

8

The Geometry of Rough Paths

8.1 Introduction

In this chapter we will discuss the (infinite-dimensional) geometric framework for rough paths and their signature. Rough path theory originated in the 1990s with the work of T. Lyons; see, for example, Lyons (1998). It seeks to establish a theory of integrals and differential equations driven by rough signals. For example, one is interested in the controlled ordinary differential equations of the following type:

$$y'_t = f(y_t) + g(y_t)X'_t. \tag{8.1}$$

Here the subscripts track the time parameter t , X is an input path with values in \mathbb{R}^d and y is the output with values in \mathbb{R}^e . We will for this exposition assume that all derivatives and integrals needed exist (e.g. if X is a smooth path). Finally, f, g are non-linear functions with values in \mathbb{R}^e and $L(\mathbb{R}^d, \mathbb{R}^e)$, i.e. $e \times d$ -matrices, respectively. Focussing on the control term in (8.1), let us consider a simple approximation to the solution in the case that $f = 0$, that is, $y'_t = g(y_t)X'_t$. For example, the first-order Euler method gives us the following approximation for the components of y :

$$y_t^i - y_s^i \approx g^i(y_s) \int_s^t dX^i = \lim_{|P| \rightarrow 0} g^i(y_s) \sum_{[t_j, t_{j+1}] \in P} (X_{t_{j+1}}^i - X_{t_j}^i), \quad i = 1, \dots, e.$$

Here superscripts denote components of maps, the integral is defined via a Riemann–Stieltjes sum and we think of $g(y_s)$ as a matrix (selecting columns and rows appropriately). The information needed for the approximation is the integral of X . In general, we would like better approximations (or even a solution), so it is natural to increase the order of the approximations. To obtain the desired formula one applies a Taylor expansion to obtain the following second-order Euler approximation for $i = 1, \dots, e$:

$$y_t^i - y_s^i \approx g^i(y_s) \int_s^t dX^i + \sum_{k=1}^e \sum_{j=1}^d g_k^i(y_s) \frac{\partial}{\partial_k} g^j(y_s) \int_s^t \int_s^r dX^i dX^j. \tag{8.2}$$

Thus higher-order approximation requires knowledge about (mixed) iterated integrals of the path X against itself. So far we have tacitly assumed that the necessary paths and derivatives exist (e.g. that all the objects in question are smooth).

Weakening this requirement, the objects we are interested in are iterated integrals of Hölder continuous paths with values in \mathbb{R}^d .¹ For convenience we will assume throughout this chapter that we are dealing with paths on the interval $[0, 1]$. This is no restriction since we can always reparametrise Hölder paths on any $[0, 1]$ to obtain corresponding paths on $[0, 1]$ (though this changes the Hölder norm). Assume we have two continuous mappings $X: [0, 1] \rightarrow \mathbb{R}^d$ and $Y: [0, 1] \rightarrow L(\mathbb{R}^d, \mathbb{R}^e)$, where the continuous linear maps $L(\mathbb{R}^d, \mathbb{R}^e)$ have been endowed with the operator norm. We would like to define an integral now of Y against the path X and these integrals should yield a continuous map $(X, Y) \mapsto \int YdX$. Setting $X_{s,t} := X_t - X_s$, we could try to define the integral as a limit of Riemann–Stieltjes sums,

$$\int_0^1 Y(t)dX(t) := \lim_{|\mathcal{P}| \rightarrow 0} \sum_{[s,t] \in \mathcal{P}} Y(s)X_{s,t},$$

where \mathcal{P} is a partition of $[0, 1]$ and the limit takes the mesh size to 0. The resulting integral is called the *Young integral* and Young (1936) showed that the Riemann–Stieltjes sum converges if X is an α -Hölder path² and Y is a β -Hölder path such that $\alpha + \beta > 1$. This result is sharp as one can construct examples of paths with $\alpha + \beta = 1$ such that the sum becomes ill defined. So in general, there is no hope for an integration theory which allows us to integrate arbitrarily rough paths against each other. However, the key insight of rough path theory is that the regularity assumption $\alpha + \beta > 1$ from Young’s theorem can be circumvented if additional structure is added to the paths. Thus a rough integral can be built if we enhance Hölder continuous paths with additional information to so-called rough paths. The point is that this information can be chosen for irregular paths such as Brownian motion (which is known to be α -Hölder for $\alpha \in]0, 1/2[$). Here rough path theory excels at clever estimates for the integrals appearing. In the present chapter we will focus on the geometric side of the picture and leave the hard analytic estimates to the rough path literature.

