A Fourier approach to pathwise stochastic integration

Massimiliano Gubinelli*  
CEREMADE & CNRS UMR 7534  
Université Paris-Dauphine  
and Institut Universitaire de France  
gubinelli@ceremade.dauphine.fr

Peter Imkeller  
Institut für Mathematik  
Humboldt-Universität zu Berlin  
imkeller@math.hu-berlin.de

Nicolas Perkowski†  
CEREMADE & CNRS UMR 7534  
Université Paris-Dauphine  
perkowski@ceremade.dauphine.fr

October 20, 2014

Abstract

We develop a Fourier 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.

Contents

1 Introduction .......................... 2

2 Preliminaries ........................... 5
   2.1 Ciesielski’s isomorphism .......... 5
   2.2 Young integration and rough paths . 7

3 Paradifferential calculus and Young integration .............. 9
   3.1 Paradifferential calculus with Schauder functions .......... 9
   3.2 Young’s integral and its different components ......... 10

M.G. is 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. is 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.
1 Introduction

The theory of rough paths [Lyo98] has recently been extended to a multiparameter setting [Hai14, GIP13]. While [Hai14] has a much wider range of applicability, both approaches allow to solve many interesting SPDEs that were well out of reach with previously existing methods; for example the continuous parabolic Anderson in dimension two [Hai14, GIP13, CF14], the three-dimensional stochastic quantization equation [Hai14, CC13], the KPZ equation [Hai13, GP14], or the three-dimensional stochastic Navier Stokes equation [ZZ14b, ZZ14a].

Our methods developed in [GIP13] are based on harmonic analysis, on Littlewood-Paley decompositions of tempered distributions, and on 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 classical rough path SDEs is not quite straightforward. That is why here we develop the approach of [GIP13] in the slightly different language of Haar / Schauder functions, which allows us to communicate our basic ideas while requiring only very basic knowledge in analysis. Moreover, in the Haar Schauder formulation the application to SDEs poses no additional technical challenges.

It is a classical result of Ciesielski [Cie60] that $C^\alpha := C^\alpha([0, 1], \mathbb{R}^d)$, the space of $\alpha$–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

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

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 $(\langle H_{pm}, df \rangle)_{p,m}$ 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 [Tri06].

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 \( \int_0^t G_{pm}(s)dG_{qn}(s) \), 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 \( L(\mathbb{R}^d, \mathbb{R}^n) \), the space of linear maps from \( \mathbb{R}^d \) to \( \mathbb{R}^n \), and \( \mathbb{R}^d \) respectively, we can formally define

\[
\int_0^t f(s)dg(s) := \sum_{p,m} \sum_{q,n} \langle H_{pm}, df \rangle \langle H_{qn}, dg \rangle \int_0^t G_{pm}(s)dG_{qn}(s).
\]

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 \( C^\beta \) if and only if \( \alpha + \beta > 1 \). Obviously, here we simply recover Young’s integral [You36]. In our study of this integral, we identify different components:

\[
\int_0^t f(s)dg(s) = S(f,g)(t) + \pi_<(f,g)(t) + L(f,g)(t),
\]

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 \( \int_0^t f(s)dg(s) - \pi_<(f,g) \in C^{\alpha+\beta} \). We will also see that for sufficiently smooth functions \( F \) we have \( F(f) \in C^\alpha \) but \( F(f) - \pi_<(DF(f),f) \in C^{2\alpha} \). So both \( \int_0^t f(s)dg(s) \) 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:

\[
\left\| L(\pi_<(f^g,g),g) - \int_0^t f^g(s)dL(g,g)(s) \right\|_{3\beta} \leq \| f^g \|_\beta \| g \|_\beta \| g \|_\beta.
\]

Therefore, the integral \( \int_0^t f(s)dg(s) \) 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’ [Ly098] 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 recall some details on 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.

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 [CKR93], Roynette [Roy93], and Rosen-
baum [Ros09]. Baldi and Roynette [BR92] 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 [BL94] for the extension to diffusions, Eddahbi, N’zi, and Ouaknine [ENO99]
for the large deviation principle for diffusions in Besov spaces, and Andresen, Imkeller, and
Perkowski [AIP13] for the large deviation principle for a Hilbert space valued Wiener process
in Hölder topology. Ben Arous, Grădinaru, and Ledoux [BGL94] use Schauder functions to
extend the Stroock-Varadhan support theorem for diffusions from the uniform to the Hölder
topology. Lyons and Zeitouni [LZ99] use Schauder functions to prove exponential moment
bounds for Stratonovich iterated integrals of a Brownian motion conditioned to stay in a
small ball. Gantert [Gan94] 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 [Roy93] constructs a version of Young’s integral on Besov spaces and shows that
in the one dimensional case the Stratonovich integral \( \int_0^t F(W_s) dW_s \), 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 [CKR93] slightly extend the Young integral of [Roy93], 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 [Kam94] extends the approach of [CKR93] to define
a multiparameter Young integral for functions in anisotropic Besov spaces. Ogawa [Oga84,
Oga85] investigates an integral for anticipating integrands he calls \textit{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 [Oga85], Nualart, Zakai [NZ89]), and used to develop a stochastic
calculus on a Brownian basis with \textit{noncausal} SDE (Ogawa [Oga07]).

Rough paths have been introduced by Lyons [Lyo98], see also [Lyo95, LQ96, LQ97] for
previous results. Lyons observed that solution flows to SDEs (or more generally ordinary dif-
fferential 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 $[p]$ 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 [Gub04, Gub10] 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 [LQ02, LCL07, FV10b, FH14].

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 $a \leq c \cdot b$, and we write $a \asymp b$ 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 $\partial^\mu = \partial^{\mu_1}_1 \partial^{\mu_2}_2 \cdots \partial^{\mu_d}_d$. $DF$ or $F'$ denote the total derivative of $F$. For $k \in \mathbb{N}$ we denote by $D^k F$ the $k$-th order derivative of $F$. We also write $\partial_x$ 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

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

When completed by $H_{00} \equiv 1$, the Haar functions are an orthonormal basis of $L^2([0,1],dt)$. For convenience 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 $G_{pm}(t) := \int_0^t H_{pm}(s) ds$ for $t \in [0,1]$, $p \in \mathbb{N}$, $0 \leq m \leq 2^p$. More explicitly, $G_{00}(t) = t$ and for $p \in \mathbb{N}$, $1 \leq m \leq 2^p$

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

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 $t_{pm}^i$, $i = 0, 1, 2$, as

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

for $p \in \mathbb{N}$ and $1 \leq m \leq 2^p$. Further, we set $t_{-10}^0 := 0$, $t_{-10}^1 := 0$, $t_{-10}^2 := 1$, and $t_{00}^0 := 0$, $t_{00}^1 := 1$, $t_{00}^2 := 1$, as well as $t_{p0}^i := 0$ for $p \geq 1$ and $i = 0, 1, 2$. The definition of $t_{-10}^i$ and $t_{00}^i$ for $i \neq 1$ is rather arbitrary, but the definition for $i = 1$ simplifies for example the statement of Lemma 2.1 below.
For \( f \in C([0,1], \mathbb{R}^d) \), \( p \in \mathbb{N} \), and \( 1 \leq m \leq 2^p \), we write
\[
\langle H_{pm}, df \rangle := 2^\frac{m}{2} [(f(t_{pm}^1) - f(t_{pm}^0)) - (f(t_{pm}^1) - f(t_{pm}^0))]
\]
\[
= 2^\frac{m}{2} [2f(t_{pm}^1) - f(t_{pm}^0) - f(t_{pm}^1)]
\]
and \( \langle H_{00}, df \rangle := f(1) - f(0) \) as well as \( \langle H_{-10}, df \rangle := f(0) \). Note that we only defined \( G_{-10} \) and not \( H_{-10} \).

**Lemma 2.1.** For \( f : [0,1] \to \mathbb{R}^d \), the function
\[
f_k := \langle H_{-10}, df \rangle G_{-10} + \langle H_{00}, df \rangle G_{00} + \sum_{p=0}^{k} \sum_{m=1}^{2^p} \langle H_{pm}, df \rangle G_{pm} = \sum_{p=-1}^{k} \sum_{m=0}^{2^p} \langle H_{pm}, df \rangle G_{pm}
\]
is the linear interpolation of \( f \) between the points \( t_{10}^p, t_{00}^p, t_{pm}^1, 0 \leq p, k \leq m \leq 2^p \). If \( f \) is continuous, then \( (f_k) \) converges uniformly to \( f \) as \( k \to \infty \).

Ciesielski [Cie60] 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
\[
\| f \|_{C^\alpha} := \| f \|_\infty + \sup_{0 \leq s \leq t \leq 1} \frac{|f(s,t)|}{|t-s|^\alpha},
\]
where we introduced the notation
\[
f_{s,t} := f(t) - f(s).
\]

**Lemma 2.2** ([Cie60]). Let \( \alpha \in (0,1) \). A continuous function \( f : [0,1] \to \mathbb{R}^d \) is in \( C^\alpha \) if and only if \( \sup_{p,m} 2^{p(\alpha-1/2)} \| \langle H_{pm}, df \rangle \| \to \infty \). In this case
\[
\| f - f_{N-1} \|_\infty = \left\| \sum_{p=N}^{\infty} \sum_{m=0}^{2^p} \langle H_{pm}, df \rangle G_{pm} \right\|_\infty \lesssim \| f \|_\infty 2^{-\alpha N}.
\]

Before we continue, let us slightly change notation. We want to get rid of the factor \( 2^{-p/2} \) in (1), and therefore we define for \( p \in \mathbb{N} \) and \( 0 \leq m \leq 2^p \) the rescaled functions
\[
\chi_{pm} := 2^\frac{m}{2} H_{pm} \quad \text{and} \quad \varphi_{pm} := 2^\frac{m}{2} G_{pm},
\]
as well as \( \varphi_{-10} := G_{-10} \equiv 1 \). Then we have for \( p \in \mathbb{N} \) and \( 1 \leq m \leq 2^p \)
\[
\| \varphi_{pm}(t) \|_\infty = \varphi_{pm}(t_{pm}) = 2^\frac{m}{2} \int_{t_{pm}^0}^{t_{pm}^1} 2^\frac{s}{2} ds = 2^p \left( \frac{2m-1}{2p+1} - \frac{2m-2}{2p+1} \right) = \frac{1}{2},
\]
so that \( \| \varphi_{pm} \|_\infty \leq 1 \) for all \( p, m \). The expansion of \( f \) in terms of \( (\varphi_{pm}) \) 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 \( m \geq 1 \)
\[
f_{pm} := 2^{-p} \langle \chi_{pm}, df \rangle = 2f(t_{pm}^1) - f(t_{pm}^0) - f(t_{pm}^1) + f(t_{pm}^0) = f_{t_{pm}^1} - f_{t_{pm}^0}.
\]
We write \( \langle \chi_{pm}, df \rangle := 2^p f_{pm} \) for all values of \( (p,m) \), despite not having defined \( \chi_{-10} \).
Definition 2.3. For $\alpha > 0$ and $f : [0, 1] \to \mathbb{R}^d$ the norm $\|\cdot\|_\alpha$ is defined as

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

and we write

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

The space $C^\alpha$ 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 $C^\alpha = C^\alpha([0, 1], \mathbb{R}^d)$. Moreover, it can be shown that $C^1$ 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 $C^\alpha$ 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 $Df$. Since the tent shaped functions $\varphi_{pm}$ are not continuously differentiable, even an $f$ with a finite Schauder expansion is generally not in $C^\alpha$.

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.

**Littlewood-Paley notation.** We will employ notation inspired from Littlewood-Paley theory. For $p \geq -1$ and $f \in C([0, 1])$ we define

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

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

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

### 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 this chapter, but we feel that it could be useful for the reader to become familiar with the basic concepts of rough paths, since it is the main inspiration for the constructions developed below.

Young’s integral [You36] allows to define $\int f \, dg$ 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 $\int_0^t f(s)dg(s)$,
and we define $I^v_t f(r) \, dg(r) := \int_0^t f(r) \, dg(r) - \int_0^s f(r) \, dg(r)$. The function $t \mapsto \int_0^t f(s) \, dg(s)$ is uniquely characterized by the fact that
\[
\left| \int_s^t f(r) \, dg(r) - f(s)(g(t) - g(s)) \right| \lesssim |t - s|^{\alpha + \beta} \| f \|_\alpha \| g \|_\beta
\]
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 $\infty$.

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 [Lyo98] 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 [Gub04].

