Electronic Journal of Probability

Electron. J. Probab. **21** (2016), no. 2, 1–37.
ISSN: 1083-6489 DOI: 10.1214/16-EJP3868

A Fourier analytic approach to pathwise stochastic integration*

Massimiliano Gubinelli† Peter Imkeller ‡ Nicolas Perkowski §

Abstract

We develop a Fourier analytic approach to rough path integration, based on the series decomposition of continuous functions in terms of Schauder functions. Our approach is rather elementary, the main ingredient being a simple commutator estimate, and it leads to recursive algorithms for the calculation of pathwise stochastic integrals, both of Itô and of Stratonovich type. We apply it to solve stochastic differential equations in a pathwise manner.

Keywords: paracontrolled calculus; Schauder functions; rough paths; stochastic integration; stochastic differential equations.
AMS MSC 2010: 60H05; 60H10.
Submitted to EJP on October 15, 2014, final version accepted on December 18, 2015.

1 Introduction

The theory of rough paths [35] has recently been extended to a multiparameter setting independently by Hairer [26] and the autors [23]. While Hairer's approach has a wider range of applicability, both allow to study many interesting problems that were well out of reach with previously existing methods; for example the continuous parabolic Anderson model in dimension two [26, 23, 10], the three-dimensional stochastic quantization equation [26, 9], the KPZ equation [25, 24], or the three-dimensional stochastic Navier Stokes equation [55, 54]. Our methods developed in [23] are based on harmonic analysis, Littlewood-Paley decompositions of tempered distributions, and a simple commutator lemma. This requires a non-negligible knowledge of Littlewood-Paley theory and Besov spaces, while at the same time the application to rough differential

*M.G. was supported by a Junior fellowship of the Institut Universitaire de France (IUF) and by the ANR Project ECRU (ANR-09-BLAN-0114-01). N.P. was supported by the Fondation Sciences Mathématiques de Paris (FSMP) and by a public grant overseen by the French National Research Agency (ANR) as part of the "Investissements d'Avenir" program (reference: ANR-10-LABX-0098), and acknowledges generous support from Humboldt-Universität zu Berlin, where a major part of this work was completed.

†CEREMADE & CNRS UMR 7534, Université Paris-Dauphine and Institut Universitaire de France, France. E-mail: gubinelli@ceremade.dauphine.fr

‡Institut für Mathematik, Humboldt-Universität zu Berlin, Germany. E-mail: imkeller@math.hu-berlin.de

§CEREMADE & CNRS UMR 7534, Université Paris-Dauphine, France. E-mail: perkowski@ceremade.dauphine.fr

equations (the classical problem that motivated Lyons' theory of rough paths) is possible but more technically involved than we would wish. That is why here we develop the approach of [23] in the slightly different language of Haar / Schauder functions, which allows us to communicate our main ideas while requiring only a very basic knowledge in analysis. Moreover, in the Haar-Schauder formulation the application to rough differential equations poses no additional technical challenges and we understand quite well the link between equations of Itô and Stratonovich type.

It is a classical result of Ciesielski [11] that $C^\alpha := C^\alpha([0,1],\mathbb{R}^d)$, the space of α–Hölder continuous functions on $[0,1]$ with values in $\mathbb{R}^d$, is isomorphic to $\ell^\infty(\mathbb{R}^d)$, the space of bounded sequences with values in $\mathbb{R}^d$. The isomorphism gives a Fourier decomposition of a Hölder-continuous function f as

$$f = \sum_{p,m} \langle H_{pm}, \mathrm{d}f\rangle G_{pm},$$

where (H_{pm}) are the Haar functions and (G_{pm}) are the Schauder functions. Ciesielski proved that a continuous function f is in $C^\alpha([0,1],\mathbb{R}^d)$ if and only if the coefficients $(\langle H_{pm}, \mathrm{d}f\rangle)_{p,m}$ decay rapidly enough. Following Ciesielski's work, similar isomorphisms have been developed for many Fourier and wavelet bases, showing that the regularity of a function is encoded in the decay of its coefficients in these bases; see for example Triebel [50].

But until this day, the isomorphism based on Schauder functions plays a special role in stochastic analysis, because the coefficients in the Schauder basis have the pleasant property that they are just rescaled second order increments of f. So if f is a stochastic process with known distribution, then also the distribution of its coefficients in the Schauder basis is known explicitly. A simple application is the Lévy-Ciesielski construction of Brownian motion. An incomplete list with further applications will be given below.

Another convenient property of Schauder functions is that they are piecewise linear, and therefore their iterated integrals $\int_0^\cdot G_{pm}(s)\mathrm{d}G_{qn}(s)$, can be easily calculated. This makes them an ideal tool for our purpose of studying integrals. Indeed, given two continuous functions f and g on $[0,1]$ with values in $\mathcal{L}(\mathbb{R}^d,\mathbb{R}^n)$, the space of linear maps from $\mathbb{R}^d$ to $\mathbb{R}^n$, and $\mathbb{R}^d$ respectively, we can formally define

$$\int_0^t f(s)\mathrm{d}g(s) := \sum_{p,m}\sum_{q,n} \langle H_{pm}, \mathrm{d}f\rangle \langle H_{qn}, \mathrm{d}g\rangle \int_0^t G_{pm}(s)\mathrm{d}G_{qn}(s).$$

In this paper we study under which conditions this formal definition can be made rigorous. We start by observing that the integral introduces a bounded operator from $C^\alpha \times C^\beta$ to C^β if and only if $\alpha+\beta > 1$. Obviously, here we simply recover Young's integral [53]. In our study of this integral, we identify different components:

$$\int_0^t f(s)\mathrm{d}g(s) = S(f,g)(t) + \pi_<(f,g)(t) + L(f,g)(t),$$

where S is the *symmetric part*, $\pi_<$ the *paraproduct*, and $L(f,g)$ the *Lévy area*. The operators S and $\pi_<$ are defined for $f\in C^\alpha$ and $g \in C^\beta$ for arbitrary $\alpha,\beta>0$, and it is only the Lévy area which requires $\alpha+\beta>1$. Considering the regularity of the three operators, we have $S(f,g)\in C^{\alpha+\beta}$, $\pi_<(f,g)\in C^\beta$, and $L(f,g)\in C^{\alpha+\beta}$ whenever the latter is defined. Therefore, in the Young regime $\int_0^\cdot f(s)\mathrm{d}g(s) - \pi_<(f,g) \in C^{\alpha+\beta}$. We will also see that for sufficiently smooth functions F we have $F(f)\in C^\alpha$ but $F(f) - \pi_<(\mathrm{D}F(f), f) \in C^{2\alpha}$. So both $\int_0^\cdot f(s)\mathrm{d}g(s)$ and $F(f)$ are given by a paraproduct plus a smoother remainder. This leads us to call a function $f\in C^\alpha$ *paracontrolled* by g if there exists a function $f^g \in C^\beta$

such that $f - \pi_<(f^g, g) \in C^{\alpha+\beta}$. Our aim is then to construct the Lévy area $L(f, g)$ for $\alpha < 1/2$ and f paracontrolled by g. If $\beta > 1/3$, then the term $L(f - \pi_<(f^g, g), g)$ is well defined, and it suffices to make sense of the term $L(\pi_<(f^g, g), g)$. This is achieved with the following commutator estimate:

$$\left\| L(\pi_<(f^g, g), g) - \int_0^\cdot f^g(s) \mathrm{d}L(g, g)(s) \right\|_{3\beta} \le \|f^g\|_\beta \|g\|_\beta \|g\|_\beta.$$

Therefore, the integral $\int_0^\cdot f(s)\mathrm{d}g(s)$ can be constructed for all f that are paracontrolled by g, provided that $L(g, g)$ can be constructed. In other words, we have found an alternative formulation of Lyons' [35] rough path integral, at least for Hölder continuous functions of Hölder exponent larger than 1/3.

Since we approximate f and g by functions of bounded variation, our integral is of Stratonovich type, that is it satisfies the usual integration by parts rule. We also consider a non-anticipating Itô type integral, that can essentially be reduced to the Stratonovich case with the help of the quadratic variation.

The last remaining problem is then to construct the Lévy area $L(g, g)$ for suitable stochastic processes g. We construct it for certain hypercontractive processes. For continuous martingales that possess sufficiently many moments we give a construction of the Itô iterated integrals that allows us to use them as integrators for our pathwise Itô integral.

Below we give some pointers to the literature, and we introduce some basic notations which we will use throughout. In Section 2 we present Ciesielski's isomorphism, and we give a short overview on rough paths and Young integration. In Section 3 we develop a paradifferential calculus in terms of Schauder functions, and we examine the different components of Young's integral. In Section 4 we construct the rough path integral based on Schauder functions. Section 5 develops the pathwise Itô integral. In Section 6 we construct the Lévy area for suitable stochastic processes. And in Section 7 we apply our integral to solve both Itô type and Stratonovich type SDEs in a pathwise way.

Acknowledgements. We would like to thank the anonymous referees for their careful reading and for their corrections and suggestions which helped us to improve the presentation and lead us to strengthen the results of Section 6.

Relevant literature. Starting with the Lévy-Ciesielski construction of Brownian motion, Schauder functions have been a very popular tool in stochastic analysis. They can be used to prove in a comparatively easy way that stochastic processes belong to Besov spaces; see for example Ciesielski, Kerkyacharian, and Roynette [12], Roynette [47], and Rosenbaum [46]. Baldi and Roynette [4] have used Schauder functions to extend the large deviation principle for Brownian motion from the uniform to the Hölder topology; see also Ben Arous and Ledoux [7] for the extension to diffusions, Eddahbi, N'zi, and Ouknine [14] for the large deviation principle for diffusions in Besov spaces, and Andresen, Imkeller, and Perkowski [1] for the large deviation principle for a Hilbert space valued Wiener process in Hölder topology. Ben Arous, Grădinaru, and Ledoux [6] use Schauder functions to extend the Stroock-Varadhan support theorem for diffusions from the uniform to the Hölder topology. Lyons and Zeitouni [33] use Schauder functions to prove exponential moment bounds for Stratonovich iterated integrals of a Brownian motion conditioned to stay in a small ball. Gantert [20] uses Schauder functions to associate to every sample path of the Brownian bridge a sequence of probability measures on path space, and continues to show that for almost all sample paths these measures converge to the distribution of the Brownian bridge. This shows that the law of the Brownian bridge can be reconstructed from a single "typical sample path".

Concerning integrals based on Schauder functions, there are three important references: Roynette [47] constructs a version of Young's integral on Besov spaces and shows that in the one dimensional case the Stratonovich integral $\int_0^{\cdot} F(W_s)\mathrm{d}W_s$, where W is a Brownian motion and $F \in C^2$, can be defined in a deterministic manner with the help of Schauder functions. Roynette also constructs more general Stratonovich integrals with the help of Schauder functions, but in that case only almost sure convergence is established, where the null set depends on the integrand, and the integral is not a deterministic operator. Ciesielski, Kerkyacharian, and Roynette [12] slightly extend the Young integral of [47], and simplify the proof by developing the integrand in the Haar basis and not in the Schauder basis. They also construct pathwise solutions to SDEs driven by fractional Brownian motions with Hurst index $H > 1/2$. Kamont [29] extends the approach of [12] to define a multiparameter Young integral for functions in anisotropic Besov spaces. Ogawa [40, 41] investigates an integral for anticipating integrands he calls *noncausal* starting from a Parseval type relation in which integrand and Brownian motion as integrator are both developed by a given complete orthonormal system in the space of square integrable functions on the underlying time interval. This concept is shown to be strongly related to Stratonovich type integrals (see Ogawa [41], Nualart, Zakai [39]), and used to develop a stochastic calculus on a Brownian basis with *noncausal* SDE (Ogawa [42]).

Rough paths have been introduced by Lyons [35], see also [34, 37, 31] for previous results. Lyons observed that solution flows to SDEs (or more generally ordinary differential equations (ODEs) driven by rough signals) can be defined in a pathwise, continuous way if paths are equipped with sufficiently many iterated integrals. More precisely, if a path has finite p–variation for some $p \geq 1$, then one needs to associate $\lfloor p \rfloor - 1$ iterated integrals to it to obtain an object which can be taken as the driving signal in an ODE, such that the solution to the ODE depends continuously on that signal. Gubinelli [21, 22] simplified the theory of rough paths by introducing the concept of controlled paths, on which we will strongly rely in what follows. Roughly speaking, a path f is controlled by the reference path g if the small scale fluctuations of f "look like those of g". Good monographs on rough paths are [32, 36, 18, 16].

Finally let us remark that, even if only quite implicitly, paraproducts based on the classical Fourier transform have already been exploited in the rough path context in the work of Unterberger on the renormalization of rough paths [51, 52], where it is referred to as "Fourier normal-ordering", and in the related work of Nualart and Tindel [38].

Notation and conventions. Throughout the paper, we use the notation $a \lesssim b$ if there exists a constant $c > 0$, independent of the variables under consideration, such that $a \leqslant c \cdot b$, and we write $a \simeq b$ if $a \lesssim b$ and $b \lesssim a$. If we want to emphasize the dependence of c on the variable x, then we write $a(x) \lesssim_x b(x)$.

For a multi-index $\mu = (\mu_1, \ldots, \mu_d) \in \mathbb{N}^d$ we write $|\mu| = \mu_1 + \ldots + \mu_d$ and $\partial^\mu = \partial^{|\mu|}/\partial_{x_1}^{\mu_1} \cdots \partial_{x_d}^{\mu_d}$. $\mathrm{D}F$ or F' denote the total derivative of F. For $k \in \mathbb{N}$ we denote by $\mathrm{D}^k F$ the k–th order derivative of F. We also write ∂_x for the partial derivative in direction x.

2 Preliminaries

2.1 Ciesielski's isomorphism

Let us briefly recall Ciesielski's isomorphism between $C^\alpha([0,1], \mathbb{R}^d)$ and $\ell^\infty(\mathbb{R}^d)$. The *Haar functions* $(H_{pm}, p \in \mathbb{N}, 1 \leq m \leq 2^p)$ are defined as

$$H_{pm}(t) := \begin{cases} \sqrt{2^p}, & t \in \left[\frac{m-1}{2^p}, \frac{2m-1}{2^{p+1}}\right), \\ -\sqrt{2^p}, & t \in \left[\frac{2m-1}{2^{p+1}}, \frac{m}{2^p}\right), \\ 0, & \text{otherwise}. \end{cases}$$

When completed by $H_{00} \equiv 1$, the Haar functions are an orthonormal basis of $L^2([0,1], \mathrm{d}t)$. For convencience of notation, we also define $H_{p0} \equiv 0$ for $p \geq 1$. The primitives of the Haar functions are called *Schauder functions* and they are given by $G_{pm}(t) := \int_0^t H_{pm}(s)\mathrm{d}s$ for $t \in [0,1]$, $p \in \mathbb{N}$, $0 \leq m \leq 2^p$. More explicitly, $G_{00}(t) = t$ and for $p \in \mathbb{N}$, $1 \leq m \leq 2^p$

$$G_{pm}(t) = \begin{cases} 2^{p/2}\left(t - \frac{m-1}{2^p}\right), & t \in \left[\frac{m-1}{2^p}, \frac{2m-1}{2^{p+1}}\right), \\ -2^{p/2}\left(t - \frac{m}{2^p}\right), & t \in \left[\frac{2m-1}{2^{p+1}}, \frac{m}{2^p}\right), \\ 0, & \text{otherwise}. \end{cases}$$

Since every G_{pm} satisfies $G_{pm}(0) = 0$, we are only able to expand functions f with $f(0) = 0$ in terms of this family (G_{pm}). Therefore, we complete (G_{pm}) once more, by defining $G_{-10}(t) := 1$ for all $t \in [0,1]$. To abbreviate notation, we define the times t^i_{pm}, $i = 0, 1, 2$, as

$$t^0_{pm} := \frac{m-1}{2^p}, \quad t^1_{pm} := \frac{2m-1}{2^{p+1}}, \quad t^2_{pm} := \frac{m}{2^p},$$

for $p \in \mathbb{N}$ and $1 \leq m \leq 2^p$. Further, we set $t^0_{-10} := 0$, $t^1_{-10} := 0$, $t^2_{-10} := 1$, and $t^0_{00} := 0$, $t^1_{00} := 1$, $t^2_{00} := 1$, as well as $t^i_{p0} := 0$ for $p \geq 1$ and $i = 0, 1, 2$. The definition of t^i_{-10} and t^i_{00} for $i \neq 1$ is rather arbitrary, but the definition for $i = 1$ simplifies for example the statement of Lemma 2.1 below.

For $f\colon [0,1] \to \mathbb{R}^d$, $p \in \mathbb{N}$, and $1 \leq m \leq 2^p$, we write

$$\begin{aligned} \langle H_{pm}, \mathrm{d}f \rangle &:= 2^{\frac{p}{2}} \left[\left(f\left(t^1_{pm}\right) - f\left(t^0_{pm}\right)\right) - \left(f\left(t^2_{pm}\right) - f\left(t^1_{pm}\right)\right)\right] \\ &= 2^{\frac{p}{2}} \left[2f\left(t^1_{pm}\right) - f\left(t^0_{pm}\right) - f\left(t^2_{pm}\right)\right] \end{aligned}$$

and $\langle H_{00}, \mathrm{d}f \rangle := f(1) - f(0)$ as well as $\langle H_{-10}, \mathrm{d}f \rangle := f(0)$. Note that we only defined G_{-10} and not H_{-10}.

Lemma 2.1. *For $f\colon [0,1] \to \mathbb{R}^d$, the function*

$$f_k := \langle H_{-10}, \mathrm{d}f \rangle G_{-10} + \langle H_{00}, \mathrm{d}f \rangle G_{00} + \sum_{p=0}^{k} \sum_{m=1}^{2^p} \langle H_{pm}, \mathrm{d}f \rangle G_{pm} = \sum_{p=-1}^{k} \sum_{m=0}^{2^p} \langle H_{pm}, \mathrm{d}f \rangle G_{pm}$$

is the linear interpolation of f between the points $t^1_{-10}, t^1_{00}, t^1_{pm}$, $0 \leq p \leq k, 1 \leq m \leq 2^p$. If f is continuous, then (f_k) converges uniformly to f as $k \to \infty$.

Ciesielski [11] observed that if f is Hölder-continuous, then the series (f_k) converges absolutely and the speed of convergence can be estimated in terms of the Hölder norm of f. The norm $\|\cdot\|_{C^\alpha}$ is defined as

$$\|f\|_{C^\alpha} := \|f\|_\infty + \sup_{0 \leq s < t \leq 1} \frac{|f_{s,t}|}{|t-s|^\alpha},$$

where we introduced the notation

$$f_{s,t} := f(t) - f(s).$$

Lemma 2.2 **([11]).** *Let $\alpha \in (0,1)$. A continuous function $f : [0,1] \to \mathbb{R}^d$ is in C^α if and only if $\sup_{p,m} 2^{p(\alpha - 1/2)} |\langle H_{pm}, \mathrm{d}f \rangle| < \infty$. In this case*

$$\sup_{p,m} 2^{p(\alpha-1/2)}|\langle H_{pm}, \mathrm{d}f\rangle| \simeq \|f\|_{C^\alpha} \textit{ and} \tag{2.1}$$

$$\|f - f_{N-1}\|_\infty = \Big\| \sum_{p=N}^\infty \sum_{m=0}^{2^p} \langle H_{pm}, \mathrm{d}f\rangle G_{pm} \Big\|_\infty \lesssim \|f\|_{C^\alpha} 2^{-\alpha N}.$$

Before we continue, let us slightly change notation. We want to get rid of the factor $2^{-p/2}$ in (2.1), and therefore we define for $p \in \mathbb{N}$ and $0 \le m \le 2^p$ the rescaled functions

$$\chi_{pm} := 2^{\frac{p}{2}} H_{pm} \qquad \text{and} \qquad \varphi_{pm} := 2^{\frac{p}{2}} G_{pm},$$

as well as $\varphi_{-10} := G_{-10} \equiv 1$. Then we have for $p \in \mathbb{N}$ and $1 \le m \le 2^p$

$$\|\varphi_{pm}(t)\|_\infty = \varphi_{pm}(t^1_{pm}) = 2^{\frac{p}{2}} \int_{t^0_{pm}}^{t^1_{pm}} 2^{\frac{p}{2}} \mathrm{d}s = 2^p \left(\frac{2m-1}{2^{p+1}} - \frac{2m-2}{2^{p+1}} \right) = \frac{1}{2},$$

so that $\|\varphi_{pm}\|_\infty \le 1$ for all p, m. The expansion of f in terms of (φ_{pm}) is given by $f_k = \sum_{p=0}^k \sum_{m=0}^{2^p} f_{pm}\varphi_{pm}$, where $f_{-10} := f(1)$, and $f_{00} := f(1) - f(0)$ and for $p \in \mathbb{N}$ and $m \ge 1$

$$f_{pm} := 2^{-p}\langle \chi_{pm}, \mathrm{d}f\rangle = 2f\left(t^1_{pm}\right) - f\left(t^0_{pm}\right) - f\left(t^2_{pm}\right) = f_{t^0_{pm}, t^1_{pm}} - f_{t^1_{pm}, t^2_{pm}}.$$

We write $\langle \chi_{pm}, \mathrm{d}f\rangle := 2^p f_{pm}$ for all values of (p, m), despite not having defined χ_{-10}.

Definition 2.3. *For $\alpha > 0$ and $f \colon [0,1] \to \mathbb{R}^d$ the norm $\|\cdot\|_\alpha$ is defined as*

$$\|f\|_\alpha := \sup_{pm} 2^{p\alpha}|f_{pm}|,$$

and we write

$$\mathcal{C}^\alpha(\mathbb{R}^d) := \left\{ f \in C([0,1], \mathbb{R}^d) : \|f\|_\alpha < \infty \right\}.$$

In case there is no ambiguity about the target set, we also write $\mathcal{C}^\alpha$ instead of $\mathcal{C}^\alpha(\mathbb{R}^d)$.

The space $\mathcal{C}^\alpha(\mathbb{R}^d)$ is isomorphic to $\ell^\infty(\mathbb{R}^d)$, in particular it is a Banach space. For $\alpha \in (0,1)$, Ciesielski's isomorphism (Lemma 2.2) states that $\mathcal{C}^\alpha(\mathbb{R}^d) = C^\alpha([0,1], \mathbb{R}^d)$. Moreover, it can be shown that $\mathcal{C}^1$ is the Zygmund space of continuous functions f satisfying $|2f(x) - f(x+h) - f(x-h)| \lesssim h$. But for $\alpha > 1$, there is no reasonable identification of $\mathcal{C}^\alpha$ with a classical function space. For example if $\alpha \in (1,2)$, the space $C^\alpha([0,1], \mathbb{R}^d)$ consists of all continuously differentiable functions f with $(\alpha - 1)$–Hölder continuous derivative $\mathrm{D}f$. Since the tent shaped functions φ_{pm} are not continuously differentiable, even an f with a finite Schauder expansion is generally not in $C^\alpha([0,1], \mathbb{R}^d)$.