¹ The basic theory extends to Banach space valued paths but this requires more technical efforts such as, for example, the introduction of tensor norms.

² The notion of α -Hölder paths is recalled in §8.3.

Our first aim is to develop a convenient geometric framework to record iterated integrals.

Notation for increments and iterated integrals We frequently encounter increments of continuous paths $X : [0, 1] \rightarrow \mathbb{R}^d$:

- $X_t := X(t)$, and
- $X_{s,t} := X_t - X_s$ for the increment and $t, s \in [0, 1]$.

With enough differentiability of X , we can define iterated integrals of X against itself. For $0 \leq s \leq t \leq 1$, we integrate componentwise in $\mathbb{R}^d \otimes \mathbb{R}^d \cong \mathbb{R}^{d^2}$ and set

$$\int_s^t \int_s^r dX \otimes dX := \int_s^t X_{s,r} \otimes dX := \int_s^t X_{s,r} \otimes dX_r := \int_s^t X_{s,r} \otimes X'_r dr.$$

The notation suppresses indices with the understanding that the objects are matrices whose components are iterated integrals. For higher iterated integrals this will quickly become impractical whence we shall use tensor notation instead.

8.2 Iterated Integrals and the Tensor Algebra

In this section we consider the tensor algebra as a continuous inverse algebra. As seen in the introduction to this chapter, we are interested in iterated integrals of the components of a path $X : [0, 1] \rightarrow \mathbb{R}^d$. For example, the second iterated integrals (8.2) yield a matrix object whose components can also be conveniently recorded using tensor notation:

$$\int_s^t \int_s^r dX^i \otimes dX^j := \int_s^t \left(\int_s^r dX^i \right) dX^j e_i \otimes e_j,$$

where e_i, e_j are standard basis vectors of \mathbb{R}^d . Iterating for higher orders, we can write the resulting integrals as elements in an iterated tensor product. Note that on the basic levels tensors are just a bookkeeping device to track the components integrated against each other. Indeed the canonical identification $\mathbb{R}^d \otimes \mathbb{R}^d \cong \mathbb{R}^{d \times d}$ maps $e_i \otimes e_j$ to the component in the i th row and j th column. Hence, detailed knowledge on tensor products is not needed and we refer to (Abraham et al., 1988, Chapter 5) for an introduction.

8.1 Definition ((Truncated) tensor algebra) For $d \in \mathbb{N}$ and $k \in \mathbb{N}$, we set

$$(\mathbb{R}^d)^{\otimes k} := \underbrace{\mathbb{R}^d \otimes \mathbb{R}^d \otimes \dots \otimes \mathbb{R}^d}_{k \text{ times}} \text{ and } (\mathbb{R}^d)^{\otimes 0} := \mathbb{R}.$$

Write $e_i, 1 \leq i \leq d$ for the standard basis vectors of \mathbb{R}^d and recall that the products $e_{i_1} \otimes \dots \otimes e_{i_k}$ form a base for $(\mathbb{R}^d)^{\otimes k}$ whence this space is isomorphic to \mathbb{R}^{d^k} . Then we define for $N \in \mathbb{N} \cup \{\infty\}$ the (truncated) tensor algebra³

$$\mathcal{T}^N(\mathbb{R}^d) := \prod_{k=0}^N (\mathbb{R}^d)^{\otimes k}.$$

Elements in $\mathcal{T}^N(\mathbb{R}^d)$ will be denoted as sequences $(x_k)_{k < N+1}$ (where $\infty + 1 = \infty$). An element concentrated in the k th factor $(\mathbb{R}^d)^{\otimes k}$ is called *homogeneous of degree k* . Then the algebra product is given by

$$(x_k)_{k < N+1} \otimes (y_k)_{k < N+1} := \left(\sum_{n+m=k} x_n \otimes y_m \right)_{k < N+1}. \tag{8.3}$$

