Introduction to Rough Paths

Marco Rehmeier

TU Berlin

Summer term 2025

final version as of 15.07.2025

Contents

1 Motivation

1.1 Ordinary differential equations

Ordinary differential equations

$$dY(t) = f(Y(t))dt, \tag{1.1.1}$$

where $f : \mathbb{R}^d \to \mathbb{R}^d$ is a given Borel measurable vector field, and a solution

$$Y : I \to \mathbb{R}^d, \quad Y : t \mapsto Y(t),$$

is defined on some interval $I \subseteq (-\infty, +\infty)$, are fundamental in pure and applied mathematics in order to model systems evolving in a parameter $t \in I$. Usually, t is considered as time and $Y(t)$ as *position*, *state* or *value* of the system. The intuition behind (1.1.1) is that from a time point $t \in I$, a small (forward) time increment $Y(t+h) - Y(t)$ behaves like $f(Y(t))(t+h-t) = f(Y(t))h$.

In words: The change of Y at time t is approximately given by the vector field's value at the current state $Y(t)$.

More generality is achieved by allowing f to depend on t, i.e. $f : I \times \mathbb{R}^d \to \mathbb{R}^d$, and considering

$$dY(t) = f(t, Y(t))dt. \tag{1.1.2}$$

Let us, however, focus here on the time-independent case. The definition of solution below renders the above intuition rigorous. Unless explicitly said differently, we always assume $I = [0, T]$, $T \in (0, \infty)$.

Definition 1.1.1. Let $f : \mathbb{R}^d \to \mathbb{R}^d$ be Borel measurable and locally bounded. An absolutely continuous curve $Y : I \to \mathbb{R}^d$ is a *solution to* (1.1.1), if

$$Y(t) = Y(0) + \int_0^t f(Y(s))\, ds, \quad \forall t \in I.$$

In this case, $t \mapsto Y(t)$ is differentiable dt-a.e. in I with

$$Y'(t) = f(Y(t)) \tag{1.1.3}$$

dt-a.s. If $t \mapsto f(Y(t))$ is continuous in I (for instance, if f is continuous), then (1.1.3) holds for all $t \in I$.

Let us briefly recall a standard result on existence and uniqueness of solutions.

Theorem 1.1.2 (Picard–Lindelöf well-posedness and stability)**.** *(i) Let $f : \mathbb{R}^d \to \mathbb{R}^d$ be Lipschitz continuous in $x \in \mathbb{R}^d$ and let $y_0 \in \mathbb{R}^d$. Then* (1.1.1) *has a unique solution $Y = Y(y_0)$ on I such that $Y(0) = y_0$.*

(ii) *Let $\mathcal{F}$ denote the set of all vector fields f as in (i) and denote the unique solution to (1.1.1) with initial value y_0 by $Y(f, y_0)$. Then the map*

$$\mathcal{F} \times \mathbb{R}^d \ni (f, y_0) \mapsto Y(f, y_0)$$

is continuous from $(\mathcal{F} \times \mathbb{R}^d, |\cdot|_\infty + |\cdot|)$ to $C(I, \mathbb{R}^d)$ with the topology of locally uniform convergence.

Remark 1.1.3. (i) *Recall that the assumptions on f in (i) of the previous theorem can be relaxed to local Lipschitz continuity plus sublinear growth, and even further to one-sided such conditions.*

(ii) *Looking at (ii) of the previous result, by fixing f it follows that solutions to (1.1.2) with a fixed vector field depend continuously on their initial datum.*

1.2 From rough signals to rough paths

Based on this very nice theory, one naturally wants to study more general differential equations

$$dY(t) = f(Y(t))dX(t), \tag{1.2.1}$$

where $X : I \to \mathbb{R}^m$ is a continuous function (usually called *signal, input, driver* or *control*) and $f : \mathbb{R}^d \to \mathbb{R}^{d\times m}$, where by the latter we denote the space of real $d \times m$-matrices. Y is then also called *observation, output* or *filtered effect*. The usual case is retrieved via $m = 1$, $X(t) = t$, and the intuition is similar as before: A solution Y to (1.2.1) should have small time increments

$$Y(t + h) - Y(t) \approx f(Y(t))(X(t + h) - X(t)). \tag{1.2.2}$$

Two immediate questions emerge:

Why and how to study equations of type (1.2.1)?

Here, *to study* means to give meaning to the equation and to obtain an analogue theory as in the classical case, i.e. well-posedness and stability results. For this, a look to Definition 1.1.1 suggests that in particular we have to give meaning to the integral $\int_0^t f(Y(s))dX(s)$. As far as stability is concerned, we now take the point of view that f is prescribed and we are interested in continuity of the solution map (also called *Itô map*)

$$S : (X, y_0) \mapsto Y(X, y_0), \tag{1.2.3}$$

where $Y(X, y_0)$ denotes the solution to (1.2.1) with signal X and initial condition y_0. At this point, it is not yet clear what "solution" means and in which topologies continuity of S should be measured.

Regarding *"why?"*, such equations are very relevant in pure and applied disciplines, e.g. finance, mathematical physics, and even algebra. Perhaps the strongest motivation for (1.2.1) is the case of X being an (arbitrary) path of a stochastic process like Brownian motion, general semi-martingales or even non-semi-martingales such as fractional Brownian motion.

For *"how?"*, we shall now see that several rather classical approaches exist, which are, however, all insufficient to obtain a sufficiently general analogous theory of well-posedness and stability as in the classical ODE case.

1.2.1 Rough signals and how to not treat them

• *Using the classical framework.* If $X \in C^1(I, \mathbb{R}^m)$, then (1.2.2) and $dX(t) = X'(t)dt$ suggest to rewrite (1.2.1) as

$$dY(t) = \tilde{f}(t, Y(t))dt, \tag{1.2.4}$$

with $\tilde{f}(t, y) := f(y)X'(t)$ and to study the latter equation in the classical ODE framework. This may appear natural, but doesn't permit generalizations to non-differentiable signals like Brownian motion: If X does not have a classical derivative X', $\tilde{f}$ is not well-defined.

• *Young integration.* If $f, g \in C(I, \mathbb{R})$ are α- and β-Hölder-continuous, respectively, such that $\alpha + \beta > 1$, one can define

$$\int_0^t f(s)\, dg(s) := \lim_{|\mathcal{P}| \to 0} \sum_{(t_k, t_{k+1}) \in \mathcal{P}} f(t_k)(g(t_{k+1}) - g(t_k)), \quad \forall t \in I, \tag{1.2.5}$$

where $\mathcal{P} = \{0 = t_0 < \cdots < t_N = t\}$ is any partition of $[0, t]$, $|\mathcal{P}| := \max_{k \leqslant N-1}(|t_{k+1} - t_k)$ its *mesh*, and $\lim_{|\mathcal{P}| \to 0}$ denotes the limit along *any* sequence $(\mathcal{P}_n)_{n \in \mathbb{N}}$ with $|\mathcal{P}_n| \to 0$. This construction is sharp in the sense that there are counterexamples for the case $\alpha + \beta \leqslant 1$. This integral is called *Young integral*, due to L.C. Young [6], also known as *Riemann-Stieltjes integral*. Based on this notion of integral, one can give meaning to (1.2.1) as

$$Y^i(t) = Y^i(0) + \int_0^t f^{ij}(Y(s))\, dX^j(s), \quad \forall t \in I, \quad \forall i \in \{1, \ldots, d\}, \tag{1.2.6}$$

where $f = (f^{ij})_{i \leqslant d, j \leqslant m}$ and $X = (X^1, \ldots, X^m)$ (using Einstein summation convention in j), and build a theory of well-posedness and stability.

However, since then $Y(t) - Y(s) \approx f(Y(s))(X(t) - X(s))$ for $|t - s| \ll 1$, one expects Y to (only) inherit the regularity of X. So, since the regularity of $t \mapsto f(Y(t))$ is at most the regularity of Y (even when $f \in C^\infty$), Young's approach is limited to $X \in C^\alpha(I, \mathbb{R}^m)$ with $\alpha > \frac{1}{2}$. Again, rougher (= insufficiently regular) signals are not allowed, in particular Brownian paths.

Both these approaches not only fail to extend the class of integrators X beyond $C^{\frac{1}{2}}$, but do not even lead to satisfying stability results when restricted to $X \in C^\infty$ (in this case Young's approach coincides with the classical one). Indeed, consider the solution map S from (1.2.3) on $\{X \in C^\infty(I, \mathbb{R}^2)\}$. Then we have the following strong negative result from [5]. We denote by τ_{pt} the topology of pointwise convergence on $C(I, \mathbb{R}^d)$.

Proposition 1.2.1. *In general, the solution map* $S : (C^\infty(I, \mathbb{R}^2) \times \mathbb{R}^2, |\cdot|_\infty + |\cdot|) \to (C(I, \mathbb{R}^2), \tau_{pt})$ *is discontinuous.*

Proof. We present a smooth vector field $f : \mathbb{R}^2 \to \mathbb{R}^2$ and a sequence of smooth signals X^n and initial conditions converging to 0 in their respective topologies for which the sequence of unique solutions $Y(f, X^n)$ to (1.2.1) does not converge pointwise to 0 (that is, to the constant path with value $(0,0) \in \mathbb{R}^2$).

Let $f : \mathbb{R}^2 \to \mathbb{R}^{2\times 2}$ be given by $f(x_1, x_2)_{11} = 1, f(x_1, x_2)_{12} = 0, f(x_1, x_2)_{21} = x_1, f(x_1, x_2)_{22} = 0$, i.e. we consider for $X = (X^1, X^2) \in C^\infty(I, \mathbb{R}^2)$ the ODE

$$\begin{cases} dY^1(t) &= dX^1(t) \\ dY^2(t) &= Y^1(t) dX^2(t), \end{cases} \tag{1.2.7}$$

with initial condition $(X^1(0), 0)$, where $Y = (Y^1, Y^2)$. Then

$$Y^1(t) = X^1(t), \quad Y^2(t) = \int_0^t X^1(r) dX^2(r).$$

Now let, for $n \in \mathbb{N}$,

$$X^{(n)}(t) = (X^{(n),1}, X^{(n),2}) := \left(\frac{\cos(n^2 t)}{n}, \frac{\sin(n^2 t)}{n} \right),$$

with $X^{(n)}(0) = (\frac{1}{n}, 0)$. Clearly $X^{(n)} \xrightarrow{n\to\infty} 0$ uniformly (and in particular $X^{(n)}(0) \to 0$). However, a straightforward calculation shows that for $n \to \infty$

$$\int_0^t X^{(n),1}(r) dX^{(n),2}(r)$$

converges to a non-zero constant proportional to t. Hence $S(X^{(n)}, X^{(n)}(0))$ does not converge to the 0-path (wrt. τ_{pt}, and hence in particular also not in the uniform topology). □

•*Stochastic integration.* The stochastic integral $\int_s^t f(s)\, dX_s$ can be defined for semi-martingales X and reasonable random fields f, in Itô- and Stratonovich sense. In particular, $X = B$ is admissible and Itô's celebrated theory of stochastic differential equations allows to study (1.2.1) via stochastic analysis. However, there are flaws to this approach:

1. The construction of $\int f(Y(s)) dX_s$ uses the whole random process X instead of a single path $X(\omega)$, i.e. stochastic integration does not give meaning to the map $\omega \mapsto \int f(Y(\omega, s))\, dX_s(\omega)$. Indeed, recall that Itô stochastic integrals are constructed via an isometry between spaces of stochastic processes rather than individual paths, crucially relying on the martingale property. A solution Y to an SDE is then only defined for a.e. path, and the set of exceptional paths depends on the initial condition y_0 and on f. Thus, there may not be a single (say, Brownian) path for which $Y(y_0, f)$ is defined for every of the uncountably many data (y_0, f).

2. For fixed (y_0, f), even on the set of paths for which Y is defined, $\omega \mapsto Y(\omega)$ is not continuous (think of $\omega \in \Omega = C(I, \mathbb{R}^d)$ with the topology of uniform convergence).

3. X needs to be a semi-martingale, i.e. fractional Brownian motion and many other rough signal processes are not permitted.

1.2.2 The missing piece: from iterated integrals to rough paths

The counterexample from the proof of Proposition 1.2.1 may be seen as evidence that (uniform) convergence of smooth $\mathbb{R}^d$-valued paths X does not imply the (pointwise, let alone uniform) convergence of its *iterated integrals*

$$\int_0^{\cdot} X(r) \otimes dX(r), \text{ where } \left(\int_0^{\cdot} X(r) \otimes dX(r) \right)_{ij} := \int_0^{\cdot} X^i(r) dX^j(r), \; i, j \leqslant d \quad (1.2.8)$$

A seemingly very naive, but in fact ingenious and correct idea will resolve this, not only for this particular example, but in general: In order to measure continuity of the solution map S, these iterated integrals should be taken into account! In fact, uniform continuity of the solution map holds once S is considered as a map not on $\{X \in C^1(I, \mathbb{R}^m)\}$, but with a suitable metric on

$$\left\{ \left(X, \int X \otimes dX\right) : X \in C^1(I, \mathbb{R}^m) \right\} \subseteq C^1(I, \mathbb{R}^m) \oplus C^1(I \times I, \mathbb{R}^{m \times m}). \quad (1.2.9)$$

The closure of this space consists of objects $\mathbf{X} = (X, \mathbb{X}) \in C(I, \mathbb{R}^m) \oplus C(I \times I, \mathbb{R}^{m \times m})$, where $\mathbb{X}$ is not necessarily the iterated integral of X (in fact, the latter need not even exist, for X may not be sufficiently smooth), but instead is considered an "abstract iterated integral". For any $\mathbf{X}$ in this closure, $S(\mathbf{X})$ (defined by uniform continuity of S) may be defined as a solution to (1.2.1) with signal $\mathbf{X}$ (we will give meaning to the integral $\int f(Y(s)) \, d\mathbf{X}_s$). $\mathbf{X}$ will be called *rough path*, and the latter integral *rough integral*. The matter is by no means trivial:

(i) For a generic element $\mathbf{X}$ in this closure, what is $\mathbb{X}$? Is it unique, is it explicitly given via X?

(ii) How large is the resulting closure, which paths X (or, better said, $\mathbf{X}$) does it accommodate?

(iii) How to give meaning to $\int Z_s \, d\mathbf{X}_s$ for an as large as possible class of integrands Z?

We will answer these and related questions in this course. Before starting with details, we stress already now that this way the solution map S factors as

$$S = \hat{S} \circ \Psi,$$

where $\Psi : X \mapsto \mathbf{X}$ is measurable and will be called *rough path lift* [still to be defined; for the moment think of the map $X \mapsto (X, \int X \otimes dX)$], and the continuous (!) map $\hat{S} : \mathbf{X} \mapsto Y(\mathbf{X})$ is the solution map for the *rough differential equation (RDE)*

$$dY_t = f(Y(t)) d\mathbf{X}_t \quad \text{(RDE)}$$

(to which we will give sense). In particular, we will show:

(i) Under suitable assumptions on f and $\mathbf{X}$, this equation has a unique solution, which coincides with the unique solution to (1.2.1) if $\mathbf{X} = (X, \mathbb{X})$ with $X \in C^1(I, \mathbb{R}^m)$. We shall see that this allows $\mathbf{X}$ with underlying path X in C^α, $\alpha < \frac{1}{2}$.

(ii) If $\mathbf{X} = \mathbf{B}(\omega) := (B(\omega), \mathbb{B}(\omega))$, where B is Brownian motion and the stochastic process $\mathbb{B}$ is yet to be defined, the solution $Y(\mathbf{B}(\omega))$ to (6.0.1) is a.s. equal to the solution to the SDE $dY(t) = f(Y(t))dB_t$ (where the RHS is a usual Itô stochastic integral).

(iii) In particular, if B^n is a smooth pathwise approximation of B such that $\mathbf{B^n}$ converges to $\mathbf{B}$ in the yet to be defined suitable sense, then the pathwise solution Y^n to the random ODE

$$dY^n = f(Y^n)dB^n$$

(where the RHS is a pathwisely defined Riemann–Stieltjes integral) converges pathwise uniformly to the SDE solution from (ii).

(iv) Everything outlined above works for general Banach spaces V, W replacing $\mathbb{R}^d$ and $\mathbb{R}^m$.

1.3 Notation

Normed spaces and tensors. For a Banach space V, we denote its norm by $|\cdot|_V$, shortly $|\cdot|$ when no confusion can occur. $|\cdot|$ is also used for the usual Euclidean norm on $\mathbb{R}^d$. For the space of continuous linear maps T between Banach spaces V, W (possibly infinite dimensional) we write $L(V, W)$, which is a Banach space itself with the usual norm $||T||_{L(V,W)} := \sup_{|v|_V \leqslant 1} |Tv|_W$. Also here we shortly write $||T||$ when non-ambiguous.

If V and W have at most countable bases $(e_i)_i$ and $(f_j)_j$, respectively, the *tensor space* $V \otimes W$ is the vector space with basis $\{e_i \otimes f_j, i, j\}$, and its elements $v \otimes w$ are called *tensors*. One can endow $V \otimes W$ with a norm, denoted $|\cdot|_{V \otimes W}$, shortly $|\cdot|$, such that $|v \otimes w| \leqslant |v|_V |w|_W$ and $|v \otimes w| = |w \otimes v|$. We say $|\cdot|_{V \otimes W}$ is *compatible* (w.r.t. $|\cdot|_V, |\cdot|_W$) and *symmetric*. Importantly, if $V = \mathbb{R}^m$, $W = \mathbb{R}^n$, then $V \otimes W \cong \mathbb{R}^{m \times n}$, where the latter denotes the space of real $m \times n$-matrices. If $\bar{V}$ is a third Banach space, one has

$$L(V, L(\bar{V}, W)) \cong L(V \otimes \bar{V}, W).$$

Function spaces. The Banach space of Banach space-valued α-Hölder continuous maps $X : [0, T] \mapsto V$ is denoted by $C^\alpha = C^\alpha([0, T], V)$, $\alpha \in (0, 1)$. It is equipped with seminorm $||\cdot||_\alpha$ and norm $||\cdot||_{C^\alpha}$

$$||X||_\alpha := \sup_{s,t \in [0,T]} \frac{|X_{s,t}|}{|t - s|^\alpha}, \quad ||X||_{C^\alpha} := |X_0| + ||X||_\alpha, \tag{1.3.1}$$

with the convention $\frac{0}{0} = 0$. Here and throughout, for a path $X : [0,T] \to V$, we write $X_{s,t} := X_t - X_s$ (not necesarily $s \leqslant t$). Note that $X \in C^\alpha$ implies $||X||_\infty < \infty$ (the latter denotes the usual L^∞-norm). In fact, $||\cdot||_{C^\alpha}$ is equivalent to $||\cdot||_\infty + ||\cdot||_\alpha$.

Similarly, we write $C_2^\beta = C_2^\beta([0,T], W)$ for the Banach space of two-parameter processes $\mathbb{X} : [0,T]^2 \to W$ with finite norm (!)

$$||\mathbb{X}||_\beta := \sup_{s,t \in [0,T]} \frac{|\mathbb{X}_{s,t}|}{|t-s|^\beta}. \tag{1.3.2}$$

For a path $X : [0,T] \mapsto V$, we write δX for the two-parameter process $\delta X : (s,t) \mapsto X_{s,t}$. Note $X \in C^\alpha$ if and only if $\delta X \in C_2^\alpha$.

2 Spaces of rough paths

Here we let $I = [0, T]$ for some $T > 0$. We define the basic spaces of α-Hölder continuous rough paths and some important subspaces. Throughout, let V be a Banach space with countable basis and with norm $|\cdot|_V$, or $|\cdot|$ when no confusion can occur.

2.1 Hölder continuous rough paths

Definition 2.1.1. Let $\alpha \in (\frac{1}{3}, \frac{1}{2}]$. A ($V$-valued) *$\alpha$-Hölder continuous rough path* $\mathbf{X}$ is a pair $\mathbf{X} = (X, \mathbb{X})$ consisting of a path $X \in C^\alpha(I, V)$ and a two-parameter process $\mathbb{X} \in C_2^{2\alpha}(I, V \otimes V)$ such that *Chen's relation*

$$\mathbb{X}_{s,t} - \mathbb{X}_{s,u} - \mathbb{X}_{u,t} = X_{s,u} \otimes X_{u,t} \tag{C}$$

holds for all $s, u, t \in I$ (not necessarily $s \leqslant u \leqslant t$). The space of α-Hölder continuous rough paths is denoted $\mathcal{C}^\alpha(I, V)$, shortly $\mathcal{C}^\alpha$.

Thus a rough path is an α-Hölder continuous path, augmented with an 2α-Hölder "second-order" process such that the algebraic relation (C) holds. A few elementary remarks and consequences of Chen's relation:

Remark 2.1.2. *(i) Call a two-parameter process $\mathbb{X} : I^2 \to W$ with values in a linear space W* additive, *if $\mathbb{X}_{s,t} = \mathbb{X}_{s,u} + \mathbb{X}_{u,t}$ for all $s \leqslant u \leqslant t$. Clearly, path increments are additive, i.e. for any path $X : I \to W$, $\mathbb{X} = \delta X$ is additive. With this viewpoint,* (C) *quantifies how (much) $\mathbb{X}$ fails to be additive.*

(ii) Since $X_{t,t} = 0$ for all $t \in I$, also $\mathbb{X}_{t,t} = 0$ for all $t \in I$ (take $s = u = t$ in (C)*).*

(iii) $\mathbb{X}_{s,t} = -\mathbb{X}_{t,s} - X_{s,t} \otimes X_{t,s}$ (let $s = t$ in (C)*), and hence also*

$$\mathbb{X}_{s,t} = X_{s,0} \otimes X_{0,t} + \mathbb{X}_{s,0} + \mathbb{X}_{0,t} = -\mathbb{X}_{0,s} + \mathbb{X}_{0,t} - X_{0,s} \otimes X_{0,t} + X_{0,s} \otimes X_{0,s},$$

which shows that the path $t \mapsto (X_{0,t}, \mathbb{X}_{0,t})$ already determines $\mathbb{X}$ (and hence $\mathbf{X}$, up to X_0). Hence we may consider the two-parameter process $\mathbb{X}$ equivalently as a one-parameter path. Which point of view is more convenient depends on the context.

(iv) Let $V = \mathbb{R}$ and $X : I \to \mathbb{R}$ continuous. Then $(X, \mathbb{X})$ with $\mathbb{X}_{s,t} := \frac{1}{2}(X_{s,t})^2$ satisfies (C)*. Since also $X \in C^\alpha \implies \mathbb{X} \in C_2^{2\alpha}$ (!), it follows that every $X \in C^\alpha(I, \mathbb{R})$ can be lifted to a rough path $\mathbf{X} \in \mathcal{C}^\alpha$. The question whether a rough path lift exists for every $X \in C^\alpha(I, \mathbb{R}^d)$, let alone $X \in C^\alpha(I, V)$, is highly nontrivial. There is, however, a positive answer, thanks to the* Lyons-Victoir extension theorem. *We will not rely on this result, but see [1]. In*

applications, for a given path X, a more or less canonical explicit choice of $\mathbb{X}$ is usually available (as we shall see). Importantly, if $X = B(\omega)$ is a Brownian path, natural (but not the only!) candidates for $\mathbb{X}_{s,t}$ are $(\int_s^t B_{s,r} \otimes dB_r)(\omega)$, i.e. the stochastic integral (in Itô- or Stratonovich sense) evaluated in ω, see Chapter 3.

(v) *Assume X is sufficiently regular (think of $X \in C^\beta, \beta > \frac{1}{2}$) in order to define $\mathbb{X}$ as the second-order iterated integral of X*

$$\mathbb{X}_{s,t} := \int_s^t X_{s,r} \otimes dX_r, \quad \forall s,t \in I.$$

Here the integral is a usual Riemann–Stieltjes integral and

$$\left(\int_s^t X_{s,r} \otimes dX_r \right)_{ij} = \int_s^t X^i_{s,r} dX^j_r,$$

where $X = e_i^*(X)$, where $\{e_i^*\}$ denotes the dual basis of the basis $\{e_i\}$ of V consisting of unit length vectors. Then (C) holds. More generally, we make the following observation.

Lemma 2.1.3. *Let $X : I \to V$ be continuous and assume $\mathbb{X}_{s,t} := \int_s^t X_{s,r} \otimes dX_r$ is defined (component wise) by any kind of integration "$\int$" such that for all $s,u,t \in I$ and c constant*

$$\int_s^t = \int_s^u + \int_u^t, \quad \int_s^t c\,dX_r = cX_{s,t}, \quad f \mapsto \int_s^t f\,dX_r \text{ linear in } f.$$

Then $(X, \mathbb{X})$ satisfies (C).

Proof. Straightforward by using the above properties for $\mathbb{X}$. □

Regarding the previous lemma, think for instance of Itô- or Stratonovich stochastic integration.

We also collect the following result for later use. The proof is a straightforward calculation.