The a priori requirement of f being continuous can be relaxed, but not much. Since the coefficients (f_{pm}) evaluate the function f only in countably many points, a general f will not be uniquely determined by its expansion. But for example it would suffice to assume that f is càdlàg.

Remark 2.4. Another way of writing the $\|\cdot\|_\alpha$ norm is

$$\|f\|_\alpha = \|(2^{p\alpha}\|(f_{pm})_m\|_{\ell_\infty})_p\|_{\ell_\infty}.$$

One could of course imagine taking other norms on the sequence spaces than the ℓ_∞ norm, and one useful definition is

$$\|f\|_{B^\beta_{r,s}} := \Big(\sum_p 2^{ps(\beta-1/s)} \Big(\sum_m |f_{pm}|^r \Big)^{s/r} \Big)^{1/s}.$$

This leads to Besov spaces with general integrability indices r and s, and in fact Ciesielski's isomorphism extends to this setting, see [12]. One can therefore develop the theory we present here also for general Besov spaces, and for the approach of [23] this was worked out in [45]. But in order to keep the presentation lighter we refrain from considering general Besov spaces here. The only exception is Theorem 6.5, where we use them in the proof.

Littlewood-Paley notation. We will employ notation inspired from Littlewood-Paley theory. For $p \geq -1$ and $f\colon [0,1] \to \mathbb{R}^d$ we define

$$\Delta_p f := \sum_{m=0}^{2^p} f_{pm}\varphi_{pm} \qquad \text{and} \qquad S_p f := \sum_{q \leq p} \Delta_q f.$$

We will occasionally refer to $(\Delta_p f)$ as the Schauder blocks of f. Note that

$$\mathcal{C}^\alpha(\mathbb{R}^d) = \{f \in C([0,1], \mathbb{R}^d) : \|(2^{p\alpha}\|\Delta_p f\|_\infty)_p\|_{\ell^\infty} < \infty\}.$$

2.2 Young integration and rough paths

Here we present the main concepts of Young integration and of rough path theory. The results presented in this section will not be applied in the remainder of the paper, but we feel that it could be useful for the reader to be familiar with the basic concepts of rough paths, since it is the main inspiration for the constructions developed below.

Young's integral [53] allows to define $\int f \mathrm{d}g$ for $f \in C^\alpha$, $g \in C^\beta$, and $\alpha + \beta > 1$. More precisely, let $f \in C^\alpha$ and $g \in C^\beta$ be given, let $t \in [0,1]$, and let $\pi = \{t_0, \ldots, t_N\}$ be a partition of $[0,t]$, i.e. $0 = t_0 < t_1 < \cdots < t_N = t$. Then it can be shown that the Riemann sums

$$\sum_{t_k \in \pi} f(t_k)(g(t_{k+1}) - g(t_k)) := \sum_{k=0}^{N-1} f(t_k)(g(t_{k+1}) - g(t_k))$$

converge as the mesh size $\max_{k=0,\ldots,N-1} |t_{k+1} - t_k|$ tends to zero, and that the limit does not depend on the approximating sequence of partitions. We denote the limit by $\int_0^t f(s)\mathrm{d}g(s)$, and we define $\int_s^t f(r)\mathrm{d}g(r) := \int_0^t f(r)\mathrm{d}g(r) - \int_0^s f(r)\mathrm{d}g(r)$. The function $t \mapsto \int_0^t f(s)\mathrm{d}g(s)$ is uniquely characterized by the fact that

$$\left| \int_s^t f(r)\mathrm{d}g(r) - f(s)(g(t) - g(s)) \right| \lesssim |t-s|^{\alpha+\beta} \|f\|_\alpha \|g\|_\beta$$

for all $s, t \in [0,1]$. The condition $\alpha + \beta > 1$ is sharp, in the sense that there exist $f, g \in C^{1/2}$, and a sequence of partitions $(\pi_n)_{n\in\mathbb{N}}$ with mesh size going to zero for which the Riemann sums $\sum_{t_k \in \pi_n} f(t_k)(g(t_{k+1}) - g(t_k))$ do not converge as n tends to ∞.

The condition $\alpha + \beta > 1$ excludes one of the most important examples: we would like to take g as a sample path of Brownian motion, and $f = F(g)$. Lyons' theory of rough paths [35] overcomes this restriction by stipulating the "existence" of basic integrals and by defining a large class of related integrals as their functionals. Here we present the approach of Gubinelli [21].

Let $\alpha \in (1/3, 1)$ and assume that we are given two functions $v, w \in C^\alpha([0,1], \mathbb{R})$, as well as an associated "Riemann integral" $I^{v,w}_{s,t} = \int_s^t v(r)\mathrm{d}w(r)$ that satisfies the estimate

$$|\Phi^{v,w}_{s,t}| \lesssim |t-s|^{2\alpha} \tag{2.2}$$

for $\Phi^{v,w}_{s,t} := I^{v,w}_{s,t} - v(s)w_{s,t}$. The remainder $\Phi^{v,w}$ is often (incorrectly) called the *area* of v and w. This name has its origin in the fact that its antisymmetric part $(\Phi^{v,w}_{s,t} - \Phi^{w,v}_{s,t})/2$

corresponds to the algebraic area spanned by the curve $((v(r), w(r)) : r \in [s,t])$ in the plane $\mathbb{R}^2$.

If $\alpha \leq 1/2$, then the integral $I^{v,w}$ cannot be constructed using Young's theory of integration, and also $I^{v,w}$ is not uniquely characterized by (2.2). But let us assume nonetheless that we are given such an integral $I^{v,w}$ satisfying (2.2). A function $f \in C^\alpha$ is *controlled* by $v \in C^\alpha$ if there exists $f^v \in C^\alpha$, such that for all $s, t \in [0,1]$

$$|f_{s,t} - f^v(s) v_{s,t}| \lesssim |t-s|^{2\alpha}. \tag{2.3}$$

Proposition 2.5 ([21], Theorem 1)**.** *Let $\alpha > 1/3$, let $v, w \in C^\alpha$, and let $\Phi^{v,w}$ satisfy (2.2). Let f and g be controlled by v and w respectively, with derivatives f^v and g^w. Then there exists a unique function $I(f,g)$ that satisfies for all $s,t \in [0,1]$*

$$|I(f,g)_{s,t} - f(s) g_{s,t} - f^v(s) g^w(s) \Phi^{v,w}_{s,t}| \lesssim |t-s|^{3\alpha}.$$

If (π_n) is a sequence of partitions of $[0,t]$, with mesh size going to zero, then

$$I(f,g)(t) = \lim_{n\to\infty} \sum_{t_k \in \pi_n} \left(f(t_k) g_{t_k,t_{k+1}} + f^v(t_k) g^w(t_k) \Phi^{v,w}_{t_k,t_{k+1}} \right).$$

Of course, all of this extends to a multidimensional setting where v, w, f, g take values in $\mathbb{R}^{d_1}, \mathbb{R}^{d_2}, \mathbb{R}^{d_3}, \mathbb{R}^{d_4}$, respectively (in which case we have to replace for example $\int_s^t v(r) \mathrm{d}w(r)$ by $\int_s^t v(r) \otimes \mathrm{d}w(r)$).

The integral $I(f,g)$ coincides with the Riemann-Stieltjes integral and with the Young integral, whenever these are defined. Moreover, the integral map is self-consistent, in the sense that if we consider v and w as paracontrolled by themselves, with derivatives $v^v = w^w \equiv 1$, then $I(v,w) = I^{v,w}$.

The only remaining problem is the construction of the integral $I^{v,w}$. This is usually achieved with probabilistic arguments. If v and w are Brownian motions, then we can for example use Itô or Stratonovich integration to define $I^{v,w}$. Already in this simple example we see that the integral $I^{v,w}$ is not unique if v and w are outside of the Young regime.

It is possible to go beyond $\alpha > 1/3$ by stipulating the existence of higher order iterated integrals. For details see [22] or any book on rough paths, such as [32, 36, 18, 16].

3 Paradifferential calculus and Young integration

In this section we develop the basic tools that will be required for our rough path integral in terms of Schauder functions, and we study Young's integral and its different components.

3.1 Paradifferential calculus with Schauder functions

Here we introduce a "paradifferential calculus" in terms of Schauder functions. Paradifferential calculus is usually formulated in terms of Littlewood-Paley blocks and was initiated by Bony [8]. For a gentle introduction see [3].

We will need to study the regularity of $\sum_{p,m} u_{pm} \varphi_{pm}$, where u_{pm} are functions and not constant coefficients. For this purpose we define the following space of sequences of functions.

Definition 3.1. *If $(u_{pm} : p \geq -1, 0 \leq m \leq 2^p)$ is a family of affine functions of the form $u_{pm} : [t^0_{pm}, t^2_{pm}] \to \mathbb{R}^d$, we set for $\alpha > 0$*

$$\|(u_{pm})\|_{\mathcal{A}^\alpha} := \sup_{p,m} 2^{p\alpha} \|u_{pm}\|_\infty,$$

where it is understood that $\|u_{pm}\|_\infty := \max_{t\in[t^0_{pm},t^2_{pm}]} |u_{pm}(t)|$*. The space* $\mathcal{A}^\alpha(\mathbb{R}^d)$ *is then defined as*

$$\mathcal{A}^\alpha(\mathbb{R}^d) := \left\{(u_{pm})_{p\geq -1, 0\leq m\leq 2^p} : u_{pm} \in C([t^0_{pm}, t^2_{pm}], \mathbb{R}^d) \textit{ is affine and } \|(u_{pm})\|_{\mathcal{A}^\alpha} < \infty\right\}.$$

In Appendix A we prove the following regularity estimate:

Lemma 3.2. *Let* $\alpha \in (0,2)$ *and let* $(u_{pm}) \in \mathcal{A}^\alpha(\mathbb{R}^d)$*. Then* $\sum_{p,m} u_{pm}\varphi_{pm} \in \mathcal{C}^\alpha(\mathbb{R}^d)$*, and*

$$\Big\|\sum_{p,m} u_{pm}\varphi_{pm}\Big\|_\alpha \lesssim \|(u_{pm})\|_{\mathcal{A}^\alpha}.$$

Let us introduce a paraproduct in terms of Schauder functions.

Lemma 3.3. *Let* $\beta \in (0,2)$*, let* $v \in C([0,1], \mathcal{L}(\mathbb{R}^d, \mathbb{R}^n))$*, and* $w \in \mathcal{C}^\beta(\mathbb{R}^d)$*. Then*

$$\pi_<(v,w) := \sum_{p=0}^\infty S_{p-1}v\Delta_p w \in \mathcal{C}^\beta(\mathbb{R}^n) \quad \textit{and} \quad \|\pi_<(v,w)\|_\beta \lesssim \|v\|_\infty \|w\|_\beta. \tag{3.1}$$

Proof. We have $\pi_<(v,w) = \sum_{p,m} u_{pm}\varphi_{pm}$ with $u_{pm} = (S_{p-1}v)|_{[t^0_{pm},t^2_{pm}]} w_{pm}$. For every (p,m), the function $(S_{p-1}v)|_{[t^0_{pm},t^2_{pm}]}$ is the linear interpolation of v between t^0_{pm} and t^2_{pm}. As $\|(S_{p-1}v)|_{[t^0_{pm},t^2_{pm}]} w_{pm}\|_\infty \leq 2^{-p\beta}\|v\|_\infty\|w\|_\beta$, the statement follows from Lemma 3.2. □

Remark 3.4. If $v \in \mathcal{C}^\alpha(\mathbb{R})$ and $w \in \mathcal{C}^\beta(\mathbb{R})$, we can decompose the product vw into three components, $vw = \pi_<(v,w) + \pi_>(v,w) + \pi_\circ(v,w)$, where $\pi_>(v,w) := \pi_<(w,v)$ and $\pi_\circ(v,w) := \sum_p \Delta_p v\Delta_p w$, and we have the estimates

$$\|\pi_>(v,w)\|_\alpha \lesssim \|v\|_\alpha\|w\|_\infty, \qquad \text{and} \qquad \|\pi_\circ(v,w)\|_{\alpha+\beta} \lesssim \|v\|_\alpha\|w\|_\beta$$

whenever $\alpha+\beta \in (0,2)$. However, we will not use this.

The paraproduct allows us to "paralinearize" nonlinear functions. We allow for a smoother perturbation, which will come in handy when constructing global in time solutions to SDEs.

Proposition 3.5. *Let* $\alpha \in (0,1/2)$*,* $\beta \in (0,\alpha]$*, let* $v \in \mathcal{C}^\alpha(\mathbb{R}^d)$*,* $w \in \mathcal{C}^{\alpha+\beta}(\mathbb{R}^d)$*, and* $F \in C^{1+\beta/\alpha}_b(\mathbb{R}^d, \mathbb{R})$*. Then*

$$\|F(v+w) - \pi_<(\mathrm{D}F(v+w), v)\|_{\alpha+\beta} \lesssim \|F\|_{C^{1+\beta/\alpha}_b}(1+\|v\|_\alpha)^{1+\beta/\alpha}(1+\|w\|_{\alpha+\beta}). \tag{3.2}$$

If $F \in C^{2+\beta/\alpha}_b$*, then* $F(v) - \pi_<(\mathrm{D}F(v), v)$ *depends on* v *in a locally Lipschitz continuous way:*

$$\begin{aligned}\|F(v) - \pi_<(\mathrm{D}F(v),v) - (F(u) - \pi_<(\mathrm{D}F(u),u))\|_{\alpha+\beta}& \\ \lesssim \|F\|_{C^{2+\beta/\alpha}_b}(1+\|v\|_\alpha+\|u\|_\alpha)^{1+\beta/\alpha}\|v-u\|_\alpha.&\end{aligned} \tag{3.3}$$

Proof. First note that $\|F(v+w)\|_\infty \leq \|F\|_\infty$, which implies the required estimate for $(p,m) = (-1,0)$ and $(p,m) = (0,0)$. For all other values of (p,m) we apply a Taylor expansion:

$$(F(v+w))_{pm} = \mathrm{D}F(v(t^1_{pm}) + w(t^1_{pm}))v_{pm} + R_{pm},$$

where $|R_{pm}| \lesssim 2^{-p(\alpha+\beta)}\|F\|_{C^{1+\beta/\alpha}_b}(\|v\|_\alpha^{1+\beta/\alpha} + \|w\|_{\alpha+\beta})$. Subtracting $\pi_<(\mathrm{D}F(v), v)$ gives

$$\begin{aligned}&F(v+w) - \pi_<(\mathrm{D}F(v+w), v)\\ &\qquad = \sum_{pm}[\mathrm{D}F(v(t^1_{pm}) + w(t^1_{pm})) - (S_{p-1}\mathrm{D}F(v+w))|_{[t^0_{pm},t^2_{pm}]}]v_{pm}\varphi_{pm} + R.\end{aligned}$$

Now $(S_{p-1}\mathrm{D}F(v+w))|_{[t^0_{pm},t^2_{pm}]}$ is the linear interpolation of $\mathrm{D}F(v+w)$ between t^0_{pm} and t^2_{pm}, so according to Lemma 3.2 it suffices to note that

$$\begin{aligned}&\|[\mathrm{D}F(v(t^1_{pm})+w(t^1_{pm}))-(S_{p-1}\mathrm{D}F(v+w))|_{[t^0_{pm},t^2_{pm}]}]v_{pm}\|_\infty\\ &\quad\lesssim 2^{-p\beta}\|\mathrm{D}F(v+w)\|_{C^\beta}2^{-p\alpha}\|v\|_\alpha\lesssim 2^{-p(\alpha+\beta)}\|F\|_{C^{1+\beta/\alpha}_b}(1+\|v\|_\alpha+\|w\|_\alpha)^{\beta/\alpha}\|v\|_\alpha.\end{aligned}$$

The local Lipschitz continuity is shown in the same way. □

Remark 3.6. Since v has compact support, it actually suffices to have $F\in C^{1+\beta/\alpha}$ without assuming boundedness. Of course, then the estimates in Proposition 3.5 have to be adapted. Similarly we can treat $F\in C^{1+\beta/\alpha}_b(\mathbb{R}^d,\mathbb{R}^m)$ by considering all m components of F separately.

Remark 3.7. The same proof shows that if f is controlled by v in the sense of Section 2.1, i.e. $f_{s,t}=f^v(s)v_{s,t}+R_{s,t}$ with $f^v\in\mathcal{C}^\alpha(\mathcal{L}(\mathbb{R}^d,\mathbb{R}^n))$ and $|R_{s,t}|\le\|R\|_{2\alpha}|t-s|^{2\alpha}$, then $f-\pi_<(f^v,v)\in\mathcal{C}^{2\alpha}(\mathbb{R}^n)$.

3.2 Young's integral and its various components

In this section we construct Young's integral using the Schauder expansion. If $v\in\mathcal{C}^\alpha(\mathcal{L}(\mathbb{R}^d,\mathbb{R}^n))$ and $w\in\mathcal{C}^\beta(\mathbb{R}^d)$, then we formally define

$$\int_0^\cdot v(s)\mathrm{d}w(s):=\sum_{p,m}\sum_{q,n}v_{pm}w_{qn}\int_0^\cdot\varphi_{pm}(s)\mathrm{d}\varphi_{qn}(s)=\sum_{p,q}\int_0^\cdot\Delta_pv(s)\mathrm{d}\Delta_qw(s).$$

We show that this definition makes sense provided that $\alpha+\beta>1$, and we identify three components of the integral that behave quite differently. This will be our starting point towards an extension of the integral beyond the Young regime.

In a first step, let us calculate the iterated integrals of Schauder functions.

Lemma 3.8. *Let $p>q\ge0$. Then*

$$\int_0^1\varphi_{pm}(s)\mathrm{d}\varphi_{qn}(s)=2^{-p-2}\chi_{qn}(t^0_{pm})\tag{3.4}$$

for all m,n. If $p=q$, then $\int_0^1\varphi_{pm}(s)\mathrm{d}\varphi_{pn}(s)=0$, except if $p=q=0$, in which case the integral is bounded by 1. If $0\le p<q$, then for all (m,n) we have

$$\int_0^1\varphi_{pm}(s)\mathrm{d}\varphi_{qn}(s)=-2^{-q-2}\chi_{pm}\left(t^0_{qn}\right).\tag{3.5}$$

If $p=-1$, then the integral is bounded by 1.

Proof. The cases $p=q$ and $p=-1$ are easy, so let $p>q\ge0$. Since $\chi_{qn}\equiv\chi_{qn}(t^0_{pm})$ on the support of φ_{pm}, we have

$$\int_0^1\varphi_{pm}(s)\mathrm{d}\varphi_{qn}(s)=\chi_{qn}(t^0_{pm})\int_0^1\varphi_{pm}(s)\mathrm{d}s=\chi_{qn}(t^0_{pm})2^{-p-2}.$$

If $0\le p<q$, then integration by parts and (3.4) imply (3.5). □

Next we estimate the coefficients of iterated integrals in the Schauder basis.

Lemma 3.9. *Let $i,p\ge-1$, $q\ge0$, $0\le j\le2^i$, $0\le m\le2^p$, $0\le n\le2^q$. Then*

$$2^{-i}\Big|\Big\langle\chi_{ij},\mathrm{d}\Big(\int_0^\cdot\varphi_{pm}\chi_{qn}\mathrm{d}s\Big)\Big\rangle\Big|\le2^{-2(i\vee p\vee q)+p+q},\tag{3.6}$$

except if $p<q=i$. In this case we only have the worse estimate

$$2^{-i}\Big|\Big\langle\chi_{ij},\mathrm{d}\Big(\int_0^\cdot\varphi_{pm}\chi_{qn}\mathrm{d}s\Big)\Big\rangle\Big|\le1.\tag{3.7}$$

Proof. We have $\langle \chi_{-10}, \mathrm{d}(\int_0^\cdot \varphi_{pm}\chi_{qn}\mathrm{d}s)\rangle = 0$ for all (p,m) and (q,n). So let $i \geq 0$. If $i < p \vee q$, then χ_{ij} is constant on the support of $\varphi_{pm}\chi_{qn}$, and therefore Lemma 3.8 gives

$$2^{-i}\,|\langle \chi_{ij}, \varphi_{pm}\chi_{qn}\rangle| \leq |\langle \varphi_{pm}, \chi_{qn}\rangle| \leq 2^{p+q-2(p\vee q)} = 2^{-2(i\vee p\vee q)+p+q}.$$

Now let $i > q$. Then χ_{qn} is constant on the support of χ_{ij}, and therefore another application of Lemma 3.8 implies that

$$2^{-i}\,|\langle \chi_{ij}, \varphi_{pm}\chi_{qn}\rangle| \leq 2^{-i}2^{q}2^{p+i-2(p\vee i)} = 2^{-2(i\vee p\vee q)+p+q}.$$

The only remaining case is $i = q \geq p$, in which

$$2^{-i}\,|\langle \chi_{ij}, \varphi_{pm}\chi_{qn}\rangle| \leq 2^{i}\int_{t^0_{ij}}^{t^2_{ij}} \varphi_{pm}(s)\mathrm{d}s \leq \|\varphi_{pm}\|_\infty \leq 1. \qquad \square$$

Corollary 3.10. *Let $i,p \geq -1$ and $q \geq 0$. Let $v \in C([0,1],\mathcal{L}(\mathbb{R}^d,\mathbb{R}^n))$ and $w \in C([0,1],\mathbb{R}^d)$. Then*

$$\Big\|\Delta_i\Big(\int_0^\cdot \Delta_p v(s)\mathrm{d}\Delta_q w(s)\Big)\Big\|_\infty \lesssim 2^{-(i\vee p\vee q)-i+p+q}\|\Delta_p v\|_\infty\|\Delta_q w\|_\infty, \tag{3.8}$$

except if $i = q > p$. In this case we only have the worse estimate

$$\Big\|\Delta_i\Big(\int_0^\cdot \Delta_p v(s)\mathrm{d}\Delta_q w(s)\Big)\Big\|_\infty \lesssim \|\Delta_p v\|_\infty\|\Delta_q w\|_\infty. \tag{3.9}$$

Proof. The case $i=-1$ is easy, so let $i \geq 0$. We have

$$\Delta_i\Big(\int_0^\cdot \Delta_p v(s)\mathrm{d}\Delta_q w(s)\Big) = \sum_{j,m,n} v_{pm}w_{qn}\langle 2^{-i}\chi_{ij}, \varphi_{pm}\chi_{qn}\rangle\varphi_{ij}.$$

For fixed j, there are at most $2^{(i\vee p\vee q)-i}$ non-vanishing terms in the double sum. Hence, we obtain from Lemma 3.9 that

$$\begin{aligned}\Big\|\sum_{m,n} v_{pm}w_{qn}\langle 2^{-i}\chi_{ij}, \varphi_{pm}\chi_{qn}\rangle\varphi_{ij}\Big\|_\infty &\\ \lesssim 2^{(i\vee p\vee q)-i}\|\Delta_p v\|_\infty\|\Delta_q w\|_\infty(2^{-2(i\vee p\vee q)+p+q} + \mathbf{1}_{i=q>p}) &\\ = (2^{-(i\vee p\vee q)-i+p+q} + \mathbf{1}_{i=q>p})\|\Delta_p v\|_\infty\|\Delta_q w\|_\infty. & \qquad \square\end{aligned}$$

Corollary 3.11. *Let $i,p,q \geq -1$. Let $v \in C([0,1],\mathcal{L}(\mathbb{R}^d,\mathbb{R}^n))$ and $w \in C([0,1],\mathbb{R}^d)$. Then for $p \vee q \leq i$ we have*

$$\|\Delta_i\,(\Delta_p v\Delta_q w)\|_\infty \lesssim 2^{-(i\vee p\vee q)-i+p+q}\|\Delta_p v\|_\infty\|\Delta_q w\|_\infty, \tag{3.10}$$

except if $i = q > p$ or $i = p > q$, in which case we only have the worse estimate

$$\|\Delta_i(\Delta_p v\Delta_q w)\|_\infty \lesssim \|\Delta_p v\|_\infty\|\Delta_q w\|_\infty. \tag{3.11}$$

If $p > i$ or $q > i$, then $\Delta_i(\Delta_p v\Delta_q w) \equiv 0$.