Let $\alpha \in (1/3, 1)$ and assume that we are given two functions $v, w \in C^\alpha$, as well as an associated “Riemann integral” $I^{v,w}_{s,t} = \int_s^t v(r) \, dw(r)$ that satisfies the estimate
\[
|\Phi^{v,w}_{s,t}| := |I^{v,w}_{s,t} - v(s)w_{s,t}| \lesssim |t - s|^{2\alpha}.
\]
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 $1/2(\Phi^{v,w}_{s,t} - \Phi^{w,v}_{s,t})$ 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). But let us assume nonetheless that we are given such an integral $I^{v,w}$ satisfying (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 $s, t \in [0, 1]$
\[
|f_{s,t} - f^v_{s,t}| \lesssim |t - s|^{2\alpha}.
\]

Proposition 2.4 ([Gub04], Theorem 1). Let $\alpha > 1/3$, let $v, w \in C^\alpha$, and let $I^{v,w}$ satisfy (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 $I(f, g) = \int_0^t f(s) \, dg(s)$ 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 $(\pi_n)$ is a sequence of partitions of $[0, t]$, with mesh size going to zero, then
\[
I(f, g)(t) = \lim_{n \to \infty} \sum_{t_k \in \pi_n} \left( f(t_k)g_{t_k,t_{k+1}} + f^v_{t_k}g^w_{t_k,t_{k+1}}\Phi^{v,w}_{s,t} \right).
\]

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 paracoordinated by themselves, with derivatives $v^w = w^w = 1$, 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 [Gub10] or any book on rough paths, such as [LQ02, LCL07, FV10b, FH14].
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 [Bon81]. For a gentle introduction see [BCD11].

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 \( u_{pm} : [t_{pm}^0, t_{pm}^2] \to \mathbb{R}^d \), we set for \( \alpha > 0 \)

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

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

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

In Appendix A we prove the following regularity estimate:

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

\[
\left\| \sum_{p,m} u_{pm} \varphi_{pm} \right\|_\alpha \lesssim \|(u_{pm})\|_{A^\alpha}.
\]

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 \( w \in C^\beta(\mathbb{R}^d) \). Then

\[
\pi_<(v,w) := \sum_{p=0}^\infty S_{p-1}v \Delta_pw \in C^\beta(\mathbb{R}^n) \quad \text{ and } \quad \|\pi_<(v,w)\|_\beta \lesssim \|v\|_\infty \|w\|_\beta.
\] (4)

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

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

\[
\|\pi_>(v,w)\|_\alpha \lesssim \|v\|_\alpha \|w\|_\infty, \quad \text{and} \quad \|\pi_o(v,w)\|_{\alpha+\beta} \lesssim \|v\|_\alpha \|w\|_\beta
\]

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 \( v \in C^\alpha(\mathbb{R}^d) \), \( w \in C^{\alpha+\beta} \), and \( F \in C^1_{\beta}C^{1+\beta/\alpha}(\mathbb{R}^d,\mathbb{R}) \). Then

\[
\|F(v + w) - \pi_<(DF(v + w),v)\|_{\alpha+\beta} \lesssim \|F\|_{C^1_{\beta}C^{1+\beta/\alpha}}(1 + \|v\|_{\alpha+\beta})(1 + \|w\|_{\alpha+\beta}).
\]

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

\[
\|F(v) - \pi_<(DF(v),v) - (F(u) - \pi_<(DF(u),u))\|_{\alpha+\beta} \\
\lesssim \|F\|_{C^2_{\beta}C^{2+\beta/\alpha}}(1 + \|v\|_{\alpha} + \|u\|_{\alpha})^{1+\beta/\alpha}\|v - u\|_{\alpha}.
\]

**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:

\[
(F(v + w))_{pm} = DF(v(t^1_{pm}) + w(t^1_{pm}))v_{pm} + R_{pm},
\]

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

\[
F(v + w) - \pi_<(DF(v + w),v) = \sum_{pm}^p [DF(v(t^1_{pm}) + w(t^1_{pm})) - (S_{p-1}DF(v + w))]_{[t^0_{pm},t^2_{pm}]}v_{pm}\varphi_{pm} + R.
\]

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

\[
\|[DF(v(t^1_{pm}) + w(t^1_{pm})) - (S_{p-1}DF(v + w))]_{[t^0_{pm},t^2_{pm}]}v_{pm}\|_\infty \\
\lesssim 2^{-p\beta}\|DF(v + w)\|_{C^\beta2^{-p\alpha}}\|v\|_\alpha \lesssim 2^{-p(\alpha+\beta)}\|F\|_{C^1_{\beta}C^{1+\beta/\alpha}}(1 + \|v\|_{\alpha} + \|w\|_{\alpha})^{\beta/\alpha}\|v\|_{\alpha}.
\]

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.

**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 \( f^v \in C^\alpha \) and \( |R_{s,t}| \leq \|R\|_{2\alpha}|t-s|^{2\alpha} \), then \( f - \pi_<(f^v,v) \in C^{2\alpha} \).

### 3.2 Young’s integral and its different components

In this section we construct Young’s integral using the Schauder expansion. If \( v \in C^\alpha \) and \( w \in C^3 \), then we formally define

\[
\int_0^t v(s)dw(s) := \sum_{p,m} \sum_{q,n} v_{pm}w_{qn} \int_0^t \varphi_{pm}(s)d\varphi_{qn}(s) = \sum_{p,q} \int_0^t \Delta_pv(s)d\Delta_qw(s).
\]

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 $p > q \geq 0$. Then
\[ \int_0^1 \varphi_{pm}(s)d\varphi_{qn}(s) = 2^{-p-2}\chi_{qn}(t^0_{pm}) \] (7)
for all $m,n$. If $p = q$, then $\int_0^1 \varphi_{pm}(s)d\varphi_{pm}(s) = 0$, except if $p = q = 0$, in which case the integral is bounded by 1. If $0 \leq p < q$, then for all $(m,n)$ we have
\[ \int_0^1 \varphi_{pm}(s)d\varphi_{qn}(s) = -2^{-q-2}\chi_{pm}(t^0_{qn}). \] (8)
If $p = -1$, then the integral is bounded by 1.

Proof. The cases $p = q$ and $p = -1$ are easy, so let $p > q \geq 0$. Since $\chi_{qn} \equiv \chi_{qn}(t^0_{pm})$ on the support of $\varphi_{pm}$, we have
\[ \int_0^1 \varphi_{pm}(s)d\varphi_{qn}(s) = \chi_{qn}(t^0_{pm}) \int_0^1 \varphi_{pm}(s)ds = \chi_{qn}(t^0_{pm})2^{-p-2}. \]
If $0 \leq p < q$, then integration by parts and (7) imply (8).

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

Lemma 3.9. Let $i,p \geq -1$, $q \geq 0$, $0 \leq j \leq 2^i$, $0 \leq m \leq 2^p$, $0 \leq n \leq 2^q$. Then
\[ 2^{-i} \left| \langle \chi_{ij}, d \left( \int_0^1 \varphi_{pm}\chi_{qn}ds \right) \rangle \right| \leq 2^{-2(i+p+q)} \] (9)
except if $p < q = i$. In this case we only have the worse estimate
\[ 2^{-i} \left| \langle \chi_{ij}, d \left( \int_0^1 \varphi_{pm}\chi_{qn}ds \right) \rangle \right| \leq 1. \] (10)

Proof. We have $\langle \chi_{-10}, d(\int_0^1 \varphi_{pm}\chi_{qn}ds) \rangle = 0$ for all $(p,m)$ and $(q,n)$. So let $i \geq 0$. If $i < p \lor q$, then $\chi_{ij}$ is constant on the support of $\varphi_{pm}\chi_{qn}$, and therefore Lemma 3.8 gives
\[ 2^{-i} \left| \langle \chi_{ij}, \varphi_{pm}\chi_{qn} \rangle \right| \leq \left| \langle \varphi_{pm}, \chi_{qn} \rangle \right| \leq 2^{p+q-2(p+q)} = 2^{-2(i+p+q)+p+q}. \]

Now let $i > q$. Then $\chi_{qn}$ is constant on the support of $\chi_{ij}$, and therefore another application of Lemma 3.8 implies that
\[ 2^{-i} \left| \langle \chi_{ij}, \varphi_{pm}\chi_{qn} \rangle \right| \leq 2^{-2i}2^{2p-i-2(p+i)} = 2^{-2(i+p+q)+p+q}. \]
The only remaining case is $i = q \geq p$, in which
\[ 2^{-i} \left| \langle \chi_{ij}, \varphi_{pm}\chi_{qn} \rangle \right| \leq 2^i \int_{t^2_{ij}} \varphi_{pm}(s)ds \leq \| \varphi_{pm} \| \infty \leq 1. \]
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
\[
\| \Delta_i \left( \int_0^1 \Delta_p v(s) d \Delta_q w(s) \right) \|_\infty \lesssim 2^{-((i+p)q) - i + p + q} \| \Delta_p v \|_\infty \| \Delta_q w \|_\infty, \quad (11)
\]
except if \( i = q > p \). In this case we only have the worse estimate
\[
\| \Delta_i \left( \int_0^1 \Delta_p v(s) d \Delta_q w(s) \right) \|_\infty \lesssim \| \Delta_p v \|_\infty \| \Delta_q w \|_\infty. \quad (12)
\]

Proof. The case \( i = -1 \) is easy, so let \( i \geq 0 \). We have
\[
\Delta_i \left( \int_0^1 \Delta_p v(s) d \Delta_q w(s) \right) = \sum_{j,m,n} v_{pn} w_{qn} (2^{-i} \chi_{ij}, \varphi_{pm} \chi_{qn}) \varphi_{ij}.
\]
For fixed \( j \), there are at most \( 2^{(i+p)q} - i \) non-vanishing terms in the double sum. Hence, we obtain from Lemma 3.9 that
\[
\left\| \sum_{m,n} v_{pn} w_{qn} (2^{-i} \chi_{ij}, \varphi_{pm} \chi_{qn}) \varphi_{ij} \right\|_\infty \lesssim 2^{((i+p)q) - i} \| \Delta_p v \|_\infty \| \Delta_q w \|_\infty \left( 2^{(i+p)q} + p + q + 1_{i=q=p} \right) \\
= (2^{((i+p)q) - i + p + q + 1_{i=q=p}}) \| \Delta_p v \|_\infty \| \Delta_q w \|_\infty.
\]

\[\square\]

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
\[
\| \Delta_i (\Delta_p v \Delta_q w) \|_\infty \lesssim 2^{-((i+p)q) - i + p + q} \| \Delta_p v \|_\infty \| \Delta_q w \|_\infty, \quad (13)
\]
except if \( i = q > p \) or \( i = p > q \), in which case we only have the worse estimate
\[
\| \Delta_i (\Delta_p v \Delta_q w) \|_\infty \lesssim \| \Delta_p v \|_\infty \| \Delta_q w \|_\infty. \quad (14)
\]
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 (11) and (12) are symmetric in \( p \) and \( q \). If for example \( p > i \), then \( \Delta_p (v(t^k)) = 0 \) for all \( k, j \), which implies that \( \Delta_i (\Delta_p v \Delta_q w) = 0 \).

The estimates (11) and (12) allow us to identify different components of the integral \( \int_0^1 v(s) d w(s) \). More precisely, (12) indicates that the series \( \sum_{p<q} \int_0^1 \Delta_p v(s) d \Delta_q w(s) \) is rougher than the remainder \( \sum_{p \geq q} \int_0^1 \Delta_p v(s) d \Delta_q w(s) \). Integration by parts gives
\[
\sum_{p<q} \int_0^1 \Delta_p v(s) d \Delta_q w(s) = \pi_<(v, w) - \sum_{p<q} \sum_{m,n} v_{pn} w_{qn} \int_0^1 \varphi_{qn}(s) d \varphi_{pm}(s).
\]
This motivates us to decompose the integral into three components, namely
\[
\sum_{p<q} \int_0^1 \Delta_p v(s) d \Delta_q w(s) = L(v, w) + S(v, w) + \pi_<(v, w).
\]
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):

$$L(v, w) := \sum_{p > q} \sum_{m,n} (v_{pm}w_{qn} - v_{qn}w_{pm}) \int_0^1 \varphi_{pm} d\varphi_{qn}$$

$$= \sum_p \left( \int_0^1 \Delta_p v dS_{p-1} w - \int_0^1 d(S_{p-1}v) \Delta_p w \right).$$

The symmetric part $S$ is defined as

$$S(v, w) := \sum_{m,n \leq 1} v_{0m}w_{0n} \int_0^1 \varphi_{0m} d\varphi_{0n} + \sum_{p \geq 1} \sum_m v_{pm}w_{pm} \int_0^1 \varphi_{pm} d\varphi_{pm}$$

$$= \sum_{m,n \leq 1} v_{0m}w_{0n} \int_0^1 \varphi_{0m} d\varphi_{0n} + \frac{1}{2} \sum_{p \geq 1} \Delta_p v \Delta_p w,$$

and $\pi_\leq$ is the paraproduct defined in (4). As we observed in Lemma 3.3, $\pi_\leq(v, w)$ 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 $C^\alpha \times C^\beta$ to $C^{\alpha+\beta}$.