Lemma 2.1.4. *Let $s = \tau_0 < \tau_1 \cdots < \tau_N = t$. Then (C) implies*

$$\mathbb{X}_{s,t} = \sum_{i=0}^{N-1} \left(\mathbb{X}_{\tau_i,\tau_{i+1}} + X_{s,\tau_i} \otimes X_{\tau_i,\tau_{i+1}} \right).$$

One should think of $\mathbb{X}_{s,t}$ as a substitute for $\int_s^t X_{s,r} \otimes dX_r$, which is in general not defined (note that in Definition 2.1.1 we imposed $\alpha \leqslant \frac{1}{2}$, i.e. the iterated integral of X cannot be defined via Young integration). Hence we stress that in general

$$\text{"} \mathbb{X}_{s,t} =: \int_s^t X_{s,r} \otimes dX_r \text{"},$$

and *not*

$$\mathbb{X}_{s,t} := \int_s^t X_{s,r} \otimes X_r.$$

A natural question is: Given $X \in C^\alpha$, to which extent is the choice of $\mathbb{X}$ such that $(X, \mathbb{X}) \in \mathcal{C}^\alpha$ unique? In other words, to which extent do (C) and the imposed 2α-regularity uniquely determine $\mathbb{X}$ from X? The answer is as follows.

Lemma 2.1.5. *Assume $X \in C^\alpha$, and $(X, \mathbb{X})$, $(X, \bar{\mathbb{X}}) \in \mathcal{C}^\alpha$. Then*

$$\mathbb{X}_{s,t} - \bar{\mathbb{X}}_{s,t} = G_{s,t},$$

where $G \in C_2^{2\alpha}(I, V \otimes V)$ satisfies $G_{s,t} = G_{s,u} + G_{u,t}$ for all $s, u, t \in I$ (i.e. G is additive) and, thus, in particular

$$G_{s,t} = G_t - G_s, \tag{2.1.1}$$

where $G_t := G_{0,t}$. Conversely, for any $G \in C^{2\alpha}(I, V \otimes V)$, we have $(X, \tilde{\mathbb{X}}) \in \mathcal{C}^\alpha$, where $\tilde{\mathbb{X}}_{s,t} := \mathbb{X}_{s,t} + G_{s,t}$.

In conclusion, for $\mathbf{X} = (X, \mathbb{X}) \in \mathcal{C}^\alpha$, the set of $\bar{\mathbb{X}}$ such that $(X, \bar{\mathbb{X}}) \in \mathcal{C}^\alpha$ is equal to $\{\mathbb{X} + \delta G : G \in C^{2\alpha}(I, V \otimes V)\}$. Put differently, $\mathbb{X}$ is uniquely determined up to path increments of 2α-continuous $V \otimes V$-valued paths.

Proof. For the first part, $G \in C_2^{2\alpha}$ is clear from the regularity of $\mathbb{X}$ and $\bar{\mathbb{X}}$. That G is additive follows since $(X, \mathbb{X})$ and $(X, \bar{X})$ satisfy (C). Then we get (2.1.1) via

$$G_{s,t} = \mathbb{X}_{s,t} - \bar{\mathbb{X}}_{s,t} = -(\mathbb{X}_{0,s} - \bar{\mathbb{X}}_{0,s} - (\mathbb{X}_{0,t} - \bar{\mathbb{X}}_{0,t})) = G_{0,t} - G_{0,s},$$

where the second equality follows from Remark 2.1.2 (iii). Conversely, it is clear that $(X, \tilde{\mathbb{X}})$ satisfies (C), since any path f has additive increments $f_{s,t} = f_{s,u} + f_{u,t}$ (apply this to $f = G \in C^{2\alpha}(I, V \otimes V)$). □

2.1.1 Canonical and smooth rough paths

Again, let $\alpha \in (\frac{1}{3}, \frac{1}{2}]$. Two obvious subsets of $\mathcal{C}^\alpha$ are the spaces of *canonical* and *smooth* rough paths, $\mathcal{L}(C^\infty)$ and $\mathcal{C}^\infty$, respectively, defined as

$$\mathcal{L}(C^\infty) := \left\{(X, \mathbb{X}) \in \mathcal{C}^\alpha : X \in C^\infty, \mathbb{X}_{s,t} = \int_s^t X_{s,r} \otimes dX_r\right\} \tag{2.1.2}$$

and

$$\mathcal{C}^\infty := \left\{(X, \mathbb{X}) \in \mathcal{C}^\alpha : X \in C^\infty, \mathbb{X} \in C_2^\infty\right\}. \tag{2.1.3}$$

Clearly

$$\mathcal{L}(C^\infty) \subseteq \mathcal{C}^\infty \subseteq \mathcal{C}^\alpha.$$

From Lemma 2.1.5 we see that the first inclusion is strict. Indeed, let $V = \mathbb{R}^d$ and consider for instance $X \equiv 0$ with its canonical rough path lift $(X, \mathbb{X}) = (0, 0) \in \mathcal{L}(C^\infty)$. But for any $G \in C^\infty(I, \mathbb{R}^{d \times d})$, the choice $\bar{\mathbb{X}} = \delta G$ gives a rough path $(0, \delta G) \in \mathcal{C}^\infty$, which is not equal to $(0, 0)$ if G is not constant, and hence not an element of $\mathcal{L}(C^\infty)$. The second inclusion is also very strict. Indeed, we already know that for $V = \mathbb{R}^1$, *any* $X \in C^\alpha$ can be lifted via $(X, \frac{1}{2}(\delta X)^2) \in \mathcal{C}^\alpha$.

Why $\alpha \in (\frac{1}{3}, \frac{1}{2}]$? Let us observe that the notion of α-Hölder continuous rough path is useful only for $\alpha \leqslant \frac{1}{2}$. Indeed, when $\alpha > \frac{1}{2}$ and $X \in C^\alpha$, the iterated integral of X is defined (Young!) and with this choice of $\mathbb{X}$, (C) holds. If $(X, \bar{\mathbb{X}}) \in C^\alpha \oplus C_2^{2\alpha}$ satisfies (C), then by Lemma 2.1.5 $\bar{\mathbb{X}} = \mathbb{X} + \delta G$ for some 2α-Hölder continuous path $G : I \to V \otimes V$. Now we use the elementary but important fact that

$$\text{any metric space-valued } \beta-\text{Hölder continuous function is constant if } \beta > 1. \tag{2.1.4}$$

Hence $\delta G \equiv 0$, which shows that in this case $\mathbb{X} = \int X \otimes dX$ is the unique candidate for a second-order rough path process for X. Consequently,

$$\mathcal{C}^\alpha = \left\{ \left(X, \int X \otimes dX \right) : X \in C^\alpha \right\}, \quad \forall \alpha > \frac{1}{2},$$

i.e. for $\alpha > \frac{1}{2}$, rough path theory cannot do better than Young's theory.

We will address the bound $\alpha > \frac{1}{3}$ later.

2.1.2 Rough path metric and norm

Definition 2.1.1 suggests to think of $\mathcal{C}^\alpha$ as a subset of $C^\alpha(I, V) \oplus C_2^{2\alpha}(I, V \otimes V)$, shortly $C^\alpha \oplus C_2^{2\alpha}$. The latter is a Banach space with semi-norm $||X||_\alpha + ||\mathbb{X}||_{2\alpha}$ and norm $||X||_{C^\alpha} + ||\mathbb{X}||_{2\alpha}$. Note that due to the constraint (C), $\mathcal{C}^\alpha \subseteq C^\alpha \oplus C_2^{2\alpha}$ is a *nonlinear* subspace and, as such, not a normed space in the usual sense. Indeed, if $(X, \mathbb{X}), (Y, \mathbb{Y}) \in C^\alpha \oplus C_2^{2\alpha}$ satisfy (C), it does not follow that $(X+Y, \mathbb{X}+\mathbb{Y})$ satisfies (C). In other words, while $Z = X + Y \in C^\alpha$ and hence there exists a rough path lift $(Z, \mathbb{Z}) \in C^\alpha \oplus C_2^{2\alpha}$ (see Remark 2.1.2 (iv)), one does *not* have $\mathbb{Z} = \mathbb{X} + \mathbb{Y}$.

Still, one may of course consider the norm on $C^\alpha \oplus C_2^{2\alpha}$ restricted to $\mathcal{C}^\alpha$ simply as a map. However, this map does not respect the natural homogeneity ($\lambda > 0$)

$$\delta_\lambda : (X, \mathbb{X}) \mapsto (\lambda X, \lambda^2 \mathbb{X}),$$

in the sense that

$$||\lambda X||_{C^\alpha} + ||\lambda^2 \mathbb{X}||_{2\alpha} \neq \lambda\big(||X||_{C^\alpha} + ||\mathbb{X}||_{2\alpha}\big).$$

That δ_λ is indeed natural can be seen via the fact that δ_λ leaves (C) invariant. Hence, it is more natural to consider the δ_λ-homogeneous norm on $C^\alpha \oplus C_2^{2\alpha}$

$$|||\mathbf{X}|||_{\mathcal{C}^\alpha} := ||X||_{C^\alpha} + \sqrt{||\mathbb{X}||_{2\alpha}}$$

and respective semi-norm

$$|||\mathbf{X}|||_\alpha := ||X||_\alpha + \sqrt{||\mathbb{X}||_{2\alpha}}$$

(which vanishes when X is constant and $\mathbb{X} \equiv 0$). Considering both as maps on $\mathcal{C}^\alpha$ will turn out useful. While not a vector space, $\mathcal{C}^\alpha$ is a metric space with the metric induced by the usual norm on $C^\alpha \oplus C_2^{2\alpha}$ (on the metric level, the (in)homogeneity property is less relevant).

Definition 2.1.6. The (inhomogeneous) α-*Hölder rough paths metric* on $\mathcal{C}^\alpha$ is defined as

$$\rho_{\mathcal{C}^\alpha}(\mathbf{X},\mathbf{Y}) := ||X-Y||_{C^\alpha} + ||\mathbb{X}-\mathbb{Y}||_{2\alpha},$$

where $\mathbf{X}=(X,\mathbb{X})$ and $\mathbf{Y}=(Y,\mathbb{Y})$ are elements in $\mathcal{C}^\alpha$, and $\mathbb{X}-\mathbb{Y}\in C_2^{2\alpha}$ is defined as $(s,t)\mapsto \mathbb{X}_{s,t}-\mathbb{Y}_{s,t}$.

Since for most of the theory of rough paths only increments of X, i.e. δX, are needed, (in particular in (C)) one often implicitly identifies $(X,\mathbb{X})$ and $(\bar{X},\mathbb{X})$, if $\delta(X-\bar{X})=0$, i.e. $X=\bar{X}+v$ for some $v\in V$. Thus there is no danger in considering instead of the true metric $\rho_{\mathcal{C}^\alpha}$ the pseudo metric

$$\rho_\alpha(\mathbf{X},\mathbf{Y}) := ||X-Y||_\alpha + ||\mathbb{X}-\mathbb{Y}||_{2\alpha},$$

for which $\rho((X,\mathbb{X}),(\bar{X},\mathbb{X}))=0$ for $X,\bar{X}$ as above. With some abuse of notation, ρ_α is also called α-*Hölder rough paths metric*, and usually ρ_α instead of ρ_{C^α} is meant when not explicitly said otherwise.

We leave as an exercise the first two parts of the following assertion.

Proposition 2.1.7. *(i)* $(\mathcal{C}^\alpha,\rho_{\mathcal{C}^\alpha})$ *is a complete metric space.*

(ii) Let $V=\mathbb{R}$. $(\mathcal{C}^\alpha(I,\mathbb{R}),\rho_{\mathcal{C}^\alpha})$ *is not separable.*

(iii) ("Interpolation") Let $\frac{1}{3}<\alpha<\beta\leqslant\frac{1}{2}$ *and* $(\mathbf{X}^n)_{n\in\mathbb{N}}\subseteq\mathcal{C}^\beta$. *If*

$$\sup_n |||\mathbf{X}^n|||_\beta < C_0 < \infty,$$

$X^n \xrightarrow{n\to\infty} X$ *and* $\mathbb{X}^n \xrightarrow{n\to\infty} \mathbb{X}$ *pointwise, then* $\mathbf{X}=(X,\mathbb{X})\in\mathcal{C}^\beta$ *and* $\rho_{\mathcal{C}^\alpha}(\mathbf{X}^n,\mathbf{X}) \xrightarrow{n\to\infty} 0$.

Similarly, if $\delta X^n \xrightarrow{n\to\infty} \delta X$ *instead of* $X^n \xrightarrow{n\to\infty} X$, *then* $\rho_\alpha(\mathbf{X}^n,\mathbf{X}) \xrightarrow{n\to\infty} 0$ *instead of* $\rho_{\mathcal{C}^\alpha}(\mathbf{X}^n,\mathbf{X}) \xrightarrow{n\to\infty} 0$.

Regarding (iii), note that $\delta X^n \xrightarrow{n\to\infty} \delta X$ pointwise is strictly weaker than $X^n \xrightarrow{n\to\infty} X$ pointwise. Indeed, clearly the latter implies the former. The converse does not hold, consider for instance $X^n\equiv 1$ for all n and $X\equiv 0$, then $\delta X^n\equiv 0$.

Proof of (iii). By definition of ρ_α and $\rho_{\mathcal{C}^\alpha}$ it is clear that the first part of the assertion follows immediately from the second. Regarding the second part, first assume both convergences hold uniformly in $(s,t)\in I^2$. Since the assumption entails (here pointwise convergence is enough)

$$|X_{s,t}| = \lim_{n\in\mathbb{N}} |X^n_{s,t}| \leqslant C_0|t-s|^\beta, \quad |\mathbb{X}_{s,t}| = \lim_{n\in\mathbb{N}} |\mathbb{X}^n_{s,t}| \leqslant C_0|t-s|^{2\beta},$$

we find $\mathbf{X}\in C^\beta\oplus C_2^{2\beta}$. Clearly, pointwise convergence implies that $\mathbf{X}$ satisfies (C), and thus $\mathbf{X}\in\mathcal{C}^\beta$.

Now, using the assumption of uniform convergence, we have uniformly in $s,t\in I$

$$|X_{s,t}-X^n_{s,t}| \leqslant \varepsilon_n, \quad |X_{s,t}-X^n_{s,t}| \leqslant 2C_0|t-s|^\beta$$

and similarly

$$|\mathbb{X}_{s,t} - \mathbb{X}^n_{s,t}| \leqslant \varepsilon_n, \quad |\mathbb{X}_{s,t} - \mathbb{X}^n_{s,t}| \leqslant 2C_0|t-s|^{2\beta}$$

for a sequence $\varepsilon_n \xrightarrow{n\to\infty} 0$. Using geometric interpolation, i.e. the inequality $a \wedge b \leqslant a^{1-\theta}b^\theta$ for all $a, b > 0$ and $0 < \theta < 1$, with $\theta = \frac{\alpha}{\beta}$ and $\theta = \frac{2\alpha}{2\beta}$, respectively, we obtain

$$|X_{s,t} - X^n_{s,t}| \leqslant C\varepsilon_n^{1-\frac{\alpha}{\beta}}|t-s|^\alpha$$

and

$$|\mathbb{X}_{s,t} - \mathbb{X}^n_{s,t}| \leqslant C\varepsilon_n^{1-\frac{\alpha}{\beta}}|t-s|^{2\alpha},$$

where $C > 0$ is some constant not depending on n, s or t. Since $\varepsilon_n^{1-\frac{\alpha}{\beta}} \to 0$, the assertion follows.

Now drop the assumption of uniform convergence and assume only pointwise convergence. All we have to prove is that, using the uniform Hölder bounds of $(\mathbf{X}^n)_{n\in\mathbb{N}}$, the latter implies the former. For given $\varepsilon > 0$, choose a finite partition $D = D(\varepsilon, \beta)$ of I, $D = \{\tau_i\}_{i\leqslant N}$, with $|D| := \max_i |\tau_{i+1} - \tau_i|$ sufficiently small to have $C_0|D|^\beta < \frac{\varepsilon}{8}$. Let now $s, t \in I$ and denote by $\bar{s}$, $\bar{t}$ the (or one of the) respective element(s) in D with the smallest distance to s and t, respectively. Then

$$|X_{s,t} - X^n_{s,t}| \leqslant |X_{\bar{s},\bar{t}} - X^n_{\bar{s},\bar{t}}| + |X_{s,\bar{s}}| + |X_{\bar{t},t}| + |X^n_{s,\bar{s}}| + |X^n_{\bar{t},t}| \leqslant |X_{\bar{s},\bar{t}} - X^n_{\bar{s},\bar{t}}| + \frac{\varepsilon}{2}.$$

By pointwise convergence and since D is finite, one can choose n sufficiently large so that the first summand on the RHS is bounded above by $\frac{\varepsilon}{2}$. This shows $\delta X \to \delta X^n$ uniformly. The argument for $\mathbb{X}$ is similar (but one has to use (C)). □

2.2 Geometric rough paths

Chen's relation (C) captures the basic algebraic relation between the components of a rough path, motivated from the smooth case $\mathbb{X}_{s,t} = \int_s^t X_{s,r} \otimes dX_r$. Of course, in the latter case, we also have the usual integration by parts formula (let for simplicity $V = \mathbb{R}^d$):

$$\begin{aligned}
\mathbb{X}^{ij}_{s,t} + \mathbb{X}^{ji}_{s,t} &= \int_s^t X^i_{s,r} dX^j_r + \int_s^t X^j_{s,r} dX^i_r \\
&= X^i_{s,t}X^j_t - \int_s^t X^j_r dX^i_{s,r} + \int_s^t X^j_{s,r} dX^i_r \\
&= X^i_{s,t}X^j_{s,t},
\end{aligned}$$

where we used the usual integration by parts formula for the second equality. Hence, in the smooth case, we find

$$\mathrm{Sym}(\mathbb{X}_{s,t}) = \frac{1}{2}X_{s,t} \otimes X_{s,t}, \tag{2.2.1}$$

where we denote by $\mathrm{Sym}(M) = \frac{1}{2}(M+M^T)$ the symmetric part of a real $d\times d$-matrix M.

This suggests that in the case of a general rough path $(X, \mathbb{X})$, postulating (2.2.1) may be valuable (we shall see: it is!). There are two natural ways to impose this geometricity condition:

Definition 2.2.1. The *space of weakly geometric α-Hölder rough paths*

$$\mathcal{C}_g^\alpha(I, V) \subseteq \mathcal{C}^\alpha(I, V)$$

consists of those $(X, \mathbb{X}) \in \mathcal{C}^\alpha(I, V)$ for which (2.2.1) holds for all $(s,t) \in I^2$.

Definition 2.2.2. The *space of geometric α-Hölder rough paths*

$$\mathcal{C}_g^{0,\alpha}(I, V) \subseteq \mathcal{C}^\alpha(I, V)$$

is defined as the closure of $\mathcal{L}(C^\infty)$ in $\mathcal{C}^\alpha$ with respect to the pseudo-metric ρ_α.

Clearly, elements in $\mathcal{C}_g^{0,\alpha} = \mathcal{C}_g^{0,\alpha}(I, V)$ satisfy (2.2.1), thus we have

$$\mathcal{C}_g^{0,\alpha} \subseteq \mathcal{C}_g^\alpha \subseteq \mathcal{C}^\alpha.$$

In fact, both inclusions are strict (an example for the strictness of the second inclusion will be given by the Itô rough path lift of Brownian motion). One can even show that, as long as V is separable, $C_g^{0,\alpha}$ is separable as well (w.r.t. ρ_α), while C_g^α is not.

Remark 2.2.3. *However, often the distinction between $\mathcal{C}_g^\alpha$ and $\mathcal{C}_g^{0,\alpha}$ is irrelevant, since*

$$\mathcal{C}_g^\beta \subseteq \mathcal{C}_g^{0,\alpha}$$

whenever $\beta > \alpha$.

In conclusion, up to here we have encountered the following spaces of rough paths, all of them subspaces of $\mathcal{C}^\alpha$:

$$\begin{aligned}\mathcal{L}(C^\infty) &\subseteq \mathcal{C}^\infty \\ &\subseteq \mathcal{C}_g^{0,\alpha} \subseteq \mathcal{C}_g^\alpha.\end{aligned}$$

3 Rough paths and Brownian motion

Throughout, $B = (B_t)_{t\in I}$ is a pathwise continuous standard $\mathbb{R}^d$-valued Brownian motion, defined on some probability space $(\Omega, \mathcal{F}, \mathbb{P})$, and we let again $I = [0, T]$ for some $T > 0$. Often, for a Brownian path $t \mapsto B_t(\omega)$, we shortly write $t \mapsto B_t$ or simply B. The components of $B = (B^1, \dots, B^d)$ are independent standard $\mathbb{R}$-valued Brownian motions.

3.1 Brownian motion and its iterated integral

Recall that for any $\alpha \in (0, \frac{1}{2})$, B may be chosen such that its paths are a.s. α-Hölder continuous, but that a.e. Brownian path is *not* Hölder continuous for $\alpha \geqslant \frac{1}{2}$. In particular, the iterated integral $\int_s^t B(\omega)^i_{s,r} dB^j_r(\omega)$ is *not* defined in the usual sense of Young or Riemann–Stieltjes integration.

However, stochastic integration allows to define the stochastic process

$$(t, \omega) \mapsto \Big(\int_0^t B^i_r dB^j_r\Big)(\omega)$$

(as such, it is defined for a.e. $\omega \in \Omega$), and hence also

$$(s, t, \omega) \mapsto \Big(\int_s^t B_{s,r} \otimes dB_r\Big)(\omega) \in \mathbb{R}^{d\times d}.$$

One may use either Itô- or Stratonovich stochastic integration, where usually we write "d" for Itô- and "$\circ d$" for Stratonovich-integration. We set

$$\mathbb{B}^I : I^2 \to \mathbb{R}^{d\times d}, \quad \mathbb{B}^I_{s,t} := \int_s^t B_{s,r} \otimes dB_r,$$

and

$$\mathbb{B}^S : I^2 \to \mathbb{R}^{d\times d}, \quad \mathbb{B}^S_{s,t} := \int_s^t B_{s,r} \otimes \circ dB_r$$

and stress that $\mathbb{B}^I$ and $\mathbb{B}^S$ are (two-parameter) stochastic processes. Since $\mathbb{B}^I$ and $\mathbb{B}^S$ satisfy the assumptions of Lemma 2.1.3, we obtain

$$\mathbf{B}^I(\omega) := (B(\omega), \mathbb{B}^I(\omega)) \text{ and } \mathbf{B}^S(\omega) := (B(\omega), \mathbb{B}^S(\omega)) \text{ satisfy (C) a.e.}$$

$\mathbf{B}^I$ and $\mathbf{B}^S$ are called *Itô-* and *Stratonovich Brownian rough path lift*.

The next natural question is whether $(B, \mathbb{B}^I)$ and $(B, \mathbb{B}^S)$ belong to the rough path spaces introduced before. Clearly, this is not the case for $\mathcal{L}(C^\infty)$ and $\mathcal{C}^\infty$.

Geometricity. Before we further address the question of regularity, we make the following observation regarding the geometricity of $\mathbf{B}^I$ and $\mathbf{B}^S$:

Lemma 3.1.1. $\mathbf{B}^S$ *satisfies* (2.2.1) *a.s., while this is a.s. not true for* $\mathbf{B}^I$.

Proof. Since Itô's product rule gives

$$\int_s^t B^i_{s,r} dB^j_r = B^i_{s,t} B^j_t - \int_s^t B^j_r dB^i_r - \langle B^i, B^j \rangle_{t-s}$$

a.s., where $\langle \cdot, \cdot \rangle_\cdot$ denotes the covariation bracket, and since due to the independence of $\{B^i\}_{i \leqslant d}$ one has $\langle B^i, B^j \rangle_{t-s} = \delta_{ij}(t-s)$, we obtain a.s.

$$\operatorname{Sym}(\mathbb{B}^I_{s,t}) = \frac{1}{2} B_{s,t} \otimes B_{s,t} - \frac{1}{2}(t-s)\,\mathrm{Id}$$

(here and below, Id denotes the $d \times d$-identity matrix). Thus, a.s. the paths of $\mathbf{B}^I$ do not satisfy (2.2.1).

Since $\int_s^t B^i_{s,r} \circ dB^i_r = \int_s^t B^i_{s,r} dB^j_r + \frac{1}{2}\langle B^i, B^i \rangle_{t-s}$, it follows immediately from the above that a.s.

$$\operatorname{Sym}(\mathbb{B}^S_{s,t}) = \operatorname{Sym}(\mathbb{B}^I_{s,t}) + \frac{1}{2}(t-s)\,\mathrm{Id} = \frac{1}{2} B_{s,t} \otimes B_{s,t}.$$

This concludes the proof. ∎

This result is not surprising, since we know that Stratonovich stochastic integration preserves the usual rules of calculus.

Remark 3.1.2. *Moreover, a straightforward calculation shows*

$$\operatorname{Ant}(\mathbb{B}^I_{s,t}) = \operatorname{Ant}(\mathbb{B}^S_{s,t}),$$

where for a $d \times d$-matrix M, $\operatorname{Ant}(M) := \frac{1}{2}\big(M - M^T\big)$ denotes its antisymmetric part.