Proof. The case $p=-1$ or $q=-1$ is easy. Otherwise we apply integration by parts and note that the estimates (3.8) and (3.9) are symmetric in p and q. If for example $p > i$, then $\Delta_p(v)(t^k_{ij}) = 0$ for all k,j, which implies that $\Delta_i(\Delta_p v\Delta_q w) = 0$. $\square$

The estimates (3.8) and (3.9) allow us to identify different components of the integral $\int_0^\cdot v(s)\mathrm{d}w(s)$. More precisely, (3.9) indicates that the series $\sum_{p<q}\int_0^\cdot \Delta_p v(s)\mathrm{d}\Delta_q w(s)$ is rougher than the remainder $\sum_{p\geq q}\int_0^\cdot \Delta_p v(s)\mathrm{d}\Delta_q w(s)$. Integration by parts gives

$$\sum_{p<q}\int_0^\cdot \Delta_p v(s)\mathrm{d}\Delta_q w(s) = \pi_<(v,w) - \sum_{p<q}\sum_{m,n} v_{pm}w_{qn}\int_0^\cdot \varphi_{qn}(s)\mathrm{d}\varphi_{pm}(s).$$

This motivates us to decompose the integral into three components, namely

$$\sum_{p,q}\int_0^\cdot \Delta_p v(s)\mathrm{d}\Delta_q w(s) = L(v,w) + S(v,w) + \pi_<(v,w).$$

Here L is defined as the antisymmetric *Lévy area* (we will justify the name below by showing that L is closely related to the Lévy area of certain dyadic martingales):

$$\begin{aligned} L(v,w) &:= \sum_{p>q}\sum_{m,n}(v_{pm}w_{qn} - v_{qn}w_{pm})\int_0^\cdot \varphi_{pm}\mathrm{d}\varphi_{qn} \\ &= \sum_p \left(\int_0^\cdot \Delta_p v \mathrm{d}S_{p-1}w - \int_0^\cdot \mathrm{d}(S_{p-1}v)\Delta_p w\right). \end{aligned}$$

The *symmetric part* S is defined as

$$\begin{aligned} S(v,w) &:= \sum_{m,n\leq 1} v_{0m}w_{0n}\int_0^\cdot \varphi_{0m}\mathrm{d}\varphi_{0n} + \sum_{p\geq 1}\sum_m v_{pm}w_{pm}\int_0^\cdot \varphi_{pm}\mathrm{d}\varphi_{pm} \\ &= \sum_{m,n\leq 1} v_{0m}w_{0n}\int_0^\cdot \varphi_{0m}\mathrm{d}\varphi_{0n} + \frac{1}{2}\sum_{p\geq 1}\Delta_p v\Delta_p w, \end{aligned}$$

and $\pi_<$ is the paraproduct defined in (3.1). As we observed in Lemma 3.3, $\pi_<(v,w)$ is always well defined, and it inherits the regularity of w. Let us study S and L.

Lemma 3.12. *Let $\alpha,\beta\in(0,1)$ be such that $\alpha+\beta>1$. Then L is a bounded bilinear operator from $\mathcal{C}^\alpha(\mathcal{L}(\mathbb{R}^d,\mathbb{R}^n))\times\mathcal{C}^\beta(\mathbb{R}^d)$ to $\mathcal{C}^{\alpha+\beta}(\mathbb{R}^n)$.*

Proof. We only argue for $\sum_p\int_0^\cdot \Delta_p v\mathrm{d}S_{p-1}w$, the term $-\int_0^\cdot \mathrm{d}(S_{p-1}v)\Delta_p w$ can be treated with the same arguments. Corollary 3.10 (more precisely (3.8)) implies that

$$\begin{aligned} &\Big\|\sum_p \Delta_i\Big(\int_0^\cdot \Delta_p v\mathrm{d}S_{p-1}w\Big)\Big\|_\infty \\ &\qquad \leq \sum_{p\leq i}\sum_{q<p}\Big\|\Delta_i\Big(\int_0^\cdot \Delta_p v\mathrm{d}\Delta_q w\Big)\Big\|_\infty + \sum_{p>i}\sum_{q<p}\Big\|\Delta_i\Big(\int_0^\cdot \Delta_p v\mathrm{d}\Delta_q w\Big)\Big\|_\infty \\ &\qquad \leq \Big(\sum_{p\leq i}\sum_{q<p}2^{-2i+p+q}2^{-p\alpha}\|v\|_\alpha 2^{-q\beta}\|w\|_\beta + \sum_{p>i}\sum_{q<p}2^{-i+q}2^{-p\alpha}\|v\|_\alpha 2^{-q\beta}\|w\|_\beta\Big) \\ &\qquad \lesssim_{\alpha+\beta} 2^{-i(\alpha+\beta)}\|v\|_\alpha\|w\|_\beta, \end{aligned}$$

where we used $1-\alpha<0$ and $1-\beta<0$ and for the second series we also used that $\alpha+\beta>1$. □

Unlike the Lévy area L, the symmetric part S is always well defined. It is also more regular than $\pi_<$.

Lemma 3.13. *Let $\alpha,\beta\in(0,1)$. Then S is a bounded bilinear operator from $\mathcal{C}^\alpha(\mathcal{L}(\mathbb{R}^d,\mathbb{R}^n))\times\mathcal{C}^\beta(\mathbb{R}^d)$ to $\mathcal{C}^{\alpha+\beta}(\mathbb{R}^n)$.*

Proof. This is shown using the same arguments as in the proof of Lemma 3.12. □

In conclusion, the integral consists of three components. The Lévy area $L(v,w)$ is only defined if $\alpha+\beta>1$, but then it is quite regular. The symmetric part $S(v,w)$ is always defined and regular. And the paraproduct $\pi_<(v,w)$ is always defined, but it is rougher than the other components. To summarize:

Theorem 3.14 (Young's integral)**.** *Let $\alpha,\beta\in(0,1)$ be such that $\alpha+\beta>1$, and let $v\in\mathcal{C}^\alpha(\mathcal{L}(\mathbb{R}^d,\mathbb{R}^n))$ and $w\in\mathcal{C}^\beta(\mathbb{R}^n)$. Then the integral*

$$I(v,\mathrm{d}w) := \sum_{p,q}\int_0^\cdot \Delta_p v \mathrm{d}\Delta_q w = L(v,w)+S(v,w)+\pi_<(v,w)\in\mathcal{C}^\beta(\mathbb{R}^n)$$

satisfies $\|I(v,\mathrm{d}w)\|_\beta \lesssim \|v\|_\alpha\|w\|_\beta$ and

$$\|I(v,\mathrm{d}w)-\pi_<(v,w)\|_{\alpha+\beta}\lesssim \|v\|_\alpha\|w\|_\beta. \tag{3.12}$$

Lévy area and dyadic martingales

Here we show that the Lévy area $L(v,w)(1)$ can be expressed in terms of the Lévy area of suitable dyadic martingales. To simplify notation, we assume that $v(0)=w(0)=0$, so that we do not have to bother with the components v_{-10} and w_{-10}.

We define a filtration $(\mathcal{F}_n)_{n\geq 0}$ on $[0,1]$ by setting

$$\mathcal{F}_n=\sigma(\chi_{pm}:0\leq p\leq n, 0\leq m\leq 2^p),$$

we set $\mathcal{F}=\bigvee_n\mathcal{F}_n$, and we consider the Lebesgue measure on $([0,1],\mathcal{F})$. On this space, the process $M_n=\sum_{p=0}^n\sum_{m=0}^{2^p}\chi_{pm}$, $n\in\mathbb{N}$, is a martingale. For any continuous function $v:[0,1]\to\mathbb{R}$ with $v(0)=0$, the process

$$M_n^v=\sum_{p=0}^n\sum_{m=0}^{2^p}\langle 2^{-p}\chi_{pm},\mathrm{d}v\rangle\chi_{pm}=\sum_{p=0}^n\sum_{m=0}^{2^p}v_{pm}\chi_{pm}=\partial_t S_n v,$$

$n\in\mathbb{N}$, is a martingale transform of M, and therefore a martingale as well. Since it will be convenient later, we also define $\mathcal{F}_{-1}=\{\emptyset,[0,1]\}$ and $M_{-1}^v=0$ for every v.

Assume now that v and w are continuous real-valued functions with $v(0)=w(0)=0$, and that the Lévy area $L(v,w)(1)$ exists. Then it is given by

$$\begin{aligned}
L(v,w)(1) &= \sum_{p=0}^\infty\sum_{q=0}^{p-1}\sum_{m,n}(v_{pm}w_{qn}-v_{qn}w_{pm})\int_0^1\varphi_{pm}(s)\chi_{qn}(s)\mathrm{d}s\\
&= \sum_{p=0}^\infty\sum_{q=0}^{p-1}\sum_{m,n}(v_{pm}w_{qn}-v_{qn}w_{pm})2^p\int_0^1\chi_{qn}(s)1_{[t^0_{pm},t^2_{pm})}(s)\mathrm{d}s\langle\varphi_{pm},1\rangle\\
&= \sum_{p=0}^\infty\sum_{q=0}^{p-1}\sum_{m,n}(v_{pm}w_{qn}-v_{qn}w_{pm})2^{-p}\int_0^1\chi_{qn}(s)\chi^2_{pm}(s)\mathrm{d}s 2^{-p-2}\\
&= \sum_{p=0}^\infty\sum_{q=0}^{p-1}2^{-2p-2}\int_0^1\sum_{m,n}\sum_{m'}(v_{pm}w_{qn}-v_{qn}w_{pm})\chi_{qn}(s)\chi_{pm}(s)\chi_{pm'}(s)\mathrm{d}s,
\end{aligned}$$

where in the last step we used that χ_{pm} and $\chi_{pm'}$ have disjoint support for $m\neq m'$. The p–th *Rademacher function* (or *square wave*) is defined for $p\geq 1$ as

$$r_p(t):=\sum_{m'=1}^{2^p}2^{-p}\chi_{pm'}(t).$$

The martingale associated to the Rademacher functions is given by $R_0 := 0$ and $R_p := \sum_{k=1}^p r_k$ for $p \geq 1$. Let us write $\Delta M^v_p = M^v_p - M^v_{p-1}$ and similarly for M^w and R and all other discrete time processes that arise. This notation somewhat clashes with the expression $\Delta_p v$ for the dyadic blocks of v, but we will only use it in the following lines, where we do not directly work with dyadic blocks. The quadratic covariation of two dyadic martingales is defined as $[M,N]_n := \sum_{k=0}^n \Delta M_k \Delta N_k$, and the discrete time stochastic integral is defined as $(M \cdot N)_n := \sum_{k=0}^n M_{k-1}\Delta N_k$. Writing $E(\cdot)$ for the integral $\int_0^1 \cdot \mathrm{d}s$, we obtain

$$
\begin{aligned}
L(v,w)(1) &= \sum_{p=0}^\infty \sum_{q=0}^{p-1} 2^{-p-2} E\left(\Delta M^v_p \Delta M^w_q \Delta R_p - \Delta M^v_q \Delta M^w_p \Delta R_p\right) \\
&= \sum_{p=0}^\infty 2^{-p-2} E\left(\left(M^w_{p-1}\Delta M^v_p - M^v_{p-1}\Delta M^w_p\right)\Delta R_p\right) \\
&= \sum_{p=0}^\infty 2^{-p-2} E\left(\Delta \left[M^w \cdot M^v - M^v \cdot M^w, R\right]_p\right).
\end{aligned}
$$

Hence, $L(v,w)(1)$ is closely related to the Lévy area $(M^w \cdot M^v - M^v \cdot M^w)/2$ of the dyadic martingale (M^v, M^w).

4 Paracontrolled paths and pathwise integration beyond Young

In this section we construct a rough path integral in terms of Schauder functions.

4.1 Paracontrolled paths

We observed in Section 3 that for $w \in \mathcal{C}^\alpha$ and $F \in C_b^{1+\beta/\alpha}$ we have $F(w) - \pi_<(\mathrm{D}F(w), w) \in \mathcal{C}^{\alpha+\beta}$. In Section 3.2 we observed that if $v \in \mathcal{C}^\alpha$, $w \in \mathcal{C}^\beta$ and $\alpha + \beta > 1$, then the Young integral $I(v, \mathrm{d}w)$ satisfies $I(v,\mathrm{d}w) - \pi_<(v,w) \in \mathcal{C}^{\alpha+\beta}$. Hence, in both cases the function under consideration can be written as $\pi_<(f^w, w)$ for a suitable f^w, plus a smoother remainder. We make this our definition of paracontrolled paths:

Definition 4.1. *Let $\alpha > 0$ and $v \in \mathcal{C}^\alpha(\mathbb{R}^d)$. For $\beta \in (0,\alpha]$ we define*

$$
\mathcal{D}^\beta_v(\mathbb{R}^n) := \left\{(f, f^v) \in \mathcal{C}^\alpha(\mathbb{R}^n) \times \mathcal{C}^\beta(\mathcal{L}(\mathbb{R}^d, \mathbb{R}^n)) : f^\sharp = f - \pi_<(f^v, v) \in \mathcal{C}^{\alpha+\beta}(\mathbb{R}^n)\right\}.
$$

When there is no ambiguity about the target set, we also write $\mathcal{D}^\beta_v$ instead of $\mathcal{D}^\beta_v(\mathbb{R}^n)$. If $(f, f^v) \in \mathcal{D}^\beta_v$, then f is called paracontrolled *by v. The function f^v is called the* derivative *of f with respect to v. Abusing notation, we write $f \in \mathcal{D}^\beta_v$ when it is clear from the context what the derivative f^v is supposed to be. We equip $\mathcal{D}^\beta_v$ with the norm*

$$
\|f\|_{v,\beta} := \|f^v\|_\beta + \|f^\sharp\|_{\alpha+\beta}.
$$

If $\tilde{v} \in \mathcal{C}^\alpha$ and $(\tilde{f}, \tilde{f}^{\tilde{v}}) \in \mathcal{D}^\beta_{\tilde{v}}$, then we also write

$$
d_{\mathcal{D}^\beta}(f, \tilde{f}) := \|f^v - \tilde{f}^{\tilde{v}}\|_\beta + \|f^\sharp - \tilde{f}^\sharp\|_{\alpha+\beta}.
$$

Example 4.2. Let $\alpha \in (0,1)$, $\beta \in (0,\alpha]$, and $v \in \mathcal{C}^\alpha$. Then Proposition 3.5 shows that $F(v) \in \mathcal{D}^\beta_v$ for every $F \in C_b^{1+\beta/\alpha}$, with derivative $\mathrm{D}F(v)$.

Example 4.3. Let $\alpha + \beta > 1$ and $v \in \mathcal{C}^\alpha$, $w \in \mathcal{C}^\beta$. Then by (3.12), the Young integral $I(v, \mathrm{d}w)$ is in $\mathcal{D}^\alpha_w$, with derivative v.

Example 4.4. If $\alpha + \beta < 1$ and $v \in \mathcal{C}^\alpha$, then $(f, f^v) \in \mathcal{D}^\beta_v$ if and only if $|f_{s,t} - f^v(s) v_{s,t}| \lesssim |t-s|^{\alpha+\beta}$ and in that case

$$
\|f^v\|_\infty + \sup_{s \neq t} \frac{|f^v_{s,t}|}{|t-s|^\beta} + \sup_{s\neq t} \frac{|f_{s,t} - f^v_s v_{s,t}|}{|t-s|^{\alpha+\beta}} \lesssim \|f\|_{v,\beta}(1 + \|v\|_\alpha).
$$

Indeed we have $|f^v(s)v_{s,t} - \pi_<(f^v,v)_{s,t}| \lesssim |t-s|^{\alpha+\beta}\|f^v\|_\beta\|v\|_\alpha$, which can be shown using similar arguments as for Lemma B.2 in [23]. In other words, for $\alpha \in (0,1/2)$ the space $\mathcal{D}_v^\alpha$ coincides with the space of controlled paths defined in Section 2.2.

The following associativity result, the analog of Theorem 2.3 of [8] in our setting, will be useful for establishing some stability properties of $\mathcal{D}_v^\beta$.

Lemma 4.5. *Let $\alpha,\beta \in (0,1)$, and let $u \in C([0,1],\mathcal{L}(\mathbb{R}^n;\mathbb{R}^m))$, $v \in \mathcal{C}^\alpha(\mathcal{L}(\mathbb{R}^d;\mathbb{R}^n))$, and $w \in \mathcal{C}^\beta(\mathbb{R}^d)$. Then*

$$\|\pi_<(u,\pi_<(v,w)) - \pi_<(uv,w)\|_{\alpha+\beta} \lesssim \|u\|_\infty\|v\|_\alpha\|w\|_\beta.$$

Proof. We have

$$\pi_<(u,\pi_<(v,w)) - \pi_<(uv,w) = \sum_{p,m}(S_{p-1}u(\pi_<(v,w))_{pm} - S_{p-1}(uv)w_{pm})\varphi_{pm}$$

and $[S_{p-1}u(\pi_<(v,w))_{pm} - S_{p-1}(uv)w_{pm}]|_{[t^0_{pm},t^2_{pm}]}$ is affine. By Lemma 3.2 it suffices to control $\|[S_{p-1}u(\pi_<(v,w))_{pm} - S_{p-1}(uv)w_{pm}]|_{[t^0_{pm},t^2_{pm}]}\|_\infty$.

The cases $(p,m)=(-1,0)$ and $(p,m)=(0,0)$ are easy, so let $p \geq 0$ and $m \geq 1$. For $r<q<p$ we denote by m_q and m_r the unique index in generation q and r respectively for which $\chi_{pm}\varphi_{qm_q} \not\equiv 0$ and similarly for r. We apply Lemma 3.9 to obtain for $q<p$

$$\begin{aligned}
|(S_{q-1}v\Delta_q w)_{pm}| &= \Big|\sum_{r<q} v_{rm_r}w_{qm_q}2^{-p}\langle\chi_{pm},\mathrm{d}(\varphi_{rm_r}\varphi_{qm_q})\rangle\Big| \\
&= \Big|\sum_{r<q} v_{rm_r}w_{qm_q}2^{-p}\langle\chi_{pm},\chi_{rm_r}\varphi_{qm_q}+\varphi_{rm_r}\chi_{qm_q}\rangle\Big| \\
&\leq \|v\|_\alpha\|w\|_\beta\sum_{r<q}2^{-r\alpha}2^{-q\beta}2^{-p}2^{-2p+r+p+q} \lesssim 2^{-2p+q(2-\alpha-\beta)}\|v\|_\alpha\|w\|_\beta.
\end{aligned}$$

Hence

$$\Big\|\Big(S_{p-1}u\sum_{q<p}(S_{q-1}v\Delta_q w)_{pm}\Big)\Big|_{[t^0_{pm},t^2_{pm}]}\Big\|_\infty \lesssim \|u\|_\infty\|v\|_\alpha\|w\|_\beta 2^{-p(\alpha+\beta)}.$$

If $p<q$, then $\Delta_q w(t^k_{pm})=0$ for all k and m, and therefore $(S_{q-1}v\Delta_q w)_{pm}=0$, so that it only remains to bound $\|[S_{p-1}u(S_{p-1}v\Delta_p w)_{pm} - S_{p-1}(uv)w_{pm}]|_{[t^0_{pm},t^2_{pm}]}\|_\infty$. We have $\Delta_p w(t^0_{pm}) = \Delta_p w(t^2_{pm}) = 0$ and $\Delta_p w(t^1_{pm}) = w_{pm}/2$. On $[t^0_{pm},t^2_{pm}]$, the function $S_{p-1}v$ is given by the linear interpolation of $v(t^0_{pm})$ and $v(t^2_{pm})$, and therefore $(S_{p-1}v\Delta_p w)_{pm} = \frac{1}{2}(v(t^0_{pm})+v(t^2_{pm}))w_{pm}$, leading to

$$\begin{aligned}
&\|[S_{p-1}u(S_{p-1}v\Delta_p w)_{pm} - S_{p-1}(uv)w_{pm}]|_{[t^0_{pm},t^2_{pm}]}\|_\infty \\
&\qquad \leq |w_{pm}| \times \Big\|\Big[\Big(u(t^0_{pm}) + \frac{\cdot - t^0_{pm}}{t^2_{pm}-t^0_{pm}}u_{t^0_{pm},t^2_{pm}}\Big)\frac{v(t^0_{pm})+v(t^2_{pm})}{2} \\
&\qquad\qquad - \Big((uv)(t^0_{pm}) + \frac{\cdot - t^0_{pm}}{t^2_{pm}-t^0_{pm}}(uv)_{t^0_{pm},t^2_{pm}}\Big)\Big]\Big|_{[t^0_{pm},t^2_{pm}]}\Big\|_\infty \\
&\qquad \lesssim \|u\|_\infty\|v\|_\alpha\|w\|_\beta 2^{-p(\alpha+\beta)},
\end{aligned}$$

where the last step follows by rebracketing. □

As a consequence, we can show that paracontrolled paths are stable under the application of sufficiently smooth functions.

