Chapter 2
The space of rough paths

We define the space of (Hölder continuous) rough paths, as well as the subspace of "geometric" rough paths which preserve the usual rules of calculus. The latter can be interpreted in a natural way as paths with values in a certain nilpotent Lie group. At the end of the chapter, we give a short discussion showing how these definitions should be generalised to treat paths of arbitrarily low regularity.

2.1 Basic definitions

In this section, we give a practical definition of the space of Hölder continuous rough paths. Our choice of Hölder spaces is chiefly motivated by our hope that most readers will already be familiar with the classical Hölder spaces from real analysis. We could in the sequel have replaced "α-Hölder continuous" by "finite p-variation" for $p = 1/\alpha$ in many statements. This choice would also have been quite natural, due to the fact that one of our primary goals will be to give meaning to integrals of the form $\int f(X)\,dX$ or solutions to controlled differential equations of the form $dY = f(Y)\,dX$ for rough paths X. The value of such an integral / solution does not depend on the parametrisation of X, which dovetails nicely with the fact that the p-variation of a function is also independent of its parametrisation. This motivated its choice in the original development of the theory. In some other applications however (like the solution theory to rough stochastic *partial* differential equations developed in [Hai11b, HW13, Hai13] and more generally the theory of regularity structures [Hai14b] exposed in the last chapters), parametrisation-independence is lost and the choice of Hölder norms is more natural.

A rough path on an interval $[0,T]$ with values in a Banach space V then consists of a continuous function $X\colon [0,T] \to V$, as well as a continuous "second order process" $\mathbb{X}\colon [0,T]^2 \to V \otimes V$, subject to certain algebraic and analytical conditions. Regarding the former, the behaviour of iterated integrals, such as (2.2) below, suggests to impose the algebraic relation ("*Chen's relation*"),

P. K. Friz, M. Hairer, *A Course on Rough Paths*, Universitext,
https://doi.org/10.1007/978-3-030-41556-3_2

$$\mathbb{X}_{s,t} - \mathbb{X}_{s,u} - \mathbb{X}_{u,t} = X_{s,u} \otimes X_{u,t} \;, \tag{2.1}$$

which we assume to hold for every triplet of times (s, u, t). Since $X_{t,t} = 0$, it immediately follows (take $s = u = t$) that we also have $\mathbb{X}_{t,t} = 0$ for every t. As already mentioned in the introduction, one should think of $\mathbb{X}$ as *postulating* the value of the quantity

$$\int_s^t X_{s,r} \otimes dX_r \stackrel{\text{def}}{=} \mathbb{X}_{s,t} \;, \tag{2.2}$$

where we take the right-hand side as a *definition* for the left-hand side. (And not the other way around!) We insist (cf. Exercise 2.4 below) that as a consequence of (2.1), knowledge of the path $t \mapsto (X_{0,t}, \mathbb{X}_{0,t})$ already determines the entire second order process $\mathbb{X}$. In this sense, the pair $(X, \mathbb{X})$ is indeed a path, and not some two-parameter object, although it is often more convenient to consider it as one. If X is a smooth function and we read (2.2) from right to left, then it is straightforward to verify (see Exercise 2.1 below) that the relation (2.1) does indeed hold. Furthermore, one can convince oneself that if $f \mapsto \int f\, dX$ denotes any form of "integration" which is linear in f, has the property that $\int_s^t dX_r = X_{s,t}$, and is such that $\int_s^t f(r)\, dX_r + \int_t^u f(r)\, dX_r = \int_s^u f(r)\, dX_r$ for any admissible integrand f, and if we use such a notion of "integral" to define $\mathbb{X}$ via (2.2), then (2.1) does automatically hold. This makes it a very natural postulate in our setting.

Note that the algebraic relations (2.1) are by themselves not sufficient to determine $\mathbb{X}$ as a function of X. Indeed, for any $V \otimes V$-valued function F, the substitution $\mathbb{X}_{s,t} \mapsto \mathbb{X}_{s,t} + F_t - F_s$ leaves the left-hand side of (2.1) invariant. We will see later on how one should interpret such a substitution. It remains to discuss what are the natural analytical conditions one should impose for $\mathbb{X}$. We are going to assume that the path X itself is α-Hölder continuous, so that $|X_{s,t}| \lesssim |t-s|^\alpha$. The archetype of an α-Hölder continuous function is one which is self-similar with index α, so that $X_{\lambda s, \lambda t} \sim \lambda^\alpha X_{s,t}$.

(We intentionally do not give any mathematical definition of self-similarity here, just think of $\sim$ as having the vague meaning of "looks like".) Given (2.2), it is then very natural to expect $\mathbb{X}$ to also be self-similar, but with $\mathbb{X}_{\lambda s, \lambda t} \sim \lambda^{2\alpha} \mathbb{X}_{s,t}$. This discussion motivates the following definition of our basic spaces of rough paths.

Definition 2.1. For $\alpha \in (\frac{1}{3}, \frac{1}{2}]$, define the space of α-Hölder rough paths (over V), in symbols $\mathscr{C}^\alpha([0,T], V)$, as those pairs $(X, \mathbb{X}) =: \mathbf{X}$ such that

$$\|X\|_\alpha \stackrel{\text{def}}{=} \sup_{s \neq t \in [0,T]} \frac{|X_{s,t}|}{|t-s|^\alpha} < \infty \;, \qquad \|\mathbb{X}\|_{2\alpha} \stackrel{\text{def}}{=} \sup_{s \neq t \in [0,T]} \frac{|\mathbb{X}_{s,t}|}{|t-s|^{2\alpha}} < \infty \;, \tag{2.3}$$

and such that the algebraic constraint (2.1) is satisfied.

The obvious example is the *canonical rough path lift* of a smooth path X, of the form $(X, \int X \otimes dX)$, and we write $\mathscr{L}(\mathcal{C}^\infty)$ for the class of rough paths obtained in this way.[1] We have the strict inclusion $\mathscr{L}(\mathcal{C}^\infty) \subset \mathscr{C}^\infty$, the class of *smooth rough paths*,[2] by which we mean a genuine rough path with the additional property that the V-valued (resp. $V \otimes V$-valued) maps $X_\cdot$ and $\mathbb{X}_{s,\cdot}$ are smooth, for every basepoint s. For instance, $\mathbf{X} \equiv (0,0)$ is the trivial canonical rough path associated to the scalar zero path, as opposed to the smooth "pure second level" rough path (over $\mathbf{R}$) given by $(s,t) \mapsto (0, t-s)$; see also Exercise 2.10 for a natural example with $\dim V > 1$.

Remark 2.2. Any *scalar* path $X \in \mathcal{C}^\alpha$ can be lifted to a rough path (over $\mathbf{R}$), simply by setting $\mathbb{X}_{s,t} := (X_{s,t})^2/2$. However, for a vector-valued path $X \in \mathcal{C}^\alpha$, with values in some Banach space V, it is far from obvious that one can find suitable "second order increments" $\mathbb{X}$ such that X lifts to a rough path $(X, \mathbb{X}) \in \mathscr{C}^\alpha$. The *Lyons–Victoir extension theorem* (Exercise 2.14) asserts that this can always be done, even in a *continuous* fashion, provided that $1/\alpha \notin \mathbf{N}$ which means $\alpha \in (\frac{1}{3}, \frac{1}{2})$ in our present discussion. (A counterexample for $\alpha = \frac{1}{2}$ is hinted on in Exercise 2.13). The reader may wonder how this continuity property dovetails with Proposition 1.1. The point is that if we define $X \mapsto \mathbf{X}$ by an application of the Lyons–Victoir extension theorem, this map restricted to smooth paths does in general *not* coincide with the Riemann–Stieltjes integral of X against itself.

Remark 2.3. In typical applications to stochastic processes with α-Hölder continuous sample paths, $\alpha \in (\frac{1}{3}, \frac{1}{2})$, such as Brownian motion, rough path lift(s) are constructed via probability, and one does not rely on the extension theorem. In many cases, one has a "canonical" (a.k.a. Stratonovich, Wong-Zakai) lift of a process given as limit (in probability and rough path topology) of canonically lifted sample path mollification of the process. Examples where such a construction works include a large class of Gaussian processes, in particular Brownian motion, and more generally fractional Brownian motion for every Hurst parameter $H > \frac{1}{4}$, cf. Section 10. However, this may not be the only meaningful construction: already in Section 3, we will discuss three natural, but different, ways to lift Brownian motion to a rough path. For a detailed discussion of Markov (with uniformly elliptic generator in divergence form) and semimartingale rough paths we refer to [FV10b].

If one ignores the nonlinear constraint (2.1), the quantities defined in (2.3) suggest to think of $(X, \mathbb{X})$ as an element of the Banach space $\mathcal{C}^\alpha \oplus \mathcal{C}_2^{2\alpha}$ with (semi-)norm $\|X\|_\alpha + \|\mathbb{X}\|_{2\alpha}$ (which vanishes when X is constant and $\mathbb{X} \equiv 0$). However, taking into account (2.1) we see that $\mathscr{C}^\alpha$ is not a linear space, although it is a closed subset of the aforementioned Banach space; see Exercise 2.7. We will need (some sort of) a norm and metric on $\mathscr{C}^\alpha$. The induced "natural" norm on $\mathscr{C}^\alpha$ given by $\|X\|_\alpha + \|\mathbb{X}\|_{2\alpha}$ fails to respect the structure of (2.1) which is homogeneous with respect to a natural *dilation* on $\mathscr{C}^\alpha$, given by $\delta_\lambda : (X, \mathbb{X}) \mapsto (\lambda X, \lambda^2 \mathbb{X})$. This suggests to introduce the α-Hölder *homogeneous rough path norm*

[1] We note immediately that "smooth" can be replaced by "sufficiently smooth", such as $\mathcal{C}^1$ and even $\mathcal{C}^\alpha$, with $\alpha > 1/2$, in view of Young integration, Section 4.1.

[2] We deviate here from the early rough path literature, including [LQ02], where smooth rough paths meant canonical rough paths. Instead, we are aligned with the terminology of regularity structures, where (canonical, smooth) *models* generalise the corresponding notions of rough paths.

$$\|\mathbf{X}\|_\alpha \stackrel{\text{def}}{=} \|X\|_\alpha + \sqrt{\|\mathbb{X}\|_{2\alpha}} \;, \tag{2.4}$$

which, although not a norm in the usual sense of normed linear spaces, is a very adequate concept for the rough path $\mathbf{X} = (X, \mathbb{X})$. On the other hand, (2.3) leads to a natural notion of rough path metric (and then rough path topology).

Definition 2.4. Given rough paths $\mathbf{X}, \mathbf{Y} \in \mathscr{C}^\alpha([0,T], V)$, we define the (inhomogeneous) α-Hölder rough path metric [3]

$$\varrho_\alpha(\mathbf{X}, \mathbf{Y}) := \sup_{s \neq t \in [0,T]} \frac{|X_{s,t} - Y_{s,t}|}{|t-s|^\alpha} + \sup_{s \neq t \in [0,T]} \frac{|\mathbb{X}_{s,t} - \mathbb{Y}_{s,t}|}{|t-s|^{2\alpha}}.$$

The perhaps easiest way to show convergence with respect to this rough path metric is based on interpolation: in essence, it is enough to establish pointwise convergence together with uniform "rough path" bounds of the form (2.3); see Exercise 2.9. Let us also note that $\mathscr{C}^\alpha([0,T], V)$ endowed with this distance is a complete metric space; the reader is asked to work out the details in Exercise 2.7.