3.2 Brownian paths as Hölder regular rough paths

Let $\alpha < \frac{1}{2}$. We are now going to show $\mathbf{B}^I, \mathbf{B}^S \in \mathcal{C}^\alpha(I, \mathbb{R}^d)$ a.s. Since we showed above that in both cases (C) is satisfied and that $B \in C^\alpha$ a.s., it remains to show $\mathbb{B}^I, \mathbb{B}^S \in C_2^{2\alpha}(I, \mathbb{R}^{d \times d})$.

3.2.1 Rough Kolmogorov continuity criterion

More generally, given random $(X, \mathbb{X}) = (X(\omega), \mathbb{X}(\omega))$ on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ such that $X : I \to V$ and $\mathbb{X} : I^2 \to V \otimes V$ satisfy (C) a.s., it is an important question whether $(X, \mathbb{X}) \in C^\alpha \oplus C_2^{2\alpha}$ a.s. (then $(X, \mathbb{X}) \in \mathcal{C}^\alpha$ a.s.). A sufficient criterion is given by the following result.

Recall that for stochastic processes X, Y on the same probability space, Y is called a *modification of* X, if $\mathbb{P}(X_t = Y_t) = 1$ for all $t \in I$. Analogously, one defines a modification of a two-parameter stochastic process.

Proposition 3.2.1 (Rough Kolmogorov continuity criterion). *Let $q \geqslant 2$, $\beta > \frac{1}{q}(\leqslant \frac{1}{2})$ and let $(X, \mathbb{X})$ be as above. Assume for all $s, t \in I$*

$$|X_{s,t}|_{L^q(\Omega)} \leqslant C|t-s|^\beta, \quad |\mathbb{X}_{s,t}|_{L^{\frac{q}{2}}(\Omega)} \leqslant C|t-s|^{2\beta} \tag{3.2.1}$$

for a constant $C < \infty$. Then, for all $\alpha \in [0, \beta - \frac{1}{q})$, there exists a modification of $(X, \mathbb{X})$ (again denoted $(X, \mathbb{X})$) and random variables $K_\alpha \in L^q(\Omega), \mathbb{K}_\alpha \in L^{\frac{q}{2}}(\Omega)$ such that for all $s, t \in I$ and $\omega \in \Omega$

$$|X_{s,t}(\omega)| \leqslant K_\alpha(\omega)|t-s|^\alpha, \quad |\mathbb{X}_{s,t}(\omega)| \leqslant \mathbb{K}_\alpha(\omega)|t-s|^{2\alpha}. \tag{3.2.2}$$

Hence, if $(X, \mathbb{X})$ satisfies (C) a.s. and (3.2.1) for some $q \geqslant 2$ and β such that $\beta - \frac{1}{q} > \frac{1}{3}$, then for any $\frac{1}{3} < \alpha < \beta - \frac{1}{q}$, there is a modification of $(X, \mathbb{X})$ belonging to $\mathcal{C}^\alpha$ a.s.

Note that when one drops all assumptions and assertions on $\mathbb{X}$, one recovers the statement and the proof (see below) of the usual Kolmogorov continuity criterion for stochastic processes.

Proof. Without loss of generality, let $I = [0,1]$. Set $D_n := \{k2^{-n}, k \in \mathbb{N}_0, k \leqslant 2^n - 1\}$, so that $\# D_n = 2^n$, and define the random variables

$$K_n := \sup_{t \in D_n} |X_{t,t+2^{-n}}|, \quad \mathbb{K}_n := \sup_{t \in D_n} |\mathbb{X}_{t,t+2^{-n}}|.$$

Then by (3.2.1)

$$\mathbb{E}[K_n^q] \leqslant \mathbb{E}\big[\sum_{t \in D_n} |X^q_{t,t+2^{-n}}|\big] \leqslant 2^n C^q 2^{-\beta q n} = C^q 2^{n(1-\beta q)} \tag{3.2.3}$$

and

$$\mathbb{E}[\mathbb{K}_n^{\frac{q}{2}}] \leqslant \mathbb{E}\big[\sum_{t \in D_n} |\mathbb{X}^{\frac{q}{2}}_{t,t+2^{-n}}|\big] \leqslant 2^n C^{\frac{q}{2}} 2^{-\beta q n} = C^{\frac{q}{2}} 2^{n(1-\beta q)}. \tag{3.2.4}$$

Fix $s < t$ in $D := \bigcup_{n \in \mathbb{N}_0} D_n$ (hence $t - s \leqslant 1$) and let m such that $2^{-m-1} < t - s \leqslant 2^{-m}$. Choose a finite partition $(\tau_i)_{1 \leqslant i \leqslant N}$ of $[s,t)$,

$$s = \tau_0 < \tau_1 < \cdots < \tau_N = t$$

such that for any $i \leqslant N-1$ there is $n \geqslant m+1$ such that $\tau_{i+1} - \tau_i = 2^{-n}$, and for each n there are at most two such i. Then, using these choices for τ_i, we find

$$|X_{s,t}| \leqslant \max_{0 \leqslant i \leqslant N-1} |X_{s,\tau_{i+1}}| \leqslant \sum_{i=0}^{N-1} |X_{\tau_i,\tau_{i+1}} \leqslant 2 \sum_{n \geqslant m+1} K_n \tag{3.2.5}$$

and similarly

$$\begin{aligned}|\mathbb{X}_{s,t}| = \left|\sum_{i=0}^{N-1} \mathbb{X}_{\tau_i,\tau_{i+1}} + X_{s,\tau_i} \otimes X_{\tau_i,\tau_{i+1}}\right| &\leqslant \sum_{i=0}^{N-1} \left(|\mathbb{X}_{\tau_i,\tau_{i+1}}| + |X_{s,\tau_i}||X_{\tau_i,\tau_{i+1}}|\right) \\ &\leqslant \sum_{i=0}^{N-1} |\mathbb{X}_{\tau_i,\tau_{i+1}}| + \max_{0\leqslant i<N} |X_{s,\tau_{i+1}}| \sum_{j=0}^{N-1} |X_{\tau_j,\tau_{j+1}}| \\ &\leqslant 2 \sum_{n\geqslant m+1} \mathbb{K}_n + \left(2 \sum_{n\geqslant m+1} K_n\right)^2, \end{aligned} \tag{3.2.6}$$

where we used (2.1.4) for the first equality and (3.2.5) for the final inequality. Thus, we obtain from (3.2.5) and the choice of m in relation to $t-s$

$$\frac{|X_{s,t}|}{|t-s|^\alpha} \leqslant 2 \sum_{n\geqslant m+1} K_n 2^{\alpha(m+1)} \leqslant 2 \sum_{n\geqslant m+1} K_n 2^{\alpha n} \leqslant 2 \sum_{n\geqslant 1} K_n 2^{\alpha n} =: K_\alpha. \tag{3.2.7}$$

Now $K_\alpha \in L^q(\Omega)$, since by (3.2.3)

$$|K_\alpha|_{L^q} \leqslant 2 \sum_{n\geqslant 1} 2^{\alpha n} |K_n|_{L^q(\Omega)} \leqslant 2C^q \sum_{n\geqslant 1} 2^{\alpha n + n(\frac{1}{q}-\beta)},$$

where the RHS is finite, since $\alpha < \beta - \frac{1}{q}$. Simillarly, using (3.2.6)

$$\frac{|\mathbb{X}_{s,t}|}{|t-s|^{2\alpha}} \leqslant 2 \sum_{n\geqslant m+1} \mathbb{K}_n 2^{2\alpha(m+1)} + \left(2 \sum_{n\geqslant m+1} K_n 2^{\alpha(m+1)}\right)^2 \leqslant \mathbb{K}_\alpha + K_\alpha^2,$$

where $\mathbb{K}_\alpha := 2\sum_{n\geqslant 1} \mathbb{K}_n 2^{2\alpha n}$. Now $\mathbb{K}_\alpha \in L^{\frac{q}{2}}(\Omega)$ by (3.2.4), and we already know $K_\alpha^2 \in L^{\frac{q}{2}}(\Omega)$ from above.

Thus, there is a $\mathbb{P}$-zero set $N \subseteq \mathcal{F}$ such that for every $\omega \in N^c$, $X(\omega)$ and $\mathbb{X}(\omega)$ are α- and 2α-Hölder continuous, respectively, for every $\alpha < \beta - \frac{1}{q}$ on $\bigcup_{n\in\mathbb{N}} D_n$. For $\omega \in N$, set $\bar{X} :\equiv 0$ and $\bar{\mathbb{X}} :\equiv 0$. For $\omega \in N^c$, define $\bar{X}$ and $\bar{\mathbb{X}}$ as the unique continuous extension of $(X_t)_{t\in D}$ and $(\mathbb{X}_{s,t})_{s,t\in D}$ (this is possible because in particular X and $\mathbb{X}$ are uniformly continuous on D).

It is left as a simple exercise to show $\bar{X} \in C^\alpha$ and $\bar{\mathbb{X}} \in C_2^{2\alpha}$ pathwise.

Then it remains to show that the process $(\bar{X}, \bar{\mathbb{X}})$ is a modification of $(X, \mathbb{X})$. We show this for $\bar{X}$, the argument for $\bar{\mathbb{X}}$ is similar. Let $t \in I$. If $t \in D$, then $\bar{X}_t = X_t$ on N^c and $\mathbb{P}(N^c) = 1$. Now let $t \in D^c$. Then there is a sequence $(t_n)_{n\in\mathbb{N}} \subseteq D$ such that $t_n \to t$, so $\bar{X}_t = \lim_n \bar{X}_{t_n} = \lim_n X_{t_n}$ a.s. In particular, X_{t_n} converges to $\bar{X}_t$ in measure (wrt. $\mathbb{P}$). On the other hand, by (3.2.1) and Chebyshev's inequality, we obtain, for all $\eta > 0$,

$$\mathbb{P}(|X_{t_n} - X_t| \geqslant \eta) \leqslant \eta^{-q} \mathbb{E}[|X_{t_n} - X_t|^q] \leqslant \eta^{-q} C |t_n - t|^\beta.$$

Hence X_{t_n} also converges to X_t in measure (wrt. $\mathbb{P}$), and so $X_t = \bar{X}_t$ a.s. This shows that $(\bar{X}, \bar{\mathbb{X}})$ is the desired modification, and the proof is complete. □

We turn to the application of the previous result to Brownian motion.

Proposition 3.2.2. *For a.e.* ω*,* $(B(\omega), \mathbb{B}^I(\omega))$ *and* $(B(\omega), \mathbb{B}^S(\omega))$ *satisfy* (3.2.1) *for* $\beta = \frac{1}{2}$ *and every* $q \in [2, \infty)$*.*

Hence, for every $\alpha \in (\frac{1}{3}, \frac{1}{2})$*,* $\mathbf{B}^I$ *and* $\mathbf{B}^S$ *belong to* $\mathcal{C}^\alpha$ *a.s.*

Proof. We leave the details of the first part of the assertion as an exercise. Let now $\alpha \in (\frac{1}{3}, \frac{1}{2})$. Choosing $q \geqslant 2$ such that $\alpha < \beta - \frac{1}{q}$, the previous proposition yields the existence of α-Hölder continuous modifications of the processes $(B, \mathbb{B}^I)$ and $(B, \mathbb{B}^S)$ satisfying (3.2.2). Since B, $\mathbb{B}^I$ and $\mathbb{B}^S$ are continuous, these modifications are $(B, \mathbb{B}^I)$ and $(B, \mathbb{B}^S)$ themselves, possibly up to a zero set of paths. Since we already know that $(B, \mathbb{B}^I)$ and $(B, \mathbb{B}^S)$ both satisfy Chen's relation pathwise, the assertion follows. □

Remark 3.2.3. *One can further show that the maps* $\Omega \ni \omega \mapsto |||\mathbf{B}^I(\omega)|||_{\mathcal{C}^\alpha}$ *and* $\omega \mapsto |||\mathbf{B}^S(\omega)|||_{\mathcal{C}^\alpha}$ *decay exponentially fast, i.e. the probability* $\mathbb{P}\big(|||\mathbf{B}^I|||_{\mathcal{C}^\alpha} \geqslant c\big)$ *decays exponentially in* c*, and similarly for* $\mathbf{B}^S$*.*

3.2.2 Approximation of Brownian rough paths

Due to $\mathbf{B}^S \in \mathcal{C}_g^\alpha$ a.s. and Remark 2.2.3, it follows that $\mathbf{B}^S \in \mathcal{C}_g^{0,\beta}$ a.s. for all $\beta < \frac{1}{2}$, hence there is a sequence $\mathbf{B}^n = (B^n, \mathbb{B}^n)$ such that $\mathbf{B}^n \in \mathcal{L}(C^\infty)$ and $\mathbf{B}^n \xrightarrow{n\to\infty} \mathbf{B}^S$ w.r.t. ρ_β. There is no general answer to the question how to construct such $\mathbf{B}^n$.

However, from probability, we know several suitable pathwise approximations of Brownian motion, and it is a natural question whether their corresponding lifts to $\mathcal{L}(C^\infty)$ work as a choice for $\mathbf{B}^n$.

The answer to this question is "in some cases yes, but not in general". In fact, we have the following positive result.

Proposition 3.2.4. *Let* B^n *be the* n*-th step piecewise linear approximation of* B*, i.e.* $B^n_t = B_t$ *when* $t = iT2^{-n}$ *for integers* i*, and linearly interpolated inbetween. Then, defining* $\mathbf{B}^n := (B^n, \mathbb{B}^n)$*, where*

$$\mathbb{B}^n_{s,t} := \int_s^t B^n_{s,r} \otimes dB^n_r,$$

where the integral is pathwise in Riemann–Stieltjes sense, we have

$$\rho_\alpha\big(\mathbf{B}^n, \mathbf{B}^S\big) \xrightarrow{n\to\infty} 0$$

for all $\alpha < \frac{1}{2}$*.*

However, we mention without further details that there exist perfectly smooth and reasonable approxmations of Brownian motion whose $\mathcal{L}(C^\infty)$-lift does not converge to $\mathbf{B}^I$ or $\mathbf{B}^S$, but to some

$$\bar{\mathbf{B}} = (B, \bar{\mathbb{B}}), \quad \bar{\mathbb{B}}_{s,t} = \mathbb{B}^S_{s,t} + (t-s)A$$

for an antisymmetric matrix A. In particular, $\bar{\mathbf{B}}$ is not the Itô rough path lift of Brownian motion.

4 Rough integrals

In this chapter we aim to define the integral $\int Y d\mathbf{X}$, where $\mathbf{X}$ is a rough path in $\mathcal{C}^\alpha$, $\alpha \in (\frac{1}{3}, \frac{1}{2}]$, and Y is a suitable integrand. In particular, we study the space of such suitable integrands. We further establish regularity of the map $(Y, \mathbf{X}) \to \int Y d\mathbf{X}$, which is one of the initial motivations for rough integrals (compare with the introduction).

Throughout, we again let $I = [0, T], T > 0$.

4.1 Integration of one-forms: intuition

For the sake of intuition, we assume temporarily $V = \mathbb{R}^d$, i.e. X and $\mathbb{X}$ take values in $\mathbb{R}^d$ and $\mathbb{R}^{d\times d}$, respectively. Let $F : \mathbb{R}^d \to \mathbb{R}^{m\times d}$ be sufficiently regular, say $F \in C_b^2(\mathbb{R}^d, \mathbb{R}^{m\times d})$, and $X \in C^\alpha$ for some $\alpha \in (0, \frac{1}{2}]$. We write $\mathcal{P} = \{\tau_i\}_{0\leqslant i\leqslant N}$ for a finite partition of the interval under consideration, for instance $s = \tau_0 < \ldots \tau_N = t$ when we consider integrals over (s, t).

In order to define $\int_s^t F(X_r)dX_r$, a natural ansatz are limits of Riemann–Stieltjes sums, i.e.

$$\int_s^t F(X_r)dX_r = \lim_{|\mathcal{P}|\to 0} \sum_{(\tau_i,\tau_{i+1})\in\mathcal{P}} F(X_{\tau_i})X_{\tau_i,\tau_{i+1}}. \tag{4.1.1}$$

However, we know from Young's integration theory that the RHS converges in general only when $X \in C^\alpha$ with $\alpha > \frac{1}{2}$. The idea behind the previous formula is to Taylor-develop $r \mapsto F(X_r)$ up to first order on each small interval (τ_i, τ_{i+1}) from $\mathcal{P}$, i.e. for $r \in (\tau_i, \tau_{i+1})$

$$F(X_r) = F(X_{\tau_i}) + DF(X_{\tau_i})X_{\tau_i,r} + R_2(X_{\tau_i}, X_r),$$

with $|R_2(X_{\tau_i}, X_r)| \leqslant |X_{\tau_i,r}|^2$. Then one expects

$$\int_s^t F(X_r)dX_r = \sum_{i=0}^{N-1} \Big(F(X_{\tau_i})X_{\tau_i,\tau_{i+1}} + DF(X_{\tau_i}) \int_{\tau_i}^{\tau_{i+1}} X_{\tau_i,r}dX_r + \int_{\tau_i}^{\tau_{i+1}} R_2(X_{\tau_i}, X_r)dX_r \Big), \tag{4.1.2}$$

where the second and third integral are defined in Riemann–Stieltjes sense (recall that for the moment we assume $X \in C^\alpha, \alpha > \frac{1}{2}$). Note that $DF(X_{\tau_i}) \in L(\mathbb{R}^d, \mathbb{R}^{m\times d}) \cong L(\mathbb{R}^{d\times d}, \mathbb{R}^m)$, so indeed $DF(X_{\tau_i})$ acts on $\int_{\tau_i}^{\tau_{i+1}} X_{\tau_i,r}dX_r \in \mathbb{R}^{d\times d}$. Since from Young integration it follows

$$\left| \int_{\tau_i}^{\tau_{i+1}} X_{\tau_i,r} \otimes dX_r \right| \lesssim |\tau_{i+1} - \tau_i|^{2\alpha} \tag{4.1.3}$$

and

$$\left| \int_{\tau_i}^{\tau_{i+1}} R_2(X_{\tau_i}, X_r) dX_r \right| \lesssim |\tau_{i+1} - \tau_i|^{3\alpha}, \tag{4.1.4}$$

we see that, letting $\mathcal{P}_n = \{s + i(t-s)2^{-n}, 0 \leqslant i \leqslant 2^n\}$ (then $|\mathcal{P}_n| = (t-s)2^{-n}$, $\#\mathcal{P}_n = 2^n + 1$, equivalently $\mathcal{P}_n$ consists of 2^n many essentially disjoint intervals $[\tau_i, \tau_{i+1}]$ of length $(t-s)2^{-n}$)

$$\left| \lim_{n\to\infty} \sum_{\tau_i \in \mathcal{P}_n} DF(X_{\tau_i}) \int_{\tau_i}^{\tau_{i+1}} X_{\tau_i, r} dX_r \right| \lesssim \limsup_n |DF|_\infty 2^n ((t-s)2^{-n})^{2\alpha} = 0 \tag{4.1.5}$$

and similarly

$$\left| \lim_{n\to\infty} \sum_{\tau_i \in \mathcal{P}_n} \int_{\tau_i}^{\tau_{i+1}} R_2(X_{\tau_i}, X_r) dX_r \right| \lesssim \limsup_n 2^n ((t-s)2^{-n})^{3\alpha} = 0 \tag{4.1.6}$$

Thus (4.1.2) is just (4.1.1).

Now, returning to our assumption $X \in C^\alpha$, $\alpha \leqslant \frac{1}{2}$, we see that (4.1.6) still holds if and only if $\alpha > \frac{1}{3}$, while (4.1.5) fails. This shows that in order to use the same ansatz to define $\int_s^t F(X_r)dX_r$ as in the regular case above, we should assume $\alpha \in (\frac{1}{3}, \frac{1}{2}]$ and may hope at best to obtain

$$\begin{aligned} \int_s^t F(X_r)dX_r &= \lim_{|\mathcal{P}|\to 0} \sum_{\tau_i \in \mathcal{P}} \left(F(X_{\tau_i}) X_{\tau_i, \tau_{i+1}} + DF(X_{\tau_i}) \int_{\tau_i}^{\tau_{i+1}} X_{\tau_i, r} \otimes dX_r \right) \\ &= \lim_{|\mathcal{P}|\to 0} \sum_{\tau_i \in \mathcal{P}} \left(F(X_{\tau_i}) X_{\tau_i, \tau_{i+1}} + DF(X_{\tau_i}) \mathbb{X}_{\tau_i, \tau_{i+1}} \right) \end{aligned} \tag{4.1.7}$$

Note that in this case, $\int_{\tau_i}^{\tau_{i+1}} X_{\tau_i, r} \otimes dX_r$ is not defined in Riemann–Stieltjes sense.

With this viewpoint, it becomes plausible why it may be a good idea to postulate the value of $\int_s^t X_{s,r} \otimes dX_r$ via an abstract process $\mathbb{X}_{s,t}$ together with some algebraic relations between $\mathbb{X}$ and X in order to facilitate cancellations in the RHS of (4.1.7) so that the sum converges.

In fact, we shall see below that the limit (4.1.7) does exist and gives rise to our desired notion of rough integral.

For later use, we collect the following result.

Lemma 4.1.1. *Let $F : \mathbb{R}^d \to \mathbb{R}^{m\times d}$ be C_b^2, $(X, \mathbb{X}) \in \mathcal{C}^\alpha$ for some $\alpha \in (\frac{1}{3}, \frac{1}{2}]$, and set*

$$Y_t := F(X_t), \quad Y'_t := DF(X_t), \quad R^Y_{s,t} := Y_{s,t} - Y'_s X_{s,t}.$$

Then $Y \in C^\alpha(I, \mathbb{R}^{m\times d}), Y' \in C^\alpha(I, L(\mathbb{R}^d, \mathbb{R}^{m\times d}))$, $R^Y \in C_2^{2\alpha}(I, \mathbb{R}^{m\times d})$, more precisely

$$||Y||_\alpha \leqslant ||DF||_\infty ||X||_\alpha, \quad ||Y'||_\alpha \leqslant ||D^2F||_\infty ||X||_\alpha, \quad ||R^Y||_{2\alpha} \leqslant \frac{1}{2} ||D^2F||_\infty ||X||_\alpha^2.$$

Proof. $Y \in C^\alpha$ and the respective bound follows immediately by the assumption on F, mean value theorem and $X \in C^\alpha$. Similarly one obtains the claims regarding Y'. Regarding R^Y, by Taylor expansion we have, for some $\xi \in (0,1)$,

$$R^Y_{s,t} = F(X_t) - F(X_s) - DF(X_s)X_{s,t} = \frac{1}{2} D^2 F(X_s + \xi X_{s,t})\big(X_{s,t} \otimes X_{s,t}\big).$$

Since $\delta X \otimes \delta X \in C_2^{2\alpha}$ (since $|X_{s,t} \otimes X_{s,t}| \leqslant |X_{s,t}|^2$), both claims regarding R^Y follow, which concludes the proof. $\square$

4.2 Sewing lemma

We shall now formulate and proof the *sewing lemma*, which provides the technical keys to make the previous intuition, leading to the expected formula (4.1.7), rigorous.

Intuition. Formulas (4.1.1) and (4.1.7) have a common structure: There is a two-parameter map $\xi : I^2 \to \mathbb{R}^m$ for which

$$\sum_{\tau_i \in \mathcal{P}} \xi_{\tau_i, \tau_{i+1}}$$

(hopefully, for the second case) converges for every sequence of partitions $\mathcal{P}$ of (s,t) such that $|\mathcal{P}| \to 0$, with limit independent of $\mathcal{P}$, and, moreover, in the first case, $\xi_{s,t}$ is a good local approximation of the LHS in (4.1.1). Indeed, $\xi_{s,t} = F(X_s)X_{s,t}$ and $\xi_{s,t} = F(X_s)X_{s,t} + DF(X_s)\mathbb{X}_{s,t}$, respectivley, and by "good local approximation" regarding (4.1.1), we mean

$$\left| \int_s^t F(X_r)dX_r - \xi_{s,t} \right| \lesssim |t-s|^{2\alpha}, \tag{4.2.1}$$

i.e. $\xi_{s,t}$ locally approximates $(s,t) \mapsto \int_s^t F(X_r)dX_r$ better than linear (in this case we had assumed $\alpha > \frac{1}{2}$), see (4.1.3)

Note that in both cases the "germ" ξ is a non-additive two-parameter map, i.e. $\xi_{s,t} \neq \xi_{s,u} + \xi_{u,t}$, whereas in the first case the map $(s,t) \mapsto \int_s^t F(X_r)dX_r$ is clearly additive, i.e. $\int_s^t F(X_r)dX_r = \int_0^t F(X_r)dX_r - \int_0^s F(X_r)dX_r$. One calls the step of "patching together" the non-additive germ ξ to the additive limit integral map *sewing*.

Statement. We shall now see that the structure mentioned above can be realized in a much more general framework, which in particular answers affirmatively the questions regarding (4.1.7). From now on, again let V, W be Banach spaces. Derivatives between these spaces are understood in the sense of Frechet differentiability.