Corollary 4.6. *Let $\alpha \in (0,1)$, $\beta \in (0,\alpha]$, $v \in \mathcal{C}^\alpha(\mathbb{R}^d)$, and $f \in \mathcal{D}_v^\beta(\mathbb{R}^n)$ with derivative f^v. Let $F \in C_b^{1+\beta/\alpha}(\mathbb{R}^n,\mathbb{R}^m)$. Then $F(f) \in \mathcal{D}_v^\beta(\mathbb{R}^m)$ with derivative $\mathrm{D}F(f)f^v$, and*

$$\|F(f)\|_{v,\beta} \lesssim \|F\|_{C_b^{1+\beta/\alpha}}(1+\|v\|_\alpha)^{1+\beta/\alpha}(1+\|f\|_{v,\beta})(1+\|f^v\|_\infty)^{1+\beta/\alpha}.$$

Moreover, there exists a polynomial P *which satisfies for all* $F \in C_b^{2+\beta/\alpha}(\mathbb{R}^n, \mathbb{R}^m)$*,* $\tilde{v} \in \mathcal{C}^\alpha(\mathbb{R}^d)$*,* $\tilde{f} \in \mathcal{D}^\beta_{\tilde{v}}(\mathbb{R}^n)$*, and*

$$M = \max\{\|v\|_\alpha, \|\tilde{v}\|_\alpha, \|f\|_{v,\beta}, \|\tilde{f}\|_{\tilde{v},\beta}\}$$

the bound

$$d_{\mathcal{D}^\beta}(F(f), F(\tilde{f})) \le P(M)\|F\|_{C_b^{2+\beta/\alpha}}(d_{\mathcal{D}^\beta}(f, \tilde{f}) + \|v - \tilde{v}\|_\alpha).$$

Proof. The estimate for $\|\mathrm{D}F(f)f^v\|_\beta$ is straightforward. For the remainder we apply Proposition 3.5 and Lemma 4.5 to obtain

$$\begin{aligned}
\|F(f)^\sharp\|_{\alpha+\beta} &\le \|F(f) - \pi_<(\mathrm{D}F(f), f)\|_{\alpha+\beta} + \|\pi_<(\mathrm{D}F(f), f^\sharp)\|_{\alpha+\beta} \\
&\quad + \|\pi_<(\mathrm{D}F(f), \pi_<(f^v, v)) - \pi_<(\mathrm{D}F(f)f^v, v)\|_{\alpha+\beta} \\
&\lesssim \|F\|_{C_b^{1+\beta/\alpha}}(1 + \|\pi_<(f^v, v)\|_\alpha)^{1+\beta/\alpha}(1 + \|f^\sharp\|_{\alpha+\beta}) \\
&\quad + \|F\|_{C_b^1}\|f\|_{v,\beta} + \|F\|_{C_b^1}\|f^v\|_\beta\|v\|_\alpha \\
&\lesssim \|F\|_{C_b^{1+\beta/\alpha}}(1 + \|f^v\|_\infty)^{1+\beta/\alpha}(1 + \|v\|_\alpha)^{1+\beta/\alpha}(1 + \|f\|_{v,\beta}).
\end{aligned}$$

The difference $F(f) - F(\tilde{f})$ is treated in the same way. □

When solving differential equations it will be crucial to have a bound which is linear in $\|f\|_{v,\beta}$. The superlinear dependence on $\|f^v\|_\infty$ will not pose any problem as we will always have $f^v = F(\tilde{f})$ for some suitable $\tilde{f}$, so that for bounded F we get $\|F(f)\|_{v,\beta} \lesssim_{F,v} 1 + \|f\|_{v,\beta}$.

4.2 A basic commutator estimate

Here we prove the commutator estimate which will be the main ingredient in the construction of the integral $I(f, \mathrm{d}g)$, where f is paracontrolled by v and g is paracontrolled by w, and where we assume that the integral $I(v, \mathrm{d}w)$ exists.

Proposition 4.7. *Let* $\alpha, \beta, \gamma \in (0, 1)$*, and assume that* $\alpha + \beta + \gamma > 1$ *and* $\beta + \gamma < 1$*. Let* $f \in \mathcal{C}^\alpha(\mathcal{L}(\mathbb{R}^n; \mathbb{R}^m))$*,* $v \in \mathcal{C}^\beta(\mathcal{L}(\mathbb{R}^d; \mathbb{R}^n))$*, and* $w \in \mathcal{C}^\gamma(\mathbb{R}^d)$*. Then the "commutator"*

$$\begin{aligned}
C(f, v, w) &:= L(\pi_<(f, v), w) - I(f, \mathrm{d}L(v, w)) \\
&:= \lim_{N\to\infty}[L(S_N(\pi_<(f, v)), S_N w) - I(f, \mathrm{d}L(S_N v, S_N w))] \\
&= \lim_{N\to\infty} \sum_{p\le N} \sum_{q<p} \Bigg[\int_0^\cdot \Delta_p(\pi_<(f, v))(s)\mathrm{d}\Delta_q w(s) - \int_0^\cdot \mathrm{d}(\Delta_q(\pi_<(f, v)))(s)\Delta_p w(s) \\
&\qquad - \Big(\int_0^\cdot f(s)\Delta_p v(s)\mathrm{d}\Delta_q w(s) - \int_0^\cdot f(s)\mathrm{d}(\Delta_q v)(s)\Delta_p w(s)\Big)\Bigg]
\end{aligned} \tag{4.1}$$

converges in $\mathcal{C}^{\alpha+\beta+\gamma-\varepsilon}(\mathbb{R}^m)$ *for all* $\varepsilon > 0$*. Moreover,*

$$\|C(f, v, w)\|_{\alpha+\beta+\gamma} \lesssim \|f\|_\alpha\|v\|_\beta\|w\|_\gamma.$$

Proof. We only argue for the first difference in (4.1), i.e. for

$$X_N := \sum_{p\le N} \sum_{q<p} \left[\int_0^\cdot \Delta_p(\pi_<(f, v))(s)\mathrm{d}\Delta_q w(s) - \int_0^\cdot f(s)\Delta_p v(s)\mathrm{d}\Delta_q w(s)\right]. \tag{4.2}$$

The second difference can be handled using the same arguments. First we prove that (X_N) converges uniformly, then we show that $\|X_N\|_{\alpha+\beta+\gamma}$ stays uniformly bounded. This will imply the desired result, since bounded sets in $\mathcal{C}^{\alpha+\beta+\gamma}$ are relatively compact in $\mathcal{C}^{\alpha+\beta+\gamma-\varepsilon}$.

To prove uniform convergence, note that

$$\begin{aligned} X_N - X_{N-1} &= \sum_{q<N} \left[\int_0^\cdot \Delta_N(\pi_<(f,v))(s)\mathrm{d}\Delta_q w(s) - \int_0^\cdot f(s)\Delta_N v(s)\mathrm{d}\Delta_q w(s) \right] \\ &= \sum_{q<N} \Bigg[\sum_{j\le N}\sum_{i<j} \int_0^\cdot \Delta_N(\Delta_i f \Delta_j v)(s)\mathrm{d}\Delta_q w(s) \\ &\qquad - \sum_{j\ge N}\sum_{i\le j} \int_0^\cdot \Delta_j(\Delta_i f \Delta_N v)(s)\mathrm{d}\Delta_q w(s) \Bigg], \end{aligned} \tag{4.3}$$

where for the second term it is possible to take the infinite sum over j outside of the integral because $\sum_j \Delta_j(\Delta_i f \Delta_N v)$ converges uniformly to $\Delta_i f \Delta_N v$ and because $\Delta_q w$ is a finite variation path. We also used that $\Delta_N(\Delta_i f \Delta_j v) = 0$ whenever $i > N$ or $j > N$. Only very few terms in (4.3) cancel. Nonetheless these cancellations are crucial, since they eliminate most terms for which we only have the worse estimate (3.11) in Corollary 3.11. We obtain

$$\begin{aligned} X_N - X_{N-1} &= \sum_{q<N}\sum_{j<N}\sum_{i<j} \int_0^\cdot \Delta_N(\Delta_i f \Delta_j v)(s)\mathrm{d}\Delta_q w(s) \\ &\quad - \sum_{q<N} \int_0^\cdot \Delta_N(\Delta_N f \Delta_N v)(s)\mathrm{d}\Delta_q w(s) \\ &\quad - \sum_{q<N}\sum_{j>N}\sum_{i<j} \int_0^\cdot \Delta_j(\Delta_i f \Delta_N v)(s)\mathrm{d}\Delta_q w(s) \\ &\quad - \sum_{q<N}\sum_{j>N} \int_0^\cdot \Delta_j(\Delta_j f \Delta_N v)(s)\mathrm{d}\Delta_q w(s). \end{aligned} \tag{4.4}$$

Note that $\|\partial_t \Delta_q w\|_\infty \lesssim 2^q \|\Delta_q w\|_\infty$. Hence, an application of Corollary 3.11, where we use (3.10) for the first three terms and (3.11) for the fourth term, yields

$$\begin{aligned} \|X_N - X_{N-1}\|_\infty &\lesssim \|f\|_\alpha \|v\|_\beta \|w\|_\gamma \Bigg[\sum_{q<N}\sum_{j<N}\sum_{i<j} 2^{-2N+i+j} 2^{-i\alpha} 2^{-j\beta} 2^{q(1-\gamma)} \\ &\qquad + \sum_{q<N} 2^{-N(\alpha+\beta)} 2^{q(1-\gamma)} + \sum_{q<N}\sum_{j>N}\sum_{i<j} 2^{-2j+i+N} 2^{-i\alpha} 2^{-N\beta} 2^{q(1-\gamma)} \\ &\qquad + \sum_{q<N}\sum_{j>N} 2^{-j\alpha} 2^{-N\beta} 2^{q(1-\gamma)} \Bigg] \\ &\lesssim \|f\|_\alpha \|v\|_\beta \|w\|_\gamma 2^{-N(\alpha+\beta+\gamma-1)}, \end{aligned} \tag{4.5}$$

where in the last step we used $\alpha, \beta, \gamma < 1$. Since $\alpha + \beta + \gamma > 1$, this gives us the uniform convergence of (X_N).

Next let us show that $\|X_N\|_{\alpha+\beta+\gamma} \lesssim \|f\|_\alpha \|v\|_\beta \|w\|_\gamma$ for all N. Similarly to (4.4) we obtain for $n \in \mathbb{N}$

$$\begin{aligned} \Delta_n X_N = \sum_{p\le N}\sum_{q<p} \Delta_n \Bigg[&\sum_{j<p}\sum_{i<j} \int_0^\cdot \Delta_p(\Delta_i f \Delta_j v)(s)\mathrm{d}\Delta_q w(s) - \int_0^\cdot \Delta_p(\Delta_p f \Delta_p v)(s)\mathrm{d}\Delta_q w(s) \\ &- \sum_{j>p}\sum_{i\le j} \int_0^\cdot \Delta_j(\Delta_i f \Delta_p v)(s)\mathrm{d}\Delta_q w(s) \Bigg], \end{aligned}$$

and therefore by Corollary 3.10

$$\|\Delta_n X_N\|_\infty \lesssim \sum_p \sum_{q<p} \Bigg[\sum_{j<p} \sum_{i<j} 2^{-(n\vee p)-n+p+q} \|\Delta_p(\Delta_i f \Delta_j v)\|_\infty \|\Delta_q w\|_\infty + 2^{-(n\vee p)-n+p+q} \|\Delta_p(\Delta_p f \Delta_p v)\|_\infty \|\Delta_q w\|_\infty + \sum_{j>p} \sum_{i\leq j} 2^{-(n\vee j)-n+j+q} \|\Delta_j(\Delta_i f \Delta_p v)\|_\infty \|\Delta_q w\|_\infty \Bigg].$$

Now we apply Corollary 3.11, where for the last term we distinguish the cases $i<j$ and $i=j$. Using that $1-\gamma>0$, we get

$$\begin{aligned}\|\Delta_n X_N\|_\infty &\lesssim \|f\|_\alpha \|v\|_\beta \|w\|_\gamma \sum_p 2^{p(1-\gamma)} \Bigg[\sum_{j<p} \sum_{i<j} 2^{-(n\vee p)-n+p} 2^{-2p} 2^{i(1-\alpha)} 2^{j(1-\beta)} \\ &\qquad + 2^{-(n\vee p)-n+p} 2^{-p\alpha} 2^{-p\beta} \\ &\qquad + \sum_{j>p} \sum_{i<j} 2^{-(n\vee j)-n+j} 2^{-2j+i(1-\alpha)+p(1-\beta)} \\ &\qquad + \sum_{j>p} 2^{-(n\vee j)-n+j} 2^{-j\alpha-p\beta} \Bigg] \\ &\lesssim \|f\|_\alpha \|v\|_\beta \|w\|_\gamma 2^{-n(\alpha+\beta+\gamma)},\end{aligned}$$

where we used both that $\alpha+\beta+\gamma>1$ and that $\beta+\gamma<1$. □

Remark 4.8. If $\beta+\gamma=1$, we can apply Proposition 4.7 with $\beta-\varepsilon$ to obtain that $C(f,v,w)\in\mathcal{C}^{\alpha+\beta+\gamma-\varepsilon}$ for every sufficiently small $\varepsilon>0$. If $\beta+\gamma>1$, then we are in the Young setting and there is no need to introduce the commutator.

For later reference, we collect the following result from the proof of Proposition 4.7:

Lemma 4.9. *Let $\alpha,\beta,\gamma,f,v,w$ be as in Proposition 4.7. Then*

$$\|C(f,v,w)-L(S_N(\pi_<(f,v)),S_N w)-I(f,\mathrm{d}L(S_N v,S_N w))\|_\infty \lesssim 2^{-N(\alpha+\beta+\gamma-1)}\|f\|_\alpha\|v\|_\beta\|w\|_\gamma.$$

Proof. Simply sum up (4.5) over N. □

4.3 Pathwise integration for paracontrolled paths

In this section we apply the commutator estimate to construct the rough path integral under the assumption that the Lévy area exists for a given reference path.

Theorem 4.10. *Let $\alpha\in(1/3,1)$, $\beta\in(0,\alpha]$ and assume that $2\alpha+\beta>1$ as well as $\alpha+\beta\neq1$. Let $v=(v^1,\dots,v^d)\in\mathcal{C}^\alpha(\mathbb{R}^d)$ and assume that the Lévy area*

$$L(v,v):=\lim_{N\to\infty}\big(L(S_N v^k,S_N v^\ell)\big)_{1\leq k\leq d,1\leq\ell\leq d}$$

converges uniformly and that $\sup_N\|L(S_N v,S_N v)\|_{2\alpha}<\infty$. Let $f\in\mathcal{D}^\alpha_v(\mathcal{L}(\mathbb{R}^d,\mathbb{R}^n))$. Then $I(S_N f,\mathrm{d}S_N v)$ converges in $\mathcal{C}^{\alpha-\varepsilon}(\mathbb{R}^n)$ for all $\varepsilon>0$. Denoting the limit by $I(f,\mathrm{d}v)$, we have

$$\|I(f,\mathrm{d}v)\|_\alpha \lesssim \|f\|_{v,\beta}\big(\|v\|_\alpha+\|v\|_\alpha^2+\|L(v,v)\|_{2\alpha}\big).$$

Moreover, $I(f,\mathrm{d}v)\in\mathcal{D}^\alpha_v(\mathbb{R}^n)$ with derivative f and

$$\|I(f,\mathrm{d}v)\|_{v,\alpha} \lesssim \|f\|_{v,\beta}\big(1+\|v\|_\alpha^2+\|L(v,v)\|_{2\alpha}\big).$$

Proof. If $\alpha + \beta > 1$, everything follows from the Young case, Theorem 3.14, so let $\alpha + \beta < 1$. We decompose

$$\begin{aligned} I(S_N f, \mathrm{d}S_N v) &= S(S_N f, S_N v) + \pi_<(S_N f, S_N v) + L(S_N f^\sharp, S_N v) \\ &\quad + [L(S_N \pi_<(f^v, v), S_N v) - I(f^v, \mathrm{d}L(S_N v, S_N v))] + I(f^v, \mathrm{d}L(S_N v, S_N v)). \end{aligned}$$

The convergence then follows from Proposition 4.7 and Theorem 3.14. The limit is given by

$$I(f, \mathrm{d}v) = S(f, v) + \pi_<(f, v) + L(f^\sharp, v) + C(f^v, v, v) + I(f^v, \mathrm{d}L(v, v)), \tag{4.6}$$

from where we easily deduce the claimed bounds. □

Remark 4.11. Since $I(f, \mathrm{d}v) = \lim_{N\to\infty} \int_0^\cdot S_N f \mathrm{d}S_N v$, the integral is a local operator in the sense that $I(f, \mathrm{d}v)$ is constant on every interval $[s, t]$ for which $f|_{[s,t]} = 0$. In particular we can estimate $I(f, \mathrm{d}v)|_{[0,t]}$ using only $f|_{[0,t]}$ and $f^v|_{[0,t]}$.

Remark 4.12. Let us advertise two nice properties of our approach. First, note that using similar arguments as in the proof of Lemma B.2 in [23] it is possible to derive the bound

$$|\pi_<(v, v)_{s,t} - v(s) \otimes v_{s,t}| \lesssim |t - s|^{2\alpha} \|v\|_\alpha^2 \tag{4.7}$$

for all $0 \le s < t \le 1$ and whenever $\alpha < 1/2$. This means that for *any* $\varphi \in \mathcal{C}^{2\alpha}([0, 1], \mathbb{R}^{d\otimes d})$ the pair $(v, \mathbb{V})$ with

$$\mathbb{V}_{s,t} = (\varphi + \pi_<(v, v) + S(v, v))_{s,t} - v(s) \otimes v_{s,t} \tag{4.8}$$

is an α-rough path, which is weakly geometric if and only if φ is antisymmetric (we refer to [16] for the definition of (weakly geometric) α-rough paths). Setting $\varphi \equiv 0$ gives a construction which is quite similar in spirit to the ones of Unterberger [51, 52] – of course only for $\alpha > 1/3$. Note that by varying φ we have a simple construction of all rough paths above v.

Next, note that the paracontrolled approach equips us with a natural approximation theory for (para)controlled paths. Indeed, if $v \in \mathcal{C}^\alpha$ and $f = \pi_<(f^v, v) + f^\sharp \in \mathcal{D}_v^\alpha$, and if (v_N) is a sequence converging to v in $\mathcal{C}^\alpha$, then we have

$$f_N := \pi_<(f^v, v_N) + f^\sharp \in \mathcal{D}_{v_N}^\alpha \tag{4.9}$$

and $d_{\mathcal{D}^\alpha}(f, f_N) = 0$ for all N. Now (4.7) and Remark 3.7 show that the paracontrolled paths are exactly the controlled paths, so we obtain an approximation of an arbitrary controlled path. In the classical approach it is not obvious how to construct a sequence of smooth functions f_N that are controlled by v_N and that approximate f in controlled path distance (see e.g. Remark 4.9 of [16]).

Let us now combine these two insights to prove that every rough path integral can be obtained as limit of Young integrals. Define for $\varphi \in \mathcal{C}^{2\alpha}([0, 1], \mathbb{R}^{d\times d})$ the bounded operator $I^\varphi \colon \mathcal{D}_v^\alpha \to \mathcal{D}_v^\alpha$ by

$$I^\varphi(f, \mathrm{d}v) := S(f, v) + \pi_<(f, v) + L(f^\sharp, v) + C(f^v, v, v) + I(f^v, \mathrm{d}\varphi),$$

in other words we replace $L(v, v)$ by φ in (4.6). Not surprisingly, what we obtain is nothing else than the rough path integral of $f \in \mathcal{D}_v^\alpha$ with respect to $(v, \mathbb{V})$, where $\mathbb{V}$ was defined in (4.8). If φ is antisymmetric, this is easy to see: Let (v_N) be a sequence of smooth paths such that $(v_N, \mathbb{V}_N)$ converges to $(v, \mathbb{V})$ in rough path topology (for $\alpha' \in (1/3, \alpha)$). Since S and $\pi_<$ are bounded operators, this is equivalent to $L(v_N, v_N)$ converging to φ in $\mathcal{C}^{2\alpha'}$ topology. Define for $f \in \mathcal{D}_v^\alpha$ the sequence (f_N) as in (4.9). Then the Young integrals $\int_0^\cdot f_N(s)\mathrm{d}v_N(s)$ converge to the controlled path integral of f with respect to $(v, \mathbb{V})$, but also to $I^\varphi(f, \mathrm{d}v)$, which proves our claim in the weakly geometric case. Otherwise, decompose $\varphi = \varphi^{\mathrm{symm}} + \varphi^{\mathrm{anti}}$ into symmetric and antisymmetric part and observe that $I^\varphi(f, \mathrm{d}v) = I^{\varphi^{\mathrm{anti}}}(f, \mathrm{d}v) + I(f^v, \mathrm{d}\varphi^{\mathrm{symm}})$, and the same relation holds for the classical rough path integral. Of course, we should point out that for general φ which are not obtained as limit of the piecewise linear dyadic approximations $(S_N v)$ we do not have the nice interpretation $I^\varphi(f, \mathrm{d}v) = \lim_N I(S_N f, \mathrm{d}S_N v)$.

For fixed v and $L(v, v)$, the map $f \mapsto I(f, \mathrm{d}v)$ is linear and bounded from $\mathcal{D}^\beta_v$ to $\mathcal{D}^\alpha_v$, and this is what we will need to solve differential equations driven by v. But we can also estimate the speed of convergence of $I(S_N f, \mathrm{d}S_N v)$ to $I(f, \mathrm{d}v)$, measured in uniform distance:

Corollary 4.13. *Let $\alpha \in (1/3, 1/2]$ and let β, v, f be as in Theorem 4.10. Then we have for all $\varepsilon \in (0, 2\alpha + \beta - 1)$*

$$\begin{aligned}\|I(S_N f, \mathrm{d}S_N v) - I(f, \mathrm{d}v)\|_\infty \lesssim_\varepsilon\, & 2^{-N(2\alpha+\beta-1-\varepsilon)}\|f\|_{v,\beta}\big(\|v\|_\alpha + \|v\|^2_\alpha\big) \\ & + \|f^v\|_\beta \|L(S_N v, S_N v) - L(v, w)\|_{2\alpha-\varepsilon}.\end{aligned}$$

Proof. We decompose $I(S_N f, \mathrm{d}S_N v)$ as described in the proof of Theorem 4.10. This gives us for example the term

$$\|\pi_<(S_N f - f, S_N v) + \pi_<(f, S_N v - v)\|_\infty \lesssim_\varepsilon \|S_N f - f\|_\infty \|v\|_\alpha + \|f\|_\infty \|f\|_\alpha \|S_N v - v\|_\varepsilon$$

for all $\varepsilon > 0$. From here it is easy to see that

$$\|\pi_<(S_N f - f, S_N g) + \pi_<(f, S_N g - g)\|_\infty \lesssim 2^{-N(\alpha-\varepsilon)}\|f\|_\alpha \|v\|_\alpha \lesssim 2^{-N(\alpha-\varepsilon)}\|f\|_{v,\beta}(\|v\|_\alpha + \|v\|^2_\alpha).$$

But now $\beta \le \alpha \le 1/2$ and therefore $\alpha \ge 2\alpha + \beta - 1$.