We conclude this part with two important remarks. First, we can ask ourselves up to which point the relations (2.1) are already sufficient to determine $\mathbb{X}$. Assume that we can associate to a given function X two different second order processes $\mathbb{X}$ and $\bar{\mathbb{X}}$, and set $G_{s,t} = \mathbb{X}_{s,t} - \bar{\mathbb{X}}_{s,t}$. It then follows immediately from (2.1) that

$$G_{s,t} = G_{u,t} + G_{s,u} \;,$$

so that in particular $G_{s,t} = G_{0,t} - G_{0,s}$. Since, conversely, we already noted that setting $\bar{\mathbb{X}}_{s,t} = \mathbb{X}_{s,t} + F_t - F_s$ for an arbitrary continuous function F does not change the left-hand side of (2.1), we conclude that $\mathbb{X}$ is in general determined only up to the increments of some function $F \in \mathcal{C}^{2\alpha}(V \otimes V)$. The choice of F *does* usually matter and there is in general no obvious canonical choice.

The second remark is that this construction can possibly be useful only if $\alpha \le \frac{1}{2}$. Indeed, if $\alpha > \frac{1}{2}$, then a canonical choice of $\mathbb{X}$ is given by reading (2.2) from right to left and interpreting the left-hand side as a simple Young integral [You36]. Furthermore, it is clear in this case that $\mathbb{X}$ must be unique, since any additional increment should be 2α-Hölder continuous by (2.3), which is of course only possible if $\alpha \le \frac{1}{2}$. Let us stress once more however that this is *not* to say that $\mathbb{X}$ is uniquely determined by X if the latter is smooth, when it is interpreted as an element of $\mathscr{C}^\alpha$ for some $\alpha \le \frac{1}{2}$. Indeed, if $\alpha \le \frac{1}{2}$, F is any 2α-Hölder continuous function with values in $V \otimes V$ and $\mathbb{X}_{s,t} = F_t - F_s$, then the path $(0, \mathbb{X})$ is a perfectly "legal" element of $\mathscr{C}^\alpha$, even though one cannot get any smoother than the function 0. The impact of perturbing $\mathbb{X}$ by some $F \in \mathcal{C}^{2\alpha}$ in the context of integration is considered

[3] As was already emphasised, $\mathscr{C}^\alpha$ is not a linear space but is naturally embedded in the Banach space $\mathcal{C}^\alpha \oplus \mathcal{C}_2^{2\alpha}$ (cf. Exercise 2.7), the (inhomogeneous) rough path metric is then essentially the induced metric. While this may not appear intrinsic (the situation is somewhat similar to using the (restricted) Euclidean metric on $\mathbf{R}^3$ on the 2-sphere), the ultimate justification is that the Itô map will turn out to be locally Lipschitz continuous in this metric.

in Example 4.14 below. In Chapter 5, we shall use this for a pathwise understanding of how exactly Itô and Stratonovich integrals differ.

Remark 2.5. There are some simple variations on the definition of a rough path, and it can be very helpful to switch from one view-point to the other. (The analytic conditions are not affected by this.)

a) From the "full increment" view point one has $(X, \mathbb{X}) : [0,T]^2 \to V \oplus V^{\otimes 2}$, $(s,t) \mapsto (X_{s,t}, \mathbb{X}_{s,t})$ subject to the "full" Chen relation

$$X_{s,t} = X_{s,u} + X_{u,t}, \quad \mathbb{X}_{s,t} = \mathbb{X}_{s,u} + \mathbb{X}_{u,t} + X_{s,u} \otimes X_{u,t} \; . \tag{2.5}$$

Every path $X : [0,T] \to V$ induces (vector) increments $X_{s,t} \equiv (\delta X)_{s,t} = X_t - X_s$ for which the first equality is a triviality. Conversely, increments determine a path modulo constants. In particular, $X_t = X_0 + X_{0,t}$ and this definition is equivalent to what we had in Definition 2.1), *if* restricted to paths with $X_0 = 0$ (or, less rigidly, by identifying paths $X, \bar{X}$ for which $\bar{X} - X$ is constant). In many situations, notably differential equations driven by $(X, \mathbb{X})$, this difference does not matter. (This increment view point is also closest to "models" (Π, Γ) in the theory of regularity structures, Section 13.3, where s is regarded as base-point and one is given a collection of functions $(X_{s,\cdot}, \mathbb{X}_{s,\cdot})$. The Chen relation (2.5) then has the interpretation of shifting the base-point.)

b) The "full path" view point starts with $\mathbf{X} : [0,T] \to \{1\} \times V \oplus V^{\otimes 2} \equiv T_1^{(2)}(V)$, a Lie group under the (truncated) tensor product, the details of which are left to Section 2.3 below. Every such path has *group increments* defined by

$$\mathbf{X}_s^{-1} \otimes \mathbf{X}_t =: \mathbf{X}_{s,t} =: (X_{s,t}, \mathbb{X}_{s,t}).$$

Chen's relation (2.5) is nothing but the trivial identity $\mathbf{X}_{s,u} \otimes \mathbf{X}_{u,t} = \mathbf{X}_{s,t}$ so that any such group-valued path $\mathbf{X}$ induces an increment map $(X, \mathbb{X})$, of the form discussed in a). Conversely, such increments determine $\mathbf{X}$ modulo constants as seen from $\mathbf{X}_t = \mathbf{X}_0 \otimes \mathbf{X}_{0,t}$. If we restrict to $\mathbf{X}_0 = \mathbf{1} = (1,0,0)$, or identify paths $\mathbf{X}, \tilde{\mathbf{X}}$ for which $\tilde{\mathbf{X}} \otimes \mathbf{X}^{-1}$ is constant, then there is no difference. (Such a "base-point free" object corresponds to "fat" $\mathbf{\Pi}$ in the theory of regularity structures and induces a model (Π, Γ) in great generality.)

c) Our Definition 2.1 is a compromise in the sense that we want to start from a familiar object, namely a path $X : [0,T] \to V$, together with minimal second level increment information to define (in Section 4.2) the prototypical rough integral $\int F(X) d(X, \mathbb{X})$. From the "increment" view point, we have thus supplied more than necessary (namely X_0), whereas from the "full path" view point, we have supplied $\mathbf{X}$, with $\mathbf{X}_0 = (1, X_0, *)$ specified on the first level only. (Of course, this affects in no way the second level increments $\mathbb{X}_{s,t}$.)

2.2 The space of geometric rough paths

While (2.1) does capture the most basic (additivity) property that one expects any decent theory of integration to respect, it does *not* imply any form of integration by parts / chain rule. Now, if one looks for a first order calculus setting, such as is valid in the context of smooth paths or the Stratonovich stochastic calculus, then for any pair e_i^*, e_j^* of elements in V^*, writing $X_t^i = e_i^*(X_t)$ and $\mathbb{X}_{s,t}^{ij} = (e_i^* \otimes e_j^*)(\mathbb{X}_{s,t})$, one would expect to have the identity

$$\begin{aligned}\mathbb{X}_{s,t}^{ij} + \mathbb{X}_{s,t}^{ji} \text{ ``=''} &\int_s^t X_{s,r}^i \, dX_r^j + \int_s^t X_{s,r}^j \, dX_r^i \\ &= \int_s^t d(X^i X^j)_r - X_s^i \, X_{s,t}^j - X_s^j \, X_{s,t}^i \\ &= (X^i X^j)_{s,t} - X_s^i \, X_{s,t}^j - X_s^j \, X_{s,t}^i = X_{s,t}^i \, X_{s,t}^j \,,\end{aligned}$$

so that the symmetric part of $\mathbb{X}$ is determined by X. In other words, for all times s, t we have the "first order calculus" condition

$$\mathrm{Sym}(\mathbb{X}_{s,t}) = \frac{1}{2} X_{s,t} \otimes X_{s,t} \,. \tag{2.6}$$

However, if we take X to be an n-dimensional Brownian path and define $\mathbb{X}$ by Itô integration, then (2.1) still holds, but (2.6) certainly does not.

There are two natural ways to define a set of "geometric" rough paths for which (2.6) holds. On the one hand, we can define the space of *weakly geometric (α-Hölder) rough paths*.

$$\mathscr{C}_g^\alpha([0,T], V) \subset \mathscr{C}^\alpha([0,T], V) \,,$$

by stipulating that $(X, \mathbb{X}) \in \mathscr{C}_g^\alpha$ if and only if $(X, \mathbb{X}) \in \mathscr{C}^\alpha$ and (2.6) holds as equality in $V \otimes V$, for every $s, t \in [0,T]$. Note that $\mathscr{C}_g^\alpha$ is a closed subset of $\mathscr{C}^\alpha$.

On the other hand, we have already seen that every smooth path can be lifted canonically to an element in $\mathscr{L}(\mathcal{C}^\infty) \subset \mathscr{C}^\alpha$ by reading the definition (2.2) from right to left. This choice of $\mathbb{X}$ then obviously satisfies (2.6) and we can define the space of *geometric (α-Hölder) rough paths*,

$$\mathscr{C}_g^{0,\alpha}([0,T], V) \subset \mathscr{C}^\alpha([0,T], V) \,,$$

as the closure of $\mathscr{L}(\mathcal{C}^\infty)$ in $\mathscr{C}^\alpha$. We leave it as exercise to the reader to see that $\mathcal{C}^\infty$ here may be replaced by $\mathcal{C}^1$ paths without changing the resulting space of geometric rough paths.

One has the obvious inclusion $\mathscr{C}_g^{0,\alpha} \subset \mathscr{C}_g^\alpha$, which turns out to be strict. In fact, $\mathscr{C}_g^{0,\alpha}$ is separable (provided V is separable), whereas $\mathscr{C}_g^\alpha$ is not, cf. Exercise 2.8 below. The situation is similar to the classical situation of the set of α-Hölder continuous functions being strictly larger than the closure of smooth functions under the α-Hölder norm. (Or the set of bounded measurable functions being strictly larger than $\mathcal{C}$, the closure of smooth functions under the supremum norm.) In practice, at least when $\dim V < \infty$, the distinction between weakly and "genuinely" geometric rough paths rarely matters for the following reason: similar to classical Hölder spaces, one has the converse inclusion $\mathscr{C}_g^\beta \subset \mathscr{C}_g^{0,\alpha}$ whenever $\beta > \alpha$, see Proposition 2.8 below and also Exercise 2.12. For this reason, we will often casually speak of "geometric rough paths", even when we mean weakly geometric rough paths. (There is no confusion in precise statements when we write $\mathscr{C}_g^{0,\alpha}$ or $\mathscr{C}_g^\alpha$.) Let us finally mention that non-geometric rough paths can always be embedded in a space of geometric rough paths at the expense of adding new components; in the present (level-2) setting this can be accomplished in terms of a *rough path bracket*, see Exercise 2.11 and also Section 5.3.