Denote by $C_2^{\alpha,\beta}(I,W)$ the set of functions $\xi : I^2 \to W$ such that $\xi_{t,t} = 0$ for all $t \in I$, and such that

$$||\xi||_{\alpha,\beta} := ||\xi||_\alpha + ||\delta\xi||_\beta < \infty,$$

where

$$\delta\xi : I^3 \to W, \quad \delta\xi_{s,u,t} := \xi_{s,t} - \xi_{s,u} - \xi_{u,t}, \quad ||\delta\xi||_\beta := \sup_{s<u<t} \frac{|\delta\xi_{s,u,t}|}{|t-s|^\beta}$$

(this is no abuse of notation compared to the previously introduced norm $||\xi||_\beta$, since the former is applied to three-parameter process, while the latter is applied to two-parameter processes). Note that $\xi = \delta G$ for some $G : I \to W$ if and only if $\delta\xi \equiv 0$. The following result is crucial for the construction of rough integrals.

Lemma 4.2.1 (Sewing Lemma)**.** *Let $0 < \alpha \leqslant 1 < \beta$. Then there exists a unique continuous, linear map $\mathcal{I} : C_2^{\alpha,\beta}(I, W) \to C^\alpha(I, W)$ such that $(\mathcal{I}\xi)_0 = 0$ and*

$$\left|(\mathcal{I}\xi)_{s,t} - \xi_{s,t}\right| \leqslant C|t-s|^\beta, \tag{4.2.2}$$

where $C > 0$ only depends on β and $||\delta\xi||_\beta$ (in fact, $C = C_0||\delta||_\beta$ for some $C_0 = C_0(\beta) > 0$).

Remark 4.2.2. *(i) Note that $\mathcal{I}\xi$ is a one-parameter map, hence $(\mathcal{I}\xi)_{s,t}$ means $(\mathcal{I}\xi)_t - (\mathcal{I}\xi)_s$. $\mathcal{I}$ is called* (abstract) integral map.

(ii) As will be seen from the proof, one constructs a unique two-parameter process $(s,t) \mapsto (\mathcal{I}\xi)_{s,t}$, and shows its additivity, so that one may define $(\mathcal{I}\xi)_t := (\mathcal{I}\xi)_{0,t}$. This construction does not specify $(\mathcal{I}\xi)_0$, so we choose *$(\mathcal{I}\xi)_0 := 0$. Without specifying the value at $t = 0$, $\mathcal{I}$ is not unique.*

Proof of Sewing Lemma. Regarding uniqueness, assume $I, \bar{I}$ are two (one-parameter) processes whose increments satisfy (4.2.2) (replacing $(\mathcal{I}\xi)_{s,t}$) and $I_0 = \bar{I}_0 = 0$. Then

$$|(I - \bar{I})_{s,t}| \lesssim |t-s|^\beta,$$

and since $\beta > 1$, we have $I \equiv \bar{I}$ (see (2.1.4)). In fact, (4.2.2) shows that the only candidate for $(\mathcal{I}\xi)_{s,t}$ is the Riemann-type limit

$$(\mathcal{I}\xi)_{s,t} = \lim_{|\mathcal{P}|\to 0} \sum_{(\tau_i,\tau_{i+1})\in\mathcal{P}} \xi_{\tau_i,\tau_{i+1}}. \tag{4.2.3}$$

with limit taken over any sequence of partitions $\mathcal{P}$ of $[s,t]$ such that $|\mathcal{P}| \to 0$. Indeed, at least for a sequence of dyadic partitions we see from (4.2.2)

$$\left|(\mathcal{I}\xi)_{s,t} - \sum_{(\tau_i,\tau_{i+1})\in\mathcal{P}} \xi_{\tau_i,\tau_{i+1}}\right| = \left|\sum_{(\tau_i,\tau_{i+1})\in\mathcal{P}} \left((\mathcal{I}\xi)_{\tau_i,\tau_{i+1}} - \xi_{\tau_i,\tau_{i+1}}\right)\right| \xrightarrow{|\mathcal{P}|\to 0} 0. \tag{4.2.4}$$

This shows linearity of $\xi \mapsto \mathcal{I}\xi$.

Regarding existence, fix $s, t \in I, s < t$, and let

$$\mathcal{P}_n := \{s + (t-s)i2^{-n}, i = 0, 1, \ldots, 2^n\},$$

i.e. $|\mathcal{P}_n| = (t-s)2^{-n}$ and $\#\mathcal{P}_n = 2^n + 1$. Equivalently, we consider $\mathcal{P}_n$ as the collection of the 2^n-many essentially disjoint intervals $[s+(t-s)i2^{-n}, s+(t-s)(i+1)2^{-n}]$, $i = 0, 1, \ldots, 2^n - 1$. Define $I^0_{s,t} = \xi_{s,t}$ and iteratively

$$I^{n+1}_{s,t} := \sum_{[u,v]\in\mathcal{P}_{n+1}} \xi_{u,v} = I^n_{s,t} - \sum_{[u,v]\in\mathcal{P}_n} \delta\xi_{u,m,v},$$

where for any $[u,v] \in \mathcal{P}_n$ we denote by m the unique point in $\mathcal{P}_{n+1}$ such that $[u,m], [m,v] \in \mathcal{P}_{n+1}$. Then we get

$$|I^{n+1}_{s,t} - I^n_{s,t}| \leqslant \underbrace{2^n}_{\#\text{ summands}} \underbrace{(2^{-n}|t-s|)^\beta}_{\text{length of each } [u,v]} ||\delta\xi||_\beta. \tag{4.2.5}$$

Since $\beta > 1$, the series $\sum_n |I^{n+1}_{s,t} - I^n_{s,t}|$ converges, i.e. $(I^n_{s,t})_{n\in\mathbb{N}}$ is Cauchy. We denote its limit $I_{s,t}$ and obtain

$$|I_{s,t} - \xi_{s,t}| \leqslant \sum_{n\geqslant 0} |I^{n+1}_{s,t} - I^n_{s,t}| \leqslant C||\delta\xi||_\beta |t-s|^\beta \tag{4.2.6}$$

for a universal constant depending only on β, hence we get (4.2.2). The existence part is concluded by letting $(\mathcal{I}\xi)_{s,t} := I_{s,t}$ for all $s,t \in I$. Moreover, since

$$|I_{s,t}| \leqslant |\xi_{s,t}| + C||\delta\xi||_\beta |t-s|^\beta,$$

we obtain $\mathcal{I}\xi \in C^\alpha_2(I,W)$ (since $\beta > \alpha$ and $\xi \in C^\alpha_2$). This inequality also yields the bound

$$||\mathcal{I}\xi||_\alpha \leqslant C\Big(||\xi||_\alpha + ||\delta\xi||_\beta T^{\beta-\alpha}\Big),$$

which shows boundedness of $\mathcal{I} : \xi \mapsto \mathcal{I}\xi$ from $C^{\alpha,\beta}_2(I,W)$ to $C^\alpha(I,W)$.

It remains to prove additivity of the two-parameter process $(s,t) \mapsto (\mathcal{I}\xi)_{s,t} = I_{s,t}$ in order to define $(\mathcal{I}\xi)_t := (\mathcal{I}\xi)_{0,t}$, which then concludes the proof. For notational simplicity suppose here $T = 1$, i.e. $I = [0,1]$. From the above construction it is straightforward to see that then additivity

$$I_{s,t} = I_{s,u} + I_{u,t}$$

holds for all $[s,t] = 2^{-k}[l, l+1]$ for $k \in \mathbb{N}$ and $l \in \{0, 1, \ldots, 2^k - 1\}$ with midpoint $u = \frac{s+t}{2}$. Indeed, on the level of the approximations we have $I^{n+1}_{s,t} = I^n_{s,u} + I^n_{u,t}$, and then it suffices to take limit in n on both sides. Likewise, for $[s,t] = 2^{-k}[l,m]$ with $l, m \in \{0, 1, \ldots, 2^k - 1\}$ with $m > l$, we obtain

$$I_{2^{-k}l, 2^{-k}m} = \sum_{j=l}^{m-1} I_{2^{-k}j, 2^{-k}(j+1)}. \tag{4.2.7}$$

Now additivity follows by approximating any $[s,t] \subseteq [0,1]$ by dyadic intervals $2^{-k}[l,m]$.

Finally, it should be noted that so far we established the convergence (4.2.3) only along dyadic partitions. With some more work, this can in fact be extended to *any* partition sequence with $|\mathcal{P}| \to 0$, but we omit the details here (for another proof which directly allows to obtain this convergence along all sequences of partitions, see [1]). □

Application to rough integration of one-forms. Next, we use the sewing lemma to construct the rough integral $\int F(X_s)d\mathbf{X}_s$ via the following result.

Proposition 4.2.3. *Let $\mathbf{X} = (X, \mathbb{X}) \in \mathcal{C}^\alpha(I, V)$, $\alpha \in (\frac{1}{3}, \frac{1}{2}]$ and $F \in C_b^2(V, L(V, W))$. Then the rough integral $\int_s^t F(X_r)d\mathbf{X}_r$ exists for all $s, t \in I$, is defined by the RHS of (4.1.7), and satisfies the local estimate*

$$\left| \int_s^t F(X_r)d\mathbf{X}_r - F(X_s)X_{s,t} - DF(X_s)\mathbb{X}_{s,t} \right| \leqslant C(\alpha)||F||_{C_b^2}\Big(||X||_\alpha^3 + ||X||_\alpha||\mathbb{X}||_{2\alpha}\Big)|t-s|^{3\alpha}, \tag{4.2.8}$$

where $C(\alpha) > 0$. Moreover, the map $t \mapsto \int_0^t F(X_r)d\mathbf{X}_r$ is α-Hölder continuous and

$$\left|\left| \int_0^\cdot F(X_r)d\mathbf{X}_r \right|\right|_\alpha \leqslant C||F||_{C_b^2}\Big(|||\mathbf{X}|||_\alpha \vee |||\mathbf{X}|||_\alpha^{\frac{1}{\alpha}}\Big), \tag{4.2.9}$$

where C depends only on α and I.

The natural attempt to use the germ $\xi_{s,t} = F(X_s)X_{s,t}$ for the application of the sewing lemma fails, since then $\delta\xi_{s,u,t} = [F(X_s) - F(X_u)]X_{u,t}$, and $||\delta\xi||_\beta < \infty$ if and only if $\beta < 2\alpha \leqslant 1$, whereas the sewing lemma requires $\beta > 1$. Instead, we see in the proof below that the germ from the RHS of (4.1.7) does the job.

Proof of Proposition 4.2.3. We use the notation from Lemma 4.1.1, i.e. $Y_t = F(X_t)$ and so on. Let

$$\xi_{s,t} := Y_s X_{s,t} + Y_s'\mathbb{X}_{s,t},$$

which clearly belongs to C^α. Then we see

$$\delta\xi_{s,u,t} = Y_sX_{s,t} + Y_s'\mathbb{X}_{s,t} - Y_sX_{s,u} - Y_s'\mathbb{X}_{s,u} - Y_uX_{u,t} - Y_u'\mathbb{X}_{u,t} \tag{4.2.10}$$

$$= -Y_{s,u}X_{u,t} + Y_s'(\mathbb{X}_{s,t} - \mathbb{X}_{s,u}) - Y'u\mathbb{X}_{u,t} \tag{4.2.11}$$

$$= -Y_{s,u}X_{u,t} + Y_s'(\mathbb{X}_{u,t} + X_{s,u} \otimes X_{u,t}) - Y_u'\mathbb{X}_{u,t} \tag{4.2.12}$$

$$= -R^Y_{s,u}X_{u,t} - Y'_{s,u}\mathbb{X}_{u,t}, \tag{4.2.13}$$

where we crucially used Chen's relation for the third equality. Since $\mathbb{X}, R^Y \in C_2^{2\alpha}$ and $X, Y' \in C^\alpha$ (see Lemma 4.1.1), it follows $||\delta\xi||_{3\alpha} < \infty$. Consequently, the sewing lemma yields that the rough integral

$$\int_s^t F(X_r)d\mathbf{X}_r = \lim_{|\mathcal{P}|\to 0} \sum_{\tau_i \in \mathcal{P}} \Big(F(X_{\tau_i})X_{\tau_i,\tau_{i+1}} + DF(X_{\tau_i})\mathbb{X}_{\tau_i,\tau_{i+1}}\Big), \quad s, t \in I,$$

exists along any sequence of partitions $\mathcal{P}$ of $[s, t]$ with $|\mathcal{P}| \to 0$. More precisely, we denote $\int_s^t F(X_r)d\mathbf{X}_r := (\mathcal{I}\xi)_t - (\mathcal{I}\xi)_s$, where $\mathcal{I}$ denotes the abstract integration map from the sewing lemma. The estimate (4.2.8) follows from the sewing lemma and (see Lemma 4.1.1)

$$|||\delta\xi|||_{3\alpha} \leqslant \frac{1}{2}||D^2F||_\infty||X||_\alpha^3 + ||D^2F||_\infty||X||_\alpha||\mathbb{X}||_{2\alpha}. \tag{4.2.14}$$

The α-Hölder regularity of $t \mapsto \mathcal{I}\xi_t = \int_0^t F(X_r)d\mathbf{X}_r$ follows from the sewing lemma. Regarding the estimate (4.2.9) we refer to [1]. □

4.3 Integration of controlled rough paths

We would like to integrate a larger class of integrands than just one-forms $F(X)$. To this end, the key observation is that for the proof of Proposition 4.2.3 not the one-form shape was crucial, but only that for the integrand Y there exists Y' with the properties of Lemma 4.1.1. This motivates the following definition of *controlled rough paths* (wrt. $\mathbf{X}$), which were first introduced and studied in [3].

The space of controlled rough paths. Here we assume Y takes values in some Banach space $\bar{W}$. For rough integration, we usually let $\bar{W} = L(V, W)$, where V is the state space of X.

Definition 4.3.1. Let $\alpha \in (0, \frac{1}{2}]$, $X \in C^\alpha(I, V)$. $Y \in C^\alpha(I, \bar{W})$ is a *controlled rough path* (wrt. X) or *controlled* (by X), if there is $Y' \in C^\alpha(I, L(V, \bar{W}))$ such that the remainder

$$R^Y_{s,t} := Y_{s,t} - Y'_s X_{s,t}, \quad s, t \in I,$$

belongs to $C_2^{2\alpha}(I, \bar{W})$. The space of all such pairs (Y, Y') is denoted by $\mathcal{D}_X^{2\alpha} = \mathcal{D}_X^{2\alpha}(I, \bar{W})$ and is called *space of controlled rough paths (wrt. X)*. Sometimes, we shortly write $Y \in \mathcal{D}_X^{2\alpha}$ instead of $(Y, Y') \in \mathcal{D}_X^{2\alpha}$.

To avoid confusion it should be stressed that the elements of $\mathcal{D}_X^{2\alpha}$ are *not* rough paths themselves.

Remark 4.3.2. *(i) Note that at this stage, X need not be considered as a rough path, but is merely a α-Hölder continuous path.*

(ii) In general, Y' is not uniquely determined from X and Y. We call any Y' as in the previous definition the (Gubinelli) derivative *of Y* (wrt. X). *See also (iv) below.*

(iii) Intuitively, one may think of elements in $\mathcal{D}_X^{2\alpha}$ as paths Y which "look like X" on small scales in the sense that

$$Y_{s,t} \approx Y'_s X_{s,t},$$

precisely

$$Y_{s,t} = Y'_s X_{s,t} + R^Y_{s,t}. \tag{4.3.1}$$

(iv) Since Y is only in C^α, but $R^Y \in C_2^{2\alpha}$, a proper cancellation is expected to take place in $Y_{s,t} - Y's X_{s,t}$ via Y'. On the other hand, note that if $Y \in C^{2\alpha}$, then one may even take $Y' \equiv 0$, and if also $X \in C^{2\alpha}$, then any continuous *Y' may be taken as Gubinelli derivative of Y. This fits the intuition (compare later) that "the rougher X (and Y), the more uniquely determined Y'".*

It should be noted that $\mathcal{D}_X^{2\alpha}$ (for fixed X) is a vector space (!). We endow it with the seminorm

$$||Y,Y'||_{X,2\alpha} := ||Y'||_\alpha + ||R^Y||_{2\alpha}$$

and norm

$$||Y,Y'||_{\mathcal{D}_X^{2\alpha}} := |Y_0| + |Y_0'| + ||Y,Y'||_{X,2\alpha}.$$

The answer to the natural question whether this norm also controls the α-Hölder norm of Y is positive:

Lemma 4.3.3. *For $(Y,Y') \in \mathcal{D}_X^{2\alpha}$, we have*

$$||Y||_\alpha \leqslant C||Y,Y'||_{\mathcal{D}_X^{2\alpha}},$$

where $C > 0$ only depends on I and $||X||_\alpha$.

Recall that $I = [0,T]$.

Proof. An elementary computation using (4.3.1) gives

$$\begin{aligned}||Y||_\alpha &\leqslant ||Y'||_\infty ||X||\alpha + T^\alpha ||R^Y||_{2\alpha} \leqslant |Y_0'|||X||_\alpha + T^\alpha\big(||Y'||_\alpha ||X||_\alpha + ||R^Y||_{2\alpha}\big)\\ &\leqslant (1 \vee T^\alpha)(1 + ||X||_\alpha)(|Y_0'| + ||Y,Y'||_{X,2\alpha}) \leqslant C||Y,Y'||_{\mathcal{D}_X^{2\alpha}},\end{aligned}$$

with $C = (1 \vee T^\alpha)(1 + ||X||_\alpha)$. □

It is then straightforward to check that $(\mathcal{D}_X^{2\alpha}, ||\cdot||_{\mathcal{D}_X^{2\alpha}})$ is a Banach space. Note that this is in contrast to the space of rough paths $\mathcal{C}^\alpha$, which, as we saw, is not even a vector space. However, the space $\mathcal{D}_X^{2\alpha}$ depends, of course, on X.

Rough integration of controlled rough paths. We now see that for $\alpha \in (\frac{1}{3}, \frac{1}{2}]$, $\mathbf{X} = (X, \mathbb{X}) \in \mathcal{C}^\alpha(I,V)$ and $(Y,Y') \in \mathcal{D}_X^{2\alpha}(I, L(V,W))$ we can define the rough integral $\int_s^t Y_r d\mathbf{X}_r$. The one-form case $Y_t = F(X_t)$, $Y_t' = DF(X_t)$ treated in the previous section suggests to try the definition

$$\int_s^t Y_r d\mathbf{X}_r = \lim_{|\mathcal{P}|\to 0} \sum_{(\tau_i,\tau_{i+1})\in\mathcal{P}} \Big(Y_{\tau_i} X_{\tau_i,\tau_{i+1}} + Y_{\tau_i}' \mathbb{X}_{\tau_i,\tau_{i+1}}\Big). \tag{4.3.2}$$

In other words, we are tempted to apply the sewing lemma to the germ

$$\xi_{s,t} := Y_s X_{s,t} + Y_s' \mathbb{X}_{s,t}. \tag{4.3.3}$$

Here we have $Y_t' \in L(V, L(V,W)) \cong L(V \otimes V, W)$, so that $Y_t' \mathbb{X}_{s,t} \in W$. Indeed, this is the correct idea as the following result shows.

Theorem 4.3.4. *Let $\mathbf{X}$ and (Y,Y') be as at the beginning of this paragraph. There exists the rough integral (4.3.2) for all $s,t \in I$ with the local estimate*

$$\left|\int_s^t Y_r d\mathbf{X}_r - Y_s X_{s,t} - Y_s' \mathbb{X}_{s,t}\right| \leqslant C\big(||X||_\alpha ||R^Y||_{2\alpha} + ||\mathbb{X}||_{2\alpha} ||Y'||_\alpha\big)|t-s|^{3\alpha}, \tag{4.3.4}$$

where $C > 0$ only depends on α. The map $t \mapsto \int_0^t Y_r d\mathbf{X}_r$ belongs to C^α.

Even more, $\int_0^\cdot Y_r d\mathbf{X}_r \in \mathcal{D}_X^{2\alpha}(I, W)$, a Gubinelli derivative is given by Y, and the map

$$(Y, Y') \mapsto \Big(\int_0^\cdot Y_r d\mathbf{X}_r, Y \Big)$$

from $\mathcal{D}_X^{2\alpha}(I, L(V, W))$ to $\mathcal{D}_X^{2\alpha}(I, W)$ is a linear continuous map between Banach spaces.

We note that it is clear from (4.3.2) that the rough integral $\int_0^\cdot Y_r d\mathbf{X}_r$ depends on Y', even though this is not visible in the notation. From this viewpoint, more accurately one should write $\int_0^\cdot (Y, Y') d\mathbf{X}_r$.

Similarly, the rough integral depends on $\mathbb{X}$, as the following example shows.

Example 4.3.5. Let $f \in C^{2\alpha}(I, V \otimes V)$, $\mathbf{X}, \bar{\mathbf{X}} \in \mathcal{C}^\alpha(I, V)$ with $\mathbf{X} = (X, \mathbb{X})$ and $\bar{\mathbf{X}} = (\bar{X}, \bar{\mathbb{X}})$ such that

$$X = \bar{X}, \quad \bar{\mathbb{X}}_{s,t} = \mathbb{X}_{s,t} + f_{s,t}.$$

Let (Y, Y') belong to $\mathcal{D}_X^{2\alpha}(I, L(V, W))$ ($= \mathcal{D}_{\bar{X}}^{2\alpha}(I, L(V, W))$, since $X = \bar{X}$). Then it follows from (4.3.2)

$$\int_s^t Y_r d\bar{\mathbf{X}}_r = \int_s^t Y_r d\mathbf{X}_r + \int_s^t Y_r' df_r,$$

where the last integral is a Riemann–Stieltjes integral.

Hence the rough integral depends on $\mathbb{X}$ (but the space of controlled rough paths wrt. X does not depend on $\mathbb{X}$).

Proof of Theorem 4.3.4. The existence of the rough integral (4.3.2) follows from the sewing lemma with the germ (4.3.3) exactly as in the construction of the rough integral for one-forms, compare Proposition 4.2.3. Indeed, that proof can simply be repeated, since we had used only regularity properties of Y, Y' and R^Y from Lemma 4.1.1, which now hold (with different bounds, not expressed in terms of some F, of course) since $(Y, Y') \in \mathcal{D}_X^{2\alpha}$. For the bound in (4.3.4), recall that the sewing lemma ($\beta = 3\alpha$) gives the upper bound $C_0 ||\delta\xi||_\beta |t - s|^\beta$, where $C_0 > 0$ only depends on α. In the present case, since

$$\delta\xi_{s,u,t} = -R_{s,u}^Y X_{u,t} - Y_{s,u}' \mathbb{X}_{u,t},$$

clearly

$$||\delta\xi||_{3\alpha} \leqslant ||X||_\alpha ||R^Y||_{2\alpha} + ||\mathbb{X}||_{2\alpha} ||Y'||_\alpha,$$

whereby (4.3.4) follows. That $t \mapsto \int_0^t Y_r d\mathbf{X}_r$ is α-Hölder continuous follows as part of the sewing lemma.

The local estimate (4.3.4) in particular implies $\int_s^t Y_r d\mathbf{X}_r - Y_s X_{s,t} = Y_s' \mathbb{X}_{s,t} + \tilde{C} |t - s|^{3\alpha}$, where $\tilde{C}$ denotes the factor on the RHS of (4.3.4). Since $\mathbb{X} \in C_2^{2\alpha}$, this implies

$$\int_0^t Y_r d\mathbf{X}_r \in \mathcal{D}_X^{2\alpha}(I, W)$$

with Gubinelli derivative Y, hence we establish the map

$$(Y, Y') \mapsto \Big(\int_0^{\cdot} Y_r d\mathbf{X}_r, Y \Big)$$

between $\mathcal{D}_X^{2\alpha}(I, L(V, W))$ and $\mathcal{D}_X^{2\alpha}(I, W)$ as claimed. The linearity of this map follows from the linearity of the map $Y \mapsto \int_0^t Y_r d\mathbf{X}_r$, which is immediate from its construction. As a linear map between Banach spaces, its continuity is equivalent to its boundedness, which follows from the estimate (we abbreviate $(Z, Z') = (\int_0^{\cdot} Y_r d\mathbf{X}_r, Y)$)

$$\begin{aligned} ||Z, Z'||_{\mathcal{D}_X^{2\alpha}} &\leqslant |Y_0| + ||Y||_\alpha + ||Y'||_\infty ||\mathbb{X}||_{2\alpha} + CT^\alpha \big(||X||_\alpha ||R^Y||_{2\alpha} + ||\mathbb{X}||_{2\alpha} ||Y'||_\alpha \big) \\ &\big(\leqslant C ||Y, Y'||_{\mathcal{D}_X^{2\alpha}} \big) \end{aligned}$$

where here $C > 0$ is a constant changing from line to line, independent from Z, Z', and in the last line depending on $X, \mathbb{X}$, T and α (recall that $(X, \mathbb{X})$ is regarded as fixed). The second equality is obvious from the definition of the norm $||\cdot||_{\mathcal{D}_X^{2\alpha}}$, hence it remains to prove the first inequality. This is left as a straightforward exercise (use the definition of the appearing norms together with (4.3.4)). $\square$

A natural question is: When X and Y are sufficiently smooth (say, C^1) and $\mathbb{X}$ is defined as the iterated integral of X (i.e. $(X, \mathbb{X}) \in \mathcal{L}(C^1)$), does the rough integral $\int Y_r d\mathbf{X}_r$ then coincide with the usual Riemann-Stieltjes integral of Y against X (and, in particular, is the rough integral in this case independent of Y'?)? The answer is positive:

Remark 4.3.6. *In the notation of the proof of the sewing lemma, if for two germs ξ and $\bar{\xi}$ we have $\xi - \bar{\xi} \in C_2^\beta$ for some $\beta > 1$, then the corresponding abstract integrals obtained by the sewing lemma coincide, i.e.*

$$\mathcal{I}\xi = \mathcal{I}\bar{\xi}.$$

Indeed, letting $t \in I$ and considering a sequence of dyadic partitions $(\mathcal{P}_n)_{n \in \mathbb{N}}$ of $[0, t]$, so that $|\mathcal{P}_n| = t2^{-n}$ and $\#\mathcal{P}_n = 2^n + 1$, it follows

$$|(\mathcal{I}\xi)_t - (\mathcal{I}\bar{\xi})_t| \leqslant \lim_n \sum_{(\tau_i, \tau_{i+1}) \in \mathcal{P}_n} \big| \xi_{\tau_i, \tau i+1} - \bar{\xi}_{\tau_i, \tau_{i+1}} \big| \leqslant C \lim_n 2^{n(1-\beta)} = 0,$$

where $C > 0$ depends only on t and $||\xi - \bar{\xi}||_\beta$.