**Proof.** We only argue for \( \sum_p \int_0^1 \Delta_p v dS_{p-1} w \), the term \(- \int_0^1 d(S_{p-1}v) \Delta_p w \) can be treated with the same arguments. Corollary 3.10 (more precisely (11)) implies that

$$\| \sum_p \Delta_i \left( \int_0^1 \Delta_p v dS_{p-1} w \right) \|_\infty \leq \sum_{p \leq i} \sum_{q < p} \| \Delta_i \left( \int_0^1 \Delta_p v d\Delta_q w \right) \|_\infty + \sum_{p > i} \sum_{q < p} \| \Delta_i \left( \int_0^1 \Delta_p v d\Delta_q w \right) \|_\infty$$

$$\leq \left( \sum_{p \leq i} \sum_{q < p} 2^{-i+q+2^{-p}} \|v\|_\alpha 2^{-q} \|w\|_\beta + \sum_{p > i} \sum_{q < p} 2^{-i+q} 2^{-p} \|v\|_\alpha 2^{-q} \|w\|_\beta \right)$$

$$\lesssim_{\alpha+\beta} 2^{-i(\alpha+\beta)} \|v\|_\alpha \|w\|_\beta,$$

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 smooth.

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

**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 smooth. The symmetric part $S(v, w)$ is always defined and smooth. And the paraproduct $\pi_\leq(v, w)$ is always defined, but it is rougher than the other components. To summarize:
\textbf{Theorem 3.14} (Young’s integral). Let $\alpha, \beta \in (0, 1)$ be such that $\alpha + \beta > 1$, and let $v \in C^\alpha$ and $w \in C^\beta$. Then the integral

$$I(v, dw) := \sum_{p,q} \int_0^1 \Delta_p v \Delta_q w = L(v, w) + S(v, w) + \pi_<(v, w) \in C^\beta$$

satisfies $\|I(v, dw)\|_\beta \lesssim \|v\|_\alpha\|w\|_\beta$ and

$$\|I(v, dw) - \pi_<(v, w)\|_{\alpha + \beta} \lesssim \|v\|_\alpha\|w\|_\beta. \quad (15)$$

\textbf{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

$$M_n^v = \sum_{p=0}^n \sum_{m=0}^{2^p} (2^{-p} \chi_{pm}, dv) \chi_{pm} = \sum_{p=0}^n \sum_{m=0}^{2^p} v_{pm} \chi_{pm} = \partial_t S_n v,$$

$n \in \mathbb{N}$, 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 $M_n^v = 0$ 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

$$L(v, w)(1) = \sum_{p=0}^{\infty} \sum_{q=0}^{p-1} \sum_{m,n} (w_{pm} w_{qn} - v_{qn} w_{pm}) \int_0^1 \varphi_{pm}(s) \chi_{qn}(s) ds$$

$$= \sum_{p=0}^{\infty} \sum_{q=0}^{p-1} \sum_{m,n} (w_{pm} w_{qn} - v_{qn} w_{pm}) 2^p \int_0^1 \chi_{qn}(s) \delta_{v_{pm}, w_{pm}}(s) ds \varphi_{pm}, 1$$

$$= \sum_{p=0}^{\infty} \sum_{q=0}^{p-1} \sum_{m,n} (w_{pm} w_{qn} - v_{qn} w_{pm}) 2^{-p} \int_0^1 \chi_{qn}(s) \delta_{v_{pm}, w_{pm}}(s) ds 2^{-p-2}$$

$$= \sum_{p=0}^{\infty} \sum_{q=0}^{p-1} 2^{-2p-2} \int_0^1 \sum_{m,n} \sum_{m'} (w_{pm} w_{qn} - v_{qn} w_{pm}) \chi_{qn}(s) \chi_{pm'}(s) ds,$$

where in the last step we used that $\chi_{pm}$ and $\chi_{pm'}$ have disjoint support for $m \neq m'$. The $p$–th Rademacher function (or \textit{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 \( \Delta M^w_p = M^w_p - M^w_{p-1} \) and similarly for \( M^u \) and \( R \) and all other discrete time processes that arise. This notation somewhat clashes with the expression \( \Delta_p u \) for the dyadic blocks of \( u \), 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 \( \int_0^1 ds \), we obtain

\[
L(v, w)(1) = \sum_{p=0}^{\infty} \sum_{q=0}^{p-1} 2^{-p-2} E \left( \Delta M^w_p \Delta M^w_q \Delta R_p - \Delta M^w_q \Delta M^w_p \Delta R_p \right)
\]

\[
= \sum_{p=0}^{\infty} 2^{-p-2} E \left( (M^w_{p-1} \Delta M^w_p - M^w_{p-1} \Delta M^w_p) \Delta R_p \right)
\]

\[
= \sum_{p=0}^{\infty} 2^{-p-2} E \left( \Delta [M^w \cdot M^u - M^v \cdot M^w, R]_p \right)
\]

Hence, \( L(v, w)(1) \) is closely related to the Lévy area \( 1/2(M^w \cdot M^u - M^v \cdot M^w) \) 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 \( w \in C^\alpha \) and \( F \in C_b^{1+\beta/\alpha} \) we have \( F(w) - \pi_c(DF(w), w) \in C^{\alpha+\beta} \). In Section 3.2 we observed that if \( v \in C^\alpha \), \( w \in C^\beta \) and \( \alpha + \beta > 1 \), then the Young integral \( I(v, dw) \) satisfies \( I(v, dw) - \pi_c(v, w) \in C^{\alpha+\beta} \). Hence, in both cases the function under consideration can be written as \( \pi_c(f^w, w) \) for a suitable \( f^w \), plus a smooth remainder. We make this our definition of paracontrolled paths:

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

\[
D_v^\beta := D_v^\beta(\mathbb{R}^n) := \left\{ (f, f^v) \in C^\alpha(\mathbb{R}^n) \times C^\beta(\mathcal{L}(\mathbb{R}^d, \mathbb{R}^n)) : f^v = f - \pi_c(f^v, v) \in C^{\alpha+\beta}(\mathbb{R}^n) \right\}
\]

If \( (f, f^v) \in D_v^\beta \), 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 \( f \in D_v^\beta \) when it is clear from the context what the derivative \( f^v \) is supposed to be. We equip \( D_v^\beta \) with the norm

\[
\| f \|_{v, \beta} := \| f^v \|_\beta + \| f^v \|_{\alpha+\beta}.
\]

If \( v \in C^\alpha \) and \( (f, \tilde{f}^v) \in D_v^\beta \), then we also write

\[
d_{D_v^\beta}(f, \tilde{f}) := \| f^v - \tilde{f}^v \|_\beta + \| f^v - \tilde{f}^v \|_{\alpha+\beta}.
\]

**Example 4.2.** Let \( \alpha \in (0, 1) \) and \( v \in C^\alpha \). Then Proposition 3.5 shows that \( F(v) \in D_v^\beta \) for every \( F \in C_b^{1+\beta/\alpha} \), with derivative \( DF(v) \).
Example 4.3. Let $\alpha + \beta > 1$ and $v \in C^\alpha, \ w \in C^\beta$. Then by (15), the Young integral $I(v, dw)$ is in $D^\alpha_\chi$, with derivative $v$.

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

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

Indeed we have $|f^r_{s,t} - \pi_\varepsilon(f^r, v)_{s,t}| \lesssim |t-s|^{\alpha+\beta}\|f^r\|_\beta\|v\|_\alpha$, which can be shown using similar arguments as for Lemma B.2 in [GIP13]. In other words, for $\alpha \in (0,1/2)$ the space $D^\alpha_\chi$ coincides with the space of controlled paths defined in Section 2.2.

The following commutator estimate, the analog of Theorem 2.3 of [Bon81] in our setting, will be useful for establishing some stability properties of $D^\beta_\chi$.

**Lemma 4.5.** Let $\alpha, \beta \in (0,1)$, and let $u \in C([0,1], \mathcal{L}(\mathbb{R}^n; \mathbb{R}^m))$, $v \in C^\alpha(\mathcal{L}(\mathbb{R}^d; \mathbb{R}^n))$, and $w \in C^\beta(\mathbb{R}^d)$. 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 $[S_{p-1}u(\pi_{<}(v, w))_{pm} - S_{p-1}(uv)w_{pm}]_{[\varphi_{pm}, t^2_{pm}]}$ is affine. By Lemma 3.2 it suffices to control

$$\|[S_{p-1}u(\pi_{<}(v, w))_{pm} - S_{p-1}(uv)w_{pm}]_{[\varphi_{pm}, t^2_{pm}]}\|_{\infty}.$$

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 $\chi_{pm} \varphi_{qm_q} \neq 0$ and similarly for $r$. We apply Lemma 3.9 to obtain for $q < p$

$$\|S_{q-1}v\Delta w\|_{pm} = \left| \sum_{r < q} v_{rm_q} \varphi_{qm_q} 2^{-p} \langle \chi_{pm}, d(\varphi_{rm_q}, \varphi_{qm_q}) \rangle \right|$$

$$\leq \|v\|_\alpha \|w\|_\beta \sum_{r < q} 2^{-r}\varphi_{2^{-q}2^{-2p-r+p+q} \leq 2^{-2p+q(2 \alpha-\beta)} \|w\|_\beta \|w\|_\alpha \|w\|_\beta.}$$

Hence

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

\[
\| [S_{p-1} u (S_{p-1} v \Delta_p \nu)_{pm} - S_{p-1} (uv) w_{pm}] \|_{\ell^2_{\nu, \ell^2_{\nu, \alpha}}(\nu)} \leq \| w_{pm} \| \times \left\| \left( (u(t^0_{pm} + \cdots + t^0_{pm} u_{pm})^2 \right) v(t^0_{pm}) + v(t^2_{pm}) - \left( (uv)(t^0_{pm} + \cdots + t^0_{pm} (uv)_{pm})^2 \right) \right\|_{\ell^2_{\nu, \ell^2_{\nu, \alpha}}(\nu)} \lesssim \| u \|_\infty \| v \|_\alpha \| w \|_\beta 2^{-p(\alpha + \beta)},
\]

where the last step follows by rebracketing.

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

\textbf{Corollary 4.6.} Let $\alpha \in (0, 1)$, $\beta \in (0, \alpha)$, $\nu \in \mathcal{C}^\alpha$, and $f \in \mathcal{D}^\beta_\nu$ with derivative $f^\nu$. Let $F \in \mathcal{C}^{\beta + \beta/\alpha}_b$. Then $F(f) \in \mathcal{D}^\beta_\nu$ with derivative $DF(f) f^\nu$, and

\[
\| F(f) \|_{\nu, \beta} \lesssim \| F \|_{\mathcal{C}^{\beta + \beta/\alpha}_b} (1 + \| v \|_\alpha)^{1+\beta/\alpha}(1 + \| f \|_{\nu, \beta})(1 + \| f^\nu \|_\infty)^{1+\beta/\alpha}.
\]

Moreover, there exists a polynomial $P$ which satisfies for all $F \in \mathcal{C}^{\beta + \beta/\alpha}_b$, $\tilde{v} \in \mathcal{C}^\alpha$, $\tilde{f} \in \mathcal{D}^\beta_\nu$, and

\[
M = \max\{\| v \|_\alpha, \| \tilde{v} \|_\alpha, \| f \|_{\nu, \beta}, \| \tilde{f} \|_{\nu, \beta}\}
\]

the bound

\[
d_{\mathcal{D}^{\beta}}(F(f), F(\tilde{f})) \leq P(M) \| F \|_{\mathcal{C}^{\beta + \beta/\alpha}_b} (d_{\mathcal{D}^{\beta}}(f, \tilde{f}) + \| u - \tilde{u} \|_\alpha).
\]

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

\[
\| F(f) \|_{\nu, \beta} \lesssim \| F(f) - \pi_\nu(DF(f), f) \|_{\nu, \beta} + \| \pi_\nu(DF(f), f^\nu) \|_{\nu, \beta} + \| \pi_\nu(DF(f), f^\nu) \|_{\alpha, \beta} \lesssim \| F \|_{\mathcal{C}^{\beta + \beta/\alpha}_b} (1 + \| v \|_\alpha)^{1+\beta/\alpha}(1 + \| f^\nu \|_{\nu, \beta}) \lesssim \| F \|_{\mathcal{C}^{\beta + \beta/\alpha}_b} (1 + \| f^\nu \|_\infty)^{1+\beta/\alpha}(1 + \| f^\nu \|_{\nu, \beta}).
\]

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 $\| f \|_{\nu, \beta}$. The superlinear dependence on $\| f^\nu \|_\infty$ will not pose any problem as we will always have $f^\nu = F(\tilde{f})$ for some suitable $\tilde{f}$, so that for bounded $F$ we get $\| F(f) \|_{\nu, \beta} \lesssim F, \nu, 1 + \| f \|_{\nu, \beta}$.