2.3 Rough paths as Lie group valued paths

We now present a very fruitful view of rough paths, taken over a Banach space V. Consider $X : [0,T] \to V$, $\mathbb{X} : [0,T]^2 \to V^{\otimes 2}$ subject to (2.1) and define, with $X_{s,t} = X_t - X_s$ as usual,

$$\mathbf{X}_{s,t} := (1, X_{s,t}, \mathbb{X}_{s,t}) \in \mathbf{R} \oplus V \oplus V^{\otimes 2} \stackrel{\text{def}}{=} T^{(2)}(V). \tag{2.7}$$

The space $T^{(2)}(V)$ is itself a Banach space, with the norm of an element (a,b,c) given by $|a| + |b| + |c|$, where in abusive notation $|\bullet|$ standards for any of the norms in $\mathbf{R}$, V and $V \otimes V$, the norm on the latter assumed compatible and symmetric, cf. Section 1.4 . More interestingly for our purposes, this space is a Banach algebra, non-commutative when $\dim V > 1$ and with unit element $(1,0,0)$, when endowed with the product

$$(a,b,c) \otimes (a',b',c') \stackrel{\text{def}}{=} (aa', ab' + a'b, ac' + a'c + b \otimes b') \; .$$

We call $T^{(2)}(V)$ the *step-*2 *truncated tensor algebra over* V. This multiplicative structure is very well adapted to our needs since Chen's relation (2.1), combined with the obvious identity $X_{s,t} = X_{s,u} + X_{u,t}$, takes the elegant form

$$\mathbf{X}_{s,t} = \mathbf{X}_{s,u} \otimes \mathbf{X}_{u,t} \; . \tag{2.8}$$

Set $T_a^{(2)}(V) \stackrel{\text{def}}{=} \{(a,b,c) : b \in V,\, c \in V \otimes V\}$. As suggested by (2.7), the affine subspace $T_1^{(2)}(V)$ will play a special role for us. We remark that each of its elements has an inverse given by

$$(1,b,c) \otimes (1,-b,-c+b\otimes b) = (1,-b,-c+b\otimes b) \otimes (1,b,c) = (1,0,0) \; , \tag{2.9}$$

so that $T_1^{(2)}(V)$ is a *Lie group*.[4] It follows that $\mathbf{X}_{s,t} = \mathbf{X}_{0,s}^{-1} \otimes \mathbf{X}_{0,t}$ are the natural increments of the group valued path $t \mapsto \mathbf{X}_{0,t} =: \mathbf{X}_t$.

[4] The Lie group $T_1^{(2)}(V)$ is finite-dimensional if and only if $\dim V < \infty$.

Identifying $1, b, c$ with elements $(1,0,0), (0,b,0), (0,0,c) \in T^{(2)}(V)$, we may write $(1,b,c) = 1+b+c$. The resulting calculus is familiar from formal power series in non-commuting indeterminates. For instance, the usual power series $(1+x)^{-1} = 1 - x + x^2 - \dots$ leads to, omitting tensors of order 3 and higher,

$$\begin{aligned}(1+b+c)^{-1} &= 1 - (b+c) + (b+c)\otimes(b+c)\\ &= 1 - b - c + b\otimes b\,,\end{aligned}$$

allowing us to recover (2.9). We also introduce the *dilation* operator δ_λ on $T^{(2)}(V)$, with $\lambda \in \mathbf{R}$, which acts by multiplication with λ^n on the nth tensor level $V^{\otimes n}$, namely

$$\delta_\lambda : (a,b,c) \mapsto \big(a, \lambda b, \lambda^2 c\big)\;.$$

Having identified $T_1^{(2)}(V)$ as the natural state space of (step-2) rough paths, we now equip it with a *homogeneous, symmetric* and *subadditive norm*. For $\mathbf{x} = (1,b,c)$,

$$|||\mathbf{x}||| \stackrel{\text{def}}{=} \tfrac{1}{2}\big(N(\mathbf{x}) + N(\mathbf{x}^{-1})\big) \quad \text{with } N(\mathbf{x}) = \max\{|b|, \sqrt{2|c|}\}\;, \tag{2.10}$$

noting $|||\delta_\lambda \mathbf{x}||| = |\lambda| |||\mathbf{x}|||$, homogeneity with respect to dilation, and $|||\mathbf{x}\otimes\mathbf{x}'||| \le |||\mathbf{x}||| + |||\mathbf{x}'|||$, a consequence of subadditivity for $N(\bullet)$ which requires a short argument left to the reader. It is clear that

$$(\mathbf{x}, \mathbf{x}') \mapsto |||\mathbf{x}^{-1}\otimes\mathbf{x}'\,||| \stackrel{\text{def}}{=} d(\mathbf{x},\mathbf{x}')$$

defines a bona fide (left-invariant) metric on the group $T_1^{(2)}(V)$. Important for us, the graded Hölder regularity (2.3) of $\mathbf{X} = (X, \mathbb{X})$, part of the definition of a rough path, can now be condensed to demand the "metric" Hölder seminorm

$$\sup_{s\neq t\in[0,T]} \frac{d(\mathbf{X}_s, \mathbf{X}_t)}{|t-s|^\alpha} \asymp \|X\|_\alpha + \sqrt{\|\mathbb{X}\|_{2\alpha}} = |||\mathbf{X}|||_{\alpha;[0,T]} \tag{2.11}$$

to be finite. To summarise, we arrived at the following appealing characterisation of (Banach space valued) rough paths.

Proposition 2.6. *(Hölder continuity is with respect to the left-invariant metric d.)*

a) Assume $(X,\mathbb{X}) \in \mathscr{C}^\alpha([0,T],V)$. *Then the path* $t\mapsto \mathbf{X}_t = (1, X_{0,t}, \mathbb{X}_{0,t})$, *with values in* $T_1^{(2)}(V)$ *is* α*-Hölder continuous.*
b) Conversely, if $[0,T]\ni t\mapsto \mathbf{X}_t$ *is a* $T_1^{(2)}(V)$*-valued and* α*-Hölder continuous path, then* $(X,\mathbb{X}) \in \mathscr{C}^\alpha([0,T],V)$ *with* $(1, X_{s,t}, \mathbb{X}_{s,t}) := \mathbf{X}_s^{-1}\otimes\mathbf{X}_t$.

The usual power series and / or basic Lie group theory suggest to define

$$\log(1+b+c) \stackrel{\text{def}}{=} b + c - \frac{1}{2} b\otimes b\,, \tag{2.12}$$

$$\exp(b+c) \stackrel{\text{def}}{=} 1 + b + c + \frac{1}{2} b\otimes b\,, \tag{2.13}$$

which allow us to identify $T_0^{(2)}(V) \cong V \oplus V^{\otimes 2}$ with $T_1^{(2)}(V) = \exp(V \oplus V^{\otimes 2})$. The following *Lie bracket* makes $T_0^{(2)}(V)$ a *Lie algebra*. For $b, b' \in V, c, c' \in V^{\otimes 2}$,

$$[b + c, b' + c'] \stackrel{\text{def}}{=} b \otimes b' - b' \otimes b \,,$$

and $T_0^{(2)}(V)$ is *step-2 nilpotent* in the sense that all iterated brackets of length 2 vanish. Define $\mathfrak{g}^{(2)}(V) \subset T_0^{(2)}(V)$ as the closed Lie subalgebra generated by $V \subset T_0^{(2)}(V)$, explicitly given by

$$\mathfrak{g}^{(2)}(V) = V \oplus [V, V] \;\text{ with } [V, V] \stackrel{\text{def}}{=} \text{cl}(\text{span}\{[v, w] : v, w \in V\}) \,,$$

called the *free step-2 nilpotent Lie algebra over V*. Note that in finite dimensions, say $V = \mathbf{R}^d$, the closing procedure is unnecessary and $[V, V]$ is nothing but the space of antisymmetric $d \times d$ matrices, with linear basis $([e_i, e_j] : 1 \le i < j \le d)$, where $(e_i : 1 \le i \le d)$ denotes the standard basis of $\mathbf{R}^d$. Thanks to step-2 nilpotency, one checks by hand the *Baker–Campbell–Hausdorff formula*

$$\exp(b + c) \otimes \exp(b' + c') = \exp(b + b' + c + c' + \tfrac{1}{2}[b, b']) \,.$$

The image of $\mathfrak{g}^{(2)}$ under the exponential map then defines a closed Lie subgroup,

$$G^{(2)}(V) \stackrel{\text{def}}{=} \exp\bigl(\mathfrak{g}^{(2)}(V)\bigr) \subset T_1^{(2)}(V) \,,$$

called the *free step-2 nilpotent group over V*. These considerations provide us with an elegant characterisation of weakly geometric rough paths. (The proof is immediate from the previous proposition and rewriting (2.6) as $\mathbb{X}_{s,t} - \frac{1}{2} X_{s,t} \otimes X_{s,t} \in [V, V]$.)

Proposition 2.7 (Weakly geometric case).

a) Assume $(X, \mathbb{X}) \in \mathscr{C}_g^\alpha([0, T], V)$. Then the path $t \mapsto \mathbf{X}_t = (1, X_{0,t}, \mathbb{X}_{0,t})$, with values in $G^{(2)}(V)$ is α-Hölder continuous (with respect to the metric d.)

b) Conversely, if $[0, T] \ni t \mapsto \mathbf{X}_t$ is a $G^{(2)}(V)$-valued and α-Hölder continuous path, then $(X, \mathbb{X}) \in \mathscr{C}_g^\alpha([0, T], V)$ with $(1, X_{s,t}, \mathbb{X}_{s,t}) := \mathbf{X}_s^{-1} \otimes \mathbf{X}_t$.

It is clear from the discussion in Section 2.2 that any sufficiently smooth path, say $\gamma \in \mathcal{C}^1([0, 1], V)$, produces an element in $G^{(2)}(V)$ by iterated integration, namely

$$S^{(2)}(\gamma) = \Bigl(1, \int_0^1 d\gamma(t), \int_0^1 \int_0^t d\gamma(s) \otimes d\gamma(t)\Bigr) \in G^{(2)}(V) \,.$$

The map $S^{(2)}$, which maps (sufficiently regular) paths on a fixed interval, here $[0, 1]$, into the above collection of tensors is know as *step-2 signature map*. We note in passing that Chen's relation here has the pretty interpretation that the signature map is a morphism from the space of paths, equipped with concatenation product, to the tensor algebra. The inclusion $S^{(2)}(\mathcal{C}^1) \subset G^{(2)}$ becomes an equality in finite dimensions,

$$\{S^{(2)}(\gamma) : \gamma \in \mathcal{C}^1([0, 1], \mathbf{R}^d)\} \;=\; G^{(2)}(\mathbf{R}^d) \,. \tag{2.14}$$