In particular, this shows that in the case where Y and X are C^1, the usual Riemann-Stieltjes integral $\int Y dX$ (with germ $\xi_{s,t} = Y_s X_{s,t}$) coincides with the rough integral $\int Y d\mathbf{X}$, where $\mathbf{X} = (X, \int X \otimes dX)$ (with germ $\xi_{s,t} = Y_s X_{s,t} + (DY)_s \int_s^t X_{s,r} \otimes dX_r$; note that the latter summand is in C_2^β for some $\beta > 1$). In this case, it follows that the rough integral is independent from Y'.

4.4 Stability of the rough integral

Again, let $\alpha \in (\frac{1}{3}, \frac{1}{2}]$. The final goal of this chapter is to study the regularity of the map $(Y, \mathbf{X}) \mapsto \int Y_r d\mathbf{X}_r$. Recall that for the usual Riemann–Stieltjes integral, this map (say, on $C^1 \times C^1$) is not continuous wrt. the uniform topology.

Let $\mathbf{X} = (X, \mathbb{X})$ and $\bar{\mathbf{X}} = (\bar{X}, \bar{\mathbb{X}})$ be rough paths in $\mathcal{C}^\alpha(I, V)$, $(Y, Y') \in \mathcal{D}_X^{2\alpha}(I, L(V, W))$, $(\bar{Y}, \bar{Y}') \in \mathcal{D}_{\bar{X}}^{2\alpha}(I, L(V, W))$, set

$$(Z, Z') = \Big(\int_0^\cdot Y_r d\mathbf{X}_r, Y \Big),$$

and similarly for $(\bar{Z}, \bar{Z}')$. As before, write

$$R^Y_{s,t} = Y_{s,t} - Y'_s X_{s,t},$$

and similarly for $R^{\bar{Y}}$.

Since (Y, Y') and $(\bar{Y}, \bar{Y}')$ live in *different* Banach spaces, the notion of "distance" between them is not meaningful in the usual sense. Nonetheless, the quantity

$$||Y, Y'; \bar{Y}, \bar{Y}'||_{X, \bar{X}, 2\alpha} := ||Y' - \bar{Y}'||_\alpha + ||R^Y - R^{\bar{Y}}||_{2\alpha} \tag{4.4.1}$$

will be useful. This is *not* a proper metric even for $X = \bar{X}$ (take (Y, Y') and $(Y + cX + \bar{c}, Y' + c))$ for any $c, \bar{c} \in \mathbb{R}$.

We state the following stability result for the rough integrals $Z, \bar{Z}$ without proof (see Thm.4.17 in [1]).

Proposition 4.4.1. *Let $M \in (0, \infty)$ such that*

$$|Y'_0| + ||Y, Y'||_{X, 2\alpha} \leqslant M, \quad ||X||_\alpha + ||\mathbb{X}||_{2\alpha} \leqslant M,$$

with identical bounds for $(\bar{Y}, \bar{Y}')$ and $(\bar{X}, \bar{\mathbb{X}})$. Then

$$||Z, Z'; \bar{Z}, \bar{Z}'||_{X, \bar{X}, 2\alpha} \leqslant C\big(\rho_\alpha(\mathbf{X}, \bar{\mathbf{X}}) + |Y'_0 - \bar{Y}'_0| + T^\alpha ||Y, Y'; \bar{Y}, \bar{Y}'||_{X, \bar{X}, 2\alpha}\big), \tag{4.4.2}$$

and also

$$||Z - \bar{Z}||_\alpha \leqslant C\Big(\rho_\alpha(\mathbf{X}, \bar{\mathbf{X}}) + |Y_0 - \bar{Y}_0| + |Y'_0 - \bar{Y}'_0| + T^\alpha ||Y, Y'; \bar{Y}, \bar{Y}'||_{X, \bar{X}, 2\alpha}\Big), \tag{4.4.3}$$

where $C > 0$ only depends on α and M.

Remark 4.4.2. *In particular, in the case $\mathbf{X} = \bar{\mathbf{X}}$ (i.e. (Y, Y') and $(\bar{Y}, \bar{Y}')$ live in the same Banach space) we have*

$$||Z - \bar{Z}||_\alpha \leqslant C\Big(|Y'_0 - \bar{Y}'_0| + ||Y' - \bar{Y}'||_\alpha + ||R^Y - R^{\bar{Y}}||_{2\alpha}\Big),$$

where $C = C(T, \alpha, M)$. It is very important to note that the constant C depends on both integrands and integrators under consideration and is hence not *uniform among $\mathbf{X} \in \mathcal{C}^\alpha$ and $(Y, Y') \in \mathcal{D}_X^{2\alpha}$.*

4.5 True roughness

Let $\alpha \in (\frac{1}{3}, \frac{1}{2}]$, $\mathbf{X} \in \mathcal{C}^\alpha(I, V)$, $(Y, Y') \in \mathcal{D}_X^{2\alpha}(I, \bar{W})$. Here we answer the question:

To which extent (or under which assumptions) is the Gubinelli derivative Y' uniquely determined by X and Y?

First, observe that Y' is in general not uniquely determined: If $Y, X \in C^{2\alpha}$, then *any* continuous Y' renders R^Y 2α-Hölder continuous. If $Y \in C^{2\alpha}$ and $X \in C^\alpha$, then $Y' \equiv 0$ works.

We shall see now that a sufficient criterion for uniqueness of Y' is that "X varies non-smoothly in all directions at all times". For the sake of intuition, let us first consider the case $V = \bar{W} = \mathbb{R}^1$:

Let $t \in I$ and assume there is a sequence $t_n \searrow t$ such that

$$\frac{|X_{t,t_n}|}{|t_n - t|^{2\alpha}} \xrightarrow{n\to\infty} \infty.$$

Then, for large n, $X_{t,t_n} \neq 0$, and by definition of R^Y, we find

$$Y'_t = \frac{Y_{t,t_n}}{X_{t,t_n}} - \frac{R^Y_{t,t_n}}{|t - t_n|^{2\alpha}} \frac{|t_n - t|^{2\alpha}}{X_{t,t_n}}.$$

Therefore, $\lim_n \frac{Y_{t,t_n}}{X_{t,t_n}}$ exists and equals Y'_t, since both factors of the second summand on the above RHS converge as $n \to \infty$, with limit 0 for the second one (recall $R^Y \in C_2^{2\alpha}$). This suggests the following multidimensional definition.

Definition 4.5.1. Let $X \in C^\alpha(I, V)$ and $t \in I$. X is called *rough at time t*, if

$$\forall v^* \in V^* \backslash \{0\} : \quad \limsup_{\tau \searrow t} \frac{|v^*(X_{t,\tau})|}{|\tau - t|^{2\alpha}} = \infty. \tag{4.5.1}$$

X is called *truly rough*, if it is rough at time t for all $t \in D$ from a dense set $D \subseteq I$.

From what we observed above, it follows that if $V = \mathbb{R}^1$ and X is truly rough, then Y' is uniquely determined from Y and X (first, Y'_t is uniquely determined on a dense set $t \in D$, and hence for all t by continuity). That this is also true in the case of general V is contained in the next result.

Proposition 4.5.2. *If X is rough at time $t \in I$, then for any $(Y, Y') \in \mathcal{D}_X^{2\alpha}(I, \bar{W})$ we have*

$$\limsup_{\tau \searrow t} \frac{|Y_{t,\tau}|}{|\tau - t|^{2\alpha}} < \infty \implies Y'_t = 0. \tag{4.5.2}$$

Consequently, if X is truly rough and $(Y, Y'), (Y, \tilde{Y}') \in \mathcal{D}_X^{2\alpha}$, then $Y' \equiv \tilde{Y}'$.

The following corollary follows from the previous proposition by considering $\left(\int_0^\cdot Y d\mathbf{X}, Y\right) \in \mathcal{D}_X^{2\alpha}(I, W)$.

Corollary 4.5.3. *If X is rough at time $t \in I$ and $(Y, Y') \in \mathcal{D}^{2\alpha}(I, L(V, W))$, then*

$$\limsup_{\tau \searrow t} \frac{\int_t^\tau Y_s d\mathbf{X}_s}{|\tau - t|^{2\alpha}} < \infty \implies Y_t = 0. \tag{4.5.3}$$

Proof of Proposition 4.5.2. The second assertion follows by applying the first part to $(0, Y' - \tilde{Y}') \in \mathcal{D}_X^{2\alpha}$ and by continuity of Y' and $\tilde{Y}'$. Regarding the first part, by assumption we have

$$\frac{Y'_t X_{t,\tau}}{|\tau - t|^{2\alpha}} = \frac{Y_{t,\tau}}{|\tau - t|^{2\alpha}} - \frac{R^Y_{t,\tau}}{|\tau - t|^{2\alpha}},$$

where the RHS is bounded as $\tau \searrow t$. For every $w^* \in \bar{W}^*$, one has an element $v^* \in V^*$ defined via $v^* := w^* \circ Y'_t$, and from above we see that

$$\limsup_{\tau \searrow t} \frac{|v^*(X_{t,\tau})}{|\tau - t|^{2\alpha}} = \limsup_{\tau \searrow t} \frac{|w^*(Y'_t X_{t,\tau})|}{|\tau - t|^{2\alpha}}$$

is finite. On the other hand, unless $v^* = 0$, roughness of X at t yields that the LHS above is equal to $+\infty$. Thus, $v^* = 0$, i.e.

$$w^*(Y'_t v) = 0, \quad \forall v \in V, w^*,$$

which clearly implies $Y'_t = 0$. □

As a consequence, we obtain a rough analogue to the classical Doob–Meyer decomposition. To this end, first recall from classical stochastic analysis, say for $V = \mathbb{R}^d$:

Theorem 4.5.4 (Doob–Meyer for continuous semimartingales)**.** *Let N be a continuous semimartingale, i.e. there exists a continuous local martingale M and a continuous adapted process A with paths of locally finite variation such that $N = M + A$. If $(\tilde{M}, \tilde{A})$ is another such pair, then $M \equiv \tilde{M}$ and $A \equiv \tilde{A}$ a.s.*

It is left as a short exercise that this implies the following: If B is a d-dimensional Brownian motion and $Y, Z, \tilde{Y}, \tilde{Z}$ are continuous stochastic processes such that a.s.

$$\int_0^\cdot Y dB + \int_0^\cdot Z dt = \int_0^\cdot \tilde{Y} dB + \int_0^\cdot \tilde{Z} dt, \quad \text{on } [0, T]. \tag{4.5.4}$$

Then $(Y_t)_{t \in [0,T]} = (\tilde{Y}_t)_{t \in [0,T]}$ and $(Z_t)_{t \in [0,T]} = (\tilde{Z}_t)_{t \in [0,T]}$ a.s. Now we prove a similar result in the rough case.

Proposition 4.5.5 (Doob–Meyer for rough paths)**.** *Let X be truly rough, $(Y, Y'), (\tilde{Y}, \tilde{Y}') \in \mathcal{D}_X^{2\alpha}(I, L(V, W))$ and $Z, \tilde{Z} \in C(I, W)$. Then*

$$\int_0^t Y_s d\mathbf{X}_s + \int_0^t Z_s ds = \int_0^t \tilde{Y}_s d\mathbf{X}_s + \int_0^t \tilde{Z}_s ds, \quad \forall t \in I \tag{4.5.5}$$

implies $(Y, Y') = (\tilde{Y}, \tilde{Y}')$ and $Z = \tilde{Z}$ on I.

Proof. Let $t \in I$ such that X is rough at time t. Then

$$\limsup_{\tau \searrow t} \frac{\int_t^\tau Y_s - \tilde{Y}_s d\mathbf{X}_s}{|\tau - t|^{2\alpha}} = \limsup_{\tau \searrow} \frac{\int_t^\tau Z_s - \tilde{Z}_s ds}{|\tau - t|^{2\alpha}} < \infty, \tag{4.5.6}$$

since $r \mapsto \int_0^r Z_s - \tilde{Z}_s ds \in C^1(I, W)$ by continuity of $t \mapsto Z_t - \tilde{Z}_t$ (recall $\alpha \leqslant \frac{1}{2}$). Hence $Y_t = \tilde{Y}_t$ by the previous corollary. Since the set of such t is dense in $[0, T]$ and $Y, \tilde{Y}$ are continuous, we get $Y = \tilde{Y}$. But then the Gubinelli derivative of $Y - \tilde{Y}$ equals 0 (and is unique) by Proposition 4.5.2. So, $(Y, Y') = (\tilde{Y}, \tilde{Y}')$, and hence

$$\int_0^\cdot Y_s d\mathbf{X}_s = \int_0^\cdot \tilde{Y}_s d\mathbf{X}_s.$$

But then (4.5.6) implies

$$\int_0^t Z_s ds = \int_0^t \tilde{Z}_s ds, \quad \forall t \in [0, T], \tag{4.5.7}$$

thus $Z_t - \tilde{Z}_t = 0$ for all $t \in [0, T]$. □

Brownian motion is truly rough. We conclude this section with the following true roughness result for d-dimensional Brownian motion. For the proof, which is obtained by using the law of iterated logarithm, we refer to [1, Sect.6.3].

Proposition 4.5.6. *Let $V = \mathbb{R}^d$ and B be a standard d-dimensional Brownian motion. Then for a.e. ω, $B(\omega)$ is truly rough with regard to any $\alpha \in [\frac{1}{4}, \frac{1}{2})$.*

5 Comparison with stochastic analysis

Having defined integrals of controlled rough paths against deterministic rough paths, with our primary example of the latter being Brownian paths, next we want to compare these rough integrals with the probabilistic construction of stochastic integrals. Moreover, we discuss a deterministic "rough Itô formula" and compare it with the usual Itô formula.

5.1 Rough and stochastic integrals

Here we let $V = \mathbb{R}^d$, $B = (B^1, \dots, B^d)$ be a d-dim. standard Brownian motion with α-Hölder continuous paths for every $\alpha \in (0, \frac{1}{2})$, and $I = [0,T]$ as before. We work on a stochastic basis $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\in[0,T]}, P)$ and denote by $N_1, N_2, \dots$ suitable P-null sets. Recall that "stochastic basis" means the filtration is right-continuous and all subsets of P-null sets belong to $\mathcal{F}$ and $\mathcal{F}_0$.

We consider the Itô Brownian rough path lift $\mathbb{B}^I$, i.e. $\mathbf{B}^I(\omega) = (B(\omega), \mathbb{B}^I(\omega)) \in \mathcal{C}^\alpha(I, \mathbb{R}^d)$ for all $\omega \in N_1^c$. Let $\alpha \in (\frac{1}{3}, \frac{1}{2}]$ and Y and Y' be stochastic processes with values in $\mathbb{R}^{m\times d}$ and $L(\mathbb{R}^{d\times d}, \mathbb{R}^m)$, respectively. Assume that $(Y(\omega), Y'(\omega)) \in \mathcal{D}^{2\alpha}_{B(\omega)}(I, \mathbb{R}^{m\times d})$ for all $\omega \in N_2^c$. From the previous chapter we know that the rough integral of $Y(\omega)$ against $\mathbb{B}(\omega)$ is defined for all $\omega \in N_3^c$, where $N_3 = N_1 \cup N_2$ as

$$\int_s^t Y_r(\omega) d\mathbf{B}_r(\omega) = \lim_{|\mathcal{P}|\to 0} \sum_{(\tau_i,\tau_{i+1})\in\mathcal{P}} \Big(Y_{\tau_i}(\omega) B_{\tau_i,\tau_{i+1}}(\omega) + Y'_{\tau_i}(\omega) \mathbb{B}^I_{\tau_i,\tau_{i+1}}(\omega) \Big).$$

On the other hand, if in addition Y is adapted, the stochastic Itô integral $\int_0^\cdot Y_r dB_r$ is defined as a stochastic process

$$(t,\omega) \mapsto \Big(\int_0^t Y_r dB_r^I \Big)(\omega).$$

The natural question arises whether these two notions of integrals coincide almost surely. The answer is positive:

Proposition 5.1.1. *Let $\mathbf{B}^I$, (Y, Y') be as above and assume in addition that (Y, Y') is $(\mathcal{F}_t)$-adapted. Then*

$$\int_0^t Y_r(\omega) d\mathbf{B}_r^I(\omega) = \Big(\int_0^t Y_r dB_r \Big)(\omega), \quad \forall t \in [0,T]$$

for all $\omega \in N_4^c$, where the RHS denotes the stochastic Itô integral.

Proof. Since both sides of the claimed equality are continuous in t for P-a.e. ω, it is sufficient to prove

$$\int_0^t Y_r(\omega) d\mathbf{B}_r^I(\omega) = \Big(\int_0^t Y_r dB_r \Big)(\omega)$$

a.s. for any fixed $t \in [0,T]$. Thus, we may without loss of generaliy take $I = [0,1]$ and $t = 1$. Recall that for any continuous adapted process Y the stochastic Itô integral $\int Y dB$ has the representation

$$\int_0^1 Y_r dB_r = \lim_{n\to\infty} \sum_{(\tau_i,\tau_{i+1})\in \mathcal{P}_n} Y_{\tau_i} B_{\tau_i,\tau_{i+1}} \tag{5.1.1}$$

along any partition sequence $(\mathcal{P}_n)_{n\in\mathbb{N}}$ of $[0,1]$ such that $|\mathcal{P}_n| \to 0$ as $n \to \infty$, where the limit is taken in probability (wrt. the measure P). Thus we may pass to a subsequence, again denoted $(\mathcal{P}_n)$, along which the convergence holds almost surely, say on N_5^c. Set $N_6 := N_3 \cup N_5$. For simplicity, let us assume there is a deterministic constant $M > 0$ such that

$$\sup_{\omega\in N_6^c} |Y'(\omega)|_\infty \leqslant M.$$

The general case then follows by localization. We have

$$\int_0^1 Y_r(\omega) d\mathbf{B}_r^I(\omega) - \Big(\int_0^1 Y_r dB_r \Big)(\omega) = \lim_{n\to\infty} \sum_{(\tau_i,\tau_{i+1})\in \mathcal{P}_n} Y'_{\tau_i}(\omega) \mathbb{B}^I_{\tau_i,\tau_{i+1}}(\omega) \tag{5.1.2}$$

for all $\omega \in N_6^c$. In particular the limit on the RHS exists a.s. Let $\mathcal{P} = \{0 = \tau_0 < \cdots < \tau_N = 1\}$ be any partition of $[0,1]$ and define a discrete-time martingale S via $S_0 := 0$ and its increments $S_{k+1} - S_k := Y'_{\tau_k} \mathbb{B}^I_{\tau_k,\tau_{k+1}}$, for $k \in \{0, \dots, N-1\}$. Since the definition of $\mathbb{B}^I$ entails that $|\mathbb{B}^I_{\tau_k,\tau_{k+1}}|^2_{L^2}$ is proportional to $|\tau_{k+1} - \tau_k|^2$, we find, using the L^2-orthogonality of martingale increments,

$$\begin{aligned}
\Big| \sum_{(\tau_i,\tau_{i+1})\in \mathcal{P}} Y'_{\tau_i} \mathbb{B}^I_{\tau_i,\tau_{i+1}} \Big|^2_{L^2} &= \Big| \sum_{k=0}^{N-1} (S_{k+1} - S_k) \Big|^2_{L^2} = \sum_{k=0}^{N-1} |S_{k+1} - S_k|^2_{L^2} \\
&\leqslant M^2 \sum_{k=0}^{N-1} |\mathbb{B}^I_{\tau_k,\tau_{k+1}}|^2_{L^2} \leqslant CM^2 \sum_{k=0}^{N-1} |\tau_{k+1} - \tau_k|^2 \leqslant CM^2 |\mathcal{P}|.
\end{aligned}$$

Thus

$$\sum_{(\tau_i,\tau i+1)\in \mathcal{P}_n} Y'_{\tau_i} \mathbb{B}^I_{\tau_i,\tau_{i+1}} \xrightarrow{n\to\infty} 0$$

in L^2 for any sequence $(\mathcal{P}_n)$ with $|\mathcal{P}_n| \to 0$. Consequently the limit on the RHS of (5.1.2) is 0 for all $\omega \in N_7^c \subseteq N_6^c$, and this concludes the proof. $\square$

A similar result is true for the Stratonovich rough path lift $\mathbb{B}^S$ of Brownian motion and the Stratonovich stochastic integral against Brownian motion. Recall that the latter integral is defined for integrands Y as above (for which in addition the quadratic covariation $\langle Y, B\rangle$ exists) as

$$\int_0^t Y_r \circ dB_r = \int_0^t Y_r dB_r + \frac{1}{2}\langle Y, B\rangle_t.$$

Then we have:

Proposition 5.1.2. *Let (Y, Y') be as in the previous proposition and assume in addition that the quadratic covariation $\langle Y, B\rangle$ exists on I. Then a.s.*

$$\int_0^t Y_r(\omega) d\mathbf{B}^S_r(\omega) = \Big(\int_0^t Y_r \circ dB_r\Big)(\omega), \quad \forall t \in [0, T].$$

We leave the proof to the reader, or see [1, Cor.5.2].