Let us treat one of the critical terms, say $L(S_N f^\sharp, S_N v) - L(f^\sharp, v)$. Since $2\alpha + \beta - \varepsilon > 1$, we can apply Lemma 3.12 to obtain

$$\begin{aligned}\|L(S_N f^\sharp, S_N v) - L(f^\sharp, v)\|_\infty &\lesssim \|L(S_N f^\sharp - f^\sharp, S_N v)\|_{1+\varepsilon} + \|L(f^\sharp, S_N v - v)\|_{1+\varepsilon} \\ &\lesssim_\varepsilon \|S_N f^\sharp - f^\sharp\|_{1+\varepsilon-\alpha}\|v\|_\alpha + \|f^\sharp\|_{\alpha+\beta}\|S_N v - v\|_{1+\varepsilon-\alpha-\beta} \\ &\lesssim 2^{-N(\alpha+\beta-(1+\varepsilon-\alpha))}\|f^\sharp\|_{\alpha+\beta}\|v\|_\alpha \\ &\quad + 2^{-N(\alpha-(1+\varepsilon-\alpha-\beta))}\|f^\sharp\|_{\alpha+\beta}\|v\|_\alpha \\ &\lesssim 2^{-N(2\alpha+\beta-1-\varepsilon)}\|f^\sharp\|_{\alpha+\beta}\|v\|_\alpha.\end{aligned}$$

Lemma 4.9 gives

$$\begin{aligned}\|L(S_N \pi_<(f^v, v), S_N v) - L(\pi_<(f^v, v), v)\|_\infty \lesssim\, & 2^{-N(2\alpha+\beta-1)}\|f^v\|_\beta \|v\|^2_\alpha \\ & + \|I(f^v, \mathrm{d}L(S_N v, S_N v)) - I(f^v, \mathrm{d}L(v, v))\|_\infty.\end{aligned}$$

The second term on the right hand side can be estimated using the continuity of the Young integral, and the proof is complete. □

Remark 4.14. In Lemma 4.9 we saw that the rate of convergence of

$$L(S_N \pi_<(f^v, v), S_N v) - I(f^v, \mathrm{d}L(S_N v, S_N v)) - (L(\pi_<(f^v, v), v) - I(f^v, \mathrm{d}L(v, v)))$$

is in fact $2^{-N(2\alpha+\beta-1)}$ when measured in uniform distance, and not just $2^{-N(2\alpha+\beta-1-\varepsilon)}$. It is possible to show that this optimal rate is attained by the other terms as well, so that

$$\begin{aligned}\|I(S_N f, \mathrm{d}S_N v) - I(f, \mathrm{d}v)\|_\infty \lesssim\, & 2^{-N(2\alpha+\beta-1)}\|f\|_{v,\beta}\big(\|v\|_\alpha + \|v\|^2_\alpha\big) \\ & + \|f^v\|_\beta \|L(S_N v, S_N w) - L(v, w)\|_{2\alpha-\varepsilon}.\end{aligned}$$

But this requires a rather lengthy calculation, so we decided not to include the arguments here.

Since we approximate f and g by the piecewise smooth functions $S_N f$ and $S_N g$ when defining the integral $I(f, \mathrm{d}g)$, it is not surprising that we obtain a Stratonovich type integral:

Proposition 4.15. *Let $\alpha \in (1/3, 1)$ and $v \in \mathcal{C}^\alpha(\mathbb{R}^d)$. Let $\varepsilon > 0$ be such that $(2+\varepsilon)\alpha > 1$ and let $F \in C^{2+\varepsilon}(\mathbb{R}^d, \mathbb{R}^n)$. Then*

$$F(v(t)) - F(v(0)) = I(\mathrm{D}F(v), \mathrm{d}v)(t) := \lim_{N\to\infty} I(S_N \mathrm{D}F(v), \mathrm{d}S_N v)(t)$$

for all $t \in [0,1]$.

Proof. The function $S_N v$ is Lipschitz continuous, so that integration by parts gives

$$F(S_N v(t)) - F(S_N v(0)) = I(\mathrm{D}F(S_N v), \mathrm{d}S_N v)(t).$$

The left hand side converges to $F(v(t)) - F(v(0))$. It thus suffices to show that $I(S_N \mathrm{D}F(v) - \mathrm{D}F(S_N v), \mathrm{d}S_N v)$ converges to zero. By the continuity of the Young integral, Theorem 3.14, it suffices to show that $\lim_{N\to\infty} \|S_N \mathrm{D}F(v) - \mathrm{D}F(S_N v)\|_{\alpha(1+\varepsilon')} = 0$ for all $\varepsilon' < \varepsilon$. Recall that $S_N v$ is the linear interpolation of v between the points (t^1_{pm}) for $p \le N$ and $0 \le m \le 2^p$, and therefore $\Delta_p \mathrm{D}F(S_N v) = \Delta_p \mathrm{D}F(v) = \Delta_p S_N \mathrm{D}F(v)$ for all $p \le N$. For $p > N$ and $1 \le m \le 2^p$ we apply a first order Taylor expansion to both terms and use the ε–Hölder continuity of $\mathrm{D}^2 F$ to obtain

$$|[S_N \mathrm{D}F(v) - \mathrm{D}F(S_N v)]_{pm}| \le C_F 2^{-p\alpha(1+\varepsilon)} \|S_N v\|_\alpha$$

for a constant $C_F > 0$. Therefore, we get for all $\varepsilon' \le \varepsilon$

$$\|S_N \mathrm{D}F(v) - \mathrm{D}F(S_N v)\|_{\alpha(1+\varepsilon')} \lesssim_F 2^{-N\alpha(\varepsilon-\varepsilon')} \|v\|_\alpha,$$

which completes the proof. □

Remark 4.16. Note that here we did not need any assumption on the area $L(v, v)$. The reason are cancellations that arise due to the symmetric structure of the derivative of $\mathrm{D}F$, the Hessian of F.

Proposition 4.15 was previously obtained by Roynette [47], except that there v is assumed to be one dimensional and in the Besov space $B^{1/2}_{1,\infty}$.

5 Pathwise Itô integration

In the previous section we saw that our pathwise integral $I(f, \mathrm{d}v)$ is of Stratonovich type, i.e. it satisfies the usual integration by parts rule. But in applications it may be interesting to have an Itô integral. Here we show that a slight modification of $I(f, \mathrm{d}v)$ allows us to treat non-anticipating Itô-type integrals.

A natural approximation of a non-anticipating integral is given for $k \in \mathbb{N}$ by

$$\begin{aligned} I^{\text{Itô}}_k(f, \mathrm{d}v)(t) &:= \sum_{m=1}^{2^k} f(t^0_{km})(v(t^2_{km} \wedge t) - v(t^0_{km} \wedge t)) \\ &= \sum_{m=1}^{2^k} \sum_{p,q} \sum_{m,n} f_{pm} v_{qn} \varphi_{pm}(t^0_{km})(\varphi_{qn}(t^2_{km} \wedge t) - \varphi_{qn}(t^0_{km} \wedge t)). \end{aligned}$$

Let us assume for the moment that $t = m2^{-k}$ for some $0 \le m \le 2^k$. In that case we obtain for $p \ge k$ or $q \ge k$ that $\varphi_{pm}(t^0_{km})(\varphi_{qn}(t^2_{km} \wedge t) - \varphi_{qn}(t^0_{km} \wedge t)) = 0$. For $p, q < k$, both

φ_{pm} and φ_{qn} are affine functions on $[t^0_{km}\wedge t, t^2_{km}\wedge t]$, and for affine u and w and $s<t$ we have

$$u(s)(w(t)-w(s)) = \int_s^t u(r)\mathrm{d}w(r) - \frac{1}{2}(u(t)-u(s))(w(t)-w(s)).$$

Hence, we conclude that for $t = m2^{-k}$

$$\begin{aligned} I^{\text{Itô}}_k(f,\mathrm{d}v)(t) &= I(S_{k-1}f,\mathrm{d}S_{k-1}v)(t) - \frac{1}{2}[S_{k-1}f,S_{k-1}v]_k(t) \\ &= I(S_{k-1}f,\mathrm{d}S_{k-1}v)(t) - \frac{1}{2}[f,v]_k(t). \end{aligned} \tag{5.1}$$

Here we write $[f,v]_k$ for the k–th dyadic approximation of the quadratic covariation $[f,v]$, i.e.

$$[f,v]_k(t) := \sum_{m=1}^{2^k}[f(t^2_{km}\wedge t)-f(t^0_{km}\wedge t)][v(t^2_{km}\wedge t)-v(t^0_{km}\wedge t)]$$

and similarly for $[S_{k-1}f,S_{k-1}v]_k$, and $[f,v]$ is the uniform limit of the $([f,v]_k)$ whenever it exists. From now on we study the right hand side of (5.1) rather than $I^{\text{Itô}}_k(f,\mathrm{d}v)$, which is justified by the following remark.

Remark 5.1. Let $\alpha\in(0,1)$. If $f\in C([0,1],\mathcal{L}(\mathbb{R}^d;\mathbb{R}^n))$ and $v\in\mathcal{C}^\alpha(\mathbb{R}^d)$, then

$$\Big\| I^{\text{Itô}}_k(f,\mathrm{d}v) - \Big(I(S_{k-1}f,\mathrm{d}S_{k-1}v) - \frac{1}{2}[S_{k-1}f,S_{k-1}v]_k\Big)\Big\|_\infty \lesssim 2^{-k\alpha}\|f\|_\infty\|v\|_\alpha.$$

This holds because both functions agree in all dyadic points of the form $m2^{-k}$, and because between those points the integrals can pick up mass of at most $\|f\|_\infty 2^{-k\alpha}\|v\|_\alpha$.

We write $[v,v] := ([v^i,v^j])_{1\le i,j\le d}$ and $L(v,v) := (L(v^i,v^j))_{1\le i,j\le d}$, and similarly for all expressions of the same type.

Theorem 5.2. *Let $\alpha\in(0,1/2)$ and let $\beta\le\alpha$ be such that $2\alpha+\beta>1$. Let $v\in\mathcal{C}^\alpha(\mathbb{R}^d)$ and $f\in\mathcal{D}^\beta_v(\mathcal{L}(\mathbb{R}^d;\mathbb{R}^n))$. Assume that $(L(S_kv,S_kv))$ converges uniformly, with uniformly bounded $\mathcal{C}^{2\alpha}$ norm. Also assume that $([v,v]_k)$ converges uniformly. Then $(I^{\text{Itô}}_k(f,\mathrm{d}v))$ converges uniformly to a limit $I^{\text{Itô}}(f,\mathrm{d}v) = I(f,\mathrm{d}v)-[f,v]/2$ which satisfies*

$$\|I^{\text{Itô}}(f,\mathrm{d}v)\|_\infty \lesssim \|f\|_{v,\beta}(\|v\|_\alpha + \|v\|^2_\alpha + \|L(v,v)\|_{2\alpha} + \|[v,v]\|_\infty),$$

and where the quadrativ variation $[f,v]$ is given by

$$[f,v] = \int_0^\cdot f^v(s)\mathrm{d}[v,v](s) := \Big(\sum_{j,\ell=1}^d \int_0^\cdot (f^{ij})^{v,\ell}(s)\mathrm{d}[v^j,v^\ell](s)\Big)_{1\le i\le n}, \tag{5.2}$$

where $(f^{ij})^{v,\ell}$ is the ℓ–th component of the v–derivative of f^{ij}. For $\varepsilon\in(0,2\alpha+\beta-1)$ the speed of convergence can be estimated by

$$\begin{aligned}\big\|I^{\text{Itô}}(f,\mathrm{d}v) - I^{\text{Itô}}_k(f,\mathrm{d}v)\big\|_\infty \lesssim_\varepsilon\ & 2^{-k(2\alpha+\beta-1-\varepsilon)}\|f\|_{v,\beta}\big(\|v\|_\alpha+\|v\|^2_\alpha\big) \\ &+ \|f^v\|_\beta\|L(S_{k-1}v,S_{k-1}v)-L(v,v)\|_{2\alpha} \\ &+ \|f^v\|_\infty\|[v,v]_k-[v,v]\|_\infty.\end{aligned}$$

Proof. By Remark 5.1, it suffices to show our claims for $I(S_{k-1}f,\mathrm{d}S_{k-1}v)-(1/2)[f,v]_k$. The statements for the integral $I(S_{k-1}f,\mathrm{d}S_{k-1}g)$ follow from Theorem 4.10 and Corollary 4.13. So let us us concentrate on the quadratic variation $[f,v]_k$. Recall from

Example 4.4 that $f \in \mathcal{D}_v^\beta$ if and only if $R^f_{s,t} = f_{s,t} - f^v(s) w_{s,t}$ satisfies $|R^f_{s,t}| \lesssim |t-s|^{\alpha+\beta}$. Hence

$$
\begin{aligned}
[f,v]^i_k(t) &= \sum_m \big(f_{t^0_{km}\wedge t, t^2_{km}\wedge t} v_{t^0_{km}\wedge t, t^2_{km}\wedge t}\big)^i \\
&= \sum_m \big(R^f_{t^0_{km}\wedge t, t^2_{km}\wedge t} v_{t^0_{km}\wedge t, t^2_{km}\wedge t}\big)^i + \sum_{j,\ell=1}^d \sum_m (f^{ij})^{v,\ell}(t^0_{km}\wedge t) v^\ell_{t^0_{km}\wedge t, t^2_{km}\wedge t} v^j_{t^0_{km}\wedge t, t^2_{km}\wedge t}.
\end{aligned}
$$

It is easy to see that the first term on the right hand side is bounded by

$$
\Big| \sum_m \big(R^f_{t^0_{km}\wedge t, t^2_{km}\wedge t} v_{t^0_{km}\wedge t, t^2_{km}\wedge t}\big)^i \Big| \lesssim 2^{-k(2\alpha+\beta-1)} \|f\|_{v,\beta} (\|v\|_\alpha + \|v\|^2_\alpha).
$$

For the second term, let us fix ℓ and j. Then the sum over m is just the integral of $(f^{ij})^{v,\ell}$ with respect to the signed measure $\mu^k_t = \sum_m \delta_{t^0_{km}} v^j_{t^0_{km}\wedge t, t^2_{km}\wedge t} v^\ell_{t^0_{km}\wedge t, t^2_{km}\wedge t}$. Decomposing μ^k_t into a positive and negative part as

$$
\mu^k_t = \frac{1}{4} \Big[\sum_m \delta_{t^0_{km}} [(v^j + v^\ell)_{t^0_{km}\wedge t, t^2_{km}\wedge t}]^2 - \sum_m \delta_{t^0_{km}} [(v^j - v^\ell)_{t^0_{km}\wedge t, t^2_{km}\wedge t}]^2 \Big]
$$

and similarly for $\mathrm{d}\mu_t = \mathrm{d}[v^j, v^\ell]_t$ we can estimate

$$
\begin{aligned}
\Big| \int_0^1 (f^{ij})^{v,\ell}(s) \mu^k_t(\mathrm{d}s) - \int_0^1 (f^{ij})^{v,\ell}(s) \mu_t(\mathrm{d}s) \Big| & \\
&\lesssim \|f^v\|_\infty \big(\big\| [v^i + v^j]_k - [v^i + v^j] \big\|_\infty + \big\| [v^i - v^j]_k - [v^i - v^j] \big\|_\infty \big) \\
&\lesssim \|f^v\|_\infty \|[v,v]_k - [v,v]\|_\infty,
\end{aligned}
$$

where we write $[u] := [u,u]$ and similarly for $[u]_k$. By assumption the right hand side converges to zero, from where we get the uniform convergence of $[f,g]_k$ to $[f,g]$. □

Remark 5.3. We calculate the pathwise Itô integral $I^{\mathrm{Itô}}(f, \mathrm{d}v)$ as limit of nonanticipating Riemann sums involving only f and v. The classical rough path integral, see Proposition 2.5, is obtained via "compensated Riemann sums" that explicitly depend on f^v and $I^{\mathrm{Itô}}(v, \mathrm{d}v)$. For applications in mathematical finance, it is more convenient to have an integral that is the limit of nonanticipating Riemann sums like $\sum_\ell f(t^0_{k\ell}) v_{t^0_{k\ell}\wedge t, t^2_{k\ell}\wedge t}$ for $k \to \infty$, because this can be interpreted as the capital process obtained by investing with the strategy f into v.

Note that $[v,v]$ is always a continuous function of bounded variation, but a priori it is not clear whether it is in $\mathcal{C}^{2\alpha}$. Under this additional assumption we have the following stronger result.

Corollary 5.4. *In addition to the conditions of Theorem 5.2, assume that also* $[v,v] \in \mathcal{C}^{2\alpha}(\mathbb{R}^{d\otimes d})$. *Then* $I^{\mathrm{Itô}}(f, \mathrm{d}v) \in \mathcal{D}^\alpha_v(\mathbb{R}^n)$ *with derivative* f, *and*

$$
\|I^{\mathrm{Itô}}(f, \mathrm{d}v)\|_{v,\alpha} \lesssim \|f\|_{v,\beta} \big(1 + \|v\|^2_\alpha + \|L(v,v)\|_{2\alpha} + \|[v,v]\|_{2\alpha} \big).
$$

Proof. This is a combination of Theorem 4.10 and the explicit representation (5.2) together with the continuity of the Young integral, Theorem 3.14. □

The term $I(S_{k-1}f, \mathrm{d}S_{k-1}v)$ has the pleasant property that if we want to refine our calculation by passing from k to $k+1$, then we only have to add the additional term $I(S_{k-1}f, \mathrm{d}\Delta_k v) + I(\Delta_k f, \mathrm{d}S_k v)$. For the quadratic variation $[f,v]_k$ this is not exactly true. But $[f,v]_k(m2^{-k}) = [S_{k-1}f, S_{k-1}v]_k(m2^{-k})$ for $m = 0, \dots, 2^k$, and there is a recursive way of calculating $[S_{k-1}f, S_{k-1}v]_k$:

Lemma 5.5. *Let $f, v \in C([0,1],\mathbb{R})$. Then*

$$[S_k f, S_k v]_{k+1}(t) = \frac{1}{2}[S_{k-1}f, S_{k-1}v]_k(t) + [S_{k-1}f, \Delta_k v]_{k+1}(t) + [\Delta_k f, S_k v]_{k+1}(t) + R_k(t) \tag{5.3}$$

for all $k \geq 1$ and all $t \in [0,1]$, where

$$R_k(t) := -\frac{1}{2} S_{k-1}f_{\llcorner t^k \lrcorner, t} S_{k-1}v_{\llcorner t^k \lrcorner, t} + S_{k-1}f_{\llcorner t^k \lrcorner, \ulcorner t^{k+1} \urcorner \wedge t} S_{k-1}v_{\llcorner t^k \lrcorner, \ulcorner t^{k+1} \urcorner \wedge t}$$
$$+ S_{k-1}f_{\ulcorner t^{k+1} \urcorner \wedge t, t} S_{k-1}v_{\ulcorner t^{k+1} \urcorner \wedge t, t}$$

and $\llcorner t^k \lrcorner := \lfloor t 2^k \rfloor 2^{-k}$ and $\ulcorner t^k \urcorner := \llcorner t^k \lrcorner + 2^{-(k+1)}$. In particular, we obtain for $t = 1$ that

$$[f,v]_{k+1}(1) = \frac{1}{2}[f,v]_k(1) + \frac{1}{2}\sum_m f_{km} v_{km} = \frac{1}{2^{k+1}} \sum_{p \leq k} \sum_m 2^p f_{pm} v_{pm}. \tag{5.4}$$

If moreover $\alpha \in (0,1)$ and $f, v \in \mathcal{C}^\alpha(\mathbb{R})$, then $\|[S_{k-1}f, S_{k-1}g]_k - [f,g]_k\|_\infty \lesssim 2^{-2k\alpha}\|f\|_\alpha \|g\|_\alpha$.

Proof. Equation (5.3) follows from a direct calculation using the fact that $S_{k-1}f$ and $S_{k-1}v$ are affine on every interval $[t^0_{k\ell}, t^1_{k\ell}]$ respectively $[t^1_{k\ell}, t^2_{k\ell}]$ for $1 \leq \ell \leq 2^k$. The formula for $[f,v]_{k+1}(1)$ follows from the that $[\Delta_p f, \Delta_q v]_{k+1}(1) = 0$ unless $p = q$, and that $[\Delta_k f, \Delta_k v]_{k+1} = (1/2)\sum_m f_{km} v_{km}$. The estimate for $\|[S_{k-1}f, S_{k-1}g]_k - [f,g]_k\|_\infty$ holds because the two functions agree in all dyadic points $m2^{-k}$. □

Remark 5.6. With the Cesàro mean formula (5.4) it becomes possible to study the existence of the quadratic variation using ergodic theory. This was previously observed by Gantert [20]. See also Gantert's thesis [19], Beispiel 3.29, where it is shown that ergodicity alone (of the distribution of v with respect to suitable transformations on path space) is not sufficient to obtain the convergence of $([v,v]_k(1))$ as k tends to ∞.

It would be more natural to assume that for the controlling path v the non-anticipating Riemann sums converge, rather than assuming that $(L(S_k v, S_k v))_k$ and $([v,v]_k)$ converge. This is indeed sufficient, as long as a uniform Hölder estimate is satisfied by the Riemann sums. We start by showing that the existence of the Itô iterated integrals implies the existence of the quadratic variation.

Lemma 5.7. *Let $\alpha \in (0, 1/2)$ and let $v \in \mathcal{C}^\alpha(\mathbb{R}^d)$. Assume that the non-anticipating Riemann sums $(I^{\mathrm{Itô}}_k(v, \mathrm{d}v))_k$ converge uniformly to $I^{\mathrm{Itô}}(v, \mathrm{d}v)$. Then also $([v,v]_k)_k$ converges uniformly to a limit $[v,v]$. If moreover*

$$\sup_k \sup_{0 \leq m < m' \leq 2^k} \frac{|I^{\mathrm{Itô}}_k(v,\mathrm{d}v)(m'2^{-k}) - I^{\mathrm{Itô}}_k(v,\mathrm{d}v)(m2^{-k}) - v(m2^{-k})(v(m'2^{-k}) - v(m2^{-k}))|}{|(m'-m)2^{-k}|^{2\alpha}}$$
$$= C < \infty, \tag{5.5}$$

then $[v,v] \in \mathcal{C}^{2\alpha}(\mathbb{R}^{d\otimes d})$ and $\|[v,v]\|_{2\alpha} \lesssim C + \|v\|^2_\alpha$.