\subsection{4.2 A basic commutator estimate}

Here we prove the commutator estimate which will be the main ingredient in the construction of the integral $I(f, dg)$, where $f$ is paracontrolled by $v$ and $g$ is paracontrolled by $w$, and where we assume that the integral $I(v, dw)$ exists.
Proposition 4.7. Let $\alpha, \beta, \gamma \in (0, 1)$, and assume that $\alpha + \beta + \gamma > 1$ and $\beta + \gamma < 1$. Let $f \in C^\alpha$, $v \in C^\beta$, and $w \in C^\gamma$. Then the "commutator"

$$C(f, v, w) := L(\pi_<(f, v), w) - I(f, dL(v, w))$$

$$= \lim_{N \to \infty} [L(S_N(\pi_<(f, v)), S_Nw) - I(f, dL(S_Nv, S_Nw))]$$

$$= \lim_{N \to \infty} \sum_{p \leq N} \sum_{q < p} \left[ \int_0^1 \Delta_p(\pi_<(f, v))(s)d\Delta_qw(s) - \int_0^1 d(\Delta_q(\pi_<(f, v)))(s)\Delta_pw(s) \right.$$

$$\left. - \int_0^1 f(s)\Delta_pv(s)d\Delta_qw(s) - \int_0^1 f(s)d(\Delta_qv(s)\Delta_pw(s)) \right]$$

converges in $C^{\alpha+\beta+\gamma-\varepsilon}$ 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 (16), i.e. for

$$X_N := \sum_{p \leq N} \sum_{q < p} \left[ \int_0^1 \Delta_p(\pi_<(f, v))(s)d\Delta_qw(s) - \int_0^1 f(s)\Delta_pv(s)d\Delta_qw(s) \right].$$

(17)

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 $C^{\alpha+\beta+\gamma}$ are relatively compact in $C^{\alpha+\beta+\gamma-\varepsilon}$.

To prove uniform convergence, note that

$$X_N - X_{N-1} = \sum_{q < N} \left[ \int_0^1 \Delta_N(\pi_<(f, v))(s)d\Delta_qw(s) - \int_0^1 f(s)\Delta_Nv(s)d\Delta_qw(s) \right]$$

$$= \sum_{q < N} \sum_{j \leq N} \sum_{i < j} \int_0^1 \Delta_N(\Delta_if\Delta_jv)(s)d\Delta_qw(s)$$

$$- \sum_{j \geq N} \sum_{i < j} \int_0^1 \Delta_j(\Delta_if\Delta_Nv)(s)d\Delta_qw(s),$$

(18)

where for the second term it is possible to take the infinite sum over $j$ outside of the integral because $\sum_j \Delta_jg$ converges uniformly to $g$ and because $\Delta_qw$ is a finite variation path. We also used that $\Delta_N(\Delta_if\Delta_jv) = 0$ whenever $i > N$ or $j > N$. Only very few terms in (18) cancel. Nonetheless these cancellations are crucial, since they eliminate most terms for which we only have the worse estimate (14) in Corollary 3.11. We obtain

$$X_N - X_{N-1} = \sum_{q < N} \sum_{j < N} \sum_{i < j} \int_0^1 \Delta_N(\Delta_if\Delta_jv)(s)d\Delta_qw(s) - \sum_{q < N} \int_0^1 \Delta_N(\Delta_Nf\Delta_Nv)(s)d\Delta_qw(s)$$

$$- \sum_{q < N} \sum_{j > N} \sum_{i < j} \int_0^1 \Delta_j(\Delta_if\Delta_Nv)(s)d\Delta_qw(s)$$

$$- \sum_{q < N} \sum_{j > N} \int_0^1 \Delta_j(\Delta_f\Delta_Nv)(s)d\Delta_qw(s).$$

(19)
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 (13) for the first three terms and (14) for the fourth term, yields

\[
\| X_N - X_{N-1} \|_\infty \lesssim \| f \|_\alpha \| v \|_\beta \| w \|_\gamma \left[ \sum_{q < N} \sum_{j < N} \sum_{i < j} 2^{-2N+i+j-2}\gamma \right. \\
+ \sum_{q < N} 2^{-N(\alpha+\beta)}2^{1-2q(1-\gamma)} + \sum_{q < N} \sum_{j > N} \sum_{i < j} 2^{-2j+i+N-2}\gamma \left. \right] \\
+ \sum_{q < N} \sum_{j > N} 2^{-j\alpha}2^{-N}\gamma \]

\[
\lesssim \| f \|_\alpha \| v \|_\beta \| w \|_\gamma 2^{-N(\alpha+\beta+1)},
\]

(20)

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 (19) we obtain for \( n \in \mathbb{N} \)

\[
\Delta_n X_N = \sum_{p \leq N} \sum_{q < p} \Delta_n \left[ \sum_{j < p} \sum_{i < j} \int_0^s \Delta_p(\Delta_i f \Delta_j v)(s) d\Delta_q w(s) - \int_0^s \Delta_p(\Delta_i f \Delta_j v)(s) d\Delta_q w(s) \\
- \sum_{j > p} \sum_{i < j} \int_0^s \Delta_j(\Delta_i f \Delta_p v)(s) d\Delta_q w(s) \right].
\]

and therefore by Corollary 3.10

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

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

\[
\| \Delta_n X_N \|_\infty \lesssim \| f \|_\alpha \| v \|_\beta \| w \|_\gamma \sum_p 2^{p(1-\gamma)} \left[ \sum_{j < p} \sum_{i < j} 2^{-(n\gamma p) - n + p - 2+2j(1-\gamma)} \\
+ 2^{-(n\gamma p) - n + p - 2+2j(1-\gamma)} \\
+ \sum_{j > p} \sum_{i < j} 2^{-(n\gamma j) - n + j - 2j+2j(1-\alpha) + p(1-\gamma)} \\
+ \sum_{j > p} 2^{-(n\gamma j) - n + j - 2j(1-\gamma)} \right] \\
\lesssim \| f \|_\alpha \| v \|_\beta \| w \|_\gamma 2^{-n(\alpha+\beta+\gamma)},
\]

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 $C(f, v, w) \in C^{\alpha+\beta+\gamma-\varepsilon}$ 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

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

Proof. Simply sum up (20) over $N$. \hfill \Box

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 $2\alpha + \beta > 0$ as well as $\alpha + \beta \neq 1$. Let $v \in C^\alpha(\mathbb{R}^d)$ and assume that the Lévy area

$$L(v) := \lim_{N \to \infty} \left( L(S_Nv^k, S_Nv^f) \right)_{1 \leq k \leq d, 1 \leq f \leq d}$$

converges uniformly and that $\sup_N\|L(S_Nv, S_Nv)\|_{2\alpha} < \infty$. Let $f \in \mathcal{D}_v^\beta(\mathcal{L}(\mathbb{R}^d, \mathbb{R}^m))$. Then $I(S_Nf, dS_Nv)$ converges in $C^{\alpha-\varepsilon}$ for all $\varepsilon > 0$. Denoting the limit by $I(f, dv)$, we have

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

Moreover, $I(f, dv) \in \mathcal{D}_v^\alpha$ with derivative $f$ and

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

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

$$I(S_Nf, dS_Nv) = S(S_Nf, S_Nv) + \pi_<(S_Nf, S_Nv) + L(S_Nf^*, S_Nv) + [L(S_Nf, S_Nv) - I(f^*, dL(S_Nv, S_Nv))] + I(f^*, dL(S_Nv, S_Nv)).$$

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

$$I(f, dv) = S(f, v) + \pi_<(f, v) + L(f^2, v) + C(f^v, v, v) + I(f^v, dL(v, v)),$$

from which we easily deduce the claimed bounds. \hfill \Box

Remark 4.11. Since $I(f, dv) = \lim_{N \to \infty} \int_0^1 S_Nf dS_Nv$, the integral is a local operator in the sense that $I(f, dv)$ is constant on every interval $[s, t]$ for which $f|[s, t] = 0$. In particular we can estimate $I(f, dv)|_{[0, t]}$ using only $f|_{[0, t]}$ and $f^v|_{[0, t]}$.

For fixed $v$ and $L(v, v)$, the map $f \mapsto I(f, dv)$ is linear and bounded from $\mathcal{D}_v^\beta$ to $\mathcal{D}_v^\alpha$, and this is what we will need to solve differential equations driven by $v$. But we can also estimate the speed of convergence of $I(S_Nf, dS_Nv)$ to $I(f, dv)$, measured in uniform distance rather than in $C^\alpha$.
Lemma 4.9 gives

But now \( \beta \parallel \) for all \( \varepsilon > 0 \). From here it is easy to see that

\[
\| \pi_<(S_N f - f, S_N v - v) \|_\infty \lesssim \| S_N f - f \|_\infty \| v \|_\alpha + \| f \|_\infty \| f \|_\alpha \| S_N v - v \|_\varepsilon
\]

for all \( \varepsilon > 0 \). From Proposition 4.14.

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

\[
\| \pi_<(S_N f - f, S_N v) + \pi_<(f, S_N v - v) \|_\infty \lesssim \| S_N f - f \|_\infty \| v \|_\alpha + \| f \|_\infty \| f \|_\alpha \| S_N v - v \|_\varepsilon
\]

But now \( \beta \leq \alpha \leq 1/2 \) and therefore \( \alpha \geq 2\alpha + \beta - 1 \).

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

\[
\| L(S_N f^2, S_N v) - L(f^2, v) \|_\infty \lesssim \| L(S_N f^2 - f^2, S_N v) \|_{1+\varepsilon} + \| L(f^2, S_N v - v) \|_{1+\varepsilon}
\]

\[
\lesssim \| S_N f^2 - f^2 \|_{1+\varepsilon} \| v \|_\alpha + \| f^2 \|_{1+\varepsilon} \| S_N v - v \|_{1+\varepsilon} \| v \|_\alpha
\]

\[
\lesssim 2^{-N(\alpha+\beta-1+\varepsilon)} \| f^2 \|_{1+\varepsilon} \| v \|_\alpha + 2^{-N(\alpha-1+\varepsilon-\alpha-\beta)} \| f^2 \|_{1+\varepsilon} \| v \|_\alpha
\]

Lemma 4.9 gives

\[
\| L(S_N \pi_<(f^v, v), S_N v) - L(\pi_<(f^v, v), v) \|_\infty \lesssim 2^{-N(2\alpha+\beta-1)} \| f^v \|_\beta \| v \|_\alpha^2
\]

\[
+ \| I(f^v, dL(S_N v, S_N v)) - I(f^v, dL(v, v)) \|_\infty.
\]

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

Remark 4.13. In Lemma 4.9 we saw that the rate of convergence of

\[
L(S_N \pi_<(f^v, v), S_N v) - I(f^v, dL(S_N v, S_N v)) - (L(\pi_<(f^v, v), v) - I(f^v, dL(v, v)))
\]

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

\[
\| I(S_N f, dS_N v) - I(f, dv) \|_\infty \lesssim 2^{-N(2\alpha+\beta-1)} \| f \|_\beta \| v \|_\alpha^2 + \| f^v \|_\beta \| L(S_N v, S_N w) - L(v, w) \|_{2\alpha-\varepsilon}.
\]

Since this requires a rather lengthy calculation, 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 \( I(f, dg) \), it is not surprising that we obtain a Stratonovich type integral:

Propostion 4.14. Let \( \alpha \in (1/3, 1/2) \) and \( v \in C^\alpha(\mathbb{R}^d) \). Let \( \varepsilon > 0 \) be such that \( (2+\varepsilon)\alpha > 1 \) and let \( F \in C^{2+\varepsilon}(\mathbb{R}^d, \mathbb{R}) \). Then

\[
F(v(t)) - F(v(0)) = I(DF(v), dv)(t) := \lim_{N \to \infty} I(S_N DF(v), dS_N v)(t)
\]

for all \( t \in [0, 1] \).
Proof. The function $S_Nv$ is Lipschitz continuous, so that integration by parts gives

$$F(S_Nv(t)) - F(S_Nv(0)) = I(DF(S_Nv), dS_Nv)(t).$$

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

$$|[S_NDF(v) - DF(S_Nv)]_{pm}| \leq C_F 2^{-p\alpha(1+\epsilon)}\|S_Nv\|_{\alpha}$$

for a constant $C_F > 0$. Therefore, we get for all $\epsilon' \leq \epsilon$

$$\|S_NDF(v) - DF(S_Nv)\|_{\alpha(1+\epsilon')} \lesssim F 2^{-N\alpha(\epsilon-\epsilon')}\|v\|_{\alpha},$$

which completes the proof. \qed

Remark 4.15. 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 $DF$, the Hessian of $F$.

Proposition 4.14 was previously obtained by Roynette [Roy93], except that there $v$ is assumed to be one dimensional and in the Besov space $B_{1,\infty}^{1/2}$.