To see this, fix $b + c \in \mathfrak{g}^{(2)}(\mathbf{R}^d)$ and try to find finitely many, say n, affine linear paths γ_i, with each signature determined by the direction $\gamma_i(1) - \gamma_i(0) = v_i \in \mathbf{R}^d$, so that

$$\exp(v_1) \otimes \ldots \otimes \exp(v_n) = \exp(b + c) \; .$$

Properly applied, the Baker–Campbell–Hausdorff formula allows to "break up" the exponential $\exp(\sum_i b^i e_i + \sum_{j,k} c^{jk}[e_j, e_k])$. In conjunction with the identity $e^{[v,w]} = e^{-w} \otimes e^{-v} \otimes e^w \otimes e^v$ it is easy to find a possible choice of $v_1, \ldots, v_n$. By concatenation of the γ_i's one has constructed a path γ with prescribed signature $S^{(2)}(\gamma) = \exp(b + c)$. This path is clearly in $\mathcal{C}^1$, the space of Lipschitz paths.[5] This gives a very natural way to introduce another (homogeneous, symmetric, subadditive) norm on $G^{(2)}(\mathbf{R}^d)$, namely

$$\|\mathbf{x}\|_{\mathrm{C}} \stackrel{\text{def}}{=} \inf\Big\{ \int_0^1 |\dot\gamma(t)|\, dt \; : \; \gamma \in \mathcal{C}^1([0,1], \mathbf{R}^d) \; , \quad S^{(2)}(\gamma) = \mathbf{x} \Big\} \; , \tag{2.15}$$

known as *Carnot–Carathéodory norm*. (In infinite dimensions, there is no guarantee for the set on the right-hand side to be non-empty.) When equipped with its Euclidean structure, $\mathbf{R}^d$ defines a "horizontal" subspace $\mathbf{R}^d \times \{0\} \subset \mathfrak{g}^{(2)}(\mathbf{R}^d)$, seen as tangent space to $G^{(2)}(\mathbf{R}^d)$ at $(1, 0, 0)$ which in turn induces a left-invariant sub-Riemannian structure on $G^{(2)}(\mathbf{R}^d)$. The associated left-invariant Carnot–Carathéodory metric d_{C} can then be seen as the minimal length of "horizontal" paths connecting two points. Any minimising sequence in (2.15), parametrised by constant speed, is equicontinuous so that by Arzela–Ascoli such minimisers, also called *sub-Riemannian geodesics*, exist and must be in $\mathcal{C}^1$. Such geodesics are a key tool in the approach of Friz–Victoir [FV10b]. The explicit computation of such geodesics (and Carnot–Carathéodory norms) is a difficult problem, with explicit formulae available for $d = 2$, noting that, as Lie groups, $G^{(2)}(\mathbf{R}^2) \cong \mathbf{H}^3$, the 3-dimensional Heisenberg group, see e.g. [Mon02]. Fortunately, a compactness argument, as detailed for example in [FV10b, Sec 7.5], shows that all continuous homogeneous norms are equivalent. Upon checking continuity of the Carnot–Carathéodory norm, one gets, for $\mathbf{x} = \exp(b + c) \in G^{(2)}(\mathbf{R}^d)$,

$$\|\mathbf{x}\|_{\mathrm{C}} \asymp_d |b| + |c|^{1/2} \asymp \max\{|b|, |c|^{1/2}\} \; , \tag{2.16}$$

which, despite its dependence on the dimension d, is sufficient for many practical purposes. As a useful application, we now state an approximation result for weakly geometric roughs over $\mathbf{R}^d$. With the preparations made, the interested reader will have no trouble to provide a full proof for

Proposition 2.8 (Geodesic approximation). *For every* $(X, \mathbb{X}) \in \mathscr{C}_g^\beta([0,T], \mathbf{R}^d)$, *there exists a sequence of smooth paths* $X^n : [0,T] \to \mathbf{R}^d$ *such that*

[5] In fact, by smoothly slowing down speed to zero whenever switching directions, the path γ can also be parametrized to be smooth. In particular, in (2.14) and (2.15) below we could have replaced $\mathcal{C}^1$ by $\mathcal{C}^\infty$.

$$(X^n, \mathbb{X}^n) := \left(X^n, \int_0^\cdot X^n_{0,t} \otimes dX^n_t\right) \to (X, \mathbb{X}) \text{ uniformly on } [0,T]$$

with uniform rough path bound $\sup_{n\geq 1} \|X^n, \mathbb{X}^n\|_\beta \lesssim \|X, \mathbb{X}\|_\beta$. *By interpolation, convergence holds in* $\mathscr{C}^\alpha$*, for any* $\alpha < \beta$.

Remark 2.9. By definition, every geometric rough path $\mathbf{X} \in \mathscr{C}_g^{0,\beta}$ is the limit of canonical rough path lifts $(X^n, \mathbb{X}^n) = \mathbf{X}^n$; trivially then, $\|\mathbf{X}^n\|_\beta \to \|\mathbf{X}\|_\beta$. This is not true for a generic weakly geometric rough path $\mathbf{X} \in \mathscr{C}_g^\beta$. However, the above proposition supplies approximations $(\mathbf{X}^n)$, which converge uniformly with uniform rough paths bounds. In such a case, $\|\mathbf{X}\|_\beta \leq \liminf_{n\geq 1} \|\mathbf{X}^n\|_\beta$ and this can be strict. This lower-semicontinuous behaviour of the rough path norm is reminiscent of norms on Hilbert spaces under weak convergence and led to the terminology of "weakly" geometric rough paths.

2.4 Geometric rough paths of low regularity

The interpretation given above gives a strong hint on how to construct geometric rough paths with α-Hölder regularity for $\alpha \leq \frac{1}{3}$: setting $N = \lfloor 1/\alpha \rfloor$, one defines the step-$N$ truncated tensor algebra over a Banach space V

$$T^{(N)}(V) \stackrel{\text{def}}{=} \bigoplus_{n=0}^{N} (V)^{\otimes n} ,$$

with the natural convention that $(V)^{\otimes 0} = \mathbf{R}$. The product in $T^{(N)}(V)$ is simply the tensor product $\otimes$, but we truncate it in a natural way by postulating that $a \otimes b = 0$ for $a \in (V)^{\otimes k}$, $b \in (V)^{\otimes \ell}$ with $k + \ell > N$. A homogeneous, symmetric and subadditive norm which generalises (2.10) to the step-N case is given by

$$\|\mathbf{x}\| \stackrel{\text{def}}{=} \tfrac{1}{2}\big(N(\mathbf{x}) + N(\mathbf{x}^{-1})\big) \quad \text{with } N(\mathbf{x}) = \max_{n=1,\ldots,N} \left(n! |\mathbf{x}^n|\right)^{1/n} , \tag{2.17}$$

where every $\mathbf{x} = (1, \mathbf{x}^1, \ldots, \mathbf{x}^N) \in T_1^{(N)}(V)$, element with scalar component 1, is invertible, and where $|\bullet|$ denotes any of the tensor norms on $(V)^{\otimes n}$, assumed compatible and symmetric (permutation invariant).[6].

Proposition 2.6 suggests the naïve definition of an *α-Hölder rough path* over V as a path $\mathbf{X}$, on $[0,T]$ say, with values in the group $T_1^{(N)}(V)$ which is α-Hölder continuous with respect to $d(\mathbf{x}, \mathbf{x}') = \|\mathbf{x}^{-1} \otimes \mathbf{x}'\|$. Modulo knowledge of $\mathbf{X}_0$ this is equivalent to a multiplicative map $(s,t) \mapsto \mathbf{X}_{s,t} \in T_1^{(N)}(V)$, *multiplicative* in the sense that Chen's relation holds,

$$\mathbf{X}_{s,t} = \mathbf{X}_{s,u} \otimes \mathbf{X}_{u,t} , \tag{2.18}$$

[6] The definitions from Section 1.4 for $N = 2$ extend easily to $N > 2$, see also [LCL07, Def 1.25]

for every triplet of times (s,u,t), and with graded Hölder regularity,

$$|\mathbf{X}^n_{s,t}| \lesssim |t-s|^{k\alpha}, \qquad n=1,\dots,N\ ,$$

uniformly over $s,t \in [0,T]$. The interpretation of rough paths discussed at length in the step-2 setting is unchanged and $\mathbf{X}^n_{s,t} \in V^{\otimes n}$ should be thought of as a substitute for the (possibly ill-defined) n-fold integral $\int dX_{u_1} \otimes \cdots \otimes dX_{u_n}$ over the n-simplex $\{s < u_1 < \ldots < u_n < t\}$. Such a notion of naïve higher oder rough path is sometimes sufficient, e.g. for solving linear rough differential equations, see also Exercise 4.18, but does not contain the necessary information to deal with non-linearities, already seen in the simple example of the form $\int_s^t (X_r - X_s)^{\otimes 2} \otimes dX_r$.

Higher order (weakly) geometric rough paths resolve this problem by imposing a chain rule. In the above example, $(\delta X)^{\otimes 2}/2 = \mathrm{Sym}(\mathbf{X}^2)$, formerly written as $\mathrm{Sym}(\mathbb{X})$, and the situation is reduced to (a linear combination of) 3-fold iterated integrals. To proceed in a systematic fashion, we first introduce the correct state space as the *free step-N nilpotent Lie group over V*

$$G^{(N)}(V) \stackrel{\text{def}}{=} \exp(\mathfrak{g}^{(N)}(V)) \subset T_1^{(N)}(V)$$

where the exponential map is defined via its power series and $\mathfrak{g}^{(N)} \subset T_0^{(N)}(V)$ is the (closed) Lie algebra generated by all elements of the form $(0,c,0,\dots,0)$ with $c \in V$ via the natural Lie bracket $[a,b] = a \otimes b - b \otimes a$. The neutral element $\mathbf{1} \in G^{(N)}(V)$ is given by $\mathbf{1} = (1,0,\dots,0)$. Given any $\alpha \in (0,1]$ and $N = \lfloor 1/\alpha \rfloor$ as the number of "levels", Proposition 2.7 now suggests the definition of a *weakly geometric α-Hölder rough path* over V as a path $\mathbf{X}$, on $[0,T]$ say, with values in the group $G^{(N)}(V)$ which is α-Hölder continuous with respect to $d(\mathbf{x},\mathbf{x}') = \|\!|\mathbf{x}^{-1} \otimes \mathbf{x}'\|\!|$. Modulo knowledge of $\mathbf{X}_0$ this is equivalent to a multiplicative map $(s,t) \mapsto \mathbf{X}_{s,t} \in G^{(N)}(V)$ with graded Hölder regularity, uniformly over $s,t \in [0,T]$,

$$|\mathbf{X}^n_{s,t}| \lesssim |t-s|^{n\alpha}, \qquad n=1,\dots,N\ .$$

Here, again multiplicative means validity of Chen's relation as spelled out in (2.18) above.

We now assume, for notationally convenience, $V = \mathbf{R}^d$, which allows us to think of components of some fixed rough path increment $\mathbf{X}_{s,t} \in T_1^{(N)}(\mathbf{R}^d)$ as being indexed by words w of length at most N with letters in the alphabet $\{1,\dots,d\}$. Similarly to before, given a word $w = w_1 \cdots w_n$, the corresponding component $\mathbf{X}^w$, which we also write as $\langle \mathbf{X}, w\rangle$, is then interpreted as the n-fold integral

$$\langle \mathbf{X}_{s,t}, w\rangle = \int_s^t \int_s^{s_n} \cdots \int_s^{s_1} dX^{w_1}_{s_1} \cdots dX^{w_n}_{s_n}\ , \tag{2.19}$$

and $\|\!|\mathbf{X}_{s,t}\|\!| \lesssim |t-s|^\alpha$ is equivalent to, for all words with length $|w| \le \lfloor 1/\alpha \rfloor$,

$$|\langle \mathbf{X}_{s,t}, w\rangle| \lesssim |t-s|^{\alpha|w|}\ . \tag{2.20}$$

In order to describe the constraints imposed on these iterated integrals by the chain rule, we define the *shuffle product* $\sqcup\!\sqcup$ between two words as the formal sum over all possible ways of interleaving them. For example, one has

$$a \sqcup\!\sqcup x = ax + xa\;, \quad ab \sqcup\!\sqcup xy = abxy + axby + xaby + axyb + xayb + xyab\;,$$

with the empty word acting as the neutral element. With this notation at hand, it was already remarked by Ree [Ree58] (see also [Che71]) that the chain rule implies the identity

$$\langle \mathbf{X}_{s,t}, v\rangle\langle \mathbf{X}_{s,t}, w\rangle = \langle \mathbf{X}_{s,t}, v \sqcup\!\sqcup w\rangle\;. \tag{2.21}$$

(The reader is asked to show this in Exercise 2.2.) It is a remarkable fact that the algebraic properties of the tensor and shuffle algebras combine in such a way that the set of elements $\mathbf{X} \in T^{(N)}$ satisfying (2.21) is not only stable under the product $\otimes$, but forms a group, which in turn was shown in [Ree58] to be nothing but the group $G^{(N)}(\mathbf{R}^d)$. In the language of Hopf algebras, this group is exactly the *character group* for the (truncated) shuffle Hopf algebra.