Proof. Let $t \in [0,1]$ and $1 \leq i,j \leq d$. Then

$$v^i(t)v^j(t) - v^i(0)v^j(0) = \sum_{m=1}^{2^k} \left[v^i(t^2_{km} \wedge t) v^j(t^2_{km} \wedge t) - v^i(t^0_{km} \wedge t) v^j(t^0_{km} \wedge t) \right]$$
$$= \sum_{m=1}^{2^k} \left[v^i(t^0_{km}) v^j_{t^0_{km}\wedge t, t^2_{km}\wedge t} + v^j(t^0_{km}) v^i_{t^0_{km}\wedge t, t^2_{km}\wedge t} + v^i_{t^0_{km}\wedge t, t^2_{km}\wedge t} v^j_{t^0_{km}\wedge t, t^2_{km}\wedge t} \right]$$
$$= I^{\mathrm{Itô}}_k(v^i, \mathrm{d}v^j)(t) + I^{\mathrm{Itô}}_k(v^j, \mathrm{d}v^i)(t) + [v^i, v^j]_k(t),$$

which implies the convergence of $([v,v]_k)_k$ as k tends to ∞. For $0 \le s < t \le 1$ this gives

$$
\begin{aligned}
([v^i, v^j]_k)_{s,t} &= \left(v^i v^j\right)_{s,t} - I_k^{\text{Itô}}(v^i, \mathrm{d}v^j)_{s,t} - I_k^{\text{Itô}}(v^j, \mathrm{d}v^i)_{s,t} \\
&= \left[v^i(s) v^j_{s,t} - I_k^{\text{Itô}}(v^i, \mathrm{d}v^j)_{s,t}\right] + \left[v^j(s) v^i_{s,t} - I_k^{\text{Itô}}(v^j, \mathrm{d}v^i)_{s,t}\right] + v^i_{s,t} v^j_{s,t},
\end{aligned}
$$

At this point it is easy to estimate $\|[v,v]\|_{2\alpha}$, where we work with the classical Hölder norm and not the $\mathcal{C}^{2\alpha}$ norm. Indeed let $0 \le s < t \le 1$. Using the continuity of $[v,v]$, we can find k and $s \le s_k = m_s 2^{-k} < m_t 2^{-k} = t_k \le t$ with $|[v,v]|_{s,s_k} + |[v,v]|_{t,t_k} \le \|v\|_\alpha^2 |t-s|^{2\alpha}$. Moreover,

$$
|[v,v]|_{s_k,t_k} \le \Big(\sup_{\ell \ge k} \sup_{0 \le m < m' \le 2^\ell} \frac{|([v,v]_\ell)_{m2^{-\ell}, m'2^{-\ell}}|}{|(m'-m)2^{-\ell}|^{2\alpha}}\Big) |t_k - s_k|^{2\alpha} \le (2C + \|v\|_\alpha^2)|t-s|^{2\alpha}. \quad \square
$$

Remark 5.8. The “coarse-grained Hölder condition” (5.5) is from [44] and has recently been discovered independently by [30].

Similarly, the convergence of $(I_k^{\text{Itô}}(v, \mathrm{d}v))$ implies the convergence of $(L(S_k v, S_k v))_k$:

Lemma 5.9. *In the setting of Lemma 5.7, assume that* (5.5) *holds. Then* $L(S_k v, S_k v)$ *converges uniformly as* k *tends to* ∞*, and*

$$
\sup_k \|L(S_k v, S_k v)\|_{2\alpha} \lesssim C + \|v\|_\alpha^2.
$$

Proof. Let $k \in \mathbb{N}$ and $0 \le m \le 2^k$, and write $t = m2^{-k}$. Then we obtain from (5.1) that

$$
\begin{aligned}
&L(S_{k-1}v, S_{k-1}v)(t) \\
&\qquad = I_k^{\text{Itô}}(v, \mathrm{d}v)(t) + \frac{1}{2}[v,v]_k(t) - \pi_<(S_{k-1}v, S_{k-1}v)(t) - S(S_{k-1}v, S_{k-1}v)(t).
\end{aligned} \tag{5.6}
$$

Let now $s, t \in [0,1]$. We first assume that there exists m such that $t^0_{km} \le s < t \le t^2_{km}$. Then we use $\|\partial_t \Delta_q v\|_\infty \lesssim 2^{q(1-\alpha)}\|v\|_\alpha$ to obtain

$$
\begin{aligned}
|L(S_{k-1}v, S_{k-1}v)_{s,t}| &\le \sum_{p<k} \sum_{q<p} \left| \int_s^t \Delta_p v(r) \mathrm{d}\Delta_q v(r) - \int_s^t \mathrm{d}\Delta_q v(r) \Delta_p v(r) \right| \\
&\lesssim \sum_{p<k} \sum_{q<p} |t-s| 2^{-p\alpha} 2^{q(1-\alpha)} \|v\|_\alpha^2 \lesssim |t-s| 2^{-k(2\alpha-1)} \|v\|_\alpha^2 \le |t-s|^{2\alpha} \|v\|_\alpha^2.
\end{aligned} \tag{5.7}
$$

Combining (5.6) and (5.7), we obtain the uniform convergence of $(L(S_{k-1}v, S_{k-1}v))$ from Lemma 5.7 and from the continuity of $\pi_<$ and S.

For s and t that do not lie in the same dyadic interval of generation k, let $\ulcorner s^k \urcorner = m_s 2^{-k}$ and $\llcorner t^k \lrcorner = m_t 2^{-k}$ be such that $\ulcorner s^k \urcorner - 2^{-k} < s \le \ulcorner s^k \urcorner$ and $\llcorner t^k \lrcorner \le t < \llcorner t^k \lrcorner + 2^{-k}$. In particular, $\ulcorner s^k \urcorner \le \llcorner t^k \lrcorner$. Moreover

$$
\begin{aligned}
|L(S_{k-1}v, S_{k-1}v)_{s,t}| \le |L(S_{k-1}v, S_{k-1}v)_{s, \ulcorner s^k \urcorner}| &+ |L(S_{k-1}v, S_{k-1}v)_{\ulcorner s^k \urcorner, \llcorner t^k \lrcorner}| \\
&+ |L(S_{k-1}v, S_{k-1}v)_{\llcorner t^k \lrcorner, t}|.
\end{aligned}
$$

Using (5.7), the first and third term on the right hand side can be estimated by $(|\ulcorner s^k \urcorner - s|^{2\alpha} + |t - \llcorner t^k \lrcorner|^{2\alpha})\|v\|_\alpha^2 \lesssim |t-s|^{2\alpha}\|v\|_\alpha^2$. For the middle term we apply (5.6) to obtain

$$
\begin{aligned}
|L(S_{k-1}v, S_{k-1}v)_{\ulcorner s^k \urcorner, \llcorner t^k \lrcorner}| &\le \left| I_k^{\text{Itô}}(v, \mathrm{d}v)_{\ulcorner s^k \urcorner, \llcorner t^k \lrcorner} - v(\ulcorner s^k \urcorner)(v(\llcorner t^k \lrcorner) - v(\ulcorner s^k \urcorner)) \right| \\
&\quad + \left| v(\ulcorner s^k \urcorner) v_{\ulcorner s^k \urcorner, \llcorner t^k \lrcorner} - \pi_<(S_{k-1}v, S_{k-1}v)_{\ulcorner s^k \urcorner, \llcorner t^k \lrcorner} \right| \\
&\quad + \frac{1}{2} \left| ([v,v]_k)_{\ulcorner s^k \urcorner, \llcorner t^k \lrcorner} \right| + \left| S(S_{k-1}v, S_{k-1}v)_{\ulcorner s^k \urcorner, \llcorner t^k \lrcorner} \right| \\
&\lesssim |\llcorner t^k \lrcorner - \ulcorner s^k \urcorner|^{2\alpha} \left(C + \|v\|_\alpha^2\right) \le |t-s|^{2\alpha} \left(C + \|v\|_\alpha^2\right),
\end{aligned}
$$

where Example 4.4, Lemma 5.7, and Lemma 3.13 have been used. $\square$

It follows from the work of Föllmer that our pathwise Itô integral satisfies Itô's formula:

Corollary 5.10. *Let* $\alpha \in (1/3, 1/2)$ *and* $v \in \mathcal{C}^\alpha(\mathbb{R}^d)$*. Assume that the non-anticipating Riemann sums* $(I_k^{\text{Itô}}(v, \mathrm{d}v))_k$ *converge uniformly to* $I^{\text{Itô}}(v, \mathrm{d}v)$ *and let* $F \in C^2(\mathbb{R}^d, \mathbb{R})$*. Then* $(I_k^{\text{Itô}}(\mathrm{D}F(v), \mathrm{d}v))_k$ *converges to a limit* $I^{\text{Itô}}(\mathrm{D}F(v), \mathrm{d}v)$ *that satisfies for all* $t \in [0, 1]$

$$F(v(t)) - F(v(0)) = I^{\text{Itô}}(\mathrm{D}F(v), \mathrm{d}v)(t) + \int_0^t \sum_{k,\ell=1}^d \partial_{x_k}\partial_{x_\ell} F(v(s)) \mathrm{d}[v^k, v^\ell](s).$$

Proof. This is Remarque 1 of Föllmer [15] in combination with Lemma 5.7. □

6 Construction of the Lévy area

To apply our theory, it remains to construct the Lévy area respectively the pathwise Itô integrals for suitable stochastic processes. In Section 6.1 we construct the Lévy area for hypercontractive stochastic processes whose covariance function satisfies a certain "finite variation" property. In Section 6.2 we construct the pathwise Itô iterated integrals for some continuous martingales.

6.1 Hypercontractive processes

Let $X\colon [0,1] \to \mathbb{R}^d$ be a centered continuous stochastic process, such that X^i is independent of X^j for $i \neq j$. We write R for its covariance function, $R\colon [0,1]^2 \to \mathbb{R}^{d\times d}$ and $R(s,t) := (E(X_s^i X_t^j))_{1\le i,j\le d}$. The increment of R over a rectangle $[s,t]\times[u,v] \subseteq [0,1]^2$ is defined as

$$R_{[s,t]\times[u,v]} := R(t,v) + R(s,u) - R(s,v) - R(t,u) := (E(X_{s,t}^i X_{u,v}^j))_{1\le i,j\le d}.$$

For a given $\rho \in [1, \infty)$ we will work under the following two assumptions:

(ρ–var) There exists $C > 0$ such that for all $0 \le s < t \le 1$ and for every partition $s = t_0 < t_1 < \cdots < t_n = t$ of $[s,t]$ we have

$$\sum_{i,j=1}^n |R_{[t_{i-1},t_i]\times[t_{j-1},t_j]}|^\rho \le C|t-s|.$$

(HC) Second order polynomials of the process X satisfy a hypercontractivity condition, i.e. for every $r \ge 1$ there exists $C_r > 0$ such that for every n and every polynomial $P : \mathbb{R}^n \to \mathbb{R}$ of degree 2, for all $i_1, \ldots, i_n \in \{1, \ldots, d\}$, and for all $p_1, \ldots, p_n \ge -1$ and $0 \le m_1 \le 2^{p_1}$, ..., $0 \le m_n \le 2^{p_n}$

$$E(|P(X^{i_1}_{p_1 m_1}, \ldots, X^{i_n}_{p_n m_n})|^r) \le C_r (E(|P(X^{i_1}_{t_1}, \ldots, X^{i_n}_{t_n})|^2))^{r/2}.$$

These conditions are taken from [17], where under even more general assumptions it is shown that it is possible to construct the iterated integrals $I(X, \mathrm{d}X)$, and that $I(X, \mathrm{d}X)$ is the limit of $(I(X^n, \mathrm{d}X^n))_{n\in\mathbb{N}}$ under a wide range of smooth approximations $(X^n)_n$ that converge to X.

Example 6.1. Condition (HC) is satisfied by all Gaussian processes. More generally, it is satisfied by every process "living in a fixed Gaussian chaos"; see [28], Theorem 3.50. Slightly oversimplifying things, this is the case if X is given by polynomials of fixed degree and iterated integrals of fixed order with respect to a Gaussian reference process.

Prototypical examples of processes living in a fixed chaos are Hermite processes. They are defined for $H \in (1/2, 1)$ and $k \in \mathbb{N}$, $k \geq 1$ as

$$Z_t^{k,H} = C(H,k) \int_{\mathbb{R}^k} \left(\int_0^t \prod_{i=1}^k (s - y_i)_+^{-\left(\frac{1}{2} + \frac{1-H}{k}\right)} \mathrm{d}s \right) \mathrm{d}B_{y_1} \ldots \mathrm{d}B_{y_k},$$

where $(B_y)_{y \in \mathbb{R}}$ is a standard Brownian motion, and $C(H,k)$ is a normalization constant. In particular, $Z^{k,H}$ lives in the Wiener chaos of order k. The covariance of $Z^{k,H}$ is

$$E(Z_s^{k,H} Z_t^{k,H}) = \frac{1}{2} \left(t^{2H} + s^{2H} + |t-s|^{2H} \right)$$

Since $Z^{1,H}$ is Gaussian, it is just the fractional Brownian motion with Hurst parameter H. For $k = 2$ we obtain the Rosenblatt process. For further details about Hermite processes see [43]. However, we should point out that it follows from Kolmogorov's continuity criterion that $Z^{k,H}$ is α–Hölder continuous for every $\alpha < H$. Since $H \in (1/2, 1)$, Hermite processes are amenable to Young integration, and it is trivial to construct $L(Z^{k,H}, Z^{k,H})$.

Example 6.2. Condition (ρ–var) is satisfied by the Brownian motion with $\rho = 1$. More generally it is satisfied by the fractional Brownian motion with Hurst index H, for which $\rho = 1/(2H)$. It is also satisfied by the fractional Brownian bridge with Hurst index H. A general criterion that implies condition (ρ–var) is the one of Coutin and Qian [13]: If $E(|X_{s,t}^i|^2) \lesssim |t-s|^{2H}$ and $|E(X_{s,s+h}^i X_{t,t+h}^i)| \lesssim |t-s|^{2H-2} h^2$ for $i = 1, \ldots, d$, then (ρ–var) is satisfied for $\rho = 1/(2H)$. For details and further examples see [18], Section 15.2.

Lemma 6.3. *Assume that the stochastic process $X : [0,1] \to \mathbb{R}$ satisfies (ρ–var). Then we have for all $p \geq -1$ and for all $M, N \in \mathbb{N}$ with $M \leq N \leq 2^p$ that*

$$\sum_{m_1, m_2 = M}^{N} |E(X_{pm_1} X_{pm_2})|^\rho \lesssim (N - M + 1) 2^{-p}. \tag{6.1}$$

Proof. The case $p \leq 0$ is easy so let $p \geq 1$. It suffices to note that

$$\begin{aligned} E(X_{pm_1} X_{pm_2}) &= E\left((X_{t_{pm_1}^0, t_{pm_1}^1} - X_{t_{pm_1}^1, t_{pm_1}^2})(X_{t_{pm_2}^0, t_{pm_2}^1} - X_{t_{pm_2}^1, t_{pm_2}^2}) \right) \\ &= \sum_{i_1, i_2 = 0,1} (-1)^{i_1 + i_2} R_{[t_{pm_1}^{i_1}, t_{pm_1}^{i_1+1}] \times [t_{pm_2}^{i_2}, t_{pm_2}^{i_2+1}]}, \end{aligned}$$

and that $\{t_{pm}^i : i = 0, 1, 2, m = M, \ldots, N\}$ partitions the interval $[(M-1)2^{-p}, N 2^{-p}]$. □

Lemma 6.4. *Let $X, Y : [0,1] \to \mathbb{R}$ be independent, centered, continuous processes, both satisfying (ρ–var) for some $\rho \in [1,2]$. Then for all $i, p \geq -1$, $q < p$, and $0 \leq j \leq 2^i$*

$$E\Big[\Big| \sum_{m \leq 2^p} \sum_{n \leq 2^q} X_{pm} Y_{qn} \langle 2^{-i} \chi_{ij}, \varphi_{pm} \chi_{qn} \rangle \Big|^2 \Big] \lesssim 2^{(p \vee i)(1/\rho - 4)} 2^{(q \vee i)(1 - 1/\rho)} 2^{-i} 2^{p(4 - 3/\rho)} 2^{q/\rho}.$$

Proof. Since $p > q$, for every m there exists exactly one $n(m)$, such that $\varphi_{pm} \chi_{qn(m)}$ is not identically zero. Hence, we can apply the independence of X and Y to obtain

$$\begin{aligned} &E\Big[\Big| \sum_{m \leq 2^p} \sum_{n \leq 2^q} X_{pm} Y_{qn} \langle 2^{-i} \chi_{ij}, \varphi_{pm} \chi_{qn} \rangle \Big|^2 \Big] \\ &\quad \leq \sum_{m_1, m_2 = 0}^{2^p} \big| E(X_{pm_1} X_{pm_2}) E(Y_{qn(m_1)} Y_{qn(m_2)}) \langle 2^{-i} \chi_{ij}, \varphi_{pm_1} \chi_{qn(m_1)} \rangle \langle 2^{-i} \chi_{ij}, \varphi_{pm_2} \chi_{qn(m_2)} \rangle \big|. \end{aligned}$$

Let us write $M_j := \{m : 0 \le m \le 2^p, \langle \chi_{ij}, \varphi_{pm}\chi_{qn(m)}\rangle \neq 0\}$. We also write ρ' for the conjugate exponent of ρ, i.e. $1/\rho + 1/\rho' = 1$. Hölder's inequality and Lemma 3.9 imply

$$\sum_{m_1,m_2\in M_j} \big|E(X_{pm_1}X_{pm_2})E(Y_{qn(m_1)}Y_{qn(m_2)})\langle 2^{-i}\chi_{ij}, \varphi_{pm_1}\chi_{qn(m_1)}\rangle\langle 2^{-i}\chi_{ij}, \varphi_{pm_2}\chi_{qn(m_2)}\rangle\big|$$
$$\lesssim \Big(\sum_{m_1,m_2\in M_j} \big|E(X_{pm_1}X_{pm_2})\big|^{\rho}\Big)^{1/\rho} \Big(\sum_{m_1,m_2\in M_j} \big|E(Y_{qn(m_1)}Y_{qn(m_2)})\big|^{\rho'}\Big)^{1/\rho'} (2^{-2(p\vee i)+p+q})^2.$$

Now write N_j for the set of n for which $\chi_{ij}\chi_{qn}$ is not identically zero. For every $\bar n \in N_j$ there are 2^{p-q} numbers $m \in M_j$ with $n(m) = \bar n$. Hence

$$\Big(\sum_{m_1,m_2\in M_j} \big|E(Y_{qn(m_1)}Y_{qn(m_2)})\big|^{\rho'}\Big)^{1/\rho'}$$
$$\lesssim (2^{2(p-q)})^{1/\rho'}\Big(\Big(\max_{n_1,n_2\in N_j}\big|E(Y_{qn_1}Y_{qn_2})\big|\Big)^{\rho'-\rho}\sum_{n_1,n_2\in N_j}\big|E(Y_{qn_1}Y_{qn_2})\big|^{\rho}\Big)^{1/\rho'},$$

where we used that $\rho \in [1,2]$ and therefore $\rho' - \rho \ge 0$ (for $\rho' = \infty$ we interpret the right hand side as $\max_{n_1,n_2\in N_j} |E(Y_{qn_1}Y_{qn_2})|$). Lemma 6.3 implies that $\big(\big|E(Y_{qn_1}Y_{qn_2})\big|^{\rho'-\rho}\big)^{1/\rho'} \lesssim 2^{-q(1/\rho-1/\rho')}$. Similarly we apply Lemma 6.3 to the sum over n_1, n_2, and we obtain

$$(2^{2(p-q)})^{1/\rho'}\Big(\Big(\max_{n_1,n_2\in N_j}\big|E(Y_{qn_1}Y_{qn_2})\big|\Big)^{\rho'-\rho}\sum_{n_1,n_2\in N_j}\big|E(Y_{qn_1}Y_{qn_2})\big|^{\rho}\Big)^{1/\rho'}$$
$$\lesssim (2^{2(p-q)})^{1/\rho'}2^{-q(1/\rho-1/\rho')}(|N_j|2^{-q})^{1/\rho'} = 2^{(q\vee i)/\rho'}2^{-i/\rho'}2^{2p/\rho'}2^{q(-2/\rho'-1/\rho)}$$
$$= 2^{(q\vee i)(1-1/\rho)}2^{i(1/\rho-1)}2^{2p(1-1/\rho)}2^{q(1/\rho-2)},$$

where we used that $|N_j| = 2^{(q\vee i)-i}$. Since $|M_j| = 2^{(p\vee i)-i}$, another application of Lemma 6.3 yields

$$\Big(\sum_{m_1,m_2\in M_j}\big|E(X_{pm_1}X_{pm_2})\big|^{\rho}\Big)^{1/\rho} \lesssim 2^{(p\vee i)/\rho}2^{-i/\rho}2^{-p/\rho}.$$

The result now follows by combining these estimates:

$$E\Big[\Big|\sum_{m\le 2^p}\sum_{n\le 2^q} X_{pm}Y_{qn}\langle 2^{-i}\chi_{ij}, \varphi_{pm}\chi_{qn}\rangle\Big|^2\Big]$$
$$\lesssim \Big(\sum_{m_1,m_2\in M_j}\big|E(X_{pm_1}X_{pm_2})\big|^{\rho}\Big)^{1/\rho}\Big(\sum_{m_1,m_2\in M_j}\big|E(Y_{qn(m_1)}Y_{qn(m_2)})\big|^{\rho'}\Big)^{1/\rho'}(2^{-2(p\vee i)+p+q})^2$$
$$\lesssim \big(2^{(p\vee i)/\rho}2^{-i/\rho}2^{-p/\rho}\big)\big(2^{(q\vee i)(1-1/\rho)}2^{i(1/\rho-1)}2^{2p(1-1/\rho)}2^{q(1/\rho-2)}\big)\big(2^{-4(p\vee i)+2p+2q}\big)$$
$$= 2^{(p\vee i)(1/\rho-4)}2^{(q\vee i)(1-1/\rho)}2^{-i}2^{p(4-3/\rho)}2^{q/\rho}. \qquad \square$$

Theorem 6.5. *Let $X\colon [0,1] \to \mathbb{R}^d$ be a continuous, centered stochastic process with independent components, and assume that X satisfies (HC) and (ρ–var) for some $\rho \in [1,2)$. Then we have for all $\alpha \in (0, 1/(2\rho))$, all $\alpha' \le \alpha$ and all $r \ge 1$*

$$E[\|S_N X - X\|_{\alpha'}^r]^{1/r} \lesssim 2^{-N(\alpha-\alpha')}$$

as well as

$$\sum_{N\ge 0} E\left(\|L(S_N X, S_N X) - L(S_{N-1}X, S_{N-1}X)\|_{2\alpha}^r\right)^{1/r} < \infty,$$

and therefore $L(X,X) = \lim_{N\to\infty} L(S_N X, S_N X)$ is almost surely 2α–Hölder continuous, where the convergence takes place both in $L^r(\Omega)$ and almost surely.