5.2 Rough Itô formula

Classical Itô formula. For the moment, let $V = \mathbb{R}^d$. Recall that for $X \in C(I, \mathbb{R}^d)$ and $G \in C_b^2(\mathbb{R}^d)$ the usual Itô formula asserts

$$G(X_t) - G(X_s) = \int_s^t (DG(X_r), dX_r) + \frac{1}{2}\int_s^t D^2G(X_r) : d\langle X, X\rangle_r, \quad \forall s, t \in I, \tag{5.2.1}$$

where we write $\langle X, Y\rangle$ for the $d \times d$-valued quadratic covariation of two $\mathbb{R}^d$-valued paths X and Y, with entries $(\langle X, Y\rangle)_{ij} = \langle X^i, Y^j\rangle$, and for two $\mathbb{R}^{d\times d}$-valued paths G, H the integral $\int_s^t F_r : dH_r$ is defined as

$$\int_s^t F_r : dH_r := \sum_{i,j\leqslant d}\int_s^t F_r^{ij} dH_r^{ij}, \tag{5.2.2}$$

and each of the summands on the RHS is assumed to be defined in Riemann–Stieltjes sense. The first integral on the RHS of (5.2.1) is defined as

$$\int_s^t (DG(X_r), dX_r) = \lim_{|\mathcal{P}|\to 0}\sum_{i\leqslant d}\sum_{(\tau_i,\tau_{i+1})\in\mathcal{P}}(\partial_i G)(X_{\tau_i})X^i_{\tau_i,\tau_{i+1}}. \tag{5.2.3}$$

Recall that for $X \in C^1(I, \mathbb{R}^d)$ one has $\langle X, X\rangle \equiv 0$ and hence retrieves the usual first-order chain rule

$$G(X_t) - G(X_s) = \int_s^t (DG(X_r), dX_r).$$

Integration of gradient one-forms. A closer look at the second integral on the RHS of (5.2.1) shows

$$\begin{aligned}\frac{1}{2}\int_s^t D^2G(X_r) : d\langle X, X\rangle_r &= \frac{1}{2}\lim_{|\mathcal{P}|\to 0}\sum_{k,l\leqslant d}\sum_{(\tau_i,\tau_{i+1})\in\mathcal{P}} \partial_{kl}G(X_{\tau_i})X^k_{\tau_i,\tau_{i+1}}X^l_{\tau_i,\tau_{i+1}} \\ &= \lim_{|\mathcal{P}|\to 0}\frac{1}{2}\sum_{(\tau_i,\tau_{i+1})\in\mathcal{P}} D^2G(X_{\tau_i})X_{\tau_i,\tau_{i+1}}\otimes X_{\tau_i,\tau_{i+1}} \\ &= \lim_{|\mathcal{P}|\to 0}\sum_{(\tau_i,\tau_{i+1})\in\mathcal{P}} D^2G(X_{\tau_i})\big(\frac{1}{2}X_{\tau_i,\tau_{i+1}}\otimes X_{\tau_i,\tau_{i+1}}\big)\end{aligned}$$

(Since $D^2G(X_{\tau_i}) \in L(\mathbb{R}^{d\times d},\mathbb{R})$ and $X_{\tau_i,\tau_{i+1}}\otimes X_{\tau_i,\tau_{i+1}} \in \mathbb{R}^{d\times d}$, the summands are $\mathbb{R}$-valued). With this in mind, plus the fact that $D^2G(x)$ is a symmetric $d\times d$-matrix for every $x\in\mathbb{R}^d$, we observe: If $\mathbf{X}\in\mathcal{C}^\alpha_g(I,\mathbb{R}^d)$ and $F = DG$ for some $G\in C^3_b(\mathbb{R}^d,\mathbb{R})$ (i.e. $F\in C^2_b(\mathbb{R}^d,\mathbb{R}^d)$), then (denoting by $(\cdot,\cdot)$ the Euclidean inner product on $\mathbb{R}^d$)

$$\begin{aligned}G(X_t) - G(X_s) &= \lim_{|\mathcal{P}|\to 0}\sum_{(\tau_i,\tau_{i+1})\in\mathcal{P}}\Big(\big(DG(X_{\tau_i}), X_{\tau_i,\tau_{i+1}}\big) + D^2G(X_{\tau_i})\big(\frac{1}{2}X_{\tau_i,\tau_{i+1}}\otimes X_{\tau_i,\tau_{i+1}}\big)\Big) \\ &= \lim_{|\mathcal{P}|\to 0}\sum_{(\tau_i,\tau_{i+1})\in\mathcal{P}}\Big(\big(DG(X_{\tau_i}), X_{\tau_i,\tau_{i+1}}\big) + D^2G(X_{\tau_i})\mathbb{X}_{\tau_i,\tau_{i+1}}\Big) \\ &= \lim_{|\mathcal{P}|\to 0}\sum_{(\tau_i,\tau_{i+1})\in\mathcal{P}}\Big(\big(F(X_{\tau_i}), X_{\tau_i,\tau_{i+1}}\big) + DF(X_{\tau_i})\mathbb{X}_{\tau_i,\tau_{i+1}}\Big) \\ &= \int_s^t F(X_r)d\mathbf{X}_r = \int_s^t DG(X_r)d\mathbf{X}_r,\end{aligned}$$

where $\mathcal{P}$ denotes any partition of $[s,t]$, for $s,t\in I$. Here the first equality is due to the usual Itô formula (see above), and the second one follows from

$$\mathbb{X}_{u,v} = \mathrm{Sym}(\mathbb{X}_{u,v}) + \mathrm{Ant}(\mathbb{X}_{u,v}),$$

$\mathrm{Sym}(\mathbb{X}_{u,v}) = \frac{1}{2}X_{u,v}\otimes X_{u,v}$ (due to the geometricity of $\mathbf{X}$) and the fact that the inner product of a symmetric ($D^2G(X_{\tau_i})$) and an antisymmetric ($\mathrm{Ant}(\mathbb{X}_{\tau_i,\tau_{i+1}})$) $d\times d$-matrix vanishes. The fourth equality is due to the definition of the rough integral $\int_s^t F(X_r)d\mathbf{X}_r$. Thus we have shown the following result, at least in the case $V=\mathbb{R}^d$, $W=\mathbb{R}$. The general Banach space-valued case follows exactly as above (note that the d-dimensional Itô formula is a simple consequence of Taylor's formula up to second order, which of course also holds in Banach spaces, where again derivatives are understood int he sense of Frechet).

Lemma 5.2.1 (Integration of gradient one forms: first order rough calculus)**.** *Let* $\mathbf{X} = (X,\mathbb{X})\in\mathcal{C}^\alpha_g(I,V)$ *for some* $\alpha\in(\frac{1}{3},\frac{1}{2}]$ *and* $G\in C^3_b(V,W)$*. Then*

$$G(X_t) - G(X_s) = \int_s^t DG(X_r)d\mathbf{X}_r, \quad \forall s,t\in I.$$

If one now drops the geometricity assumption for $\mathbf{X}$ and repeats the previous computations, one obviously arrives at the following formula:

$$G(X_t) - G(X_s) = \int_s^t DG(X_r) d\mathbf{X}_r + \int_s^t D^2 G(X_r) d\Big[\frac{1}{2} X_r \otimes X_r - \mathrm{Sym}(\mathbb{X}_r)\Big],$$

where the final integral is defined in Riemann-Stieltjes (or Young) sense, since $r \mapsto \frac{1}{2} X_r \otimes X_r - \mathrm{Sym}(\mathbb{X}_r)$ is 2α-Hölder continuous and thus the sum of the regularities of integrand and integrator is $\alpha + 2\alpha > 1$.

This observation and the fact that the previous computations do not involve the antisymmetric part $\mathrm{Ant}(\mathbb{X})$ motivates the following definition.

Definition 5.2.2. Let $\alpha \in (\frac{1}{3}, \frac{1}{2}]$, $X \in C^\alpha(I, V)$ and $\mathbb{S} \in C_2^{2\alpha}(I, \mathrm{Sym}(V \otimes V))$. $\mathbf{X} = (X, \mathbb{S})$ is a (α-continuous) *reduced rough path* if the *reduced Chen relation*

$$\mathbb{S}_{s,t} - \mathbb{S}_{s,u} - \mathbb{S}_{u,t} = \mathrm{Sym}(X_{s,u} \otimes X_{u,t}), \quad \forall s, u, t \in I \tag{5.2.4}$$

holds. The space of α-continuous reduced rough paths is denoted by $\mathcal{C}_r^\alpha(I, V)$, shortly $\mathcal{C}_r^\alpha$.

Remark 5.2.3. *(i) If $\mathbf{X} = (X, \mathbb{X}) \in \mathcal{C}^\alpha$, then $(X, \mathrm{Sym}(\mathbb{X})) \in \mathcal{C}_r^\alpha$. Indeed, in this case (5.2.4) is obtained by considering the symmetric part of both sides of the full Chen relation (C).*

(ii) A key feature of reduced rough paths is the following: For any path $X \in C^\alpha$, the choice $\mathbb{S}_{s,t} = \frac{1}{2} X_{s,t} \otimes X_{s,t}$ yields a reduced rough path $(X, \mathbb{S}) \in \mathcal{C}_r^\alpha$. In other words, there is a trivial reduced rough path lift for any $X \in C^\alpha$. This is in stark contrast to the "full" case, where a general trivial *lift $X \mapsto \mathbf{X} \in \mathcal{C}^\alpha$ does not exist.*

(iii) This trivial reduced rough path lift is natural with regard to geometricity. Indeed, defining the reducing operator

$$\Lambda : \mathcal{C}^\alpha \to \mathcal{C}_r^\alpha, \quad \Lambda : (X, \mathbb{X}) \mapsto (X, \mathrm{Sym}(\mathbb{X})),$$

we see that for $(X, \mathbb{X}) \in \mathcal{C}_g^\alpha$ one has

$$\Lambda(X, \mathbb{X}) = (X, \frac{1}{2} X \otimes X).$$

In analogy to Lemma 2.1.5, we have the following lemma

Lemma 5.2.4. *Let $X \in C^\alpha(I, V)$. Then $(X, \bar{\mathbb{S}}) \in \mathcal{C}_r^\alpha(I, V)$ if and only if $\bar{\mathbb{S}} \in C_2^{2\alpha}(I, \mathrm{Sym}(V \times V))$ with*

$$\bar{\mathbb{S}}_{s,t} = \frac{1}{2} X_{s,t} \otimes X_{s,t} + \gamma_{s,t} \tag{5.2.5}$$

for some $\gamma \in C^{2\alpha}(I, \mathrm{Sym}(V \otimes V))$.

Proof. We only have to check that $(X, \bar{\mathbb{S}}) \in C^\alpha(I, V) \oplus C_2^{2\alpha}(I, \mathrm{Sym}(V \otimes V))$ satisfies the reduced Chen relation if and only if $\bar{\mathbb{S}}$ is as in (5.2.5). But this is obvious, since $(X, \frac{1}{2} X \otimes X)$ satisfies the reduced Chen relation, hence a perturbation of $\bar{\mathbb{S}}$ preserves that relation if and only if it is additive. $\square$

Definition 5.2.5. For $\mathbf{X} = (X, \mathbb{S}) \in \mathcal{C}_r^\alpha$, we define the *bracket* (of $(\mathbf{X})$)

$$[\mathbf{X}] : I \to \mathrm{Sym}(V \otimes V), \quad t \mapsto [\mathbf{X}]_t := X_{0,t} \otimes X_{0,t} - 2\mathbb{S}_{0,t},$$

and (as usual)

$$[\mathbf{X}]_{s,t} := (\delta[\mathbf{X}])_{s,t}$$

(i.e. $[\mathbf{X}]_{s,t} = X_{s,t} \otimes X_{s,t} - 2\mathbb{S}_{s,t}$ (!)). In particular $\delta[\mathbf{X}] \in C_2^{2\alpha}(I, \mathrm{Sym}(V \otimes V))$.

Remark 5.2.6. *(i) Comparing with Lemma 5.2.4, we have* $[\mathbf{X}]_t = -2\gamma_{0,t}$ *and* $[\mathbf{X}]_{s,t} = -2\gamma_{s,t}$.

The above contents now culminate in the following rough Itô formula.

Proposition 5.2.7. *[Rough Itô formula] Let* $\alpha \in (\frac{1}{3}, \frac{1}{2}]$, $G \in C_b^3(V, W)$ *and* $\mathbf{X} = (X, \mathbb{S}) \in \mathcal{C}_r^\alpha(I, V)$. *Then for all* $s, t \in I$

$$G(X_t) - G(X_s) = \int_s^t DG(X_r) d\mathbf{X}_r + \frac{1}{2} \int_s^t D^2 G(X_r) d[\mathbf{X}]_r, \tag{5.2.6}$$

where, denoting by $\mathcal{P}$ *any partition of* $[s, t]$,

$$\int_s^t DG(X_r) d\mathbf{X}_r := \lim_{|\mathcal{P}| \to 0} \sum_{(\tau_i, \tau_{i+1}) \in \mathcal{P}} \Big(DG(X_{\tau_i}) X_{\tau_i, \tau_{i+1}} + D^2 G(X_{\tau_i}) \mathbb{S}_{\tau_i, \tau_{i+1}} \Big)$$

(this integral is well-defined by the Sewing Lemma and the reduced Chen relation).

Remark 5.2.8. *(i) In particular, for* $(X, \mathbb{X}) \in \mathcal{C}^\alpha$*, the previous result applies to* $\mathbf{X} = (X, \mathrm{Sym}(\mathbb{X}))$ *(i.e. one first throws away the irrelevant* $\mathrm{Ant}(\mathbb{X})$ *and then applies the rough Itô formula). In this case, the notation* $\int_s^t DG(X_r) d\mathbf{X}_r$ *is not ambigous, since the "full" rough integral coincides with the one in* (5.2.6) *(since* $D^2 G(X_{\tau_i}) \mathrm{Ant}(\mathbb{X}_{\tau_i, \tau_{i+1}}) = 0$*).*

(ii) In the geometric case $\mathbb{S}_{s,t} = \frac{1}{2} X_{s,t} \otimes X_{s,t}$*, the bracket vanishes* $[\mathbf{X}] \equiv 0$ *and we retrieve Lemma 5.2.1.*

Proof of Proposition 5.2.7. All aspects of the proof have been considered above. By Itô formula one has, denoting by $\mathcal{P}$ any partition of $[s, t]$,

$$G(X_t) - G(X_s) = \lim_{|\mathcal{P}| \to 0} \sum_{(\tau_i, \tau_{i+1}) \in \mathcal{P}} \Big(DG(X_{\tau_i}) X_{\tau_i, \tau_{i+1}} + \frac{1}{2} D^2 G(X_{\tau_i}) (X_{\tau_i, \tau_{i+1}} \otimes X_{\tau_i, \tau_{i+1}}) \Big),$$

and, using the definition of $[\mathbf{X}]$, the RHS equals

$$\lim_{|\mathcal{P}| \to 0} \sum_{(\tau_i, \tau_{i+1}) \in \mathcal{P}} \Big(DG(X_{\tau_i}) X_{\tau_i, \tau_{i+1}} + \frac{1}{2} D^2 G(X_{\tau_i}) [\mathbf{X}]_{\tau_i, \tau_{i+1}} + D^2 G(X_{\tau_i}) \mathbb{S}_{\tau_i, \tau_{i+1}} \Big)$$

$$=$$

$$\lim_{|\mathcal{P}|\to 0} \sum_{(\tau_i,\tau_{i+1})\in\mathcal{P}} \Big(DG(X_{\tau_i})X_{\tau_i,\tau_{i+1}} + D^2G(X_{\tau_i})\mathbb{S}_{\tau_i,\tau_{i+1}} \Big) + \frac{1}{2} \lim_{|\mathcal{P}|\to 0} \sum_{(\tau_i,\tau_{i+1})\in\mathcal{P}} D^2G(X_{\tau_i})[\mathbf{X}]_{\tau_i,\tau_{i+1}}$$

Now the first limit equals, by definition, $\int_s^t DG(X_r)d\mathbf{X}_r$ (as said in the assertion, its well-definedness follows from the sewing lemma and the reduced Chen relation (!)), and the second limit is the Riemann–Stieltjes integral $\frac{1}{2}\int_s^t D^2G(X_r)d[\mathbf{X}]_r$, which completes the proof. □

6 Rough differential equations

We are now prepared to study rough differential equations

$$dY = f(Y)d\mathbf{X}, \tag{6.0.1}$$

which are the key motivation for this lecture. Solutions to such an equation will be controlled rough paths (Y, Y') (w.r.t. X, where $\mathbf{X} = (X, \mathbb{X})$). We need to ensure that with (Y, Y') also $f(Y)$ is controlled w.r.t. X, i.e. in particular we need to find a Gubinelli derivative of $f(Y)$. This is possible since the spaces of controlled rough path are stable under composition with sufficiently regular maps. We investigate this in more detail in the next section.

6.1 Composition of regular functions and controlled rough paths

Let $X \in C^\alpha(I, V)$, where again $I = [0, T]$, and let $(Y, Y') \in \mathcal{D}_X^{2\alpha}(I, W)$ as well as $\varphi : W \to \bar{W}$ with $\varphi \in C_b^2(W, \bar{W})$. For a moment, think of Y' as a classical derivative of Y. Then one expects to find $(\varphi(Y))'$ as

$$\varphi(Y)'_t = D\varphi(Y_t)Y'_t$$

(note $D\varphi(y) \in L(W, \bar{W})$ for all $y \in W$ and hence $D\varphi(Y_t)Y'_t \in L(V, \bar{W})$).

Lemma 6.1.1. *In the situation above, one has $(\varphi(Y), \varphi(Y)') \in \mathcal{D}_X^{2\alpha}(I, \bar{W})$.*

Proof. $\varphi \in C_b^2(W, \bar{W})$ and $(Y, Y') \in \mathcal{D}_X^{2\alpha}(I, W)$ implies $\varphi(Y) \in C^\alpha(I, \bar{W})$ and $\varphi(Y)' \in C^\alpha(I, L(V, \bar{W}))$ (recall $\varphi(Y)'_t := D\varphi(Y_t)Y'_t$). It remains to prove

$$R^{\varphi(Y)} \in C_2^{2\alpha}(I, W),$$

where

$$R^{\varphi(Y)}_{s,t} := \varphi(Y)_{s,t} - \varphi(Y)'_s X_{s,t} = \varphi(Y)_{s,t} - D\varphi(Y_s)Y'_s X_{s,t}.$$

This is true, since

$$R^{\varphi(Y)}_{s,t} = \varphi(Y)_{s,t} - D\varphi(Y_s)Y_{s,t} + D\varphi(Y_s)[Y_{s,t} - Y'_s X_{s,t}] \tag{6.1.1}$$

and $(s,t) \mapsto \varphi(Y)_{s,t} - D\varphi(Y_s)Y_{s,t}$ as well as $(s,t) \mapsto Y_{s,t} - Y'_s X_{s,t}$ belong to $C_2^{2\alpha}(I, \bar{W})$ (by Taylor formula) and $C_2^{2\alpha}(I, W)$, respectively, while $s \mapsto D\varphi(Y_s)$ is continuous. $\square$

A more refined result which will be very important for the construction of solutions to rough differential equations is the following.

Lemma 6.1.2. *(i) Let $\varphi \in C_b^2(W, \bar{W})$, $(Y, Y') \in \mathcal{D}_X^{2\alpha}(I, W)$ for some $X \in C^\alpha(I, V)$, $\alpha \in (\frac{1}{3}, \frac{1}{2}]$ and let $M > 1$ such that*

$$|Y_0'| + ||Y, Y'||_{X,2\alpha} \leqslant M. \tag{6.1.2}$$

Then $(\varphi(Y), \varphi(Y)') \in \mathcal{D}_X^{2\alpha}(I, \bar{W})$, and there is a constant $C = C(T, \alpha)$ such that

$$||\varphi(Y), \varphi(Y)'||_{X,2\alpha} \leqslant C(M+1)||\varphi||_{C_b^2}\big(1+||X||_\alpha\big)^2\big(|Y_0'|+|Y,Y'|_{X,2\alpha}\big), \tag{6.1.3}$$

where $\varphi(Y)'$ is again defined as $\varphi(Y)_t' := D\varphi(Y_t)Y_t'$. Moreover, C is locally bounded in T.

(ii) If in addition $\bar{X} \in C^\alpha$ and $(\bar{Y}, \bar{Y}') \in \mathcal{D}_{\bar{X}}^{2\alpha}$ such that $(\bar{Y}, \bar{Y}')$ also satisfies (6.1.2), then

$$||\varphi(Y), \varphi(Y)'; \varphi(\bar{Y}), \varphi(\bar{Y})'||_{X,\bar{X},2\alpha} \leqslant C_M\big(||X-\bar{X}||_\alpha+|Y_0-\bar{Y}_0|+|Y_0'-\bar{Y}_0'|+||Y,Y';\bar{Y},\bar{Y}'||_{X,\bar{X},2\alpha}\big), \tag{6.1.4}$$

where $C_M > 0$ only depends on φ, α and M, on the latter in a multiplicative, hence monotone, way.

Proof. We only proof (i), for (ii) see [1, Theorem 7.6]. Regarding (i), from Lemma 6.1.1 we already know $(\varphi(Y), \varphi(Y)') \in \mathcal{D}_X^{2\alpha}(I, \bar{W})$, so it remains to prove (6.1.3). By definition of $||\cdot,\cdot||_{X,2\alpha}$ and $\varphi(Y)'$, we find

$$\begin{aligned} ||\varphi(Y), \varphi(Y)'||_{X,2\alpha} \leqslant{}& ||D\varphi(Y_\cdot)||_\infty ||Y'||_\alpha + ||Y_\cdot'||_\infty ||D^2\varphi(Y_\cdot)||_\infty ||Y_\cdot||_\alpha \\ &+ \frac{1}{2}||D^2\varphi||_\infty ||Y||_\alpha^2 + ||D\varphi||_\infty ||R^Y||_{2\alpha}, \end{aligned}$$

where we also used (6.1.1) and Taylor formula and we estimated $||\varphi(Y)'||_\alpha$ by the sum in the first line of the RHS of the previous inequality. Clearly, this RHS is further bounded from above by

$$\begin{aligned} &||\varphi||_{C_b^2}\Big(||Y'||_\alpha + ||Y'||_\infty ||Y||_\alpha + ||Y||_\alpha^2 + ||R^Y||_{2\alpha}\Big) \\ &\leqslant C_{\alpha,T}||\varphi||_{C_b^2}(1+||X||_\alpha)^2\big(1+|Y_0'|+||Y,Y'||_{X,2\alpha}\big)\big(|Y_0'+||Y,Y'||_{X,2\alpha}\big), \end{aligned}$$

where we used $||Y||_\alpha \leqslant (1+||X||_\alpha)(|Y_0'| + T^\alpha||Y,Y'||_{X,2\alpha})$, compare the proof of Lemma 4.3.3. □

6.2 Solutions to RDEs

We now study the rough differential equation (RDE) (6.0.1). As before, let $I = [0, T]$, V and W be Banach spaces, $\mathbf{X} = (X, \mathbb{X}) \in \mathcal{C}^\alpha(I, V)$ for some $\alpha \in (\frac{1}{3}, \frac{1}{2}]$, and let $\xi \in W$. Let $f \in C_b^2(W, L(V, W))$.

While V is the state space of our rough path $\mathbf{X}$, W is the state space of the equation, i.e. of its solutions. ξ is a given initial datum (a point from the state space W).

Definition 6.2.1. A solution to (6.0.1) with initial datum ξ is a controlled rough path $(Y, Y') \in \mathcal{D}_X^{2\alpha}(I, W)$ such that

$$Y_t = \xi + \int_0^t f(Y_s) d\mathbf{X}_s, \quad \forall t \in I, \tag{6.2.1}$$

where $\int_0^t f(Y_s) d\mathbf{X}_s = \int_0^t \big(f(Y)_s, f(Y)'_s\big) d\mathbf{X}_s$, with $f(Y)' := Df(Y)Y'$ as in the previous section.

Remark 6.2.2. *(i) Note that due to $f \in C_b^2(W, L(V, W))$ and thanks to Lemma 6.1.1, the above rough integral is well defined (with regard to Lemma 6.1.1, here we have $\bar{W} = L(V, W)$).*

(ii) If $V = \mathbb{R}^1$, $W = \mathbb{R}^d$, $X \in C^1(\mathbb{R}^1)$ and (Y, Y') a solution to (6.0.1), then by Remark 4.3.6 $\int_0^t f(Y_s) d\mathbf{X}_s = \int_0^t f(Y_s) dX_s$, where the latter is a Riemann–Stieltjes integral and equals $\int_0^t f(Y_s) X'_s ds$. In particular, (6.2.1) does not depend on Y' or $f(Y)'$. Since in this case $t \mapsto \int_0^t f(Y_s) X'_s ds$ is differentiable, so is Y, and consequently $(Y, g) \in \mathcal{D}_X^{2\alpha}(I, \mathbb{R}^d)$ for all $g \in C^\alpha(I, \mathbb{R}^d)$ and each such (Y, g) satisfies (6.2.1). This shows that in this situation the set of solutions for a given initial datum is either empty or uncountable. Uniqueness can only be achieved by postulating the choice of Y'.

(iii) Continuing the previous part, note that for any solution Y to (6.0.1), we have $(Y, f(Y)) \in \mathcal{D}_X^{2\alpha}(I, W)$:

$$|Y_{s,t} - f(Y_s) X_{s,t}| = \left| \int_s^t f(Y_u) d\mathbf{X}_u - f(Y_s) X_{s,t} \right| \leqslant |f(Y_s)' \mathbb{X}_{s,t}| + C|t - s|^{3\alpha}, \tag{6.2.2}$$

where the inequality follows from the construction of the rough integral, see Theorem 4.3.4. Since $\mathbb{X} \in C_2^{2\alpha}(I, V \otimes V)$, we see $R_{s,t}^{Y,f} := Y_{s,t} - f(Y_s) X_{s,t} \in C_2^{2\alpha}(I, W)$, and so $(Y, f(Y)) \in \mathcal{D}_2^{2\alpha}(I, W)$ as claimed. We will use this natural Gubinelli derivative as a choice to formulate uniqueness results.

The next theorem on existence and uniqueness for rough differential equations is our main goal for this chapter.

Theorem 6.2.3. *[Local well-posedness of rough differential equation] Let $I = [0, T]$, $\xi \in W$, $f \in C^3(W, L(V, W))$, $\mathbf{X} = (X, \mathbb{X}) \in \mathcal{C}^\alpha(I, V)$ for some $\alpha \in (\frac{1}{3}, \frac{1}{2}]$.*

Then there exists $0 < T_0 \leqslant T$ and a unique solution $(Y, Y') \in \mathcal{D}_X^{2\alpha}([0, T_0], W)$ to (6.0.1) on $I_0 := [0, T_0]$ (i.e. Definition 6.2.1 is satisfied with I_0 instead of $[0, T]$) with $Y' = f(Y)$ and initial datum ξ. If $f \in C_b^3$, then we may take $T_0 = T$.

Again, it should be stressed that $Y' = f(Y)$ is crucial for the uniqueness part of the assertion. We make the following remark on the interval of existence:

Remark 6.2.4. *Let ξ and f be as in the previous theorem.*

(i) *Let Y be the solution on some interval I_0. Then Y can either be extended to a solution on the whole interval I or only on some maximal interval $[0,\tau) \subset I$. Here* maximal *means that Y cannot be extended to any interval $[\tau, \tau+\varepsilon]$, $\varepsilon > 0$.*

(ii) *τ depends in general on ξ, f and $\mathbf{X}$. In fact, for fixed ξ and f, the map $\mathbf{X} \mapsto \tau(\mathbf{X})$ is lower semicontinuous:*

$$\liminf_n \tau(\mathbf{X}_n) \geqslant \tau(\mathbf{X})$$

whenever $\mathbf{X}_n \to \mathbf{X}$ in $\mathcal{C}^\alpha$.