5 Pathwise Itô integration

In the previous section we saw that our pathwise integral $I(f,dv)$ 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 $I(f,dv)$ 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

$$I_k^{\text{Itô}}(f,dv)(t) := \sum_{m=1}^{2^k} f(t_{km}^0)(v(t_{km}^0 \wedge t) - v(t_{km}^0 \wedge t))$$

$$= \sum_{m=1}^{2^k} \sum_{p,q} \sum_{m,n} f_{pm} v_{qn} \varphi_{pm}(t_{km}^0)(\varphi_{qn}(t_{km}^0 \wedge t) - \varphi_{qn}(t_{km}^0 \wedge t)).$$

Let us assume for the moment that $t = m2^{-k}$ for some $0 \leq m \leq 2^k$. In that case we obtain for $p \geq k$ or $q \geq k$ that $\varphi_{pm}(t_{km}^0)(\varphi_{qn}(t_{km}^0 \wedge t) - \varphi_{qn}(t_{km}^0 \wedge t)) = 0$. For $p,q < k$, both $\varphi_{pm}$ and $\varphi_{qn}$ are affine functions on $[t_{km}^0 \wedge t, t_{km}^0 \wedge t]$, and for affine $u$ and $w$ and $s < t$ we have

$$u(s)(w(t) - w(s)) = \int_s^t u(r)dw(r) - \frac{1}{2}[u(t) - u(s)][w(t) - w(s)].$$

22
Hence, we conclude that for \( t = m2^{-k} \)
\[
I_{k}^{t\beta}(f, dv)(t) = I(S_{k-1}f, dS_{k-1}v)(t) - \frac{1}{2}[f, v]_{k}(t),
\]
where \([f, v]_{k}\) is the \( k\)-th dyadic approximation of the quadratic covariation \([f, v]\), i.e.
\[
[f, v]_{k}(t) := \sum_{m=1}^{2^{k}} [f(t_{km}^{2} \wedge t) - f(t_{km}^{1} \wedge t)][v(t_{km}^{2} \wedge t) - v(t_{km}^{1} \wedge t)].
\]
From now on we study the right hand side of (21) rather than \( I_{k}^{t\beta}(f, dv) \), which is justified by the following remark.

**Remark 5.1.** Let \( \alpha \in (0, 1) \) and \( f \in C([0, 1]) \) and \( v \in C^{\alpha} \), then
\[
\left\| I_{k}^{t\beta}(f, dv) - \left( I(S_{k-1}f, dS_{k-1}v) - \frac{1}{2}[S_{k-1}f, S_{k-1}v]_{k}\right) \right\|_{\infty} \lesssim 2^{-k\alpha} \| f \|_{\infty} \| v \|_{\alpha}.
\]
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 \([v, v] := ([v^{i}, v^{j}])_{1\leq i,j\leq d} \) and \( L(v, v) := (L(v^{i}, v^{j}))_{1\leq i,j\leq d} \), and similarly for all expressions of the same type.

**Theorem 5.2.** Let \( \alpha \in (0, 1/2) \) and let \( \beta \leq \alpha \) be such that \( 2\alpha + \beta > 1 \). Let \( v \in C^{\alpha}(\mathbb{R}^{d}) \) and \( f \in \mathcal{D}_{0}^{\alpha}(\mathcal{L}(\mathbb{R}^{d}; \mathbb{R}^{n})) \). Assume that \( (L(S_{k}v, S_{k}v)) \) converges uniformly, with uniformly bounded \( C^{2\alpha} \) norm. Also assume that \((v, v)_{k}\) converges uniformly. Then \((I_{k}^{t\beta}(f, dv))\) converges uniformly to a limit \( I^{t\beta}(f, dv) = I(f, dv) - 1/2[f, v]\) which satisfies
\[
\| I^{t\beta}(f, dv) \|_{\infty} \lesssim \| f \|_{\alpha,\beta}(\| v \|_{\alpha} + \| v \|_{\alpha}^{2} + \| L(v, v) \|_{2\alpha} + \| [v, v] \|_{\infty}),
\]
and where the quadratic variation \([f, v]\) is given by
\[
[f, v] = \int_{0}^{t} f^{v}(s)dv(s) := \left( \sum_{j,\ell=1}^{d} \int_{0}^{t} (f^{ij})^{v,\ell}(s)dv(s)dv^{\ell}(s) \right)_{1\leq i\leq n}, \quad (22)
\]
where \((f^{ij})^{v,\ell}\) is the \( \ell\)-th component of the \( v\)-derivative of \( f^{ij} \). For \( \varepsilon \in (0, 3\alpha - 1) \) the speed of convergence can be estimated by
\[
\| I^{t\beta}(f, dv) - I_{k}^{t\beta}(f, dv) \|_{\infty} \lesssim \varepsilon 2^{-k(2\alpha + \beta - 1 - \varepsilon)} \| f \|_{\alpha,\beta}(\| v \|_{\alpha} + \| v \|_{\alpha}^{2})
\]
\[
+ \| f \|_{\beta} \| L(S_{k-1}v, S_{k-1}v) - L(v, v) \|_{2\alpha}
\]
\[
+ \| f \|_{\infty}\| [v, v]_{k} - [v, v] \|_{\infty}.
\]

**Proof.** By Remark 5.1, it suffices to show our claims for \( I(S_{k-1}f, dS_{k-1}v) - 1/2[f, v]_{k} \). The statements for the integral \( I(S_{k-1}f, dS_{k-1}g) \) follow from Theorem 4.10 and Corollary 4.12. So let us us concentrate on the quadratic variation \([f, v]_{k}\). Recall from Example 4.4 that \( f \in \mathcal{D}_{0}^{\alpha}(\mathcal{L}(\mathbb{R}^{d}; \mathbb{R}^{n})) \) if and only if \( R_{s,t}^{f} = f_{s,t} - f^{v}(s)w_{s,t} \) satisfies \( |R_{s,t}^{f}| \lesssim |t - s|^{\alpha + \beta} \). Hence
\[
[f, v]_{k}(t) = \sum_{m} \left( f_{0k^{2}m,t}^{0k^{2}m,t}v_{0k^{2}m,t}^{0k^{2}m,t} \right)_{i}^{i}
\]
\[
= \sum_{m} \left( R_{0k^{2}m,t}^{i}v_{0k^{2}m,t}^{i} \right)_{i}^{i} + \sum_{j,\ell=1}^{d} \sum_{m} \left( (f^{ij})^{v,\ell}v_{0k^{2}m,t}^{i} \right)_{i}^{i} + \sum_{j,\ell=1}^{d} \sum_{m} \left( (f^{ij})^{v,\ell}v_{0k^{2}m,t}^{i} \right)_{i}^{i}.
\]

23
It is easy to see that the first term on the right hand side is bounded by
\[
\left| \sum_m \left( R^f_{i,j,m}(t^2_{j,m},\ell^2_{j,m}) \int^1_0 (f^{i,j})_{\ell}(s) \right) \right| \lesssim 2^{-(2n+\beta-1)} \| f \|_{v,\beta} (\| v \|_\alpha + \| v \|_\alpha^2).
\]

For the second term, let us fix \( \ell \) and \( j \). Then the sum over \( m \) is just the integral of \((f^{i,j})_{\ell}\) with respect to the signed measure \( \mu^k_{\ell} = \sum_m \delta^\ell_{m,j_m}(v^i_{j_m}^m \land t^2_{j_m,m} \land t^2_{j_m,m}) \). Decomposing \( \mu^k_{\ell} \) into a positive and negative part as
\[
\mu^k_{\ell} = \frac{1}{4} \left\| \sum_m \int_0^1 \delta^\ell_{m,j_m}[(v^i + v^j)]_{j_m}^m \land t^2_{j_m,m} \land t^2_{j_m,m} \right\|^2 - \sum_m \int_0^1 \delta^\ell_{m,j_m}[(v^i - v^j)]_{j_m}^m \land t^2_{j_m,m} \land t^2_{j_m,m} \right\|^2
\]
and similarly for \( d\mu_t = d[v^i, v^j] \), we can estimate
\[
\left| \int_0^1 (f^{i,j})_{\ell}(s) \mu^k_{\ell}(ds) - \int_0^1 (f^{i,j})_{\ell}(s) \mu_t(ds) \right| \lesssim \| f \|_{v,\alpha} \left( \| [v^i + v^j] \|_\alpha + \| [v^i - v^j] \|_\alpha \right) \lesssim \| f \|_{v,\alpha} \| [v, v] \|_\alpha.
\]

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 \([f, g]_k\).

**Remark 5.3.** We calculate the pathwise Itô integral \( I^{\text{Itô}}(f, dv) \) as limit of nonanticipating Riemann sums involving only \( f \) and \( v \). This is interesting for applications of mathematical finance, because the integral process has a natural interpretation as capital obtained from investing. The classical rough path integral, see Proposition 2.4, is obtained via “compensated Riemann sums” that explicitly depend on \( f^v \) and \( I^{\text{Itô}}(v, dv) \).

**Remark 5.4.** We calculate the pathwise Itô integral \( I^{\text{Itô}}(f, dv) \) as limit of nonanticipating Riemann sums involving only \( f \) and \( v \). The classical rough path integral, see Proposition 2.4, is obtained via “compensated Riemann sums” that depend explicitly on the derivative \( f^v \) and the iterated integrals of \( v \). For applications in mathematical finance, it is more convenient to have an integral that is the limit of nonanticipating Riemann sums, because this can be interpreted as capital process obtained from investing.

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

**Corollary 5.5.** In addition to the conditions of Theorem 5.2, assume that also \([v, v]_k \in C^{2\alpha} \). Then \( I^{\text{Itô}}(f, dv) \in D^{\alpha}_v \) with derivative \( f \), and
\[
\| I^{\text{Itô}}(f, dv) \|_{v,\alpha} \lesssim \| f \|_{v,\beta} (1 + \| v \|_\alpha^2 + \| L(v, v) \|_{2\alpha} + \| [v, v] \|_{2\alpha}).
\]

**Proof.** This is a combination of Theorem 4.10 and the explicit representation (22) together with the continuity of the Young integral, Theorem 3.14.

The term \( I(S_{k-1}f, dS_{k-1}v) \) 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 \( I(S_{k-1}f, d\Delta v) + I(\Delta f, dS_{k}v) \). For the quadratic variation \([f, v]_k\), this is not exactly true. But \([f, v]_k(m^{-2k}) = [S_{k-1}f, S_{k-1}v]_k(m^{-2k}) \) for \( m = 0, \ldots, 2^k \), and there is a recursive way of calculating \([S_{k-1}f, S_{k-1}v]_k\).
Lemma 5.6. Let \( f, v \in C([0, 1], \mathbb{R}) \). Then
\[
[S_k f, S_k v]_{k+1}(t) = \frac{1}{2} [S_{k-1} f, S_{k-1} v](t) + [S_{k-1} f, \Delta_k v]_{k+1}(t) + [\Delta_k f, S_k v]_{k+1}(t) + R_k(t)
\]
for all \( k \geq 1 \) and all \( t \in [0, 1] \), where
\[
R_k(t) := -\frac{1}{2} f(t_k^+, t_{k-1}^+) + \frac{1}{2} \sum_{m} f_m v_m
\]
and \( t_k^+ := \lceil t^k \rceil 2^{-k} \) and \( \tau_k^+ := \lceil t^k \rceil + 2^{-(k+1)} \). In particular, we obtain for \( t = 1 \) that
\[
[f, v]_{k+1}(1) = \frac{1}{2} [f, v](1) + \frac{1}{2} \sum_{m} f_m v_m = \frac{1}{2^{k+1}} \sum_{p \leq k} \sum_{m} 2^p f_m v_{pm}.
\]
If moreover \( \alpha \in (0, 1) \) and \( f, v \in C^\alpha \), then \( \| [S_k f, S_k v]_{k+1} - [f, g]_k \|_\infty \lesssim 2^{-2k\alpha} \| f \|_\alpha \| g \|_\alpha \).

Proof. Equation (23) follows from a direct calculation using the fact that \( S_{k-1} f \) and \( S_{k-1} v \) are affine on every interval \([t^k, t_{k-1}^+]\) respectively \([t^k, t_{k-1}^-] \) for \( 1 \leq k \leq 2^k \). The formula for \( [f, v]_{k+1}(1) \) follows from the that \( [\Delta_k f, \Delta_k v]_{k+1}(1) = 0 \) unless \( p = q \), and that \( [\Delta_k f, \Delta_k v]_{k+1} = 1/2 \sum f_m v_m \). The estimate for \( \| [S_k f, S_k v]_{k+1} - [f, g]_k \|_\infty \) holds because the two functions agree in all dyadic points \( m2^{-k} \).

Remark 5.7. The Cesàro mean formula (24) makes the study of existence of the quadratic variation accessible to ergodic theory. This was previously observed by Gantert [Gan91], 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 convergence of \( (v, v)_k(1) \) as \( k \) tends to \( \infty \).