Proof. The statement about the Hölder norm of X follows from Kolmogorov's continuity criterion because $\|S_N X - X\|_{\alpha'}^r \lesssim 2^{-N(\alpha-\alpha')r}\|X\|_\alpha^r$.

For the Lévy area note that L is antisymmetric, and in particular the diagonal of the matrix $L(S_N X, S_N X)$ is constantly zero. To treat the off-diagonal terms it will be convenient to introduce general Besov spaces: let $\beta \in \mathbb{R}$, $r, s \geq 1$, and define for $f \in C([0,1], \mathbb{R}^d)$

$$\|f\|_{B^\beta_{r,s}} := \Big(\sum_i 2^{is(\beta-1/s)}\Big(\sum_j |f_{ij}|^r\Big)^{s/r}\Big)^{1/s}.$$

Depending on r, s, β, the norm might be finite also for non-continuous functions, but we do not need this. All we need is that $\|f\|_{B^\beta_{\infty,\infty}} = \|f\|_\beta$ and the trivial observation that

$$\|f\|_{B^{\beta-(1/r_1-1/r_2)}_{r_2,s_2}} \leq \|f\|_{B^\beta_{r_1,s_1}} \tag{6.2}$$

whenever $r_1 \leq r_2$ and $s_1 \leq s_2$.

Let now $k, \ell \in \{1, \dots, d\}$ with $k \neq \ell$ and let $\beta \in \mathbb{R}$ and $r \geq 1$. Then

$$\begin{aligned} &E\big[\|L(S_N X^k, S_N X^\ell) - L(S_{N-1}X^k, S_{N-1}X^\ell)\|^r_{B^\beta_{r,r}}\big] \\ &\quad = E\Big[\Big\|\sum_{q<N}\sum_{m,n}(X^k_{Nm}X^\ell_{qn} - X^k_{qn}X^\ell_{Nm})\int_0^\cdot \varphi_{Nm}(s)\mathrm{d}\varphi_{qn}(s)\Big\|^r_{B^\beta_{r,r}}\Big] \\ &\quad = \sum_i 2^{ir(\beta-1/r)}\sum_j E\Big[\Big|\Big(\sum_{q<N}\sum_{m,n}(X^k_{Nm}X^\ell_{qn} - X^k_{qn}X^\ell_{Nm})\int_0^\cdot \varphi_{Nm}(s)\mathrm{d}\varphi_{qn}(s)\Big)_{ij}\Big|^r\Big], \end{aligned} \tag{6.3}$$

and Minkowski's inequality yields

$$\begin{aligned} &E\Big[\Big|\Big(\sum_{q<N}\sum_{m,n}(X^k_{Nm}X^\ell_{qn} - X^k_{qn}X^\ell_{Nm})\int_0^\cdot \varphi_{Nm}(s)\mathrm{d}\varphi_{qn}(s)\Big)_{ij}\Big|^r\Big] \\ &\qquad \leq \Big(\sum_{q<N} E\Big[\Big|\Big(\sum_{m,n}(X^k_{Nm}X^\ell_{qn} - X^k_{qn}X^\ell_{Nm})\int_0^\cdot \varphi_{Nm}(s)\mathrm{d}\varphi_{qn}(s)\Big)_{ij}\Big|^r\Big]^{1/r}\Big)^r. \end{aligned}$$

Observe that the term inside the expectation is a second order polynomial of the hypercontractive process X, and therefore we can estimate its L^r norm by its L^2 norm raised to the $r/2$–th power. In combination with Lemma 6.4, this gives

$$\begin{aligned} &E\Big[\Big|\Big(\sum_{q<N}\sum_{m,n}(X^k_{Nm}X^\ell_{qn} - X^k_{qn}X^\ell_{Nm})\int_0^\cdot \varphi_{Nm}(s)\mathrm{d}\varphi_{qn}(s)\Big)_{ij}\Big|^r\Big] \\ &\qquad \lesssim \Big(\sum_{q<N} E\Big[\Big|\Big(\sum_{m,n}(X^k_{Nm}X^\ell_{qn} - X^k_{qn}X^\ell_{Nm})\int_0^\cdot \varphi_{Nm}(s)\mathrm{d}\varphi_{qn}(s)\Big)_{ij}\Big|^2\Big]^{1/2}\Big)^r \\ &\qquad \lesssim \Big(\sum_{q<N}\big(2^{(N\vee i)(1/\rho-4)}2^{(q\vee i)(1-1/\rho)}2^{-i}2^{N(4-3/\rho)}2^{q/\rho}\big)^{1/2}\Big)^r. \end{aligned}$$

Plugging this back into (6.3), we get

$$\begin{aligned} &E\big[\|L(S_N X^k, S_N X^\ell) - L(S_{N-1}X^k, S_{N-1}X^\ell)\|^r_{B^\beta_{r,r}}\big] \\ &\quad \lesssim \sum_i 2^{ir(\beta-1/r)}2^i\Big(\sum_{q<N}\big(2^{(N\vee i)(1/\rho-4)}2^{(q\vee i)(1-1/\rho)}2^{-i}2^{N(4-3/\rho)}2^{q/\rho}\big)^{1/2}\Big)^r \\ &\quad = \sum_{i\leq N} 2^{ir\beta}\Big(\sum_{q<i} 2^{-N/\rho}2^{-i/(2\rho)}2^{q/(2\rho)} + \sum_{i\leq q<N} 2^{-N/\rho}2^{-i/2}2^{q/2}\Big)^r \\ &\qquad + \sum_{i>N} 2^{ir\beta}\Big(\sum_{q<N} 2^{N(2-3/(2\rho))}2^{-2i}2^{q/(2\rho)}\Big)^r \end{aligned}$$

$$\lesssim \sum_{i \le N} 2^{ir\beta}\big(2^{-Nr\rho} + 2^{Nr(1/2-\rho)}2^{-ir/2}\big) + \sum_{i>N} 2^{ir(\beta-2)}2^{Nr(2-1/\rho)}.$$

If $\beta \in (1/2, 2)$, then the right hand side is $\simeq 2^{Nr(\beta-1/\rho)}$, and if $\beta < 1/\rho$ (which requires $\rho < 2$), then

$$\sum_N (2^{Nr(\beta-1/\rho)})^{1/r} < \infty.$$

In conclusion,

$$\sum_N E\big[\|L(S_N X^k, S_N X^\ell) - L(S_{N-1}X^k, S_{N-1}X^\ell)\|^r_{B^\beta_{r,r}}\big]^{1/r} < \infty$$

whenever $\beta \in (1/2, 1/\rho)$ and $r \ge 1$, and it only remains to apply the Besov embedding result (6.2) to conclude. □

6.2 Continuous martingales

Here we assume that $(X_t)_{t\in[0,1]}$ is a d–dimensional continuous martingale. Of course in that case it is no problem to construct the Itô integral $I^{\mathrm{Itô}}(X, \mathrm{d}X)$. But to apply the results of Section 5, we still need the pathwise convergence of $I^{\mathrm{Itô}}_k(X, \mathrm{d}X)$ to $I^{\mathrm{Itô}}(X, \mathrm{d}X)$ and the uniform Hölder continuity of $I^{\mathrm{Itô}}_k(X, \mathrm{d}X)$ along the dyadics.

Recall that for a d–dimensional semimartingale $X = (X^1, \dots, X^d)$, the quadratic variation is defined as $[X] = ([X^i, X^j])_{1\le i,j\le d}$. We also write $X_s X_{s,t} := (X^i_s X^j_{s,t})_{1\le i,j\le d}$ for $s, t \in [0, 1]$.

Theorem 6.6. *Let $X = (X^1, \dots, X^d)$ be a d–dimensional continuous martingale. Assume that there exist $p \ge 1$, $\beta > 0$, such that $2p\beta > 1$, and $C > 0$ with*

$$E(|[X]_{s,t}|^p) \le C^{2p}|t-s|^{2p\beta} \tag{6.4}$$

for all $s, t \in [0, 1]$. Then $I^{\mathrm{Itô}}_k(X, \mathrm{d}X)$ converges to $I^{\mathrm{Itô}}(X, \mathrm{d}X)$ in $\mathcal{C}^\alpha(\mathbb{R}^{d\otimes d})$, both almost surely and in $L^p(\Omega)$. Furthermore, for all $\alpha \in (0, \beta - 1/(2p))$ we have

$$E(\|X\|^{2p}_\alpha) + E(M^p_\alpha) + E(D^p_\alpha) \lesssim C^{2p}, \tag{6.5}$$

where

$$M_\alpha := \sup_k \sup_{0\le \ell<\ell'\le 2^k} \frac{|I^{\mathrm{Itô}}_k(X, \mathrm{d}X)_{\ell 2^{-k},\ell' 2^{-k}} - X_{\ell 2^{-k}} X_{\ell 2^{-k},\ell' 2^{-k}}|}{|(\ell'-\ell)2^{-k}|^{2\alpha}}$$

and

$$D_\alpha := \sup_k 2^{k\alpha}\|I^{\mathrm{Itô}}_k(X, \mathrm{d}X) - I^{\mathrm{Itô}}(X, \mathrm{d}X)\|_\alpha.$$

Proof. The Hölder continuity of X follows from Kolmogorov's continuity criterion. Indeed, applying the Burkholder-Davis-Gundy inequality and (6.4) we have

$$E(|X_{s,t}|^{2p}) \lesssim \sum_{i=1}^d E(|X^i_{s,t}|^{2p}) \lesssim \sum_{i=1}^d E(|[X^i]_{s,t}|^p) \lesssim E(|[X]_{s,t}|^p) \le C^{2p}|t-s|^{2p\beta},$$

so that $E(\|X\|^{2p}_\alpha) \lesssim C^{2p}$ for all $\alpha \in (0, \beta - 1/(2p))$. Since we will need it below, let us also study the regularity of the Itô integral $I^{\mathrm{Itô}}(X, \mathrm{d}X)$: A similar application of the Burkholder-Davis-Gundy inequality gives

$$\begin{aligned} E(|I^{\mathrm{Itô}}(X, \mathrm{d}X)_{s,t} - X_s X_{s,t}|^p) &\lesssim E\Big(\Big|\int_s^t |X_r - X_s|^2 \mathrm{d}|[X]|_r\Big|^{\frac{p}{2}}\Big) \\ &\le E\Big(\sup_{r\in[s,t]} |X_r - X_s|^p \times |[X]|^{\frac{p}{2}}_{s,t}\Big) \end{aligned}$$

$$\leq \sqrt{E\Big(\sup_{r\in[s,t]}|X_r - X_s|^{2p}\Big)}\sqrt{E(|[X]|^p_{s,t})}$$
$$\lesssim E(|[X]_{s,t}|^p) \leq C^{2p}|t-s|^{2p\beta}.$$

The Kolmogorov criterion for rough paths, Theorem 3.1 of [16], now implies that for all $\alpha \in (0, \beta - 1/(2p))$

$$E(|M^{(1)}_\alpha|^p) \lesssim C^{2p}, \qquad \text{where } M^{(1)}_\alpha = \sup_{0\leq s<t\leq 1} \frac{|I^{\text{Itô}}(X, \mathrm{d}X)_{s,t} - X_s X_{s,t}|}{|t-s|^{2\alpha}}. \tag{6.6}$$

Moreover, $M_\alpha \leq D_\alpha + M^{(1)}_\alpha$, so it only remains to prove the bound for the p–th moment of D_α. First assume that $s = \ell 2^{-k}$ and $t = \ell' 2^{-k}$. As before, we have

$$\begin{aligned}
&E(|I^{\text{Itô}}(X, \mathrm{d}X)_{\ell 2^{-k},\ell' 2^{-k}} - I^{\text{Itô}}_k(X, \mathrm{d}X)_{\ell 2^{-k},\ell' 2^{-k}}|^p) \qquad (6.7)\\
&\qquad = E\Big(\Big|\int_{\ell 2^{-k}}^{\ell' 2^{-k}} \sum_{m=\ell}^{\ell'-1} \mathbf{1}_{[m2^{-k},(m+1)2^{-k})}(r) X_{m2^{-k},r}\mathrm{d}X_r\Big|^p\Big)\\
&\qquad \leq \sqrt{E\Big(\max_{m=\ell,\dots,\ell'} \sup_{r\in[m2^{-k},(m+1)2^{-k}]} |X_r - X_{m2^{-k}}|^{2p}\Big)}\sqrt{E(|[X]|^p_{\ell 2^{-k},\ell' 2^{-k}})}\\
&\qquad \lesssim C^{2p}\sqrt{(\ell'-\ell)(2^{-k})^{2p\beta}}\sqrt{|(\ell'-\ell)2^{-k}|^{2p\beta}} = ((\ell'-\ell)2^{-k})^{1/2+p\beta}2^{-k(p\beta-1/2)}.
\end{aligned}$$

From here we use a chaining argument to conclude. Define

$$M^{(2)}_{k,m} = \max_{\ell\leq 2^m} |I^{\text{Itô}}(X, \mathrm{d}X)_{\ell 2^{-m},(\ell+1)2^{-m}} - I^{\text{Itô}}_k(X, \mathrm{d}X)_{\ell 2^{-m},(\ell+1)2^{-m}}|,$$

and observe that (6.7) yields

$$E(|M^{(2)}_{k,m}|^p) \lesssim C^{2p}2^m 2^{-m(1/2+p\beta)}2^{-k(p\beta-1/2)} = C^{2p}2^{-m(p\beta-1/2)}2^{-k(p\beta-1/2)}$$

whenever $m \leq k$. We now write $[s,t]$ as a finite union of intervals of the form $[m2^{-i},(m+1)2^{-i}]$ for $i \leq k$ and $i \geq i_0$ where $2^{-i_0-1} < |t-s| \leq 2^{-i_0}$, with at most two intervals in every dyadic generation. Thus, we get

$$\frac{|I^{\text{Itô}}_k(X, \mathrm{d}X)_{s,t} - I^{\text{Itô}}(X, \mathrm{d}X)_{s,t}|}{|t-s|^{2\alpha}} \lesssim \frac{\sum_{m=i_0}^k M^{(2)}_{k,m}}{|t-s|^\alpha} \lesssim \sum_{m=i_0}^k 2^{m\alpha}M^{(2)}_{k,m} \leq \sum_{m=0}^k 2^{m\alpha}M^{(2)}_{k,m}, \tag{6.8}$$

and therefore we have for

$$M^{(2)}_\alpha = \sup_k 2^{k\alpha} \sup_{0\leq\ell<\ell'\leq 2^k} \frac{|I^{\text{Itô}}_k(X, \mathrm{d}X)_{\ell 2^{-k},\ell' 2^{-k}} - I^{\text{Itô}}(X, \mathrm{d}X)_{\ell 2^{-k},\ell' 2^{-k}}|}{|(\ell'-\ell)2^{-k}|^\alpha}$$

the estimate

$$\begin{aligned}
\|M^{(2)}_\alpha\|_{L^p(\Omega)} &\leq \sum_k 2^{k\alpha}\sum_{m\leq k}\|2^{m\alpha}M^{(2)}_{k,m}\|_{L^p(\Omega)}\\
&\lesssim C^2 \sum_k 2^{k\alpha}\sum_{m\leq k} 2^{m\alpha}2^{-m(\beta-1/(2p))}2^{-k(\beta-1/(2p))}.
\end{aligned}$$

The sum on the right hand side is finite as long as $\alpha < \beta - 1/(2p)$. For general $0 \leq s < t \leq 1$ define $\tau_k(r) \in 2^{-k}\mathbb{N}_0$ for $r \in [0,1]$ such that $\tau_k(r) \leq r < \tau_k(r) + 2^{-k}$. If $\tau_k(s) = \tau_k(t)$, then

$$\begin{aligned}
|I^{\text{Itô}}_k(X, \mathrm{d}X)_{s,t} - I^{\text{Itô}}(X, \mathrm{d}X)_{s,t}| &\leq |X_{\tau_k(s),s}X_{s,t}| + |X_s X_{s,t} - I^{\text{Itô}}(X, \mathrm{d}X)_{s,t}|\\
&\leq 2^{-k\alpha}\|X\|^2_\alpha|t-s|^\alpha + 2^{-k\alpha}M^{(1)}_\alpha|t-s|^\alpha.
\end{aligned}$$

If $\tau_k(s) < \tau_k(t)$, then

$$\begin{aligned}|I_k^{\mathrm{Itô}}(X,\mathrm{d}X)_{s,t} - I^{\mathrm{Itô}}(X,\mathrm{d}X)_{s,t}| &\le |I_k^{\mathrm{Itô}}(X,\mathrm{d}X)_{s,\tau_k(s)+2^{-k}} - I^{\mathrm{Itô}}(X,\mathrm{d}X)_{s,\tau_k(s)+2^{-k}}| \\ &\quad + |I_k^{\mathrm{Itô}}(X,\mathrm{d}X)_{\tau_k(s)+2^{-k},\tau_k(t)} - I^{\mathrm{Itô}}(X,\mathrm{d}X)_{\tau_k(s)+2^{-k},\tau_k(t)}| \\ &\quad + |I_k^{\mathrm{Itô}}(X,\mathrm{d}X)_{\tau_k(t),t} - I^{\mathrm{Itô}}(X,\mathrm{d}X)_{\tau_k(t),t}| \\ &\lesssim 2^{-k\alpha}|t-s|^{\alpha}\big(\|X\|_\alpha^2 + M_\alpha^{(1)} + M_\alpha^{(2)}\big),\end{aligned}$$

which shows that $D_\alpha \lesssim \|X\|_\alpha^2 + M_\alpha^{(1)} + M_\alpha^{(2)}$ and thus the proof is complete. □

Remark 6.7. We actually showed that

$$E\Big(\sup_{\substack{0\le s<t\le 1\\ |t-s|\ge 2^{-k}}}\Big|\frac{|I_k^{\mathrm{Itô}}(X,\mathrm{d}X)_{s,t} - X_s X_{s,t}|}{|t-s|^{2\alpha}}\Big|^p\Big) \lesssim C^{2p},$$

which is obviously stronger than $E(M_\alpha^p) \lesssim C^{2p}$.

Example 6.8. The conditions of Theorem 6.6 are satisfied by all Itô martingales of the form $X_t = X_0 + \int_0^t \sigma_s \mathrm{d}W_s$, as long as σ satisfies $E(\sup_{s\in[0,1]}|\sigma_s|^{2p}) < \infty$

for some $p > 1$. Then we can take $\beta = 1/2$ so that for $p > 3$ we have $\beta - 1/(2p) > 1/3$ and X and $I^{\mathrm{Itô}}(X,\mathrm{d}X)$ are sufficiently regular to apply the results of Section 5.

7 Pathwise stochastic differential equations

We are now ready to solve SDEs of the form

$$\mathrm{d}y(t) = b(y(t))\mathrm{d}t + \sigma(y(t))\mathrm{d}v(t), \qquad y(0) = y_0, \tag{7.1}$$

pathwise, where the "stochastic" integral $\mathrm{d}v$ will be interpreted as $I(\sigma(y),\mathrm{d}v)$ or $I^{\mathrm{Itô}}(\sigma(y), \mathrm{d}v)$.

Assume for example that $(v, L(v,v)) \in \mathcal{C}^\alpha(\mathbb{R}^d) \times \mathcal{C}^{2\alpha}(\mathbb{R}^{d\otimes d})$ for some $\alpha \in (1/3, 1/2)$ are given, and that $b\colon \mathbb{R}^n \to \mathbb{R}^n$ is Lipschitz continuous whereas $\sigma \in C_b^{1+\varepsilon}(\mathbb{R}^n, \mathcal{L}(\mathbb{R}^d;\mathbb{R}^n))$ for some ε with $\alpha(2+\varepsilon) > 1$. Then Corollary 4.6 implies that $\sigma(y) \in \mathcal{D}_v^{\varepsilon\alpha}(\mathcal{L}(\mathbb{R}^d;\mathbb{R}^n))$ for every $y \in \mathcal{D}_v^\alpha(\mathbb{R}^n)$, and Theorem 4.10 then shows that $y_0 + \int_0^\cdot b(y(t))\mathrm{d}t + I(\sigma(y),\mathrm{d}v) \in \mathcal{D}_v^\alpha(\mathbb{R}^n)$ with derivative $\sigma(y)$. Moreover, if we restrict ourselves to the set

$$\mathcal{M}_\sigma = \{y \in \mathcal{D}_v^\alpha(\mathbb{R}^n) : \|y^v\|_\infty \le \|\sigma\|_\infty\},$$

then the map $\mathcal{M}_\sigma \ni (y, y^v) \mapsto \Gamma(y) = (y_0 + \int_0^\cdot b(y(t))\mathrm{d}t + I(\sigma(y),\mathrm{d}v), \sigma(y)) \in \mathcal{M}_\sigma$ satisfies the bound

$$\begin{aligned}\|\Gamma(y)\|_{v,\alpha} &\lesssim |y_0| + |b(0)| + \|b\|_{\mathrm{Lip}}\|y\|_\infty + \|\sigma(y)\|_{v,\varepsilon\alpha}(\|v\|_\alpha + \|v\|_\alpha^2 + \|L(v,v)\|_{2\alpha}) + \|\sigma(y)\|_\alpha \\ &\lesssim |y_0| + |b(0)| + (1+\|b\|_{\mathrm{Lip}})(1+\|\sigma\|_{C_b^{1+\varepsilon}}^{2+\varepsilon})(1+\|v\|_\alpha^2 + \|L(v,v)\|_{2\alpha})(1+\|y\|_{v,\varepsilon\alpha}),\end{aligned}$$

where we wrote $\|b\|_{\mathrm{Lip}}$ for the Lipschitz norm of b.