(iii) *For finite-dimensional V, W, one can show: If Y can only be extended to a maximal interval $[0,\tau)$, then $\limsup_{t\nearrow\tau} |Y_t| = +\infty$, and τ is also called explosion time of the solution Y.*

For later use, we first state and prove the following a priori bound for RDE-solutions.

Proposition 6.2.5. *[A priori bound for RDE-solutions] Let $\xi \in W$, $f \in C_b^2(W, L(V,W))$, $\mathbf{X}$ as in the previous theorem and let $(Y,Y') \in \mathcal{D}_X^{2\alpha}(I,W)$ be a solution to (6.0.1) on I with initial condition ξ, and with $Y' = f(Y)$. Then*

$$||Y||_\alpha \leqslant C\Big[\big(||f||_{C_b^2}|||\mathbf{X}|||_\alpha\big) \vee \big(||f||_{C_b^2}|||\mathbf{X}|||_\alpha\big)^{\frac{1}{\alpha}}\Big], \tag{6.2.3}$$

where $C > 0$ only depends on α.

Proof. By a scaling argument, we may assume $||f||_{C_b^2} \leqslant 1$. Let $J = [s,t] \subseteq I$, then (denoting by "$\lesssim$" that we suppress multiplicative constants depending only on f and α)

$$\begin{aligned}
|R^Y_{s,t}| &= |Y_{s,t} - f(Y_s)X_{s,t}| \\
&\leqslant \left| \int_s^t f(Y_u)d\mathbf{X}_u - f(Y_s)X_{s,t} - Df(Y_s)f(Y_s)\mathbb{X}_{s,t} \right| + |Df(Y_s)f(Y_s)\mathbb{X}_{s,t}| \\
&\lesssim \big(||X||_\alpha ||R^{f(Y)}||_{2\alpha} + ||\mathbb{X}||_{2\alpha}||f(Y)||_\alpha\big)|t-s|^{3\alpha} + ||\mathbb{X}||_{2\alpha}|t-s|^{2\alpha}.
\end{aligned}$$

The second inequality above follows from Theorem 4.3.4.

While $||\cdot||_\alpha$ denotes Hölder-seminorms considered on the full interval I, we write $||\cdot||_{\alpha;J}$ to denote Hölder-seminorms for maps on J and, moreover, for $0 < h < |I|$, set $||\cdot||_{\alpha;h} := \sup_{J\subseteq I, |J|\leqslant h} ||\cdot||_{\alpha;J}$. With this notation the previous estimate yields

$$||R^Y||_{2\alpha;h} \lesssim ||\mathbb{X}||_{2\alpha} + \big(||X||_{\alpha;h}||R^{f(Y)}||_{2\alpha;h} + ||\mathbb{X}||_{2\alpha;h}||f(Y)||_{\alpha;h}\big)h^\alpha. \tag{6.2.4}$$

Now we relate R^Y and $R^{f(Y)}$: Since

$$\begin{aligned}
R^{f(Y)}_{s,t} &= f(Y_t) - f(Y_s) - Df(Y_s)Y'_s X_{s,t} \\
&= f(Y_t) - f(Y_s) - Df(Y_s)Y_{s,t} + Df(Y_s)R^Y_{s,t},
\end{aligned}$$

we get by Taylor formula

$$||R^{f(Y)}||_{2\alpha;h} \leqslant \frac{1}{2}|D^2 f|_\infty ||Y||^2_{\alpha;h} + |Df|_\infty ||R^Y||_{2\alpha;h} \lesssim ||Y||^2_{\alpha;h} + ||R^Y||_{2\alpha;h}.$$

Since also $||f(Y)||_{\alpha;h} \lesssim ||Y||_{\alpha;h}$, there is $c_1 > 0$ only depending on α and f such that, by (6.2.4),

$$||R^Y||_{2\alpha;h} \leqslant c_1||\mathbb{X}||_{2\alpha}+c_1||X||_{\alpha;h}h^\alpha||Y||^2_{\alpha;h}+c_1||X||_{\alpha;h}h^\alpha||R^Y||_{2\alpha;h}+c_1||\mathbb{X}||_{2\alpha;h}h^\alpha||Y||_{\alpha;h}. \tag{6.2.5}$$

We now restrict to sufficiently small $h > 0$, more precisely let h such that

$$c_1||X||_\alpha h^\alpha \leqslant \frac{1}{2}, \quad c_1||\mathbb{X}||^{\frac{1}{2}}_{2\alpha} h^\alpha \leqslant \frac{1}{2}, \tag{6.2.6}$$

and note that of course this choice of h only depends on $\mathbf{X}$, f and α. Then the previous inequality can be further estimated as

$$||R^Y||_{2\alpha;h} \leqslant c_1||\mathbb{X}||_{2\alpha} + \frac{1}{2}||Y||^2_{\alpha;h} + \frac{1}{2}||R^Y||_{2\alpha;h} + \frac{1}{2}||\mathbb{X}||^{\frac{1}{2}}_{2\alpha;h}||Y||_{\alpha;h},$$

which implies, using Young inequality for the second estimate below,

$$||R^Y||_{2\alpha;h} \leqslant 2c_1||\mathbb{X}||_{2\alpha}+||Y||^2_{\alpha;h}+||\mathbb{X}||^{\frac{1}{2}}_{2\alpha;h}||Y||_{\alpha;h} \leqslant c_2||\mathbb{X}||_{2\alpha;h}+2||Y||^2_{\alpha;h}, \tag{6.2.7}$$

where $c_2 := 2c_1 + 1$. But from $Y_{s,t} = f(Y_s)X_{s,t} - R^Y_{s,t}$ and boundedness of f we also get

$$||Y||_{\alpha;h} \lesssim ||X||_\alpha + ||R^Y||_{2\alpha;h}h^\alpha,$$

whereby together with (6.2.6) we get

$$\begin{aligned} ||Y||_{\alpha;h} &\leqslant c_3||X||_\alpha + c_3||\mathbb{X}||_{2\alpha}h^\alpha + c_3||Y||^2_{\alpha;h}h^\alpha \\ &\leqslant c_3||X||_\alpha + c_4||\mathbb{X}||^{\frac{1}{2}}_{2\alpha} + c_3||Y||^2_{\alpha;h}h^\alpha \end{aligned}$$

for some $c_3, c_4 > 0$. Multiplying the previous inequality with c_3h^α and setting

$$\psi_h := c_3||Y||_{\alpha;h}h^\alpha, \quad \lambda_h := c_5|||\mathbf{X}|||_\alpha h^\alpha,$$

with $c_5 := c_3^2 + c_3c_4$, we have for all $h > 0$ as above

$$\psi_h \leqslant \lambda_h + \psi_h^2. \tag{6.2.8}$$

For h sufficiently small (depending on Y) such that $\psi_h \leqslant \frac{1}{2}$, we have $\psi_h \leqslant \lambda_h + \frac{\psi_h}{2}$ and hence $\psi_h \leqslant 2\lambda_h$, i.e.

$$||Y||_{\alpha;h} \leqslant c_6|||\mathbf{X}|||_\alpha.$$

For reasons to be seen below let us even take h sufficiently small such that $\psi_h < \frac{1}{6}$. In fact (and importantly) the previous inequality holds for all h sufficiently small *without* dependence on Y: Let h_0 such that $\lambda_{h_0} < \frac{1}{4}$ (note λ_h is independent from Y). Then, by (6.2.8), for each $h < h_0$ one of the following regimes holds:

$$\psi_h \geqslant \psi_+ := \frac{1}{2} + \sqrt{\frac{1}{4} - \lambda_h} \, (\geqslant \frac{1}{2},), \quad \psi_h \leqslant \psi_- := \frac{1}{2} - \sqrt{\frac{1}{4} - \lambda_h} \ (\searrow 0 \text{ as } h \to 0). \tag{6.2.9}$$

We want to show that the second regime holds for all $h < h_0$. We know that for $h < h_0$ sufficiently small depending on Y, we have $\psi_h \leqslant 2\lambda_h \xrightarrow{h \to 0} 0$, so for $h \to 0$ we are in the second regime. Since $\psi_h < \frac{1}{6}$ in the second and $\psi_h \geqslant \frac{1}{2}$ in the first regime, respectively, we only need to rule out that the increasing function $h \mapsto \psi_h$ has a jump of relative size more than 3. But $\psi_h \leqslant 3 \lim_{g \nearrow h} \psi_g$, since

$$||Y||_{\alpha;h} \leqslant 3||Y||_{\alpha;\frac{h}{3}} \leqslant 3 \lim_{g \nearrow h} ||Y||_{\alpha;g}$$

(similarly $\lim_{g \searrow h} \psi_g \leqslant 3\psi_h$). Hence we never jump from the second into the first regime, whereby we conclude that the second regime holds for all $h < h_0$. By elementary considerations the definitions of ψ_- and λ_h entail the existence of a constant $c_6 > 0$ such that $\psi_h \leqslant c_6 \lambda_h$ for all $h < h_0$, and so

$$||Y||_{\alpha;h} \leqslant c_6 |||\mathbf{X}|||_\alpha, \quad \forall h < h_0.$$

Noting that $h_0 \leqslant c_7 |||\mathbf{X}|||_\alpha^{-\frac{1}{\alpha}}$, hence the following lemma yields

$$||Y||_\alpha \leqslant c_6 |||\mathbf{X}|||_\alpha (1 \vee 2c_7 |||\mathbf{X}|||_\alpha^{\frac{1-\alpha}{\alpha}}) \leqslant c_8 (|||\mathbf{X}|||_\alpha \vee |||\mathbf{X}|||_\alpha^{\frac{1}{\alpha}}), \tag{6.2.10}$$

where $c_8 > 0$ only depends on α. As said above, for general f the claim follows by a scaling argument. □

At the end of the previous proof, we used the following lemma:

Lemma 6.2.6. *Let $\alpha \in (0,1]$, $h > 0$, $M > 0$ and consider $Z : I \to V$ such that*

$$||Z||_{\alpha;h} \leqslant M.$$

Then

$$||Z||_{\alpha;I} \leqslant M\big(1 \vee 2h^{-(1-\alpha)}\big).$$

Proof. A short exercise, or see Exercise 4.5. in [1]. □

Remark 6.2.7. *Often one is interested in more general equations with a time-dependent vector field f and an additional (time-dependent) nonlinear drift, i.e.*

$$dY_t = g_t(Y_t)dt + f_t(Y_t)d\mathbf{X}_t, \tag{6.2.11}$$

where $f : I \times W \ni (t,w) \mapsto f_t(w) \in L(V,W)$ and $g : I \times W \ni (t,w) \mapsto g_t(w) \in W$. One may recast this equation in the form (6.0.1) by considering

$$d\bar{Y} = \bar{f}(\bar{Y})d\bar{\mathbf{X}},$$

where $\bar{Y}_t = (Y_t, t) \in W \times I$*,* $\bar{\mathbf{X}} = (\bar{X}, \bar{\mathbb{X}})$ *with*

$$\bar{X}_t := (X_t, t),$$

$\bar{\mathbb{X}}$ *takes values in* $(V \times I) \otimes (V \times I)$ *with components* $\mathbb{X}$*,* $\int_s^t X_{s,r} dr$*,* $\int_s^t (r-s) dX_r$ *and* $\int_s^t (r-s) dr$ *(all defined as Riemann–Stieltjes integrals), and* $\bar{f} \in L(I \times W; L(I \times V, I \times W))$ *given by its components*

$$\bar{f}^{1,1}(t,w) := g(t,w), \quad \bar{f}^{1,2}(t,w) := 0 =: \bar{f}^{2,1}(t,w), \quad \bar{f}^{2,2}(t,w) := f(t,w).$$

One may then apply Theorem 6.2.3 (and, more generally, the machinery of RDEs). However, this leads to non-optimal assumptions for and bounds of solutions (for instance, clearly one will not want to assume $g \in C_b^3$*). A more sophisticated approach for (6.2.11) is more instance developed in [2].*

Proof of Theorem 6.2.3. We consider the case $f \in C_b^3$ and construct a unique global solution. At the end of the proof, we point out what may go wrong if f is not bounded (leading to a local solution with blowup).

Without loss of generality, let $T = 1$. Let $\frac{1}{3} < \beta < \alpha (\leqslant \frac{1}{2})$. Of course $\mathbf{X} \in \mathcal{C}^\beta(I, V)$. By Lemma 6.1.2, for any $(Y, Y') \in \mathcal{D}_X^{2\beta}(I, W)$ we have

$$(\Xi, \Xi') := (f(Y), f(Y)') := (f(Y), Df(Y)Y') \in \mathcal{D}_X^{2\beta}(I, L(V, W)).$$

For any $0 < \mathcal{T} \leqslant 1$, consider the map $\mathcal{M}_\mathcal{T} : \mathcal{D}_X^{2\beta}([0, \mathcal{T}], W) \to \mathcal{D}_X^{2\beta}([0, \mathcal{T}], L(V, W))$, defined by

$$\mathcal{M}_\mathcal{T} : (Y, Y') \mapsto \left(\xi + \int_0^\cdot \Xi_s d\mathbf{X}_s, \Xi\right).$$

Clearly, a fixed point of $\mathcal{M}_\mathcal{T}$ is a solution to (6.0.1) in $\mathcal{D}_X^{2\beta}([0, \mathcal{T}], W)$ on $[0, \mathcal{T}]$ with initial datum ξ. Since $\mathbf{X} \in \mathcal{C}^\alpha$, it turns out that this solution even belongs to $\mathcal{D}_X^{2\alpha}(I, W)$. Indeed, by the construction of the rough integral,

$$|Y_{s,t}| = \left| \int_s^t f(Y_r) d\mathbf{X}_r \right| \lesssim |X_{s,t}| + ||Y'||_\infty |\mathbb{X}_{s,t}| + |t-s|^{3\beta},$$

and similarly one shows $R^Y \in C_2^{2\alpha}$. By definition, $(Y_0, Y_0') = (\xi, f(\xi))$ implies $\mathcal{M}_\mathcal{T}(Y, Y')$ also equals $(\xi, f(\xi))$ at 0, and hence $\mathcal{M}_\mathcal{T}$ can be considered on the space of controlled rough paths started at $(\xi, f(\xi))$, i.e.

$$\mathcal{D}_X^{2\beta}(I, W; \xi) := \{(Y, Y') \in \mathcal{D}_X^{2\beta}([0, \mathcal{T}], W) : Y_0 = \xi, Y_0' = f(\xi)\}.$$

As an affine subspace of the Banach space $\mathcal{D}_X^{2\beta}(I, W)$, this space is a complete metric space under the induced metric. The same is true for its subset defined as the closed unit ball $B_\mathcal{T}$ centered at the element

$$[t \mapsto (\xi + f(\xi) X_{0,t}, f(\xi))] \in \mathcal{D}_X^{2\beta}([0, \mathcal{T}], W; \xi)$$

(it is tempting to use as a center $[t \mapsto (\xi, f(\xi))]$ instead, but it is a short exercise that this element in general does not belong to $\mathcal{D}_X^{2\beta}(I, W; \xi)$). Put differently, $B_{\mathcal{T}}$ consists of those controlled rough paths such that $Y_0 = \xi, Y_0' = f(\xi)$ and

$$|Y_0 - \xi| + |Y_0' - f(\xi)| + ||(Y_\cdot - (\xi + f(\xi)X_{0,\cdot}), Y_\cdot' - f(\xi))||_{X,2\beta} = ||(Y_\cdot - f(\xi)X_{0,\cdot}, Y_\cdot' - f(\xi))||_{X,2\beta} \leqslant 1. \tag{6.2.12}$$

By triangle inequality and since $||(f(\xi)X_{0,\cdot}, f(\xi))||_{X,2\beta} = ||f(\xi)||_\beta + ||0||_{2\beta} = 0$, it follows that

$$||(Y - f(\xi)X_{0,\cdot}, Y_\cdot' - f(\xi))||_{X,2\beta} = ||Y, Y'||_{X,2\beta},$$

and consequently

$$B_{\mathcal{T}} = \Big\{(Y, Y') \in \mathcal{D}_X^{2\beta}([0, \mathcal{T}], W) : Y_0 = \xi, Y_0' = f(\xi), ||Y, Y'||_{X,2\beta} \leqslant 1\Big\}. \tag{6.2.13}$$

Also note the trivial further bound

$$|Y_0'| + ||(Y, Y')||_{X,2\beta} \leqslant |f|_\infty + 1 =: M, \tag{6.2.14}$$

where we used the boundedness of f. Now our goal is to show that a) for $\mathcal{T} > 0$ sufficiently small, $\mathcal{M}_{\mathcal{T}}(B_{\mathcal{T}}) \subseteq B_{\mathcal{T}}$, i.e. $\mathcal{M}_{\mathcal{T}}$ leaves $B_{\mathcal{T}}$ invariant, and b) that for such (or a suitably smaller) $\mathcal{T}$, $\mathcal{M}_{\mathcal{T}}$ is a contraction on $B_{\mathcal{T}}$. With a)+b), we then apply Banach's fixed point theorem to the complete metric space $B_{\mathcal{T}}$ to obtain a unique fixed point of $\mathcal{M}_{\mathcal{T}}$. Constants appearing below depend on $\alpha, \beta, \mathbf{X}$, but not on $\mathcal{T}$. Dependence on f will be indicated.

a) Invariance: From Lemma 6.1.2 and Theorem 4.3.4, we have

$$||\Xi, \Xi'||_{X,2\beta} \leqslant C(M+1)||f||_{C_b^2}(|Y_0'| + ||Y, Y'||_{X,2\beta}) \tag{6.2.15}$$

and

$$\begin{aligned}
&\left|\left| \int_0^\cdot \Xi_s d\mathbf{X}_s, \Xi \right|\right|_{X,2\beta} \\
&\qquad \leqslant ||\Xi||_\beta + ||\Xi'||_\infty ||\mathbb{X}||_{2\beta} + C(||X||_\beta ||||R^\Xi||_{2\beta} + ||\mathbb{X}||_{2\beta}||\Xi'||_\beta) \\
&\qquad \leqslant ||\Xi||_\beta + C(|\Xi_0'| + ||\Xi, \Xi'||_{X,2\beta})\mathcal{T}^{\alpha-\beta},
\end{aligned} \tag{6.2.16}$$

where for the final inequality we used $||X||_\beta + ||\mathbb{X}||_{2\beta} \leqslant \mathcal{T}^{\alpha-\beta}||X||_\alpha + \mathcal{T}^{2\alpha-2\beta}||\mathbb{X}||_{2\alpha} \leqslant C\mathcal{T}^{\alpha-\beta}$, where C depends only on $\mathbf{X}$. For $(Y, Y') \in B_{\mathcal{T}}$, since $||\Xi||_\beta \leqslant ||f||_{C_b^1}||Y||_\beta$ and $|\Xi_0'| = |Df(Y_0)Y_0'| \leqslant ||f||_{C_b^1}^2$, we find, using (6.2.15) and the previous chain of inequalities,

$$\begin{aligned}
||\mathcal{M}_{\mathcal{T}}(Y, Y')||_{X,2\beta} &= \left|\left| \int_0^\cdot \Xi_s d\mathbf{X}_s, \Xi \right|\right|_{X,2\beta} \\
&\leqslant ||\Xi||_\beta + C(|\Xi_0'| + ||\Xi, \Xi'||_{X,2\beta})\mathcal{T}^{\alpha-\beta} \\
&\leqslant ||f||_{C_b^1}||Y||_\beta + C\big(||f||_{C_b^1}^2 + C(M+1)||f||_{C_b^2}(|Y_0'| + ||Y, Y'||_{X,2\beta})\big)\mathcal{T}^{\alpha-\beta} \\
&\leqslant ||f||_{C_b^1}\mathcal{T}^{\alpha-\beta} + C(M+1)\big(||f||_{C_b^1}^2 + ||f||_{C_b^2}M\big)\mathcal{T}^{\alpha-\beta},
\end{aligned}$$

where $C = C(\alpha, \beta, \mathbf{X}) > 0$ changes from line to line, and we used the definition of M above. We also used

$$|Y_{s,t}| \leqslant |Y'|_\infty |X_{s,t}| + ||R^Y||_{2\beta} |t-s|^{2\beta} \leqslant (|Y'_0| + ||Y'||_\beta) ||X||_\alpha |t-s|^\alpha + ||R^Y||_{2\beta} |t-s|^{2\beta}, \tag{6.2.17}$$

which yields

$$||Y||_{\beta;[0,\mathcal{T}]} \leqslant C_f \mathcal{T}^{\alpha-\beta},$$

since $||R^Y||_{2\beta} \leqslant ||Y, Y'||_{X,2\beta} \leqslant 1$, $|Y'_0| + ||Y'||_\beta \leqslant M$ and $|t-s|^{2\beta-\beta} \leqslant \mathcal{T}^\beta \leqslant \mathcal{T}^{\alpha-\beta}$ (the latter since $2\beta > \alpha$ and $\mathcal{T} \leqslant 1$). Hence

$$||\mathcal{M}_\mathcal{T}(Y, Y')||_{X,2\beta} = ||\mathcal{M}_\mathcal{T}(Y, Y')||_{X,2\beta;[0,\mathcal{T}]} \leqslant C_f \mathcal{T}^{\alpha-\beta}, \tag{6.2.18}$$

with $C_f > 0$ independent from $\mathcal{T}$. Thus we can choose $\mathcal{T} \in (0, 1)$ sufficiently small such that

$$||\mathcal{M}_\mathcal{T}(Y, Y')||_{X,2\beta} \leqslant 1.$$

But by (6.2.13) this means

$$\mathcal{M}_\mathcal{T}(B_\mathcal{T}) \subseteq B_\mathcal{T}.$$

b) Contraction: For $(Y, Y'), (\tilde{Y}, \tilde{Y}') \in B_\mathcal{T}$ with $\mathcal{T} \in (0, 1)$ as chosen at the end of step (a), we aim to show

$$||\mathcal{M}_\mathcal{T}(Y, Y') - \mathcal{M}_\mathcal{T}(\tilde{Y}, \tilde{Y}')||_{X,2\beta} \leqslant C_f ||Y - \tilde{Y}, Y' - \tilde{Y}'||_{X,2\beta} \mathcal{T}^{\alpha-\beta}, \tag{6.2.19}$$

for a constant $C_f > 0$ not depending on $\mathcal{T}$, (Y, Y') and $(\tilde{Y}, \tilde{Y}')$. Then we can shrink $\mathcal{T}$ further such that

$$C_f \mathcal{T}^{\alpha-\beta} < 1 \tag{6.2.20}$$

and it follows that for such $\mathcal{T}$, $\mathcal{M}_\mathcal{T}$ is a contraction on $B_\mathcal{T}$. So, let us prove (6.2.19). Let us use the notation

$$h_s := f(Y_s) - f(\tilde{Y}_s),$$

then we have

$$\begin{aligned} ||\mathcal{M}_\mathcal{T}(Y, Y') - \mathcal{M}_\mathcal{T}(\tilde{Y}, \tilde{Y}')||_{X,2\beta} &= \left|\left| \int_0^\cdot h_s d\mathbf{X}_s, h. \right|\right|_{X,2\beta} \\ &\leqslant ||h||_\beta + C\big(|h'_0| + ||(h, h')||_{X,2\beta}\big)\mathcal{T}^{\alpha-\beta} \\ &\leqslant C||f||_{C_b^2} ||Y - \tilde{Y}||_\beta + C||(h, h')||_{X,2\beta} \mathcal{T}^{\alpha-\beta}. \end{aligned} \tag{6.2.21}$$

Here the first inequality follows as in (6.2.16) and the second follows from $h'_0 = Df(Y_0)Y'_0 - Df(\tilde{Y}_0)\tilde{Y}'_0 = 0$. Now consider (6.2.17) with Y replaced by $Y - \tilde{Y}$ to get

$$||Y - \tilde{Y}||_\beta \leqslant ||Y' - \tilde{Y}'||_\beta ||X||_\alpha \mathcal{T}^{\alpha-\beta} + ||R^Y - R^{\tilde{Y}}||_{2\beta} \mathcal{T}^{\alpha-\beta} \leqslant C||Y - \tilde{Y}, Y' - \tilde{Y}'||_{X,2\beta} \mathcal{T}^{\alpha-\beta}.$$

$$(6.2.22)$$

Next, write $h_s = G_s(Y_s - \tilde{Y}_s)$, where $G_s := g(Y_s, \tilde{Y}_s)$, where

$$g(x,y) := \int_0^1 Df(tx + (1-t)y)dt$$

is defined on $W \times W$ with values in $L(W, L(V, W))$ and belongs to C_b^2, with $||g||_{C_b^2} \leqslant ||f||_{C_b^3}$. By Lemma 6.1.1, this regularity of g yields that G is again a β-Hölder continuous controlled rough path (w.r.t. X) with Gubinelli derivative $G' = (D_x g)Y' + (D_y g)\tilde{Y}'$, where $D_x g$ and $D_y g$ denote the first derivative of g w.r.t. its first and second variable, respectively. It is in fact straightforward to show

$$||G, G'||_{X,2\beta} \leqslant C||f||_{C_b^3}, \tag{6.2.23}$$

uniformly over $(Y, Y'), (\tilde{Y}, \tilde{Y}') \in B_{\mathcal{T}}$ and $\mathcal{T} \leqslant 1$. To continue, note that $\mathcal{D}_X^{2\beta}$ is an algebra in the sense that $(GH, (GH)') \in \mathcal{D}_X^{2\beta}([0, \mathcal{T}], L(V, W))$ with $(GH)' := G'H + GH'$ and it is straightforward to check

$$||GH, (GH)'||_{X,2\beta} \leqslant C\big(|G_0| + |G_0'| + ||G, G'||_{X,2\beta}\big)\big(|H_0| + |H_0'| + ||H, H'||_{X,2\beta}\big) \tag{6.2.24}$$

Choosing $H = Y - \tilde{Y}$ yields $H_0 = Y_0 - \tilde{Y}_0 = \xi - \xi = 0$, and similarly $H_0' = 0$. Hence for each $(Y, Y'), (\tilde{Y}, \tilde{Y}) \in B_{\mathcal{T}}$:

$$\begin{aligned} ||h, h'||_{X,2\beta} &\leqslant C_f\big(||g||_\infty + ||g||_{C_b^1}(|Y_0'| + |\tilde{Y}_0'|) + ||f||_{C_b^3}\big)||Y - \tilde{Y}, Y' - \tilde{Y}'||_{X,2\beta} \\ &\leqslant C_f||Y - \tilde{Y}, Y' - \tilde{Y}'||_{X,2\beta} \end{aligned}$$

where we used (6.2.23) and $||H, H'||_{X,2\beta} \leqslant C_f||Y - \tilde{Y}, Y' - \tilde{Y}'||_{X,2\beta}$ for the first, and $||g||_{C_b^1} \leqslant ||f||_{C_b^3}$ and $|Y_0'| = |\bar{Y}_0'| = |f(\xi)| \leqslant |f|_\infty$ for the second inequality.