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 5.8. Let \( \alpha \in (0, 1/2) \) and let \( v \in C^\alpha(\mathbb{R}^d) \). Assume that the non-anticipating Riemann sums \( I_k^{1\alpha}(v, dv) \) converge uniformly to \( I^{1\alpha}(v, dv) \). Then also \( (v, v)_k \) converges uniformly to a limit \( [v, v] \). If moreover
\[
\sup_k \sup_{0 \leq m < m' \leq 2^k} \frac{|I_k^{1\alpha}(v, dv)(m'2^{-k}) - I_k^{1\alpha}(v, dv)(m2^{-k}) - v(m2^{-k})v(m'2^{-k}) - v(m2^{-k})v(m'2^{-k})|}{(m' - m)2^{-k}2^{2\alpha}} < C < \infty,
\]
then \( [v, v] \in C^{2\alpha} \) and \( \| [v, v] \|_{2\alpha} \lesssim C + \| v \|_{\alpha}^2 \).

Proof. Let \( t \in [0, 1] \) and \( 1 \leq i, j \leq d \). Then
\[
v^j(t)v^j(t) - v^j(0)v^j(0) = \sum_{m=1}^{2^k} [v^j(t^j_{km} \wedge t)v^j(t^j_{km} \wedge t) - v^j(t^j_{km} \wedge t)v^j(t^j_{km} \wedge t)]
\]
\[
= \sum_{m=1}^{2^k} \left[ v^j(t^j_{km})v^j(t^j_{km} \wedge t) + v^j(t^j_{km})v^j(t^j_{km} \wedge t) + v^j(t^j_{km} \wedge t)v^j(t^j_{km} \wedge t) \right]
\]
\[
= I_k^{1\alpha}(v^j, dv^j)(t) + I_k^{1\alpha}(v^j, dv^j)(t) + [v^j, v^j]_k(t),
\]
25
which implies the convergence of \([v, v]_k\) as \(k\) tends to \(\infty\). For \(0 \leq s < t \leq 1\) this gives
\[
([v', v']_s, t) = (v'v')_{s, t} - I^\Pi_0(v', dv')_{s, t} - I^\Pi_0(v', dv')_{s, t} = \left[ v'(s)v'_{s, t} - I^\Pi_0(v', dv')_{s, t} \right] + v'_{s, t}v'_{s, t},
\]
At this point it is easy to estimate \(\|[v, v]_\alpha\|^2\), where we work with the classical Hölder norm and not the \(C^{2\alpha}\) norm. Indeed let \(0 \leq s < t \leq 1\). Using the continuity of \([v, v]\), we can find \(k\) and \(s \leq k = m_s2^{-k} < m_t2^{-k} = t_k \leq t\) with \(\|[v, v]_s, s_k\| + \|[v, v]_{t_k, t}\| \leq \|[v]_\alpha^2|t - s|^{2\alpha}\). Moreover,
\[
\|[v, v]_s, t_k\| \leq \left( \sup_{\ell \geq k} \sup_{0 \leq m \leq 2^\ell} \frac{|(v, v)|_{m2^{-\ell}, m2^{-\ell} + 2\ell}}{|(m' - m)2^{-\ell}} \right) t_k - |s - t|^{2\alpha} \leq (2C + \|[v]_\alpha^2|t - s|^{2\alpha}.
\]

**Remark 5.9.** The “coarse-grained Hölder condition” (25) is from [PP13] and has recently been discovered independently by [Kel14].

Similarly convergence of \((I^\Pi_0(v, dv))\) implies convergence of \((L(S_kv, S_kv))_k\):

**Lemma 5.10.** In the setting of Lemma 5.8, assume that (25) holds. Then \((L(S_kv, S_kv))_k\) converges uniformly as \(k\) tends to \(\infty\), and
\[
\sup_k \|L(S_kv, S_kv)\|_\alpha \lesssim C + \|[v]_\alpha^2.
\]

**Proof.** Let \(k \in \mathbb{N}\) and \(0 \leq m \leq 2^k\), and write \(t = m2^{-k}\). Then we obtain from (21) that
\[
L(S_k-1v, S_k-1v)(t) = I^\Pi_0(v, dv)(t) + \frac{1}{2}[v, v]_k(t) - \pi_<(S_k-1v, S_k-1v)(t) - S(S_k-1v, S_k-1v)(t).
\]
Let now \(s, t \in [0, 1]\). We first assume that there exists \(m\) such that \(t^0_{km} \leq s < t \leq t^0_{km}\). Then we use \(\|\partial_t \Delta_q v\|_\alpha \lesssim 2^{q(1-\alpha)}\|v\|_\alpha\) to obtain
\[
|L(S_k-1v, S_k-1v)|_{s, t} \leq \sum_{p < q < p} \int_s^t \Delta_q v(r) d\Delta_q v(r) - \int_s^t d\Delta_q v(r) \Delta_q v(r) \leq \sum_{p < q < p} |t - s| 2^{-p\alpha} 2^{q(1-\alpha)} \|v\|_\alpha^2 \lesssim |t - s| 2^{-k(2\alpha - 1)} \|v\|_\alpha^2 \leq |t - s| 2\alpha \|v\|_\alpha^2.
\]
Combining (26) and (27), we obtain the uniform convergence of \((L(S_k-1v, S_k-1v))_k\) from Lemma 5.8 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 \(r^{k\gamma} = m_s2^{-k}\) and \(r^{k\gamma} = m_t2^{-k}\) be such that \(r^{k\gamma} - 2^{-k} < s \leq r^{k\gamma}\) and \(t^{k\gamma} = t < t^{k\gamma} + 2^{-k}\). In particular, \(r^{k\gamma} \leq t^{k\gamma}\). Moreover
\[
|L(S_k-1v, S_k-1v)|_{s, t} \leq |L(S_k-1v, S_k-1v)_{s, r^{k\gamma}}| + |L(S_k-1v, S_k-1v)_{r^{k\gamma}, t^{k\gamma}}| + |L(S_k-1v, S_k-1v)_{t^{k\gamma}, t}|
\]
Using (27), the first and third term on the right hand side can be estimated by $(|r s^k - s|^{2\alpha} + |t - r t^k|^{2\alpha})||v||^2_\alpha \lesssim |t - s|^{2\alpha} ||v||^2_\alpha$. For the middle term we apply (26) to obtain

$$
|L(S_{k-1}v, S_{k-1}v)_{r, s^k, t^k}| \leq |I^{\text{Itô}}(v, dv)_{r, s^k, t^k} | - v(r s^k) - v(r s^k) - \pi_c(S_{k-1}v, S_{k-1}v)_{r, s^k, t^k} + \frac{1}{2} |([v, v]_k)_{r, s^k, t^k}| + |S(S_{k-1}v, S_{k-1}v)_{r, s^k, t^k}|
$$

$$
\lesssim \|v\|_\alpha^2 \left( C + \|v\|_\alpha^2 \right) \leq |t - s|^{2\alpha} \left( C + \|v\|_\alpha^2 \right),
$$

where Example 4.4, Lemma 5.8, and Lemma 3.13 have been used.

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

**Corollary 5.11.** Let $\alpha \in (1/3, 1/2)$ and $v \in C^\alpha(\mathbb{R}^d)$. Assume that the non-anticipating Riemann sums $(I_k^{\text{Itô}}(v, dv))_k$ converge uniformly to $I^{\text{Itô}}(v, dv)$ and let $F \in C^2(\mathbb{R}^d, \mathbb{R})$. Then $(I_k^{\text{Itô}}(DF(v), dv))_k$ converges to a limit $I^{\text{Itô}}(DF(v), dv)$ that satisfies for all $t \in [0, 1]$

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

**Proof.** This is Remarque 1 of Föllmer [Föl79] in combination with Lemma 5.8.

---

### 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 integrals for some continuous martingales.

#### 6.1 Hypercontractive processes

Let $X_t : [0, 1] \to \mathbb{R}^d$ 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, $R : [0, 1]^2 \to \mathbb{R}^{d \times d}$ and $R(s, t) := (E(X_i X_j))_{1 \leq i, j \leq d}$. The increment of $R$ over a rectangle $[s, t] \times [u, v] \subseteq [0, 1]^2$ is defined as

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

Let us make the following two assumptions.

(\(\rho\text{-var}\)) There exists $C > 0$ such that for all $0 \leq s < t \leq 1$ and for every partition $s = t_0 < t_1 < \cdots < t_n = t$ of $[s, t]$ we have

$$
\sum_{i, j = 1}^n |R_{[t_{i-1}, t_i] \times [t_{j-1}, t_j]}|^\rho \leq C|t - s|.
$$
The process $X$ is hypercontractive, i.e. for every $m,n \in \mathbb{N}$ and every $r \geq 1$ there exists $C_{r,m,n} > 0$ such that for every polynomial $P : \mathbb{R}^n \to \mathbb{R}$ of degree $m$, for all $i_1, \ldots, i_n \in \{1, \ldots, d\}$, and for all $t_1, \ldots, t_n \in [0,1]$

$$E(|P(X_{t_1}^{i_1}, \ldots, X_{t_n}^{i_n})|^{2r}) \leq C_{r,m,n}E(|P(X_{t_1}^{i_1}, \ldots, X_{t_n}^{i_n})|^2)^r.$$

These conditions are taken from [FV10a], where under even more general assumptions it is shown that it is possible to construct the iterated integrals $I(X, dX)$, and that $I(X, dX)$ is the limit of $(I(X^n, dX^n))_{n \in \mathbb{N}}$ 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 [Jan97], 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 $k \in \mathbb{N}$, $k \geq 1$

$$Z^{k,H}_t = C(H, k) \int_{\mathbb{R}^k} \left( \int_0^t \prod_{i=1}^k (s - y_i)^{-H} \frac{1}{t^{H+k}} ds \right) dB_{y_1} \cdots dB_{y_k},$$

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

$$E(Z^{k,H}_s Z^{k,H}_t) = \frac{1}{2} \left( t^{2H} + s^{2H} + |t-s|^{2H} \right).$$

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 [PT11]. However, we should point out that it follows from Kolmogorov’s continuity criterion that $Z^{k,H}$ is 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 ($\rho$-var) is satisfied by 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 ($\rho$-var) is the one of Coutin and Qian [CQ02]: If $E(|X_{s,t}^{i}|^2) \lesssim |t-s|^{2H}$ and $|E(X_{s,t}^{i} X_{s,t}^{j})| \lesssim |t-s|^{2H-2}h^2$ for $i = 1, \ldots, d$, then ($\rho$-var) is satisfied for $\rho = 1/(2H)$. For details and further examples see [FV10b], Section 15.2.

**Lemma 6.3.** Assume that the stochastic process $X : [0, 1] \to \mathbb{R}$ satisfies ($\rho$-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

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

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

$$E(X_{p m_1} X_{p m_2}) = E \left( (X_{p m_1}^{i_1, m_1} - X_{p m_1}^{i_1, m_1})(X_{p m_2}^{i_2, m_2} - X_{p m_2}^{i_2, m_2}) \right)$$

$$= \sum_{i_1, i_2 = 0}^1 (-1)^{i_1 + i_2} R_{(i_1, i_1 + 1) \times (i_2, i_2 + 1)},$$

and that $\{i_m : i = 0, 1, 2, m = M, \ldots, N\}$ partitions the interval $[(M - 1)2^{-p}, N2^{-p}]$. 

28
Lemma 6.4. Let \( X, Y : [0, 1] \to \mathbb{R} \) be independent, centered, continuous processes, both satisfying \( (p, \var) \)-var for some \( \rho \in [1, 2] \). Then for all \( i, p \geq -1, q < p, \) and \( 0 \leq j \leq 2^q \)

\[
E \left[ \left| \sum_{m \leq 2^p} \sum_{n \leq 2^q} X_{pm} Y_{qn} \langle 2^{-i} \chi_{ij}, \varphi_{pm} \chi_{qn} \rangle \right|^2 \right] \lesssim 2^{(p \var + (1/\rho) - 2q)} 2^{-2q(q(1 - 1/\rho))}. \]

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

\[
E \left[ \left| \sum_{m \leq 2^p} \sum_{n \leq 2^q} X_{pm} Y_{qn} \langle 2^{-i} \chi_{ij}, \varphi_{pm} \chi_{qn} \rangle \right|^2 \right] \lesssim \sum_{m_1, m_2 = 0}^{2^p} \left| E(X_{pm_1} Y_{m_1}) E(Y_{qn(m_1)} Y_{qn(m_2)}) \langle 2^{-i} \chi_{ij}, \varphi_{pm_1} \chi_{qn(m_1)} \rangle \langle 2^{-i} \chi_{ij}, \varphi_{pm_2} \chi_{qn(m_2)} \rangle \right|.
\]

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

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

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

\[
\left( \sum_{m_1, m_2 \in M_j} \left| E(Y_{qn(m_1)} Y_{qn(m_2)}) \right|^{\rho'} \right)^{1/\rho'} \lesssim \left( 2^{(p - q) \rho'} \right)^{1/\rho'} \left( \max_{n_1, n_2 \in N_j} \left| E(Y_{qn_1} Y_{qn_2}) \right| \right)^{\rho' - \rho} \sum_{n_1, n_2 \in N_j} \left| E(Y_{qn_1} Y_{qn_2}) \right|^\rho \lesssim 2^{-q(1 - 1/\rho')}.
\]