To pick up a small factor λ that turns Γ into a contraction, we apply a scaling argument. For $\lambda \in (0,1]$ we introduce the map $\Lambda_\lambda\colon \mathcal{C}^\beta \to \mathcal{C}^\beta$ defined by $\Lambda_\lambda f(t) = f(\lambda t)$. Then for $\lambda = 2^{-k}$ and on the interval $[0,\lambda]$ equation (7.1) is equivalent to

$$\mathrm{d}y^\lambda(t) = \lambda b(y^\lambda(t))\mathrm{d}t + \lambda^\alpha \sigma(y^\lambda(t))\mathrm{d}v^\lambda(t), \qquad y^\lambda(0) = y_0, \tag{7.2}$$

where $y^\lambda = \Lambda_\lambda y$, $v^\lambda = \lambda^{-\alpha}\Lambda_\lambda v$. To see this, note that

$$\Lambda_\lambda I(f,\mathrm{d}v) = \lim_{N\to\infty}\int_0^{\lambda\cdot} S_N f(t)\partial_t S_N v(t)\mathrm{d}t = \lim_{N\to\infty}\int_0^{\cdot} (\Lambda_\lambda S_N f)(t)\partial_t(\Lambda_\lambda S_N v)(t)\mathrm{d}t.$$

But now $\Lambda_{2^{-k}} S_N g = S_{N-k} \Lambda_\lambda g$ for all sufficiently large N, and therefore

$$\Lambda_\lambda I(f, \mathrm{d}v) = \lambda^\alpha I(\Lambda_\lambda f, \mathrm{d}v^\lambda).$$

For the quadratic covariation we have

$$\Lambda_\lambda [f, v] = [\Lambda_\lambda f, \Lambda_\lambda v] = \lambda^\alpha [\Lambda_\lambda f, v^\lambda],$$

from where we get (7.2) also in the Itô case. In other words we can replace b by λb, σ by $\lambda^\alpha \sigma$, and v by v^λ.

It now suffices to show that v^λ, $L(v^\lambda, v^\lambda)$, and $[v^\lambda, v^\lambda]$ are uniformly bounded in λ. Since only increments of v appear in (7.1) we may suppose $v(0) = 0$, in which case it is easy to see that $\|\Lambda_\lambda v\|_\alpha \lesssim \lambda^\alpha \|v\|_\alpha$ and $\|[v^\lambda, v^\lambda]\|_{2\alpha} \lesssim \|[v, v]\|_{2\alpha}$. As for the Lévy area, we have

$$\begin{aligned} L(v^\lambda, v^\lambda) &= I(v^\lambda, \mathrm{d}v^\lambda) - \pi_<(v^\lambda, v^\lambda) - S(v^\lambda, v^\lambda) = \lambda^{-2\alpha} \Lambda_\lambda I(v, \mathrm{d}v) - \pi_<(v^\lambda, v^\lambda) - S(v^\lambda, v^\lambda) \\ &= \lambda^{-2\alpha} \big\{ \Lambda_\lambda L(v, v) + [\Lambda_\lambda \pi_<(v, v) - \pi_<(\Lambda_\lambda v, \Lambda_\lambda v)] + [\Lambda_\lambda S(v, v) - S(\Lambda_\lambda v, \Lambda_\lambda v)] \big\}, \end{aligned}$$

and therefore

$$\|L(v^\lambda, v^\lambda)\|_{2\alpha} \lesssim \|L(v, v)\|_{2\alpha} + \|S(v, v)\|_{2\alpha} + \|v\|_\alpha^2 + \lambda^{-2\alpha} \|\Lambda_\lambda \pi_<(v, v) - \pi_<(v^\lambda, v^\lambda)\|_{2\alpha}.$$

But now

$$\begin{aligned} |\Lambda_\lambda \pi_<(v, v)_{s,t} - \pi_<(v^\lambda, v^\lambda)_{s,t}| &\leq |\pi_<(v, v)_{\lambda s, \lambda t} - v(\lambda s) v_{\lambda s, \lambda t}| \\ &\quad + |\Lambda_\lambda v(s) (\Lambda_\lambda v)_{s,t} - \pi_<(\Lambda_\lambda v, \Lambda_\lambda v)_{s,t}| \\ &\lesssim \|v\|_\alpha^2 |\lambda(t-s)|^{2\alpha} + \|\Lambda_\lambda v\|_\alpha |t-s|^{2\alpha} \\ &\lesssim \lambda^{2\alpha} \|v\|_\alpha^2 |(t-s)|^{2\alpha}. \end{aligned}$$

From here we obtain the uniform boundedness of $\|v^\lambda\|_{v^\lambda, \alpha}$ for small λ, depending only on $b, \sigma, v, L(v, v)$ and possibly $[v, v]$, but not on y_0. If $\sigma \in C_b^{2+\varepsilon}$, similar arguments give us a contraction for small λ, and therefore we obtain the existence and uniqueness of solutions to (7.2). Since all operations involved depend on $(v, L(v, v), y_0)$ and possibly $[v, v]$ in a locally Lipschitz continuous way, also y^λ depends locally Lipschitz continuously on this extended data.

Then $y = \Lambda_{\lambda^{-1}} y^\lambda$ solves (7.1) on $[0, \lambda]$, and since λ can be chosen independently of y_0, we obtain the global in time existence and uniqueness of a solution which depends locally Lipschitz continuously on $(v, L(v, v), y_0)$ and possibly $[v, v]$.

Theorem 7.1. *Let $\alpha \in (1/3, 1)$ and let $(v, L(v, v))$ satisfy the assumptions of Theorem 4.10. Let $y_0 \in \mathbb{R}^n$ and $\varepsilon > 0$ be such that $\alpha(2+\varepsilon) > 1$ and let $\sigma \in C_b^{2+\varepsilon}(\mathbb{R}^n, \mathcal{L}(\mathbb{R}^d; \mathbb{R}^n))$ and $b \colon \mathbb{R}^n \to \mathbb{R}^n$ be Lipschitz continuous. Then there exists a unique $y \in \mathcal{D}_v^\alpha(\mathbb{R}^n)$ such that*

$$y = y_0 + \int_0^\cdot b(y(t)) \mathrm{d}t + I(\sigma(y), \mathrm{d}v).$$

The solution y depends locally Lipschitz continuously on $(v, L(v, v), y_0)$. If furthermore $[v, v]$ satisfies the assumptions of Corollary 5.4, then there also exists a unique solution $x \in \mathcal{D}_v^\alpha(\mathbb{R}^n)$ to

$$\begin{aligned} x &= y_0 + \int_0^\cdot b(x(t)) \mathrm{d}t + I^{\mathrm{It\hat{o}}}(\sigma(x), \mathrm{d}v) \\ &= y_0 + \int_0^\cdot b(x(t)) \mathrm{d}t + I(\sigma(x), \mathrm{d}v) - \frac{1}{2} \int_0^\cdot \mathrm{D}\sigma(x(t)) \sigma(x(t)) \mathrm{d}[v, v]_t \end{aligned}$$

and x depends locally Lipschitz continuously on $(v, L(v, v), [v, v], y_0)$.

Remark 7.2. Since our integral is pathwise continuous, we can easily consider anticipating initial conditions and coefficients. Such problems arise naturally in the study of random dynamical systems; see for example [27, 2]. There are various approaches, for example filtration enlargements, Skorokhod integrals, or the noncausal Ogawa integral. While filtration enlargements are technically difficult, Skorokhod integrals have the disadvantage that in the anticipating case the integral is not always easy to interpret and can behave pathologically; see [5]. With classical rough path theory these technical problems disappear. But then the integral is given as the limit of compensated Riemann sums (see Proposition 2.5). With our formulation of the integral it is clear that we can indeed consider usual Riemann sums. An approach to pathwise integration which allows to define anticipating integrals without many technical difficulties while retaining a natural interpretation of the integral is the stochastic calculus via regularization of Russo and Vallois [48, 49]. The integral notion studied by Ogawa [40, 41] for anticipating stochastic integrals with respect to Brownian motion is based on Fourier expansions of integrand and integrator, and therefore related to our and the Stratonovich integral (see Nualart, Zakai [39]). Similarly as the classical Itô integral, it is interpreted in an L^2 limit sense, not a pathwise one.

A Regularity for Schauder expansions with affine coefficients

Here we study the regularity of series of Schauder functions that have affine functions as coefficients. First let us establish an auxiliary result.

Lemma A.1. *Let $s < t$ and let $f : [s,t] \to \mathcal{L}(\mathbb{R}^d, \mathbb{R}^n)$ and $g : [s,t] \to \mathbb{R}^d$ be affine functions. Then for all $r \in (s,t)$ and for all $h > 0$ with $r - h \in [s,t]$ and $r + h \in [s,t]$ we have*

$$|(fg)_{r-h,r} - (fg)_{r,r+h}| \leq 4|t-s|^{-2} h^2 \|f\|_\infty \|g\|_\infty. \tag{A.1}$$

Proof. For $f(r) = a_1 + (r-s)b_1$ and $g(r) = a_2 + (r-s)b_2$ we have

$$|(fg)_{r-h,r} - (fg)_{r,r+h}| = |2f(r)g(r) - f(r-h)g(r-h) - f(r+h)g(r+h)| = |-h^2 b_1 b_2|.$$

Now $f_{s,t} = b_1(t-s)$ so that $|b_1| \leq 2|t-s|^{-1}\|f\|_\infty$, and similarly for b_2. □

Now we are ready to prove the regularity estimate.

Lemma A.2. *Let $\alpha \in (0,2)$ and let $(u_{pm}) \in \mathcal{A}^\alpha(\mathbb{R}^d)$. Then $\sum_{p,m} u_{pm}\varphi_{pm} \in \mathcal{C}^\alpha(\mathbb{R}^d)$ and*

$$\Big\|\sum_{p,m} u_{pm}\varphi_{pm}\Big\|_\alpha \lesssim \|(u_{pm})\|_{\mathcal{A}^\alpha}.$$

Proof. We need to examine the coefficients $2^{-q}\langle \chi_{qn}, \mathrm{d}(\sum_{pm} u_{pm}\varphi_{pm})\rangle$. The cases $(q,n) = (-1,0)$ and $(q,n) = (0,0)$ are easy, so let $q \geq 0$ and $1 \leq n \leq 2^q$. If $p > q$, then $\varphi_{pm}(t^i_{qn}) = 0$ for $i = 0,1,2$ and for all m, and therefore

$$2^{-q}\Big\langle \chi_{qn}, \mathrm{d}\Big(\sum_{p,m} u_{pm}\varphi_{pm}\Big)\Big\rangle = 2^{-q}\sum_{p\leq q}\sum_m \langle \chi_{qn}, \mathrm{d}(u_{pm}\varphi_{pm})\rangle.$$

If $p < q$, there is at most one m_0 with $\langle \chi_{qn}, \mathrm{d}(u_{pm}\varphi_{pm})\rangle \neq 0$. The support of χ_{qn} is then contained in $[t^0_{pm_0}, t^1_{pm_0}]$ or in $[t^1_{pm_0}, t^2_{pm_0}]$ and u_{pm} and φ_{pm} are affine on these intervals, so (A.1) yields

$$\begin{aligned}\sum_m |2^{-q}\langle \chi_{qn}, \mathrm{d}(u_{pm}\varphi_{pm})\rangle| &= \sum_m |(u_{pm}\varphi_{pm})_{t^0_{qn},t^1_{qn}} - (u_{pm}\varphi_{pm})_{t^1_{qn},t^2_{qn}}| \\ &\lesssim 2^{2p}2^{-2q}\|u_{pm}\|_\infty\|\varphi_{pm}\|_\infty \lesssim 2^{p(2-\alpha)-2q}\|(u_{pm})\|_{\mathcal{A}^\alpha}.\end{aligned}$$

For $p = q$ we have $\varphi_{qn}(t^0_{qn}) = \varphi_{qn}(t^2_{qn}) = 0$ and $\varphi_{qn}(t^1_{qn}) = 1/2$, and thus

$$\sum_m |2^{-q}\langle \chi_{qn}, \mathrm{d}(u_{qm}\varphi_{qm})\rangle| = \Big|(u_{qn}\varphi_{qn})_{t^0_{qn},t^1_{qn}} - (u_{qn}\varphi_{qn})_{t^1_{qn},t^2_{qn}}\Big| = |u(t^1_{qn})| \lesssim 2^{-\alpha q}\|(u_{pm})\|_{\mathcal{A}^\alpha}.$$

Combining these estimate and using that $\alpha < 2$, we obtain

$$2^{-q}\Big|\Big\langle \chi_{qn}, \mathrm{d}\Big(\sum_{pm} u_{pm}\varphi_{pm}\Big)\Big\rangle\Big| \lesssim \sum_{p\le q} 2^{p(2-\alpha)-2q}\|(u_{pm})\|_{\mathcal{A}^\alpha} \simeq 2^{-\alpha q}\|(u_{pm})\|_{\mathcal{A}^\alpha},$$

which completes the proof. □

References

[1] Andreas Andresen, Peter Imkeller, and Nicolas Perkowski, *Large Deviations for Hilbert-Space-Valued Wiener Processes: A Sequence Space Approach*, Malliavin Calculus and Stochastic Analysis, Springer, 2013, pp. 115–138.

[2] Ludwig Arnold and Peter Imkeller, *Rotation numbers for linear stochastic differential equations*, Ann. Probab. **27** (1999), no. 1, 130–149. MR-1681134

[3] Hajer Bahouri, Jean-Yves Chemin, and Raphael Danchin, *Fourier analysis and nonlinear partial differential equations*, Springer, 2011.

[4] Paolo Baldi and Bernard Roynette, *Some exact equivalents for the Brownian motion in Hölder norm*, Probab. Theory Related Fields **93** (1992), no. 4, 457–484.

[5] Martin T. Barlow and Peter Imkeller, *On some sample path properties of Skorohod integral processes*, Séminaire de Probabilités, XXVI, Lecture Notes in Math., vol. 1526, Springer, Berlin, 1992, pp. 70–80. MR-1231984

[6] Gérard Ben Arous, Mihai Grădinaru, and Michel Ledoux, *Hölder norms and the support theorem for diffusions*, Ann. Inst. H. Poincaré Probab. Statist. **30** (1994), no. 3, 415–436. MR-1288358

[7] Gérard Ben Arous and Michel Ledoux, *Grandes déviations de Freidlin-Wentzell en norme hölderienne*, Séminaire de Probabilités, XXVIII, Lecture Notes in Math., vol. 1583, Springer, Berlin, 1994, pp. 293–299. MR-1329119

[8] Jean-Michel Bony, *Calcul symbolique et propagation des singularites pour les équations aux dérivées partielles non linéaires*, Ann. Sci. Éc. Norm. Supér. (4) **14** (1981), 209–246.

[9] Rémi Catellier and Khalil Chouk, *Paracontrolled distributions and the 3-dimensional stochastic quantization equation*, arXiv preprint arXiv:1310.6869 (2013).

[10] Khalil Chouk and Peter K. Friz, *Support theorem for a singular semilinear stochastic partial differential equation*, arXiv preprint arXiv:1409.4250 (2014).

[11] Zbigniew Ciesielski, *On the isomorphisms of the spaces H_α and m*, Bull. Acad. Polon. Sci. Sér. Sci. Math. Astronom. Phys. **8** (1960), 217–222. MR-0132389

[12] Zbigniew Ciesielski, Gérard Kerkyacharian, and Bernard Roynette, *Quelques espaces fonctionnels associés à des processus gaussiens*, Studia Math. **107** (1993), no. 2, 171–204.

[13] Laure Coutin and Zhongmin Qian, *Stochastic analysis, rough path analysis and fractional Brownian motions.*, Probab. Theory Relat. Fields **122** (2002), no. 1, 108–140.

[14] M'hamed Eddahbi, Modeste N'zi, and Youssef Ouknine, *Grandes déviations des diffusions sur les espaces de Besov-Orlicz et application*, Stochastics Stochastics Rep. **65** (1999), no. 3-4, 299–315. MR-1687624

[15] Hans Föllmer, *Calcul d'Itô sans probabilités*, Séminaire de Probabilités XV **80** (1979), 143–150.

[16] Peter Friz and Martin Hairer, *A course on rough paths*, Springer, 2014.

[17] Peter Friz and Nicolas Victoir, *Differential equations driven by Gaussian signals*, Ann. Inst. Henri Poincaré Probab. Stat. **46** (2010), no. 2, 369–413. MR-2667703

[18] Peter Friz and Nicolas Victoir, *Multidimensional stochastic processes as rough paths. Theory and applications*, Cambridge University Press, 2010.

[19] Nina Gantert, *Einige große Abweichungen der Brownschen Bewegung*, Ph.D. thesis, Universität Bonn, 1991.

[20] Nina Gantert, *Self-similarity of Brownian motion and a large deviation principle for random fields on a binary tree*, Probab. Theory Related Fields **98** (1994), no. 1, 7–20. MR-1254822

[21] Massimiliano Gubinelli, *Controlling rough paths*, J. Funct. Anal. **216** (2004), no. 1, 86–140.

[22] Massimiliano Gubinelli, *Ramification of rough paths*, J. Differential Equations **248** (2010), no. 4, 693–721.

[23] Massimiliano Gubinelli, Peter Imkeller, and Nicolas Perkowski, *Paracontrolled distributions and singular PDEs*, Forum Math. Pi **3** (2015), e6, 75. MR-3406823

[24] Massimiliano Gubinelli and Nicolas Perkowski, *KPZ reloaded*, arXiv preprint arXiv:1508.03877 (2015).

[25] Martin Hairer, *Solving the KPZ equation*, Ann. Math. **178** (2013), no. 2, 559–664.

[26] Martin Hairer, *A theory of regularity structures*, Invent. Math. **198** (2014), no. 2, 269–504. MR-3274562

[27] Peter Imkeller, *The smoothness of laws of random flags and Oseledets spaces of linear stochastic differential equations*, Potential Anal. **9** (1998), no. 4, 321–349. MR-1667026

[28] Svante Janson, *Gaussian Hilbert spaces*, Cambridge Tracts in Mathematics, vol. 129, Cambridge University Press, Cambridge, 1997. MR-1474726

[29] Anna Kamont, *Isomorphism of some anisotropic Besov and sequence spaces*, Studia Math. **110** (1994), no. 2, 169–189. MR-1279990

[30] David TB Kelly, *Rough path recursions and diffusion approximations*, Ann. Appl. Probab. **26** (2016), no. 1, 425–461.

[31] Terry Lyons and Zhongmin Qian, *Flow equations on spaces of rough paths*, J. Funct. Anal. **149** (1997), no. 1, 135–159. MR-1471102

[32] Terry Lyons and Zhongmin Qian, *System control and rough paths*, Oxford University Press, 2002.

[33] Terry Lyons and Ofer Zeitouni, *Conditional exponential moments for iterated Wiener integrals*, Ann. Probab. **27** (1999), no. 4, 1738–1749. MR-1742886

[34] Terry J. Lyons, *Uncertain volatility and the risk-free synthesis of derivatives*, Appl. Math. Finance **2** (1995), no. 2, 117–133.

[35] Terry J. Lyons, *Differential equations driven by rough signals*, Rev. Mat. Iberoam. **14** (1998), no. 2, 215–310.

[36] Terry J. Lyons, Michael Caruana, and Thierry Lévy, *Differential equations driven by rough paths*, Lecture Notes in Mathematics, vol. 1908, Springer, Berlin, 2007. MR-2314753

[37] Terry J. Lyons and Zhongmin Qian, *Calculus for multiplicative functionals, Itô's formula and differential equations*, Itô's stochastic calculus and probability theory, Springer, Tokyo, 1996, pp. 233–250. MR-1439528

[38] David Nualart and Samy Tindel, *A construction of the rough path above fractional Brownian motion using Volterra's representation*, The Annals of Probability **39** (2011), no. 3, 1061–1096.

[39] David Nualart and Moshe Zakai, *On the Relation Between the Stratonovich and Ogawa Integrals*, Ann. Probab. **17** (1989), no. 4, 1536–1540.

[40] Shigeyoshi Ogawa, *Quelques Propriétés de l'Intégrale Stochastique du Type Noncausal*, Japan J. Appl. Math. **1** (1984), 405–416.

[41] Shigeyoshi Ogawa, *The stochastic Integral of Noncausal Type as an Extension of the Symmetric Integrals*, Japan J. Appl. Math. **2** (1985), 229–240.

[42] Shigeyoshi Ogawa, *Noncausal stochastic calculus revisited—around the so-called Ogawa integral*, Advances in deterministic and stochastic analysis, World Sci. Publ., Hackensack, NJ, 2007, pp. 297–320. MR-2381859

[43] Giovanni Peccati and Murad S. Taqqu, *Wiener chaos: Moments, cumulants and diagrams. A survey with computer implementation.*, Springer, 2011.

[44] Nicolas Perkowski and David J Prömel, *Pathwise stochastic integrals for model free finance*, to appear in Bernoulli, arXiv:1311.6187 (2016+).

[45] David J. Prömel and Mathias Trabs, *Rough differential equations on Besov spaces*, arXiv preprint arXiv:1506.03252 (2015).

[46] Mathieu Rosenbaum, *First order p-variations and Besov spaces*, Statist. Probab. Lett. **79** (2009), no. 1, 55–62. MR-2483397

[47] Bernard Roynette, *Mouvement Brownien et espaces de Besov*, Stochastics Stochastics Rep. (1993), 221–260.

[48] Francesco Russo and Pierre Vallois, *Forward, backward and symmetric stochastic integration*, Probab. Theory Related Fields **97** (1993), no. 3, 403–421. MR-1245252

[49] Francesco Russo and Pierre Vallois, *Elements of stochastic calculus via regularization*, Séminaire de Probabilités XL, Lecture Notes in Math., vol. 1899, Springer, Berlin, 2007, pp. 147–185. MR-2409004

[50] Hans Triebel, *Theory of function spaces. III*, Monographs in Mathematics, vol. 100, Birkhäuser Verlag, Basel, 2006. MR-2250142

[51] Jérémie Unterberger, *A rough path over multidimensional fractional Brownian motion with arbitrary Hurst index by Fourier normal ordering*, Stochastic Processes and their Applications **120** (2010), no. 8, 1444–1472.

[52] Jérémie Unterberger, *Hölder-Continuous Rough Paths by Fourier Normal Ordering*, Comm. Math. Phys. **298** (2010), no. 1, 1–36.

[53] Laurence C. Young, *An inequality of the Hölder type, connected with Stieltjes integration*, Acta Math. **67** (1936), no. 1, 251–282.

[54] Rongchan Zhu and Xiangchan Zhu, *Approximating three-dimensional Navier-Stokes equations driven by space-time white noise*, arXiv preprint arXiv:1409.4864 (2014).

[55] Rongchan Zhu and Xiangchan Zhu, *Three-dimensional Navier-Stokes equations driven by space-time white noise*, J. Differential Equations **259** (2015), no. 9, 4443–4508. MR-3373412