In general, one may decide to forego the chain rule (after all, it doesn't hold in the context of Itô integration, as is manifest in Itô's formula) in which case there is no reason to impose (2.21). In this case, considering a rough path as an enhancement of a path X by iterated integrals of the type (2.19) no longer provides sufficient additional data. Indeed, in order to solve differential equations driven by X, one would like to give meaning to expressions like for example

$$\int_s^t \Big(\int_s^r dX^i_u\Big)\Big(\int_s^r dX^j_v\Big) dX^k_r =: \langle \mathbf{X}_{s,t}, {}^{i}\!\curlyvee^{j}_{k}\, \rangle\;. \tag{2.22}$$

We already remarked earlier, that in the (weakly) geometric case, the assumed chain rule (now in the form of (2.21)) allows to reduce such expressions to linear combinations of iterated integrals. In general, one should define a rough path as the enhancement of a path X with additional functions that are interpreted as the various formal expressions that can be formed by the two operations "multiplication" and "integration against X". The resulting algebraic construction is more involved and gives rise to the concept of *branched rough path* X due to Gubinelli [Gub10]. The terminology comes from the fact that the natural way of indexing the components of such an object is no longer given by words, but by labelled trees, as suggested in (2.22) above with labels $i, j, k \in \{1, \ldots, d\}$. As detailed in [Gub10], see also [HK15, BCFP19], branched rough paths take values in the character group of the *Connes–Kreimer Hopf algebra* of trees [CK00], also known as the *Butcher group* [But72]. A concise description of the branched rough path regularity via an explicit homogeneous subadditive norms on this Lie group, similar to (2.17), can be found in [TZ18], cf. also [HS90].

2.5 Exercises

♯ **Exercise 2.1** *Let X be a smooth V-valued path.*

a) Show that $\mathbb{X}_{s,t} := \int_s^t X_{s,r} \otimes \dot{X}_r \, dr$ satisfy Chen's relation (2.1).
b) Consider the collection of all iterated integrals over $[s,t]$,

$$\mathbf{X}_{s,t} := \left(1, X_{s,t}, \mathbb{X}_{s,t}, \int_{\Delta^{(3)}_{s,t}} dX_{u_1} \otimes dX_{u_2} \otimes dX_{u_3}, \dots \right) \in T((V)) \;, \tag{2.23}$$

where $\Delta^{(3)}_{s,t} = \{u \,:\, s < u_1 < u_2 < u_3 < t\}$ and $T((V)) \stackrel{def}{=} \prod_{k=0}^{\infty} V^{\otimes k}$ is the space of tensor series over V, equipped with the obvious algebra structure (cf. Section 2.4). Show that the following general form of Chen's relation holds:

$$\mathbf{X}_{s,t} = \mathbf{X}_{s,u} \otimes \mathbf{X}_{u,t} \;.$$

The element $\mathbf{X}_{s,t} \in T((V))$ is known as the signature of X on the interval $[s,t]$.
c) Show that the indefinite signature $\mathbf{S} := \mathbf{X}_{0,\cdot}$ solves the linear differential equation

$$d\mathbf{S} = \mathbf{S} \otimes dX \;, \qquad \mathbf{S}_0 = \mathbf{1} \;.$$

We will see later (Exercises 4.6 and 8.9) that the signature can be defined for every rough path.

Hint: *For point (b), it suffices to consider the projection of $\mathbf{X}_{s,t}$ to $V^{\otimes n}$, for an arbitrary integer n, given by the n-fold integral of $dX_{u_1} \otimes \ldots \otimes dX_{u_n}$ over the simplex $\{s < u_1 < \ldots < u_n < t\}$.*

♯ **Exercise 2.2 (Shuffle)** *Let $V = \mathbf{R}^d$. As discussed in (2.19), the collection $\mathbf{X}_{s,t}$ of all iterated integrals over a fixed interval $[s,t]$ can also be viewed as*

$$\left\{ \mathbf{X}^w_{s,t} = \langle \mathbf{X}_{s,t}, w \rangle : w \text{ word on } \mathcal{A} \right\} ,$$

with alphabet $\mathcal{A} = \{1, \dots, d\}$, where we recall that a word *on $\mathcal{A}$ is a finite sequence of elements of $\mathcal{A}$, including the empty sequence $\emptyset$, called the* empty word. *By convention, $\mathbf{X}^{\emptyset}_{s,t} = 1$. Write uv for the concatenation of two words u and v, and accordingly ui for attaching a letter $i \in \mathcal{A}$ to the right of u. The linear span of such words (which can be identified with polynomials in d non-commuting indeterminates) carries an important commutative product known as the* shuffle product. *It is defined recursively by requiring $\emptyset$ to be the neutral element, ie. $u \sqcup\!\sqcup \emptyset = \emptyset \sqcup\!\sqcup u = u$, and then*

$$ui \sqcup\!\sqcup vj = (u \sqcup\!\sqcup vj)i + (ui \sqcup\!\sqcup v)j \;.$$

Let $\mathbf{X}_{s,t}$ be the signature of a smooth path X, as given in (2.23). Show that, for all words u, v,

$$\langle \mathbf{X}_{s,t}, u \sqcup\!\sqcup v \rangle = \langle \mathbf{X}_{s,t}, u \rangle \langle \mathbf{X}_{s,t}, v \rangle \;. \tag{2.24}$$

The case of single letter words $w = i, v = j$ gives $i \sqcup\!\sqcup\, j = ij + ji$ and expresses precisely the product rule from calculus, which leads us to the level-2 geometricity condition (2.6).

Hint: *Proceed by induction in joint length: express $\langle \mathbf{X}_{s,t}, ui\rangle\langle \mathbf{X}_{s,t}, vj\rangle$ by the product rule as an integral over $[s,t]$ and use the hypothesis for words of joint length $|u| + |v| + 1 < |ui| + |vj|$.*

$*$ **Exercise 2.3** *Call a tensor series $\mathbf{x} \in T((\mathbf{R}^d))$* group-like, *in symbols $\mathbf{x} \in G((\mathbf{R}^d))$, if for all words u, v,*

$$\langle \mathbf{x}, u \sqcup\!\sqcup\, v\rangle = \langle \mathbf{x}, u\rangle\langle \mathbf{x}, v\rangle \ . \tag{2.25}$$

An element in $T((\mathbf{R}^d))$ is called a Lie series *if, for all $N \in \mathbf{N}$, its projection to $T^{(N)} = T^{(N)}(\mathbf{R}^d)$ is a* Lie polynomial, *i.e. an element of $\mathfrak{g}^{(N)}$, which was defined in Section 2.4 as the Lie algebra generated by $\mathbf{R}^d \subset T_0^{(N)}$. Given $\mathbf{x} \in T((\mathbf{R}^d))$, show that $\mathbf{x}$ is group-like, i.e. $\mathbf{x} \in G((\mathbf{R}^d))$, if and only if $\log \mathbf{x}$ is a Lie series.*

$\sharp$ **Exercise 2.4**

a) It is common to define the $(V \otimes V)$-valued map $\mathbb{X}$ on $\Delta_{0,T} := \{(s,t) : 0 \le s \le t \le T\}$ rather than $[0,T]^2$. There is no difference however: if $\mathbb{X}_{s,t}$ is only defined for $s \le t$, show that the relation (2.1) implies

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

b) In fact, show that knowledge of the path $t \mapsto (X_{0,t}, \mathbb{X}_{0,t})$ already determines the entire second order process $\mathbb{X}$. In this sense $(X, \mathbb{X})$ is indeed a path, and not some two-parameter object, cf. Remark 2.5.

c) Specialise to the case of geometric rough path and show the identity $\mathbb{X}_{t,s} = \mathbb{X}_{t,s}^T$ where $(\dots)^T$ denotes the transpose. (When $\dim V = 1$, so that X is scalar valued, this is a trivial consequence of $\mathbb{X}_{s,t} = X_{s,t}^2/2$.)

Exercise 2.5 *Consider $s \equiv \tau_0 < \tau_1 < \dots < \tau_N \equiv t$. Show that (2.1) implies*

$$\begin{aligned}\mathbb{X}_{s,t} &= \sum_{0\le i<N} \mathbb{X}_{\tau_i,\tau_{i+1}} + \sum_{0\le j<i<N} X_{\tau_j,\tau_{j+1}} \otimes X_{\tau_i,\tau_{i+1}} \\ &= \sum_{i=0}^{N-1} \big(\mathbb{X}_{\tau_i,\tau_{i+1}} + X_{s,\tau_i} \otimes X_{\tau_i,\tau_{i+1}}\big).\end{aligned} \tag{2.26}$$

This identity effectively compares $\mathbb{X}_{s,t}$ with a left-point Riemann-Stieltjes approximation $\sum_{i=0}^{N-1} X_{s,\tau_i} \otimes X_{\tau_i,\tau_{i+1}}$ of the "motivating" integral expression in (2.2).

Exercise 2.6 *Following Section 2.3 and Exercise 2.4, view $\mathbf{X} \in \mathscr{C}^\alpha([0,T], V)$ as a one-parameter path and define the (time T) time reversal of $\mathbf{X}$ in the "naïve" way as*

$$\overleftarrow{\mathbf{X}}_t = \mathbf{X}_{T-t} \ , \qquad 0 \le t \le T \ .$$

Verify that $\overleftarrow{\mathbf{X}}$ is again a rough path, i.e. $\overleftarrow{\mathbf{X}} \in \mathscr{C}^\alpha$. Show furthermore that $\overleftarrow{\mathbf{X}}$ is geometric if and only if $\mathbf{X}$ is geometric.

♯ **Exercise 2.7** *Let V be a Banach space.*

a) Let $\alpha \in (0,1]$. Show that the linear space of all continuous maps $\mathbb{X} : [0,T]^2 \to V \otimes V$ s.t. $\|\mathbb{X}\| := \sup |\mathbb{X}_{s,t}|/|t-s|^{2\alpha} < \infty$ is a Banach space, denoted by $\mathcal{C}_2^{2\alpha}$. Deduce that $\mathcal{C}^\alpha \oplus \mathcal{C}_2^{2\alpha}$ is also Banach, with seminorm $\|\bullet,\bullet\|_{\alpha,2\alpha} = \|\bullet\|_\alpha + \|\bullet\|_{2\alpha}$. (A genuine norm is given by $(X,\mathbb{X}) \mapsto |X_0| + \|X,\mathbb{X}\|_{\alpha,2\alpha}$.)

b) Show that the rough path spaces $\mathscr{C}_g^\alpha$ and $\mathscr{C}^\alpha$ are complete metric spaces. In fact, both are closed subspaces, defined through (nonlinear) algebraic relations, of the infinite-dimensional Banach space $\mathcal{C}^\alpha \oplus \mathcal{C}_2^{2\alpha}$.

c) Show that the rough path spaces $\mathscr{C}_g^\alpha$ and $\mathscr{C}^\alpha$ over $V = \mathbf{R}$ (and a fortiori *every $V \neq 0$) are* **not** *separable. (You may use the well-known fact that the Hölder spaces $\mathcal{C}^\alpha([0,T],\mathbf{R})$ are non-separable.)*

Exercise 2.8 (Separable rough path spaces) *Let V be a separable Banach space and $\alpha \in (\frac{1}{3}, \frac{1}{2}]$.*

a) Show separability of the space of geometric (α-Hölder) rough paths

$$\mathscr{C}_g^{0,\alpha}([0,T],V) \stackrel{def}{=} \mathrm{cl}(\mathscr{L}(\mathcal{C}^\infty)) \subset \mathscr{C}^\alpha([0,T],V)\ .$$

Together with Exercise 2.7, b), this shows that $\mathscr{C}_g^{0,\alpha}$ is Polish.

b) Show that the closure of smooth rough paths,

$$\mathscr{C}^{0,\alpha}([0,T],V) \stackrel{def}{=} \mathrm{cl}(\mathscr{C}^\infty) \subset \mathscr{C}^\alpha([0,T],V)\ ,$$

is also separable (and hence Polish).