Thus, returning to (6.2.21), we arrive at

$$||\mathcal{M}_{\mathcal{T}}(Y, Y') - \mathcal{M}_{\mathcal{T}}(\tilde{Y}, \tilde{Y}')||_{X,2\beta} \leqslant C_f||Y - \tilde{Y}, Y' - \tilde{Y}'||_{X,2\beta}\mathcal{T}^{\alpha-\beta},$$

i.e. we have proven (6.2.19) and obtain the contraction property of $\mathcal{M}_{\mathcal{T}}$ on $B_{\mathcal{T}}$ for sufficiently small $\mathcal{T} > 0$ as described at the beginning of part (b).

Conclusion: For sufficiently small $\mathcal{T} \in (0, 1)$, $\mathcal{M}_{\mathcal{T}}$ is a contraction on the complete metric space $B_{\mathcal{T}}$. Thus, there is a unique fixed point $(Y, Y') \in B_{\mathcal{T}}$, i.e. $\mathcal{M}_{\mathcal{T}}((Y, Y')) = (Y, Y')$. By definition of $\mathcal{M}_{\mathcal{T}}$, it follows that (Y, Y') is the unique solution to the rough differential equation (6.0.1) in the sense of Definition 6.2.1 on $[0, \mathcal{T}]$ with initial datum ξ.

Finally, let us observe that this solution can be extended to $[0, 1]$ thanks to the additional assumption $f \in C_b^3$. To this end, consider the end point of the solution as a new initial condition ξ_1, i.e. $\xi_1 := Y(\mathcal{T})$. We may repeat the above proof for this initial condition and find a unique solution $(Y^1, {Y^1}')$ on an interval $[\mathcal{T}, \mathcal{T}_1]$ for

some $\mathcal{T}_1 \in (\mathcal{T}, \mathcal{T}+1)$. It is straightforward that the concatenated path $(\tilde{Y}, \tilde{Y}')$, defined by
$$\tilde{Y} \equiv Y \text{ on } [0, \mathcal{T}], \quad \tilde{Y} \equiv Y^1 \text{ on } [\mathcal{T}, \mathcal{T}_1]$$
(and similarly for $\tilde{Y}'$) is a solution to (6.0.1) on $[0, \mathcal{T}_1]$ with initial datum ξ. Hence it only remains to argue that the number $\mathcal{T}$ in the previous proof can be chosen uniformly in the initial datum $\xi \in W$, since then we can take $\mathcal{T}_1 = \mathcal{T}$, and iteratively define a concatenated solution on $[0, n\mathcal{T}]$ for all $n \in \mathbb{N}$ (this in particular shows that there is no difference in contructing solutions on $[0,1]$ or any $[0,T]$, $T > 1$). Indeed, $\mathcal{T}$ chosen in (a) and (b) above only depends on $||f||_{C_b^3}$ (see (6.2.18) and (6.2.20)). On the other hand, if f was not bounded, then the constant M from (6.2.14) would depend on ξ, thus $\mathcal{T}$ will in general depend on ξ as well. In this case, it cannot be ruled out that the iterative construction indicated above yields time intervals $[\mathcal{T}_i, \mathcal{T}_{i+1}]$ with $\lim_i \mathcal{T}_i < 1$. In this case, we cannot extend the solution to $[0,1]$. ☐

Continuity of the Itô-Lyons map. Our next goal is to investigate regularity of the solution map $\mathbf{X} \mapsto Y = Y(\mathbf{X})$ to a rough differential equation. More precisely, assume we are given $f \in C_b^3$ and fix an initial datum $\xi \in W$. Theorem 6.2.3 implies the existence of a unique solution $Y = Y(\mathbf{X})$ of (6.0.1) with initial condition ξ for each $\mathbf{X} \in \mathcal{C}^\alpha(I, V)$ on any $I = [0,T]$ (more precisely, as seen above, the solution is a pair (Y, Y') and uniqueness holds only under the additional condition $Y' = f(Y)$). Since

$$\hat{S} : \mathcal{C}^\alpha(I, V) \to C^\alpha(I, W), \quad \hat{S} : \mathbf{X} \mapsto Y(\mathbf{X})$$

is a map between metric spaces, it is a natural question whether $\hat{S}$ is continuous, and this is the content of the following result.

Remark 6.2.8. *(i) $\hat{S}$ is often called* Itô-Lyons map *(compare the introduction, where $\hat{S}$ was already mentioned).*

(ii) Of course we can also consider $\hat{S}(\mathbf{X}) := (Y, f(Y))(\mathbf{X})$, i.e. $\hat{S}$ then is not path-, but controlled rough path-valued. It should be recalled that for $\mathbf{X} = (X, \mathbb{X})$, $\bar{\mathbf{X}} = (\bar{X}, \bar{\mathbb{X}})$ with $X \neq \bar{X}$, $\hat{S}(\mathbf{X})$ and $\hat{S}(\bar{\mathbf{X}})$ live in different Banach spaces, i.e. $\mathcal{D}_X^{2\alpha}$ and $\mathcal{D}_{\bar{X}}^{2\alpha}$,and hence not in a common metric space. However, in (4.4.1) we introduced the suitable "distance" (not a metric in the usual sense) $||\cdot, \cdot||_{X, \bar{X}, 2\alpha}$. We are going to use this distance to obtain a more general estimate in Theorem 6.2.9 below.

Theorem 6.2.9. *[Local Lipschitz continuity of Itô-Lyons map] Let $f \in C_b^3$ as above, $\alpha \in (\frac{1}{3}, \frac{1}{2}]$, $\mathbf{X}, \bar{\mathbf{X}} \in \mathcal{C}^\alpha(I, V)$, and $(Y, f(Y)), (\bar{Y}, f(\bar{Y}))$ the respective unqiue RDE-solutions on $I = [0,T]$ with initial data $\xi, \bar{\xi}$. Let $|||\mathbf{X}|||_\alpha, |||\bar{\mathbf{X}}|||_\alpha \leqslant M < \infty$. Then*

$$||(Y, f(Y)); (\bar{Y}, f(\bar{Y}))||_{X, \bar{X}; 2\alpha} \leqslant C\big(|\xi - \bar{\xi}| + \rho_\alpha(\mathbf{X}, \bar{\mathbf{X}})\big), \tag{6.2.25}$$

and also

$$||Y - \bar{Y}||_\alpha \leqslant C\big(|\xi - \bar{\xi}| + \rho_\alpha(\mathbf{X}, \bar{\mathbf{X}})\big), \tag{6.2.26}$$

where $C = C(\alpha, f, T, M) > 0$ depends on M in a monotone way.

Thus, in particular the Itô-Lyons map $\hat{S}$ (for a given initial datum ξ) is locally Lipschitz continuous from $(\mathcal{C}^\alpha, \rho_\alpha)$ to $(C^\alpha, ||\cdot||_\alpha)$. For $\mathbf{X} = \bar{\mathbf{X}}$, (6.2.26) gives (global) Lipschitz continuity of solutions to a single RDE with respect to the initial datum.

As we shall see below, the proof only relies on the a priori bound from Proposition 6.2.5 and the usual estimates for rough integrals from Chapter 4.

Proof. We write $(\Xi, \Xi') = (f(Y), f(Y)')$ and similarly for $(\bar{\Xi}, \bar{\Xi}')$; also set

$$(Z_\cdot, Z'_\cdot) := \left(\xi + \int_0^\cdot f(Y_s)d\mathbf{X}_s, f(Y_\cdot)\right)$$

and similarly for $(\bar{Z}, \bar{Z}')$. By (4.4.2) and since $Y = Z, Y' = f(Y)$ (similarly for $\bar{Y}$ and $\bar{Y}'$) we have

$$\begin{aligned}||Y, Y'; \bar{Y}, \bar{Y}'||_{X,\bar{X},2\alpha} &= ||Z, Z'; \bar{Z}, \bar{Z}'||_{X,\bar{X},2\alpha} \\ &\leqslant C_0\big(\rho_\alpha(\mathbf{X}, \bar{\mathbf{X}}) + |Df(\xi)f(\xi) - Df(\bar{\xi})f(\bar{\xi})| + T^\alpha||\Xi, \Xi'; \bar{\Xi}, \bar{\Xi}'||_{X,\bar{X},2\alpha}\big) \\ &\leqslant C_1\big(\rho_\alpha(\mathbf{X}, \bar{\mathbf{X}}) + |\xi - \bar{\xi}| + T^\alpha||\Xi, \Xi'; \bar{\Xi}, \bar{\Xi}'||_{X,\bar{X},2\alpha}\big),\end{aligned}$$

where

$$\begin{aligned}C_0 := \max\big(|||\mathbf{X}|||_\alpha, |||\bar{\mathbf{X}}|||_\alpha, &|Df(\xi)f(\xi)| + ||\Xi, \Xi'||_{X,2\alpha}, \\ &|Df(\bar{\xi})f(\bar{\xi})| + +||\bar{\Xi}, \bar{\Xi}'||_{\bar{X},2\alpha}\big)\end{aligned}$$

and $C_1 = C_f \max\big(|||\mathbf{X}|||_\alpha, |||\bar{\mathbf{X}}|||_\alpha, 1 + ||\Xi, \Xi'||_{X,2\alpha}, 1 + ||\bar{\Xi}, \bar{\Xi}'|||_{\bar{X},2\alpha}\big)$ for a suitable constant $C_f > 0$ only depending on $||f||_{C_b^2}$. Lemma 6.1.2 (ii) yields

$$||\Xi, \Xi'; \bar{\Xi}, \bar{\Xi}'||_{X,\bar{X},2\alpha} \leqslant C_2\big(\rho_\alpha(\mathbf{X}, \bar{\mathbf{X}}) + 2|\xi - \bar{\xi}| + ||Y, f(Y); \bar{Y}, f(\bar{Y})||_{X,\bar{X},2\alpha}\big),$$

where

$$C_2 := C_{f,\alpha} \max\big(1 + ||Y, Y'||_{X,2\alpha}, 1 + ||\bar{Y}, \bar{Y}'||_{\bar{X},2\alpha}\big).$$

Consequently, for $C_3 := 2T^\alpha C_2$, we have

$$||Y, Y'; \bar{Y}, \bar{Y}'||_{X,\bar{X},2\alpha} \leqslant C_1(C_3 + 1)\big(\rho_\alpha(\mathbf{X}, \bar{\mathbf{X}}) + |\xi - \bar{\xi}| + T^\alpha||Y, f(Y); \bar{Y}, f(\bar{Y})||_{X,\bar{X},2\alpha}\big).$$

Note that at this stage C_1 and C_3 depend on Y and $\bar{Y}$. But Lemma 6.1.2 (i) gives

$$||\Xi, \Xi'||_{X,2\alpha} \leqslant C(T, \alpha, f)(2 + ||Y, Y'||_{X,2\alpha})||f||_{C_b^2}(1 + ||X||_\alpha)^2(1 + ||Y, Y'||_{X,2\alpha})$$

and similarly for $(\bar{\Xi}, \bar{\Xi}')$. Therefore, we are now going to show that $||Y, Y'||_{X,2\alpha}$ and $||\bar{Y}, \bar{Y}'||_{\bar{X},2\alpha}$ are bounded by a constant only depending on α, f, T and $|||\mathbf{X}|||_\alpha$ and $|||\bar{\mathbf{X}}|||_\alpha$ (on the latter two quantities in a monotone way), at least for $T > 0$ sufficiently small. Then it follows from the above that

$$||Y, Y'; \bar{Y}, \bar{Y}'||_{X,\bar{X},2\alpha} \leqslant C(\alpha, f, T, |||\mathbf{X}|||_\alpha, |||\bar{\mathbf{X}}|||_\alpha)\big(\rho_\alpha(\mathbf{X}, \bar{\mathbf{X}}) + |\xi - \bar{\xi}| + T^\alpha||Y, f(Y); \bar{Y}, f(\bar{Y})||_{X,\bar{X},2\alpha}\big),$$

and by possibly shrinking T further to a value T_0 such that $T_0^\alpha C(\alpha, f, T_0, |||\mathbf{X}|||_\alpha, |||\bar{\mathbf{X}}|||_\alpha) \leqslant \frac{1}{2}$, we conclude (6.2.25) on $[0, T_0]$. For $T > T_0$, one iterates the above procedure finitely many times to obtain (6.2.25) on $[0,T]$ (note that all constants are independent from the initial data ξ, $\bar{\xi}$). We now focus on (Y,Y'), the proof for $(\bar{Y},\bar{Y}')$ is similar. Below, constants may change from line to line, but we indicate their dependence on absolute parameters. Recall $||Y,Y'||_{X,2\alpha} = ||Y'||_\alpha + ||R^Y||_{2\alpha}$. Since $Y' = f(Y)$, by the apriori bound from Proposition 6.2.5, we have

$$||Y'||_\alpha \leqslant C_f||Y||_\alpha \leqslant C(f,\alpha,|||\mathbf{X}|||_\alpha) < \infty.$$

Furthermore

$$R^Y_{s,t} = Y_{s,t} - Y'_s X_{s,t} = \int_s^t f(Y_r)d\mathbf{X}_r - f(Y_s)X_{s,t},$$

thus the standard estimates for rough integrals yield

$$|R^Y_{s,t}| \leqslant |f(Y)'_s \mathbb{X}_{s,t}| + C\big(||X||_\alpha ||R^{f(Y)}||_{2\alpha} + ||\mathbb{X}||_{2\alpha}||f(Y)'||_\alpha\big)|t-s|^{3\alpha},$$

where $C > 0$ only depends on α. Since $f(Y)' = Df(Y)f(Y)$, we have

$$||f(Y)'||_\alpha \leqslant C_f||Y||_\alpha \leqslant C(f,\alpha,|||\mathbf{X}|||_\alpha).$$

Finally, we note

$$||R^{f(Y)}||_{2\alpha} \leqslant C_f(||Y||^2_\alpha + ||R^Y||_{2\alpha},)$$

(we observed this estimate in the proof of Proposition 6.2.5), so we find

$$\begin{aligned} ||R^Y||_{2\alpha} &\leqslant C(f,\alpha,|||\mathbf{X}|||_\alpha)\big(1 + (C(f,\alpha,|||\mathbf{X}|||_\alpha) + ||R^Y||_{2\alpha} + 1)T^\alpha\big) \\ &\leqslant C(f,\alpha,T,|||\mathbf{X}|||_\alpha)\big[1 + C(f,\alpha,T,|||\mathbf{X}|||_\alpha) + ||R^Y||_{2\alpha}T^\alpha\big]. \end{aligned}$$

Shrinking $T > 0$ such that $C(f,\alpha,T,|||\mathbf{X}|||_\alpha)T^\alpha \leqslant \frac{1}{2}$, we get

$$||R^Y||_{2\alpha} \leqslant 2C(f,\alpha,T,|||\mathbf{X}|||_\alpha)(1 + C(f,\alpha,T,|||\mathbf{X}|||_\alpha)),$$

and so

$$||Y,Y'||_{X,2\alpha} \leqslant C(f,\alpha,T,|||\mathbf{X}|||_\alpha).$$

It is obvious that all dependencies of the above constants on $|||\mathbf{X}|||_\alpha$ and $|||\bar{\mathbf{X}}|||_\alpha$ are monotone. As explained above, this concludes the proof. Finally, (6.2.26) follows from (6.2.25) similarly as (4.4.3) follows from (4.4.2).

□

6.3 RDEs and SDEs

We saw in Chapter 3 that a.e. path of a standard $\mathbb{R}^d$-valued Brownian motion $B = (B_t)_{t\geqslant 0}$ on a stochastic basis $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geqslant 0}, \mathbb{P})$ lifts to a rough path, either in Itô- or Stratonovich-sense, i.e.

$$\mathbf{B}^I = (B, \mathbb{B}^I), \quad \mathbf{B}^S = (B, \mathbb{B}^S) \quad \in \mathcal{C}^\alpha(I, \mathbb{R}^d)$$

pathwise for any $\alpha \in (\frac{1}{3}, \frac{1}{2}]$ and $I = [0, T]$, where the two-parameter $\mathbb{R}^d \otimes \mathbb{R}^d$-valued stochastic processes $\mathbb{B}^I$ and $\mathbb{B}^S$ are defined as the following Itô- and Stratonovich-integrals, respectively:

$$\mathbb{B}^I_{s,t} = \int_s^t B_{s,r} \otimes dB_r, \quad \mathbb{B}^S_{s,t} = \int_s^t B_{s,r} \otimes \circ dB_r.$$

Here we focus on the Itô-case, the Stratonovich-case can be treated similarly. Standard theory of stochastic differential equations (SDEs) asserts that if $f \in C_b(\mathbb{R}^m, L(\mathbb{R}^d, \mathbb{R}^m))$ is Lipschitz continuous and $\xi \in \mathbb{R}^m$, then the Itô-SDE

$$dX_t = f(X_t)dB_t, \quad X_0 = \xi \tag{6.3.1}$$

has a unique probabilistically strong solution $X = X(\xi)$, $X = (X_t)_{t \in I}$, on any $I = [0, T]$. On the other hand, for a.e. $\omega \in \Omega$, thanks to Chapter 6, we can also study the RDE on $\mathbb{R}^m$

$$dY_t = f(Y_t)d\mathbf{B}^I(\omega)_t, \quad Y_0 = \xi, \tag{6.3.2}$$

at least if we assume $f \in C_b^3(\mathbb{R}^d, L(\mathbb{R}^d, \mathbb{R}^m))$. Here $\omega \in \Omega$ is arbitrary, but fixed, i.e. we consider the deterministic rough path $\mathbf{B}^I(\omega)$, and the solution $Y = Y(\mathbf{B}^I(\omega)) = \hat{S}(\mathbf{B}^I(\omega))$. Now we investigate the natural question whether

$$\hat{S}(\mathbf{B}^I(\omega)) = X(\omega)$$

a.s. The answer is positive:

Theorem 6.3.1. *Let $f \in C_b^3(\mathbb{R}^m, L(\mathbb{R}^m, \mathbb{R}^m))$ and $\xi \in \mathbb{R}^m$. Then $Y = (Y_t(\omega))_{t \in I, \omega \in \Omega}$, where $Y(\omega) = \hat{S}(\mathbf{B}^I(\omega))$ is the unique RDE-solution to (6.3.2) for a.e. $\omega \in \Omega$, is the unique probabilistically strong solution to (6.3.1).*

Remark 6.3.2. *In particular, the unique strong solution X to (6.3.1) factorizes as*

$$\omega \mapsto X(\omega) = \hat{S} \circ \Psi(B(\omega)),$$

where $\Psi(B(\omega)) = \mathbf{B}^I(\omega)$ is the measurable Itô-rough path lift of Brownian paths, and $\hat{S}$ is the continuous solution map of the RDE (6.3.2) (the continuity holds due to Theorem 6.2.9). In particular, this shows that the dependence of X via B is pathwise - which is not at all clear on the level of the SDE. Note that indeed the map Ψ maps $B(\omega)$ to $(B(\omega), \mathbb{B}^I(\omega))$, since by Itô's product rule, $\mathbb{B}^I(\omega)$ depends only on $B(\omega)$ (and not on the whole process B).

Remark 6.3.3. *Since the zero set of those ω for which (6.3.2) does not have a unique RDE-solution only depends on B, but not on ξ, it follows immediately from the previous theorem that there is a zero set N* independent *from $\xi \in \mathbb{R}^d$ such that a unique strong solution $X(\xi)$ to (6.3.1) is defined pathwise in $\omega \in N^c$ for* all *intial data $\xi \in \mathbb{R}^m$. Hence, in particular $(\xi, \omega, t) \mapsto X_t(\xi, \omega)$ is well-defined as a* flow*. This is in contrast to standard SDE-theory, where the construction of a flow usually requires additional considerations. However, the price to pay here is the C_b^3-regularity of f, which is of course much stronger than in the standard theory.*

Proof of Theorem 6.3.1. We only need to show that $Y = (Y(\omega))_{\omega\in\Omega}$ as in the assertion is a (weak) solution to (6.3.1) which is adapted to the Brownian filtration $(\mathcal{F}_r^B)_{r\leqslant T}$, $\mathcal{F}_r^B := \sigma(B_u, 0 \leqslant u \leqslant r)$. By Proposition 5.1.1, it suffices to show that $(Y(\omega), f(Y(\omega)))$, where $Y(\omega) = \hat{S}(\mathbf{B}^I(\omega))$ is $(\mathcal{F}_t)$-adapted. But this follows since

$$Y = \hat{S} \circ \Psi(B)$$

as mentioned in Remark 6.3.2, and $\Psi : B(\omega)_{|[0,t]} \mapsto (B(\omega), \mathbb{B}^I(\omega))_{|[0,t]}$ is clearly $\mathcal{F}_t^B/\mathcal{B}(\mathcal{C}^\alpha([0,t], \mathbb{R}^d))$-measurable for each $t \in I$, and since $\hat{S} : \mathbf{B}^I(\omega) \mapsto Y(\omega)$ is continuous between $\mathcal{C}^\alpha([0,t], \mathbb{R}^d)$ and $C^\alpha([0,t], \mathbb{R}^m)$. Thus the required measurability holds and Proposition 5.1.1 applies. $\square$

The proof for the Stratonovich-case is similar.

A simple proof of the classical Wong-Zakai approximation-result. A classical result, see for instance [4, p.392], states: If B is a Brownian motion and B^n piecewise linear pathwise approximations of B, then for sufficiently regular vector fields f, the solutions X^n to the random ODE

$$dX_t^n = f(X_t^n) dB_t^n$$

converge to the solution of the Stratonovich SDE

$$dX_t = f(X_t) \circ dB_t.$$

This and related results are usually called *Wonk-Zakai approximations*. With the theory from above, we can give a simple proof of this result: We only need

$$\mathbf{B}^n := \big(B^n, \int B^n \circ dB^n\big) \xrightarrow{n\to\infty} \mathbf{B}^S \tag{6.3.3}$$

in $(\mathcal{C}^\alpha, \rho_\alpha)$. Indeed, then, since for $f \in C_b^3$ X^n coincides pathwise with the RDE-solution $Y^n(\omega)$ of

$$dY_t^n = f(Y_t^n) d\mathbf{B}_t^n(\omega)$$

and X, by Theorem 6.3.1, pathwise coincides with the RDE-solution $Y(\omega)$ of

$$dY_t = f(Y_t) d\mathbf{B}_t^S(\omega),$$

pathwise a.s. convergence $X^n \to X$ follows from Theorem 6.2.9. But (6.3.3) was shown in Proposition 3.2.4.

Bibliography

[1] Peter K. Friz and Martin Hairer. *A course on rough paths*. Universitext. Springer, Cham, second edition, [2020] ©2020. With an introduction to regularity structures.

[2] Peter K. Friz and Nicolas B. Victoir. *Multidimensional Stochastic Processes as Rough Paths: Theory and Applications*. Cambridge Studies in Advanced Mathematics. Cambridge University Press, 2010.

[3] M Gubinelli. Controlling rough paths. *Journal of Functional Analysis*, 216(1):86–140, 2004.

[4] N. Ikeda and S. Watanabe. *Stochastic Differential Equations and Diffusion Processes*. North-Holland, 1989.

[5] Terry J. Lyons. Differential equations driven by rough signals. *Revista Matemática Iberoamericana*, 14(2):215–310, 1998.

[6] L. C. Young. An inequality of the Hölder type, connected with Stieltjes integration. *Acta Mathematica*, 67(none):251 – 282, 1936.