Similarly we apply Lemma 6.3 to the sum over \( n_1, n_2 \), and we obtain

\[
\left( 2^{(p - q) \rho'} \right)^{1/\rho'} \left( \max_{n_1, n_2 \in N_j} \left| E(Y_{qn_1} Y_{qn_2}) \right| \right)^{\rho' - \rho} \sum_{n_1, n_2 \in N_j} \left| E(Y_{qn_1} Y_{qn_2}) \right|^\rho \lesssim \left( 2^{(p - q) \rho'} \right)^{1/\rho'} \left( 2^{q \var} \right)^{1/\rho'} \left( 2^{-q(1/\rho') - 1} \right) = 2^{(q \var - 2 - q(1/\rho') - 1) \rho'} 2^{q \var} 2^{2q(1/\rho') 2q(1 - 1/\rho')},
\]

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

\[
\left( \sum_{m_1, m_2 \in M_j} \left| E(X_{pm_1} X_{pm_2}) \right|^{\rho} \right)^{1/\rho} \lesssim 2^{(p \var - i) / (\rho - i)} 2^{-i/\rho} 2^{2q(1 - 1/\rho')}.
\]
The result now follows by combining these estimates:

\[
E \left[ \left| \sum_{m \leq 2^p} \sum_{n \leq 2^q} X_{pm} Y_{qn} (2^{-i} \phi_{ij}, \varphi_{pm} \chi_{qn}) \right|^2 \right] \\
\leq \left( \sum_{m_1, m_2 \in M_f} |E(X_{pm_1} X_{pm_2})|^\rho \right)^{1/\rho} \left( \sum_{m_1, m_2 \in M_f} |E(Y_{qm(m_1)} Y_{qn(m_2)})|^\rho' \right)^{1/\rho'} (2^{-2(p \rho \vee) + p + q})^2 \\
\leq \left( 2^{(p \rho \vee)/(p - 1)} \right) \left( 2^{(q \rho \vee) (1 - 1/p)} \right) 2^{p(1 - 1/p) 2^{2p(1 - 1/p)}} (2^{-4(p \rho \vee) + 2p + 2q}) \\
= 2^{(p \rho \vee)(1 - 1/p) - 1} 2^{(q \rho \vee)(1 - 1/p) - 1} 2^{(4 - 3/p) \rho / p}.
\]

\[\square\]

**Theorem 6.5.** Let \(X : [0, 1] \to \mathbb{R}^d\) be a continuous, centered stochastic process with independent components, and assume that \(X\) satisfies (HC) and \((\rho \text{-var})\) for some \(\rho \in (1, 2)\). Then for every \(\alpha \in (0, 1/\rho)\) almost surely

\[
\sum_{N \geq 0} \|L(S_N X, S_N X) - L(S_{N-1} X, S_{N-1} X)\|_\alpha < \infty,
\]

and therefore \(L(X, X) = \lim_{N \to \infty} L(S_N X, S_N X)\) is almost surely \(\alpha\)-Hölder continuous.

**Proof.** First note that \(L\) is antisymmetric, and in particular the diagonal of the matrix \(L(S_N X, S_N X)\) is constantly zero. For \(k, \ell \in \{1, \ldots, d\}\) with \(k \neq \ell\) we have

\[
\|L(S_N X^k, S_N X^\ell) - L(S_{N-1} X^k, S_{N-1} X^\ell)\|_\alpha \\
= \left\| \sum_{m,n} \sum_{q < N} (X_{km}^k X_{qn}^\ell - X_{km}^\ell X_{qn}^k) \int_0^\infty \varphi_{Nm} \varphi_{qn} (s) q^2 q^{-2}\right\|_\alpha \\
\leq \sum_{m,n} \sum_{q < N} \|X_{km}^k X_{qn}^\ell \int_0^\infty \varphi_{Nm} \varphi_{qn} (s) q^2 q^{-2}\|_\alpha + \sum_{m,n} \sum_{q < N} \|X_{km}^\ell X_{qn}^k \int_0^\infty \varphi_{Nm} \varphi_{qn} (s) q^2 q^{-2}\|_\alpha
\]

Let us argue for the first term on the right hand side, the arguments for the second one being identical. Let \(r \geq 1\). Using the hypercontractivity condition (HC), we obtain

\[
\sum_{i,N} \sum_{j \leq 2^i} \sum_{q < N} \mathbb{P} \left( \left| \sum_{m,n} X_{Nm}^k X_{qn}^k (2^{-i} \phi_{ij}, \varphi_{Nm \chi_{qn}}) \right| > 2^{-i \alpha 2^{-N/(2r)}} 2^{-q/(2r)} \right) \\
\leq \sum_{i,N} \sum_{j \leq 2^i} \sum_{q < N} E \left( \left| \sum_{m,n} X_{Nm}^k X_{qn}^k (2^{-i} \phi_{ij}, \varphi_{Nm \chi_{qn}}) \right|^2 \right)^{2r / 2} 2^{i \alpha 2^r N + q} \\
\leq \sum_{i,N} \sum_{j \leq 2^i} \sum_{q < N} E \left( \left| \sum_{m,n} X_{Nm}^k X_{qn}^k (2^{-i} \phi_{ij}, \varphi_{Nm \chi_{qn}}) \right|^2 \right)^r 2^{i \alpha 2^r N + q}.
\]
Now we can apply Lemma 6.4 to bound this expression by
\[
\sum_{i,N} \sum_{j \leq 2^i} \sum_{q < N} (2^{N\alpha})(1)2^{q\alpha}(1-1/p)\frac{2^{-i}2^{N(1-3/p)}2^{q/r}}{r}(2a2rN+q) \\
\lesssim \sum_{i} 2^i \sum_{N \leq q < N} 2^{i(2a-4)}2^{N\alpha}(1-1/p)2^{q(r(1/p)+1/r)} \\
+ \sum_{i} 2^i \sum_{N > q \geq i} 2^{i(2a-1)}2^{N\alpha}(1-1/p)2^{q(r(1/p)+1/r)} \\
+ \sum_{i} 2^i \sum_{N > i < q < N} 2^{i(2a-1)}2^{N\alpha}(1-1/p)2^{q(r(1/p)+1/r)} \\
\lesssim \sum_{i} 2^{i(2a+3/2-2/p)} + \sum_{i} 2^{i(2a+2/r)}2^{N\alpha}(1-1/p)2^{q(r(1/p)+1/r)} \\
+ \sum_{i} 2^{i(2a+1/r-1)}2^{N\alpha}(1-1/p)2^{q(r(1/p)+1/r)}.
\]

For \( r \geq 1 \) we have \( 1/r - 2/\rho < 0 \), because \( \rho < 2 \). Therefore, the sum over \( N \) in the second term on the right hand side converges. If now we choose \( r > 1 \) large enough so that \( 1+3/r - 2/\rho < 0 \) (and then also \( 2a + 3/r - 2/\rho < 0 \)), then all three series on the right hand side are finite. Hence, Borel-Cantelli implies the existence of \( C(\omega) > 0 \), such that for almost all \( \omega \in \Omega \) and for all \( N, i, j \) and \( q < N \)
\[
\left| \sum_{m,n} X_{N}^{f}(\omega)X_{q}^{g}(\omega)(2^{-i}\chi_{ij}N\chi_{qn}) \right| \leq C(\omega)^{2^{-i}N(2\rho)}2^{-N(2\rho)}/\rho < q/(2\rho).
\]

From here it is straightforward to see that for these \( \omega \) we have
\[
\lim_{N \to \infty} \left\| L(S_{N}X(\omega), S_{N}X(\omega)) - L(S_{N-1}X(\omega), S_{N-1}X(\omega)) \right\|_{\alpha} < \infty.
\]

\[\square\]

### 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 \( \int_{t}^{\omega}(X_{d}X) \). But to apply the results of Section 5, we still need the pathwise convergence of \( \int_{t}^{\omega}(X_{d}X) \) to \( \int_{t}^{\omega}(X_{d}X) \) and the uniform Hölder continuity of \( H_{t}^{\omega}(X_{d}X) \) along the dyadics.

Recall that for a \( d \)-dimensional semimartingale \( X = (X^{1}, \ldots, X^{d}) \), the quadratic variation is defined as \( [X] = ([X^{i}, X^{j}])_{1 \leq i, j \leq d} \). We also write \( X_{s,t} := (X^{i}_{s,t})_{1 \leq i, j \leq d} \) for \( s, t \in [0,1] \).

#### Theorem 6.6

Let \( X = (X^{1}, \ldots, X^{d}) \) be a \( d \)-dimensional continuous martingale. Assume that there exist \( p \geq 2 \) and \( \beta > 0 \), such that \( p \beta > 7/2 \), and such that
\[
E([|X|_{s,t}|]) \lesssim |t - s|^{2p\beta}
\]
for all \( s, t \in [0,1] \). Then \( \int_{t}^{\omega}(X_{d}X) \) almost surely converges uniformly to \( \int_{t}^{\omega}(X_{d}X) \). Furthermore, for all \( \alpha \in (0, \beta - 1/p) \) we have \( X \in \mathcal{C}^{\alpha} \) and almost surely
\[
\sup_{k} \sup_{0 \leq \ell \leq 2^{k}} \frac{\left| \int_{t}^{\omega}(X_{d}X)_{\ell,2^{-k}} - X_{\ell,2^{-k}}X_{\ell,2^{-k}} \right|}{(2^{k})^{2\alpha}} < \infty.
\]
Proof. The Hölder continuity of $X$ follows from Kolmogorov’s continuity criterion. Indeed, applying the Burkholder-Davis-Gundy inequality and (29) we have

$$E(|X_{s,t}|^{2p}) \lesssim \sum_{i=1}^{d} E(|X_{s,t}^i|^{2p}) \lesssim \sum_{i=1}^{d} E(|\langle X \rangle_{s,t}^i|^p) \lesssim E(|\langle X \rangle_{s,t}|^p) \lesssim |t-s|^{2p\beta},$$

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

$$E(|I^\text{Itô}(X, dX)_{s,t} - X_s X_{s,t}|^p) \lesssim E\left(\left| \int_s^t |X_r - X_s|^2 \, d|X|_s \right|^p\right).$$

We apply Hölder’s inequality (here we need $p \geq 2$) to obtain

$$E\left(\left| \int_s^t |X_r - X_s|^2 \, d|X|_s \right|^\frac{p}{2}\right) \lesssim E\left(\left| \int_s^t |X_r - X_s|^p \, d|X|_s \right|^\frac{2}{p}\right).$$

Now the inequalities by Cauchy-Schwarz and then by Burkholder-Davis-Gundy yield

$$E\left(\left| \int_s^t |X_r - X_s|^p \, d|X|_s \right|^\frac{2}{p}\right) \lesssim E\left(\left| \int_s^t |X_r - X_s|^p \, d|X|_s \right|^\frac{2}{p}\right)^{\frac{p}{2}} \lesssim E\left(\left| X_{s,t} - X_{s,s}|^p\right) \lesssim |t-s|^{2p\beta}.$$ 

The Kolmogorov criterion for rough paths, Theorem 3.1 of [FH14], now implies that

$$|I^\text{Itô}(X, dX)_{s,t} - X_s X_{s,t}| \lesssim |t-s|^{2\alpha}$$

almost surely for all $\alpha \in (0, \beta - 1/p)$.