Solution. (a) Let Q be a countable, dense subset of V and consider the space Λ_n of paths which are piecewise linear between level-n dyadic rationals $\mathbb{D}^n := \{kT/2^n : 0 \le k \le 2^n\}$, and, at level-$n$ dyadic points, take values in Q. Clearly $\Lambda = \cup \Lambda_n$ is countable for each Λ_n is in one-to-one correspondence with the (2^n+1)-fold Cartesian product of Q. It is easy to see that each smooth X is the limit in $\mathcal{C}^1$ of some sequence $(X^n) \subset \Lambda$. Indeed, one can take X^n to be the piecewise linear dyadic approximation, modified such that $X^n|_{\mathbb{D}^n}$ takes values in Q and such that $|(X^n - X)|_{\mathbb{D}^n}| < 1/n$. By continuity of the map $X \in \mathcal{C}^1 \mapsto \big(X, \int X \otimes dX\big) \in \mathscr{C}^\alpha$ in the respective topologies (we could even take $\alpha = 1$), we have more than enough to assert that every lifted smooth path, $\big(X, \int X \otimes dX\big)$, is the limit in $\mathscr{C}^\alpha$ of lifted paths in Λ. It is then easy to see that every limit point of lifted smooth paths is also the limit of lifted paths in Λ.

♯ **Exercise 2.9 (Interpolation)** *Assume that $\mathbf{X}^n \in \mathscr{C}^\beta$, for $1/3 < \alpha < \beta$, with uniform bounds*

$$\sup_n \|X^n\|_\beta < \infty \qquad \text{and} \qquad \sup_n \|\mathbb{X}^n\|_{2\beta} < \infty$$

and uniform convergence $X^n_{s,t} \to X_{s,t}$ and $\mathbb{X}^n_{s,t} \to \mathbb{X}_{s,t}$, i.e. uniformly over $s,t \in [0,T]$. Show that this implies $\mathbf{X} \in \mathscr{C}^\beta$ and $\mathbf{X}^n \to \mathbf{X}$ in $\mathscr{C}^\alpha$. Show furthermore that the assumption of uniform convergence can be weakened to pointwise convergence:

$$\forall t \in [0,T]: \quad X^n_{0,t} \to X_{0,t} \quad \textit{and} \quad \mathbb{X}^n_{0,t} \to \mathbb{X}_{0,t} \ .$$

Solution. Using the uniform bounds and *pointwise* convergence, there exists C such that uniformly in s,t

$$|X_{s,t}| = \lim_n \left|X^n_{s,t}\right| \le C|t-s|^\beta \ , \qquad |\mathbb{X}_{s,t}| = \lim_n \left|\mathbb{X}^n_{s,t}\right| \le C|t-s|^{2\beta} \ .$$

It readily follows that $\mathbf{X} = (X,\mathbb{X}) \in \mathscr{C}^\beta$. In combination with the assumed uniform convergence, there exists $\varepsilon_n \to 0$, such that, uniformly in s,t,

$$\begin{aligned} |X_{s,t} - X^n_{s,t}| &\le \varepsilon_n \ , \qquad |X_{s,t} - X^n_{s,t}| \le 2C|t-s|^\beta \ , \\ |\mathbb{X}^n_{s,t} - \mathbb{X}_{s,t}| &\le \varepsilon_n \ , \qquad |\mathbb{X}^n_{s,t} - \mathbb{X}_{s,t}| \le 2C|t-s|^{2\beta} \ . \end{aligned}$$

By geometric interpolation ($a \wedge b \le a^{1-\theta} b^\theta$ when $a,b > 0$ and $0 < \theta < 1$) with $\theta = \alpha/\beta$ we have

$$|X_{s,t} - X^n_{s,t}| \lesssim \varepsilon_n^{1-\alpha/\beta} |t-s|^\alpha \ , \qquad |\mathbb{X}^n_{s,t} - \mathbb{X}_{s,t}| \lesssim \varepsilon_n^{1-\alpha/\beta} |t-s|^{2\alpha} \ ,$$

and the desired convergence in $\mathscr{C}^\alpha$ follows.

It remains to weaken the assumption to pointwise convergence. By Chen's relation, pointwise convergence of $\mathbf{X}^n_{0,t}$ for all t actually implies pointwise convergence of $\mathbf{X}^n_{s,t}$ for all s,t. We claim that, thanks to the uniform Hölder bounds, this implies uniform convergence. Indeed, given $\varepsilon > 0$, pick a (finite) dissection D of $[0,T]$ with small enough mesh so that $C|D|^\beta < \varepsilon/8$. Given $s,t \in [0,T]$ write $\hat{s}, \hat{t}$ for the nearest points in D and note that

$$\begin{aligned} |X_{s,t} - X^n_{s,t}| &\le |X_{\hat{s},\hat{t}} - X^n_{\hat{s},\hat{t}}| + |X_{s,\hat{s}}| + |X^n_{s,\hat{s}}| + |X_{t,\hat{t}}| + |X^n_{t,\hat{t}}| \\ &\le |X_{\hat{s},\hat{t}} - X^n_{\hat{s},\hat{t}}| + \varepsilon/2 \ . \end{aligned}$$

By picking n large enough, $|X_{\hat{s},\hat{t}} - X^n_{\hat{s},\hat{t}}|$ can also be bounded by $\varepsilon/2$, uniformly over the (finitely many!) points in D, so that $X^n \to X$ uniformly. Although the second level is handled similarly, the non-additivity of $(s,t) \mapsto \mathbb{X}_{s,t}$ requires some extra care, (2.1). For simplicity of notation only, we assume $s < \hat{s} < t = \hat{t}$ so that

$$|\mathbb{X}_{s,t} - \mathbb{X}^n_{s,t}| \le |\mathbb{X}_{s,\hat{s}} - \mathbb{X}^n_{\hat{s},t}| + |\mathbb{X}_{\hat{s},t}| + |X_{s,\hat{s}} \otimes X_{\hat{s},t} - X^n_{s,\hat{s}} \otimes X^n_{\hat{s},t}|.$$

It remains to write the last summand as $|X_{s,\hat{s}} \otimes (X_{\hat{s},t} - X^n_{\hat{s},t}) - (X^n_{s,\hat{s}} - X_{s,\hat{s}}) \otimes X^n_{\hat{s},t}|$ and to repeat the same reasoning as in the first level.

♯ **Exercise 2.10 (Pure area rough path)** *Identify* $\mathbf{R}^2$ *with the complex numbers and consider*

$$[0,1] \ni t \mapsto n^{-1} \exp\left(2\pi i n^2 t\right) \equiv X^n.$$

a) Set $\mathbb{X}^n_{s,t} := \int_s^t X^n_{s,r} \otimes dX^n_r$. *Show that, for fixed* $s < t$,

$$X^n_{s,t} \to 0, \qquad \mathbb{X}^n_{s,t} \to \pi(t-s)\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}. \tag{2.27}$$

b) *Establish the uniform bounds* $\sup_n \|X^n\|_{1/2} < \infty$ *and* $\sup_n \|\mathbb{X}^n\|_1 < \infty$.

c) *Conclude that* $(X^n, \mathbb{X}^n)$ *converges in* $\mathscr{C}^\alpha$, *any* $\alpha < 1/2$.

Solution. a) Obviously, $X^n_{s,t} = O(1/n) \to 0$ uniformly in s, t. Then

$$\mathbb{X}^n_{s,t} = \frac{1}{2} X^n_{s,t} \otimes X^n_{s,t} + A^n_{s,t} = O\big(1/n^2\big) + A^n_{s,t}$$

where $A^n_{s,t} \in \mathfrak{so}(2)$ is the antisymmetric part of $\mathbb{X}^n_{s,t}$. To avoid cumbersome notation, we identify

$$\begin{pmatrix} 0 & a \\ -a & 0 \end{pmatrix} \in \mathfrak{so}(2) \leftrightarrow a \in \mathbf{R}.$$

$A^n_{s,t}$ then represents the signed area between the curve $(X^n_r : s \le r \le t)$ and the straight chord from X^n_t to X^n_s. (This is a simple consequence of Stokes theorem: the exterior derivative of the 1-form $\frac{1}{2}(x\,dy - y\,dx)$ which vanishes along straight chords, is the volume form $dx \wedge dy$.) With $s < t$, $(X^n_r : s \le r \le t)$ makes $\lfloor n^2(t-s) \rfloor$ full spins around the origin, at radius $1/n$. Each full spin contributes area $\pi(1/n)^2$, while the final incomplete spin contributes some area less than $\pi(1/n)^2$. The total signed area, with multiplicity, is thus

$$A^n_{s,t} = \big(n^2(t-s) + O(1)\big)\frac{\pi}{n^2} = \pi(t-s) + \frac{C_{s,t}}{n^2}\,,$$

where $|C_{s,t}| \le \pi$ uniformly in s, t. It follows that

$$\mathbb{X}^n_{s,t} = \pi(t-s)\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} + O\big(1/n^2\big) \tag{2.28}$$

and the claimed uniform convergence follows.

b) The following two estimates for path increments of $n^{-1}\exp\big(2\pi i n^2 t\big) \equiv X^n_t$ hold true:

$$\big|X^n_{s,t}\big| \le \big|\dot{X}^n\big|_\infty |t-s| \le n|t-s|\,, \qquad \big|X^n_{s,t}\big| \le 2|X^n|_\infty = 2/n\,.$$

Since $a \wedge b \le \sqrt{ab}$, it immediately follows that

$$\big|X^n_{s,t}\big| \le \sqrt{2|t-s|}\,,$$

uniformly in n, s, t. In other words, $\sup_n \|X^n\|_{1/2} < \infty$. The argument for the uniform bounds on $\mathbb{X}_{s,t}$ is similar. On the one hand, we have the bound (2.28). On the other hand, we also have

$$\left|\mathbb{X}^n_{s,t}\right| = \left|\int\int_{s<u<v<t} \dot{X}^n_u \otimes \dot{X}^n_v \,du\,dv\right| \le \left|\dot{X}^n\right|^2_\infty \frac{|t-s|^2}{2} \le \frac{n^2}{2}|t-s|^2 \,.$$

The required uniform bound on $\|\mathbb{X}\|_1$ follows by using (2.28) for $n^2|t-s| > 1$ and the above bound for $n^2|t-s| \le 1$.

c) The interpolation argument is left to the reader.