Since tensor products over \mathbb{R} with elements of \mathbb{R} contract, that is, $\lambda \otimes v = \lambda v$, we see that the tensor algebra has a unit element $\mathbf{1} = (1, 0, 0, \dots) \in \mathcal{T}^N(\mathbb{R}^d)$ (i.e. $\mathbf{1}$ is homogeneous of degree 0) and the map $\pi_0^N : \mathcal{T}^N(\mathbb{R}^d) \rightarrow (\mathbb{R}^d)^{\otimes 0} = \mathbb{R}$ is an algebra morphism (and, in particular, the homogeneous elements of degree 0 form a subalgebra of the (truncated) tensor algebra). Note, however, that the algebra is almost always non-commutative (Exercise 8.2.2).

In the introduction we saw that iterated integrals can be conveniently identified with elements in the tensor algebra. From this point of view it would be enough to treat the tensor algebra as locally convex space which simply stores information. However, it turns out that iterated integrals satisfy several natural identities which can be expressed using the product in the tensor algebra. For example, if $X : [0, 1] \rightarrow \mathbb{R}^d$ is a smooth path and $\mathbf{X}_{s,t} := \int_s^t X_{s,r} \otimes dX \in \mathbb{R}^d \otimes \mathbb{R}^d$ its iterated integral, it is easy to see that these satisfy *Chen’s relation*

$$\mathbf{X}_{s,t} - \mathbf{X}_{s,u} - \mathbf{X}_{u,t} = X_{s,u} \otimes X_{u,t}, \quad \text{for all } u \in [s, t]. \tag{8.4}$$

To get a feeling for these identities and for the truncated tensor algebra, we recommend Exercise 8.2.2. Summing up, we should be interested in the algebra structure of the tensor algebra as well. Our next result discusses the topological structure of these algebras.

8.2 Lemma Fix $N \in \mathbb{N} \cup \{\infty\}$ and $d \in \mathbb{N}$.

- (a) An element a in $\mathcal{T}^N(\mathbb{R}^d)$ is invertible in the tensor algebra if and only if the associated element of degree 0, $a_0 := \pi_0^N(a)$ is invertible.

³ The tensor algebra discussed in this section differs from what is usually called the tensor algebra. In the literature, the tensor algebra usually denotes the direct sum of iterated tensor products (i.e. finite sequences of iterated tensors). From this perspective, the tensor algebra $\mathcal{T}^\infty(\mathbb{R}^d)$ should rather be called the completed tensor algebra.

- (b) The algebra $\mathcal{T}^N(\mathbb{R}^d)$ is a continuous inverse algebra (CIA). Moreover, if $N < \infty$, then $\mathcal{T}^N(\mathbb{R}^d)$ is a Banach algebra and if $N = \infty$, $\mathcal{T}^N(\mathbb{R}^d)$ is a Fréchet algebra.

Proof As a first step, let us topologise $\mathcal{T}^N(\mathbb{R}^d)$.

If $N < \infty$, we obtain a finite-dimensional algebra and thus it is a Banach algebra and a CIA (alternatively see Exercise 8.2.1).

If $N = \infty$, we observe first that $\mathcal{T}^\infty(\mathbb{R}^d)$ is a countable product of finite dimensional locally convex spaces, whence it is a Fréchet space. Note that this topology and the Banach topology for $N < \infty$ turn the projection π_0^N into a continuous map. Exploiting the product topology, we see that the product $\mathcal{T}^\infty(\mathbb{R}^d) \times \mathcal{T}^\infty(\mathbb{R}^d) \rightarrow \mathcal{T}^\infty(\mathbb{R}^d)$ is continuous if and only if the component maps $P_k : \mathcal{T}^\infty(\mathbb{R}^d) \times \mathcal{T}^\infty(\mathbb{R}^d) \rightarrow (\mathbb{R}^d)^{\otimes k}, k \in \mathbb{N}$ are continuous. However, as the algebra product (8.3) respects the homogeneous degree of elements, it is clear that P_k factors through the continuous inclusion $\mathcal{T}^k(\mathbb{R}^d)$ and the algebra product of $\mathcal{T}^k(\mathbb{R}^d)$. Both the inclusion and the algebra product are continuous, and so P_k and consequently the product on $\mathcal{T}^\infty(\mathbb{R}^d)$ are continuous. We deduce that $\mathcal{T}^\infty(\mathbb{R}^d)$ is a locally convex algebra. To see that it is also a CIA, let us prove first the claim on invertibility of elements.

