas*, Electronic Journal of Probability **20** (2015), no. 46, 1–15. MR3339866

[56] Nicolas Perkowski and David J. Prömel, *Pathwise stochastic integrals for model free finance*, Bernoulli **22** (2016), no. 4, 2486–2520. MR3498035

[57] Maurizio Pratelli, *Integration stochastique et geometrie des espaces de Banach*, Séminaire de Probabilités XXII (Jacques Azéma, Marc Yor, and Paul André Meyer, eds.), Springer Berlin Heidelberg, 1988, pp. 129–137. MR960517

[58] Francesco Russo and Pierre Vallois, *Intégrales progressive, rétrograde et symétrique de processus non adaptés*, Comptes Rendus de l'Académie des Sciences. Série I. Mathématique **312** (1991), no. 8, 615–618. MR1101044

[59] Francesco Russo and Pierre Vallois, *Forward, backward and symmetric stochastic integration*, Probability Theory and Related Fields **97** (1993), no. 3, 403–421. MR1245252

[60] Francesco Russo and Pierre Vallois, *Noncausal stochastic integration for làd làg processes*, Stochastic analysis and related topics (Oslo, 1992), Stochastics Monographs, no. 8, Gordon and Breach, Montreux, 1993, pp. 227–263. MR1268016

[61] Francesco Russo and Pierre Vallois, *The generalized covariation process and ito formula*, Stochastic Processes and their Applications **59** (1995), no. 1, 81–104. MR1350257

[62] Francesco Russo and Pierre Vallois, *Itô formula for C^1-functions of semimartingales*, Probability Theory and Related Fields **104** (1996), no. 1, 27–41. MR1367665

[63] Francesco Russo and Pierre Vallois, *Stochastic calculus with respect to continuous finite quadratic variation processes*, Stochastics and Stochastic Reports **70** (2000), no. 1–2, 1–40. MR1785063

[64] Francesco Russo and Pierre Vallois, *Elements of stochastic calculus via regularization*, Séminaire de Probabilités XL (Catherine Donati-Martin, Michel Émery, Alain Rouault, and Christophe Stricker, eds.), Springer Berlin Heidelberg, 2007, pp. 147–185. MR2409004

[65] Raymond A. Ryan, *Introduction to tensor products of Banach spaces*, Springer Monographs in Mathematics, Springer-Verlag London, 2002. MR1888309

[66] H. H. Schaefer and M. P. Wolff, *Topological vector spaces*, 2 ed., Graduate Texts in Mathematics, no. 3, Springer-Verlag New York, 1999. MR1741419

[67] Alexander Schied, *Model-free CPPI*, Journal of Economic Dynamics & Control **40** (2014), 84–94. MR3166493

[68] Alexander Schied, *On a class of generalized Takagi functions with linear pathwise quadratic variation*, Journal of Mathematical Analysis and Applications **433** (2016), no. 2, 974–990. MR3398747

[69] Alexander Schied, Leo Speiser, and Iryna Voloshchenko, *Model-free portfolio theory and its functional master formula*, SIAM Journal on Financial Mathematics **9** (2018), no. 3, 1074–1101. MR3856206

[70] Dieter Sondermann, *Introduction to stochastic calculus for finance*, Lecture Notes in Economics and Mathematical Systems, no. 579, Springer-Verlag Berlin Heidelberg, 2006. MR2254170

[71] J. M. A. M. van Neerven, M. C. Veraar, and L. Weis, *Stochastic integration in UMD Banach spaces*, The Annals of Probability **35** (2007), no. 4, 1438–1478. MR2330977

[72] Mark Veraar and Ivan Yaroslavtsev, *Cylindrical continuous martingales and stochastic integration in infinite dimensions*, Electronic Journal of Probability **21** (2016), no. Paper No. 59, 53. MR3563887

[73] Vladimir Vovk, *Continuous-time trading and the emergence of volatility*, Electronic Communications in Probability **13** (2008), 319–324. MR2415140

[74] Vladimir Vovk, *Rough paths in idealized financial markets*, Lithuanian Mathematical Journal **51** (2011), no. 2, 274–285. MR2805744

[75] Vladimir Vovk, *Continuous-time trading and the emergence of probability*, Finance and Stochastics **16** (2012), no. 4, 561–609. MR2972235

[76] Vladimir Vovk, *Itô calculus without probability in idealized financial markets*, Lithuanian Mathematical Journal **55** (2015), no. 2, 270–290. MR3347596

[77] Vladimir Vovk, *Purely pathwise probability-free Itô integral*, Matematychni Studii **46** (2016), no. 1, 96–110. MR3649053

[78] Walter Willinger and Murad S. Taqqu, *Pathwise approximations of processes based on the fine structure of their filtrations*, Séminaire de Probabilités XXII (Jacques Azéma, Marc Yor, and Paul André Meyer, eds.), Springer Berlin Heidelberg, Berlin, Heidelberg, 1988, pp. 542–599. MR960547

[79] Walter Willinger and Murad S. Taqqu, *Pathwise stochastic integration and applications to the theory of continuous trading*, Stochastic Processes and their Applications **32** (1989), no. 2, 253–280. MR1014453

[80] Ivan Yaroslavtsev, *Burkholder–Davis–Gundy inequalities in UMD Banach spaces*, Communications in Mathematical Physics **379** (2020), 417–459. MR4156214

[81] Marc Yor, *Sur les intégrales stochastiques à valeurs dans un espace de Banach*, Annales de l'I.H.P. Probabilités et statistiques **10** (1974), no. 1, 31–36. MR358986

Acknowledgments. The author thanks Professor Jun Sekine, his doctoral supervisor, and Professor Masaaki Fukasawa for their helpful comments and encouragement. The author is also grateful to the anonymous referees for their various useful comments on an earlier version of this paper.