Exercise 2.11 (Second order translation and bracket) *Fix* $\alpha \in (\frac{1}{3}, \frac{1}{2}]$ *and* $\mathbf{X} = (X, \mathbb{X}) \in \mathscr{C}^\alpha([0,T], V)$. *Define the (second order)* translation *of* $\mathbf{X}$ *in direction* $\mathbb{H} \in \mathcal{C}^{2\alpha}([0,T], V \otimes V)$ *by*

$$T_{\mathbb{H}}(\mathbf{X}) \overset{def}{=} \big(X, \mathbb{X} + \delta\mathbb{H}\big)\,,$$

where $(\delta\mathbb{H})$ *denotes the map* $(s,t) \mapsto \mathbb{H}_t - \mathbb{H}_s$.

a) *Show that* $T_{\mathbb{H}}(\mathbf{X}) \in \mathscr{C}^\alpha$. *In fact, show that the (linear) space* $\mathcal{C}^{2\alpha}$ *acts freely on the (nonlinear) rough path space* $\mathscr{C}^\alpha$ *in the sense that, for all* $\mathbb{G}, \mathbb{H} \in \mathcal{C}^{2\alpha}$, *we have*
$$T_{\mathbb{G}}\big(T_{\mathbb{H}}(\mathbf{X})\big) = (T_{\mathbb{G}} \circ T_{\mathbb{H}})(\mathbf{X}) = T_{\mathbb{G}+\mathbb{H}}(\mathbf{X})\,.$$
Fix $\mathbf{X} \in \mathscr{C}^\alpha$. *Is* $\mathbb{H} \mapsto T_{\mathbb{H}}(\mathbf{X})$ *is injective?*

b) *When does* $T_{\mathbb{H}}$ *preserve the space* $\mathscr{C}^\alpha_g([0,T], V)$*?*

c) *Show that any* $\mathbf{X} = (X, \mathbb{X}) \in \mathscr{C}^\alpha([0,T], V)$ *can be written, in a unique way, as* $T_{\mathbb{H}}(\mathbf{X}_g)$, *where* $\mathbf{X}_g \in \mathscr{C}^\alpha_g([0,T], V)$ *for some* $\mathbb{H} \in \mathcal{C}^{2\alpha}([0,T], \mathrm{Sym}(V \otimes V))$, *so that we have the bijection*

$$\mathscr{C}^\alpha([0,T], V) \leftrightarrow \mathscr{C}^\alpha_g([0,T], V) \times \mathcal{C}^{2\alpha}([0,T], \mathrm{Sym}(V \otimes V)).$$

Show that $2\delta\mathbb{H} = (\delta X)^{\otimes 2} - 2\,\mathrm{Sym}(\mathbb{X}) =: [\mathbf{X}]$, *called* bracket *of the rough path* $\mathbf{X}$, *further studied in Section 5.3.*

Exercise 2.12 (Vanishing Hölder oscillation) *a) Let* $X \in \mathcal{C}^\alpha([0,T], V)$ *with Hölder exponent* $\alpha \in (0,1]$. *Define the space of Hölder path with "vanishing Hölder oscillation",*

$$\mathcal{C}^{\mathrm{van},\alpha} \overset{def}{=} \left\{ X \in \mathcal{C}^\alpha : \sup_{s,t:|t-s|<\varepsilon} \frac{|X_{s,t}|}{|t-s|^\alpha} \to 0, \text{ as } \varepsilon \to 0 \right\}.$$

Show that for $\alpha \in (0,1)$ *we have* $\mathcal{C}^{\mathrm{van},\alpha} = \mathcal{C}^{0,\alpha}$, *the closure of smooth paths in* $\mathcal{C}^\alpha$. *(For* $\alpha = 1$ *this fails, why?) Show by explicit example that the inclusion* $\mathcal{C}^{0,\alpha} \subset \mathcal{C}^\alpha$ *is strict. (Hint: consider the function* $t \mapsto t^\alpha$*.)*

b) *Let* $\mathbf{X} = (X, \mathbb{X}) \in \mathscr{C}^\alpha_g([0,T], V)$ *with* $\alpha \in (\frac{1}{3}, \frac{1}{2}]$. *Define the space of Hölder rough paths with "vanishing Hölder oscillation",*

$$\mathscr{C}^{\mathrm{van},\alpha}_g \overset{def}{=} \left\{ \mathbf{X} \in \mathscr{C}^\alpha_g : \sup_{|t-s|<\varepsilon} \frac{|X_{s,t}|}{|t-s|^\alpha} + \sup_{|t-s|<\varepsilon} \frac{|\mathbb{X}_{s,t}|}{|t-s|^{2\alpha}} \to 0 \text{ as } \varepsilon \to 0 \right\}.$$

i) Show the inclusions $\mathscr{C}_g^{0,\alpha} \subset \mathscr{C}_g^{\text{van},\alpha}$ *and also* $\mathscr{C}_g^{\beta} \subset \mathscr{C}_g^{\text{van},\alpha}$*, whenever* $\alpha < \beta$*. Show that the inclusion* $\mathscr{C}_g^{\text{van},\alpha} \subset \mathscr{C}_g^{\alpha}$ *is strict.*

ii) Assume $\dim V < \infty$ *from here on. Show* $\mathscr{C}_g^{0,\alpha} = \mathscr{C}_g^{\text{van},\alpha}$ *(Hint: use the "geodesic" approximations from Proposition 2.8.)*

iii) From ii) we have $\mathscr{C}_g^{\beta} \subset \mathscr{C}_g^{0,\alpha} \subset \mathscr{C}_g^{\alpha}$*, whenever* $\frac{1}{3} < \alpha < \beta \leq \frac{1}{2}$*. Show that one has the compact embedding (Hint: Arzela–Ascoli)*

$$\mathscr{C}_g^{\beta} \hookrightarrow \mathscr{C}_g^{0,\alpha}.$$

c) Discuss similar statements for non-geometric rough path spaces. In particular, discuss the validity of

$$\mathscr{C}^{0,\alpha} \stackrel{\textit{def}}{=} \text{cl}(\mathscr{C}^{\infty}) = \mathscr{C}^{\text{van},\alpha}\,,$$

and also, cf. Exercise 2.11, c),

$$\mathscr{C}^{0,\alpha} \leftrightarrow \mathscr{C}_g^{0,\alpha} \times \mathcal{C}^{0,2\alpha}\,;$$

for $\alpha = 1/2$ *this fails, why?*

Remark: *This is essentially taken from [FV06a], for a recent extension to*

$*$ **Exercise 2.13** *Show that for every geometric* $1/2$*-Hölder rough path,* $\mathbf{X} \in \mathscr{C}_g^{0,1/2}$*,* $\mathbb{X}$ *is necessarily the iterated Riemann–Stieltjes integral of the underlying path* $X \in \mathcal{C}^{0,1/2}$*. Show also that there exists* $X \in \mathcal{C}^{0,1/2}$ *(with values in* $\mathbf{R}^2$*) such that the iterated Riemann–Stieltjes integrals do not exist. This further shows that the Lyons–Victoir extension (Exercise 2.14, part d) can fail for* α*-Hölder rough paths when* $1/\alpha \in \mathbf{N}$*.*

Solution. We use $\mathscr{C}_g^{0,\alpha} \subset \mathscr{C}_g^{\text{van},\alpha}$ Exercise 2.12, for $\alpha = 1/2$. Consider a dissection $\{s = \tau_0 < \tau_1 < \ldots < \tau_N = t\}$ with mesh $\leq \varepsilon$. It follows from Chen's relation (2.1), in the form (2.26),

$$\begin{aligned}\Big|\mathbb{X}_{s,t} - \sum_{0\leq i<N} X_{s,\tau_i} \otimes X_{\tau_i,\tau_{i+1}}\Big| &= \Big|\sum_{0\leq i<N} \mathbb{X}_{\tau_i,\tau_{i+1}}\Big| \\ &\leq C(\varepsilon) \sum_{0\leq i<N} |\tau_{i+1} - \tau_i|^{2\alpha} = TC(\varepsilon).\end{aligned}$$

It follows that $\mathbb{X}_{s,t}$ is the limit of the above Riemann–Stieltjes sum.

Regarding the second question, a counterexample is found in [FV10b, Ex.9.14 (iii)].

$\sharp *$ **Exercise 2.14 (Lyons–Victoir extension [LV07])** *Let* $\alpha \in (0, 1/2)$ *and consider* $X \in \mathcal{C}^{\alpha}([0,T], L(V,W)), Y \in \mathcal{C}^{\alpha}([0,T], V)$ *and* $\mathbb{Z} \in \mathcal{C}_2^{2\alpha}([0,T], W)$*. We omit* $[0,T]$ *and the precise target space in what follows. We here say that Chen's relation holds if, for every triple of times* (s,t,u)*,*

$$\mathbb{Z}_{s,u} = \mathbb{Z}_{s,t} + \mathbb{Z}_{t,u} + Y_{s,t}X_{t,u}.$$

(This is the algebraic relation satisfied by $(s,t) \mapsto \int_s^t Y_{s,r} dX_r$ *whenever* $X \in \mathcal{C}^1$*.)*

a) Show that here exists a bilinear continuous map $\Phi : \mathcal{C}^\alpha \times \mathcal{C}^\alpha \to \mathcal{C}_2^{2\alpha}$,

$$(Y, X) \mapsto \mathbb{Z} := \Phi(Y, X)$$

such that Chen's relation holds.

b) Show that the restriction of Φ *to Hölder paths with exponent* $\beta \in (1/2, 1)$ *cannot possibly be a continuous as map* $\mathcal{C}^\beta \times \mathcal{C}^\beta \to \mathcal{C}_2^{2\beta}$*. (Hint: the Chen relation would force* $\Phi(Y, X)$ *to coincide with the Young integral* $\int Y dX$*. In particular,* $\Phi_{0,\cdot}$ *would have to coincide with* $\int_0^\cdot Y(t)\dot{X}(t)dt$ *in case of smooth path. Proposition 1.1 then allows to conclude.)*

c) Show however that Φ *can be constructed such that its restriction to a map* $\mathcal{C}^\beta \times \mathcal{C}^\beta \to \mathcal{C}^\beta$*, where the image is now regarded as path* $t \mapsto \Phi(Y, X)_{0,t}$*, is a bilinear continuous map.*

d) Let $\alpha \in (1/3, 1/2)$*. Show that every path* $X \in \mathcal{C}^\alpha([0, T], V)$ *admits a (if so desired: geometric) rough path lift* $(X, \mathbb{X}) \in \mathscr{C}^\alpha([0, T], V)$*.*

e) Conclude that the nonlinear rough path space $\mathscr{C}^\alpha([0, T], V)$ *is in (non-canonical) one-one correspondence with the linear space* $\mathcal{C}^\alpha([0, T], V) \oplus \mathcal{C}^{2\alpha}([0, T], V \otimes V)$*. (For a generalisation of this to rough paths of low regularity see [TZ18].)*