Let us get to the convergence of $I^\text{Itô}_K(X, dX)$. As before, we have

$$E(|I^\text{Itô}_K(X, dX)_{\ell^2-k, \ell^2-k} - I^\text{Itô}_K(X, dX)_{\ell^2-k, \ell^2-k}|^p)$$

$$= E\left(\left| \int_{\ell^2-k}^{\ell^2-k} \sum_{m=\ell}^{\ell^2-k} 1_{\{m^2-k, (m+1)^2-k\}}(r) X_{m^2-k, r} \, dX_r \right|^p\right)$$

$$\lesssim E\left(\left| \int_{\ell^2-k}^{\ell^2-k} \sum_{m=\ell}^{\ell^2-k} 1_{\{m^2-k, (m+1)^2-k\}}(r) X_{m^2-k, r} \, d|X|_s \right|^p\right).$$

Since the terms in the sum all have disjoint support, we can pull the exponent $p/2$ into the
sum, from where we conclude that

\[
E\left(\|X\|_{\ell_2^{-k}, \ell_2^{2-k}}^2 \right) \lesssim \sum_{m=\ell}^{2^{-k}} E\left(\sup_{r \in [s,t]} \left| \sum_{m=\ell}^{2^{-k}} 1_{[m2^{-k},(m+1)2^{-k}]}(r) X_{m2^{-k},r} \right|^p \right)
\]

Hence, we obtain for \(\alpha \in \mathbb{R}\) that

\[
P\left( |I_k^{\text{Itô}}(X,dX)_{\ell_2^{-k}, \ell_2^{2-k}} - I_k^{\text{Itô}}(X,dX)_{\ell_2^{-k}, \ell_2^{2-k}}| > \|(\ell' - \ell)2^{-k}\|^{2\alpha} \right) \lesssim \left( \frac{((\ell' - \ell)2^{-k})^{\frac{1}{2} + p\beta - 2p\alpha}}{(\ell' - \ell)^{2p\alpha - 2k\beta}} \right).
\]

If we set \(\alpha = \beta - 1/(2p) - \varepsilon\), then \(1/2 + p\beta - 2p\alpha = 3/2 - p\beta + 2p\varepsilon\). Now by assumption \(p\beta > 7/2\) and therefore we can find \(\alpha \in (0, \beta - 1/(2p))\) such that

\[
1/2 + p\beta - 2p\alpha < -2.
\]

Estimating the double sum by a double integral, we easily see that for all \(\gamma < -2\)

\[
\sum_{\ell=1}^{2^k} \sum_{\ell'=\ell+1}^{2^k} (\ell' - \ell)^\gamma \lesssim 2^k.
\]

Therefore, we have for \(\alpha \in (0, \beta - 1/(2p))\) satisfying (32)

\[
\sum_{\ell=1}^{2^k} \sum_{\ell'=\ell+1}^{2^k} P\left( |I_k^{\text{Itô}}(X,dX)_{\ell_2^{-k}, \ell_2^{2-k}} - I_k^{\text{Itô}}(X,dX)_{\ell_2^{-k}, \ell_2^{2-k}}| > \|(\ell' - \ell)2^{-k}\|^{2\alpha} \right) \lesssim 2^{k2^{-k}2p\beta(\beta - \alpha)}.
\]

Since \(\alpha < \beta - 1/(2p)\), this is summable in \(k\), and therefore Borel-Cantelli implies that

\[
\sup_k \sup_{0 \leq \ell < \ell' \leq 2^k} \frac{|I_k^{\text{Itô}}(X,dX)_{\ell_2^{-k}, \ell_2^{2-k}} - I_k^{\text{Itô}}(X,dX)_{\ell_2^{-k}, \ell_2^{2-k}}|}{\|(\ell' - \ell)2^{-k}\|^{2\alpha}} < \infty \tag{33}
\]

almost surely. We only proved this for \(\alpha\) close enough to \(\beta - 1/(2p)\), but of course then it also holds for all \(\alpha' \leq \alpha\). The estimate (30) now follows by combining (31) and (33). The uniform convergence of \(I_k^{\text{Itô}}(X,dX)\) to \(I_k^{\text{Itô}}(X,dX)\) follows from (33) in combination with the H"older continuity of \(X\).

**Example 6.7.** The conditions of Theorem 6.6 are satisfied by all Itô martingales of the form \(X_t = X_0 + \int_0^t \sigma_s dW_s\), as long as \(\sigma\) satisfies \(E(\sup_{s \in [0,1]} |\sigma_s|^{p}) < \infty\) for some \(p > 7\). In that case we can take \(\beta = 1/2\) so that in particular \(\beta - 1/p > 1/3\), which means that \(X\) and \(I_k^{\text{Itô}}(X,dX)\) 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

$$dy(t) = b(y(t))dt + \sigma(y(t))dv(t), \quad y(0) = y_0,$$  \hfill (34)

pathwise, where the “stochastic” integral \(dv\) will be interpreted as \(I(\sigma(y), dv)\) or \(I^\text{Itô}(\sigma(y), dv)\).

Assume for example that \((v, L(v, v)) \in C^\alpha \times C^{2\alpha}\) for some \(\alpha \in (1/3, 1/2)\) are given, and that \(b\) is Lipschitz continuous whereas \(\sigma \in C_b^{1+\epsilon}\) for some \(\epsilon \in (2(\alpha + \epsilon) > 1)\). Then Corollary 4.6 implies that \(\sigma(y) \in D_v^{\alpha}\) for every \(y \in D_v^{\alpha}\), and Theorem 4.10 then shows that \(y_0 + \int_0^\lambda b(y(t))dt + I(\sigma(y), dv, \sigma(y)) \in D_v^{\alpha}\).

Moreover, if we restrict ourselves to the set

$$\mathcal{M}_\sigma = \{y \in D_v^{\alpha} : \|y\|_\infty \leq \|\sigma\|_\infty\},$$

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

$$\|\Gamma(y)\|_{v,\alpha} \lesssim \|y_0\| + \|b(0)\| + \|b\|_{\text{Lip}}\|y\|_\infty + \|\sigma(y)\|_{v,\alpha}(\|\sigma\|_\alpha + \|v\|_\alpha^2 + \|L(v, v)\|_{2\alpha}) + \|\sigma(y)\|_\alpha$$

$$\lesssim \|y_0\| + \|b(0)\| + (1 + \|b\|_{\text{Lip}})(1 + \|\sigma\|_{C^\alpha}^2 + \|L(v, v)\|_{2\alpha})(1 + \|y\|_{v,\alpha}),$$

where we wrote \(\|b\|_{\text{Lip}}\) for the Lipschitz norm of \(b\).

To pick up a small factor we apply a scaling argument. For \(\lambda \in (0, 1]\) we introduce the map \(\Lambda_\lambda : C^\beta \to C^\beta\) defined by \(\Lambda_\lambda f(t) = f(\lambda t)\). Then for \(\lambda = 2^{-k}\) and on the interval \([0, \lambda]\) equation (34) is equivalent to

$$dy^\lambda(t) = \lambda b(y^\lambda(t))dt + \lambda^\alpha \sigma(y^\lambda(t))dv^\lambda(t), \quad y^\lambda(0) = y_0,$$  \hfill (35)

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

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

But now \(\Lambda_2^{-k} S_N g = S_{N-k} \Lambda_\lambda g\) for all sufficiently large \(N\), and therefore

$$\Lambda_\lambda I(f, dv) = \lambda^\alpha I(\Lambda_\lambda f, dv^\lambda).$$

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 (35) also in the Itô case. In other words we can replace \(b\) by \(\lambda b\), \(\sigma\) by \(\lambda^\alpha \sigma\), and \(v\) by \(v^\lambda\).

It now suffices to show that \(v^\lambda, L(v^\lambda, v^\lambda), [v^\lambda, v^\lambda]\) are uniformly bounded in \(\lambda\). Since only increments of \(v\) appear in (34) 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

$$L(v^\lambda, v^\lambda) = I(v^\lambda, dv^\lambda) - \pi_<(v^\lambda, v^\lambda) - S(v^\lambda, v^\lambda) = \lambda^{-2\alpha} \Lambda_\lambda I(v, dv) - \pi_<(v^\lambda, v^\lambda) - S(v^\lambda, v^\lambda)$$

$$= \lambda^{-2\alpha} \{ \Lambda_\lambda L(v, v) + [\Lambda_\lambda \pi_<(v, v) - \pi_<(\Lambda_\lambda v, \Lambda_\lambda v)] + [\Lambda_\lambda S(v, v) - S(\Lambda_\lambda v, \Lambda_\lambda v)] \},$$

34
and therefore
\[ \|L(v^\lambda, v^\lambda)\|_{2\alpha} \lesssim \|L(v, v)\|_{2\alpha} + \|S(v, v)\|_{2\alpha} + \|v\|_\alpha^2 + \lambda^{-2\alpha}\|A_\lambda \pi_<(v, v) - \pi_<(v^\lambda, v^\lambda)\|_{2\alpha}. \]

But now
\[
|\Lambda_\lambda \pi_<(v, v)_{s,t} - \pi_<(v^\lambda, v^\lambda)_{s,t}| \leq \|\pi_<(v, v)_{\lambda s, \lambda t} - v(\lambda s)v_{\lambda s, \lambda t}\|
+ |\Lambda_\lambda v(s)(\Lambda_\lambda v)_{s,t} - \pi_<(\Lambda_\lambda v, \Lambda_\lambda v)_{s,t}|
\lesssim \|v\|_{\alpha}^2 \lambda |t-s|^{2\alpha} + \|\Lambda_\lambda v\|_\alpha |t-s|^{2\alpha}
\lesssim \lambda^{2\alpha} \|v\|_\alpha^2 |t-s|^{2\alpha}.
\]

From here we obtain the uniform boundedness of \(\|v^\lambda\|_{v^\lambda, \alpha}\) for small \(\lambda\), 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 \(\lambda\) and therefore we obtain the existence and uniqueness of solutions to (35). Since all operations involved depend on \((v, L(v, v), y_0)\) and possibly \([v, v]\) in a locally Lipschitz continuous way, also \(y^\lambda\) depends locally Lipschitz continuously on this extended data.

Then \(y = \Lambda_{\lambda^{-1}} y^\lambda\) solves (34) on \([0, \lambda]\), and since \(\lambda\) 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 \(y_0 \in \mathbb{R}^d\) and \(\varepsilon > 0\) be such that \(\alpha(2 + \varepsilon) > 2\) and let \(\sigma \in C_b^{2+\varepsilon}\) and \(b\) be Lipschitz continuous. Then there exists a unique \(y \in D_\alpha^\varepsilon\) such that
\[ y = y_0 + \int_0^\cdot b(y(t))dt + I(\sigma(y), dv). \]

The solution \(y\) depends locally Lipschitz continuously on \((v, L(v, v), y_0)\). If furthermore \([v, v]\) satisfies the assumptions of Corollary 5.5, then there also exists a unique solution \(x \in D_\alpha^\varepsilon\) to
\[ x = y_0 + \int_0^\cdot b(x(t))dt + I^{10}(\sigma(x), dv) \]
\[ = y_0 + \int_0^\cdot b(x(t))dt + I(\sigma(x), dv) - \frac{1}{2} \int_0^\cdot D\sigma(x(t))\sigma(x(t))d[v, v] \]
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 of course consider anticipating initial conditions and coefficients. Such problems arise naturally in the study of random dynamical systems; see for example [Imk98, A199]. 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 [BI92]. With classical rough path theory these technical problems disappear. But then the integral is given as limit of compensated Riemann sums (see Proposition 2.4). 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 [RV93, RV07]. The integral notion studied by Ogawa [Oga84, Oga85] 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 [NZ89]). 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

\[
| (fg)_{r-h,r} - (fg)_{r,r+h} | \leq 4 |t-s|^{-2} h^2 \| f \|_\infty \| g \|_\infty. \tag{36}
\]

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

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

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

Now we are ready to prove the regularity estimate.

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

\[
\left\| \sum_{p,m} u_{pm} \varphi_{pm} \right\|_\alpha \lesssim \|(u_{pm})\|_{A^\alpha}.
\]

**Proof.** We need to examine the coefficients $2^{-q} \langle \chi_{qn}, d(\sum_{p,m} u_{pm} \varphi_{pm}) \rangle$. 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 $\varphi_{pm}(t_{qn}) = 0$ for $i = 0, 1, 2$ and for all $m$, and therefore

\[
2^{-q} \langle \chi_{qn}, d\left( \sum_{p,m} u_{pm} \varphi_{pm} \right) \rangle = 2^{-q} \sum_{p \leq q} \sum_m \langle \chi_{qn}, d(u_{pm} \varphi_{pm}) \rangle.
\]

If $p < q$, there is at most one $m_0$ with $\langle \chi_{qn}, d(u_{pm} \varphi_{pm}) \rangle \neq 0$. The support of $\chi_{qn}$ is then contained in $[t_{qn}^{0, m_0}, t_{qn}^1]$ or in $[t_{qn}^0, t_{pm_0}^1]$ and $u_{pm}$ and $\varphi_{pm}$ are affine on these intervals, so (36) yields

\[
\sum_m |2^{-q} \langle \chi_{qn}, d(u_{pm} \varphi_{pm}) \rangle| = \sum_m \left| (u_{pm} \varphi_{pm})_{t_{qn}^0} - (u_{pm} \varphi_{pm})_{t_{qn}^1} \right| \lesssim 2^{2p-2q} \| u_{pm} \|_\infty \| \varphi_{pm} \|_\infty \lesssim 2^{p(2-\alpha)-2q} \| (u_{pm}) \|_{A^\alpha}.
\]

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

\[
\sum_m |2^{-q} \langle \chi_{qn}, d(u_{pm} \varphi_{qn}) \rangle| = \left| (u_{pm} \varphi_{qn})_{t_{qn}^0} - (u_{pm} \varphi_{qn})_{t_{qn}^1} \right| = |u(t_{qn})| \lesssim 2^{-\alpha q} \| (u_{pm}) \|_{A^\alpha}.
\]

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

\[
2^{-q} \langle \chi_{qn}, d\left( \sum_{p,m} u_{pm} \varphi_{pm} \right) \rangle \lesssim \sum_{p \leq q} 2^{p(2-\alpha)-2q} \| (u_{pm}) \|_{A^\alpha} \lesssim 2^{-\alpha q} \| (u_{pm}) \|_{A^\alpha},
\]

which completes the proof. \qed

36
References


[LQ02], *System control and rough paths*, Oxford University Press, 2002.