Solution. We show a) and c) together; d) is really a variation / consequence of a) and we leave b) and e) to the reader. Without loss of generality, $T = 1$. Write $\mathbb{Z}_{(s,t]} \equiv \mathbb{Z}_{s,t}$ and similarly for the path increments of Y, X. We want to construct $\mathbb{Z}$ such that

$$\mathbb{Z}_I = \mathbb{Z}_L + \mathbb{Z}_R + Y_L \otimes X_R$$

whenever $I = (s, t]$ is the union of two adjacent "left and right" intervals L and R, and such that

$$|\mathbb{Z}_I| \lesssim |I|^{2\alpha} \qquad (\star)$$

where $|I| = |t - s|$. By a continuity and chaining argument (see the proof of Theorem 3.1 below), it is enough to do so for dyadic times, i.e. $s, t \in \bigcup_{n \geqslant 0} \mathbf{D}_n$ where $\mathbf{D}_0 = \{(0, 1]\}$, $\mathbf{D}_1 = \{(0, 1/2], (1/2, 1]\}$ and so on. We start with the (ad-hoc!) choice $\mathbb{Z}_{0,1} \equiv \mathbb{Z}_{(0,1]} = 0$ and note its (trivial) bilinearity in (Y, X). Assume now $\mathbb{Z}_I$ for $I \in \mathbf{D}_{n-1}$ has been constructed. Write I as the union of two nth level dyadic intervals, $I = L \cup R$. Make the (ad-hoc) imposition $\mathbb{Z}_L = \mathbb{Z}_R$ which leads to

$$\mathbb{Z}_L = \mathbb{Z}_R = \frac{1}{2}(\mathbb{Z}_I - Y_L \otimes X_R).$$

(Note that bilinear dependence in Y, X is preserved.) On the analytic side, we have

$$|\mathbb{Z}_L| = |\mathbb{Z}_R| = \frac{1}{2}|\mathbb{Z}_I - Y_L \otimes X_R| \leqslant \frac{1}{2}|\mathbb{Z}_I| + \frac{1}{2}|Y_L| \cdot |X_R|$$

and, setting $a_n := \sup_{J \in \mathbf{D}_n} |\mathbb{Z}_J| / |J|^{2\alpha} = 2^{2n\alpha} \sup_{J \in \mathbf{D}_n} |\mathbb{Z}_J|$, it follows that

$$a_n \leqslant 2^{-(1-2\alpha)} a_{n-1} + \frac{1}{2}\|Y\|_\alpha \|X\|_\alpha,$$

so that the sequence (a_n) is bounded since $1 - 2\alpha > 0$. In fact, one easily obtains the bound

$$\sup_{n\geqslant 0} |a_n| \lesssim \|Y\|_\alpha \|X\|_\alpha,$$

with proportionality constant only depending on $\alpha < 1/2$. This implies the estimate ($\star$) and also settles continuity of $\Phi = \Phi(Y, X)$. It remains to show that $t \mapsto \mathbb{Z}_{0,t} \in \mathcal{C}^\beta$ whenever $Y, X \in \mathcal{C}^\beta$ and $\beta \in (1/2, 1)$. But this is an immediate consequence of the bound

$$|\mathbb{Z}_{0,t} - \mathbb{Z}_{0,s}| \leqslant |\mathbb{Z}_{s,t}| + |X_{0,s}| \cdot |X_{s,t}|,$$

noting that, thanks to the first part of the theorem, $|\mathbb{Z}_{s,t}| \lesssim |t-s|^{2\alpha}$ for all $2\alpha < 1$.

Exercise 2.15 (Translation of rough paths) *Fix $\alpha \in (\frac{1}{3}, \frac{1}{2}]$ and $\mathbf{X} = (X, \mathbb{X}) \in \mathscr{C}^\alpha([0,T], \mathbf{R}^d)$. For sufficiently smooth $h : [0,T] \to \mathbf{R}^d$, the translation of $\mathbf{X}$ in direction h is given by*

$$T_h(\mathbf{X}) \stackrel{\text{def}}{=} \left(X^h, \mathbb{X}^h\right),$$

where $X^h := X + h$ and

$$\mathbb{X}^h_{s,t} := \mathbb{X}_{s,t} + \int_s^t h_{s,r} \otimes dX_r + \int_s^t X_{s,r} \otimes dh_r + \int_s^t h_{s,r} \otimes dh_r \,. \tag{2.29}$$

a) Assume $h \in \mathcal{C}^1$. (In particular, the last three integrals above are well-defined Riemann–Stieltjes integrals.) Show that for fixed h, the translation operator $T_h : \mathbf{X} \mapsto T_h(\mathbf{X})$ is a continuous map from $\mathscr{C}^\alpha$ into itself.

b) By convention, $h \in \mathcal{C}^1$ means Lipschitz or equivalently $h \in W^{1,\infty}$, where $W^{1,q}$ denotes the space of absolutely continuous paths h with derivative $\dot{h} \in L^q$. Weaken the assumption on h by only requiring $\dot{h} \in L^q$, for suitable $q = q(\alpha)$. Show that $q = 2$ (“Cameron–Martin paths of Brownian motion”) works for all $\alpha \leq 1/2$. (As a matter of fact, the integrals appearing in (2.29) make sense for every $q \geq 1$, but the resulting translated “rough path” falls out of the class of Hölder rough paths. One can resolve this issue by switching to $(1/\alpha)$-variation rough paths.)

c) Call any $\mathbf{h} = (h, \mathbb{H}) : [0,T] \to \mathbf{R}^d \oplus (\mathbf{R}^d)^{\otimes 2} = T_0^{(2)}$, with $h \in W^{1,2}$ and $\mathbb{H} \in \mathcal{C}^{2\alpha}$ an admissible perturbation. With some notational overloading, T is also used for the second order translation introduced in Exercise 2.11, show that

$$T_{\mathbf{h}} := T_h \circ T_{\mathbb{H}} = T_{\mathbb{H}} \circ T_h$$

is a well-defined action on $\mathscr{C}^\alpha$, in the sense of $T_{\mathbf{g}} \circ T_{\mathbf{h}} = T_{\mathbf{g}+\mathbf{h}}$. Show that for any fixed $(a, b) \in T_0^{(2)}$, the constant speed perturbation $t \mapsto (at, bt)$ is admissible, which then yields an action of $T_0^{(2)}$ with its additive structure on $\mathscr{C}^\alpha$. Show that these statements remain true for $\mathscr{C}_g^\alpha$ provided admissible perturbations take values in the Lie algebra $\mathfrak{g}^{(2)} = \mathbf{R}^d \oplus \mathfrak{so}(d)$ as introduced in Section 2.3.

Remark: *Some far-reaching extensions of this are found in [BCFP19]. Constant speed perturbations respect stationarity of the noise (stationary increments of the process) and thus serve as elementary examples of (algebraic) renormalisation of models in regularity structures. The (abelian) groups* $(\mathfrak{g}^{(2)}, +)$ *and* $(T_0^{(2)}, +)$ *together with their action* $\mathbf{h} \mapsto T_{\mathbf{h}}$*, are examples of a renormalisation group in the sense of Section 15.5.1.*

2.6 Comments

Many early works in stochastic analysis starting from Itô (and then in no particular order Kunita, Yamato, Sugita, Azencott, Ben Arous [BA89], etc) and in control theory (Magnus, Brocket, Sussmann, Fliess [FNC82], etc) have recognised the importance of iterated integrals of the driving noise / signal; many references are given [Lyo98] and the books [LQ02, LCL07, FV10b].

The notion of rough path is due to Lyons and was introduced in [Lyo98] in p-variation sense, $p \in [1, \infty)$, and over Banach spaces. Earlier notes [Lyo94, Lyo95] already dealt with α-Hölder rough paths for $\alpha \in \left(\frac{1}{3}, \frac{1}{2}\right]$.

The analytical aspects of rough paths are related to Young's seminal work [You36], revisited in Chapter 4. On the algebraic side, Chen's relation is rooted in [Che54, Che57] and encodes abstractly basic additivity properties of iterated integrals. A key observation of Chen [Che57, Che58] was that log signatures are Lie series, the description via shuffles (cf. Section 2.4) is due to Ree [Ree58] (see also [Che71]). It follows from the works of Chow and Rashevskii [Cho39, Ras38], also [Che57, Che58], that this map is, upon truncation, onto: for every element in $\mathbf{x} \in G^{(N)}(\mathbf{R}^d) := \exp(\mathfrak{g}^{(N)}(\mathbf{R}^d))$ there exists a smooth path $\gamma : [0, 1] \to \mathbf{R}^d$ with prescribed signature $\mathbf{x} = S^{(N)}(\gamma)$. The shortest such path can be viewed as sub-Riemannian geodesic, concatenation of such geodesics is then a natural way to approximate weakly geometric rough paths (cf. Proposition 2.8) and underlies the geometric approach of Friz–Victoir [FV05, FV10b], surveyed from a sub Riemannian perspective in [FG16a]. The polynomial nature of (truncated) shuffle relations and log Lie conditions recently led Améndola, Friz and Sturmfels [AFS19] to the study of *signature varieties* in computational algebraic geometry.

Up to equivalence under a generalised notion of reparameterisation of paths known as *treelike equivalence*, the "full" signature map $\gamma \mapsto S(\gamma) \in G((V)) \subset T((V))$ was shown to be injective by Chen [Che58] in case of piecewise smooths paths, Hambly–Lyons [HL10] in case of rectifiable paths, and Boedihardjo et al. [BGLY16] in case of weakly geometric rough paths of arbitrarily low regularity, see also Boedihardjo, Ni and Qian [BNQ14]. The inversion problem "signature $\mapsto$ path" is studied by Lyons–Xu [LX17, LX18] and [AFS19]. All this is part of the mathematical justification of the *signature method* in machine learning, see e.g. Lyons' ICM article [Lyo14] and the survey [CK16].

For some constructions of level-2 geometric rough paths motivated from harmonic analysis see Hara–Lyons [HL07] and Lyons–Yang [LY13], see also the comments

Section 3.8 for some martingale constructions related to harmonic analysis. Lyons–Qian, in their monograph [LQ02] work with geometric rough paths (over a Banach space V), per definition limits of canonically lifted smooth paths. The strict inclusion "geometric $\subset$ weakly geometeric" was somewhat blurred in the earlier rough paths literature. For $\dim V < \infty$, matters were clarified in [FV06a]. For a discussion of weakly geometric rough paths over Banach spaces in their own right, see e.g. in [CDLL16], see also the supplementary appendix [BGLY15] of [BGLY16]. The discussion in Section 2.4, the "shuffle" view on weakly geometric rough paths and then Gubinelli's branched rough paths [Gub10], also extends from $V = \mathbf{R}^d$ to infinite dimension but setting up basis-independent notations is somewhat more involved. See for example [CW16, CCHS20] for some recent results in this direction.

"Naïve" higher order non-geometric rough paths with values in $T_1^{(N)}(V)$ are called in [Lyo98] *multiplicative functionals* (with α-Hölder or p-variation regularity, $\lfloor p \rfloor = N$), insisting on their inability to handle nonlinearities when $N \geq 3$. The notion of branched rough path, for any $\alpha \in (0,1]$, further studied in [HK15, FZ18, BCFP19, BC19, TZ18] provides the required extra information when $N \geq 3$; for $N = \lfloor 1/\alpha \rfloor = 2$ there is no difference. It is possible to embed spaces of non-geometric rough paths of low regularity into suitable spaces of geometric rough paths, see [LV06] or Exercise 2.11 part c) when $N = 2$. The case of very low regularities, with N large, is much more involved and studied by Hairer–Kelly [HK15] and later Boedihardjo–Chevyrev [BC19].

Rough paths with jumps, in p-variation scale, are studied in [Wil01, FS17, FZ18, CF19], previously introduced *discrete rough paths* [Kel16] are also accomodated e.g. by the *càdlàg rough path* setting of [FZ18]. See also the comment Sections 4.8, 5.6 and 9.6. Rough paths in a geometric ambient space have been studied by Cass, Driver, Litterer and Lyons in [CLL12, CDL15], see also Bailleul [Bai19] for rough paths on Banach manifolds.