Let $a \in \mathcal{T}^N(\mathbb{R}^d)$ and write $a = a_0 + b$, where $a_0 = \pi_0^N(a)$ and $b = a - a_0$. Since π_0^N is an algebra homomorphism, a_0 must be invertible if a is invertible. For the converse, we observe that a is invertible if and only if $a_0^{-1}a = \mathbf{1} + a_0^{-1}b$ is invertible. Plugging $a_0^{-1}a$ into the Neumann inversion formula (see Werner, 2000, Theorem II.1.11)

$$(\mathbf{1} + X)^{-1} = \sum_{k=0}^{\infty} (-1)^k X^{\otimes k}, \tag{8.5}$$

we obtain the desired inverse $a_0^{-1}(\mathbf{1} + b)^{-1}$ if the series converges. For this we observe that $\mathbf{1} - a_0^{-1}a$ has no part which is homogeneous of degree 0. Taking products of this element with itself, we only obtain contributions by homogeneous parts of higher degrees. Hence, if $N < \infty$ and we truncate, the series is just a polynomial. For $N = \infty$ we see similarly that after projecting to the factors $(\mathbb{R}^d)^{\otimes k}$ we again obtain only a polynomial. Hence also in this case the series converges as it (trivially) converges in every factor. Thus $a_0^{-1}a$ is invertible and the inverse depends continuously on a . Moreover, the set of units $\mathcal{T}^\infty(\mathbb{R}^d)^\times = (\pi_0^\infty)^{-1}(\mathbb{R} \setminus \{0\})$ is open. Summing up, this proves that $\mathcal{T}^\infty(\mathbb{R}^d)$ is a CIA. □

8.3 Corollary For every $N \in \mathbb{N} \cup \{\infty\}$ and $d \in \mathbb{N}$, the group $(\mathcal{T}^N(\mathbb{R}^d))^\times = \{v \in \mathcal{T}^N(\mathbb{R}^d) \mid \pi_0^N(v) \neq 0\}$ is a regular locally exponential Lie group whose Lie algebra is $\mathcal{T}^N(\mathbb{R}^d)$ with the commutator bracket. Moreover, its Lie group exponential is given by

$$\exp_{\otimes} : \mathcal{T}^N(\mathbb{R}^d) \rightarrow (\mathcal{T}^N(\mathbb{R}^d))^{\times}, \quad \exp_{\otimes}(v) = \sum_{k=0}^{\infty} \frac{v^{\otimes k}}{k!},$$

where $v^{\otimes k}$ is the k th power of v with respect to the algebra product.

Proof The Lie group structure was established in Example 3.4 and the Lie algebra computed in Exercise 3.2.8. Note that thanks to Lemma 8.2 the (truncated) tensor algebra is complete, whence its unit group is regular by Example 3.35. There we also mentioned that the evolution map Evol is given by the Volterra series (3.6). Plugging a constant path $t \mapsto v$ into the Volterra series we immediately obtain that \exp_{\otimes} is the Lie group exponential (check this!). If $N < \infty$ the CIA is finite dimensional, and so $(\mathcal{T}^N(\mathbb{R}^d))^{\times}$ is a finite-dimensional Lie group and thus locally exponential. The case $N = \infty$ is much more involved and we refer to Glöckner and Neeb (2012, Theorem 5.6) for details. Note, however, that we will establish the convergence of the Lie group exponential and its inverse on a certain closed subspace in Lemma 8.5. \square

Returning briefly to iterated integrals of a smooth path $X : [0, 1] \rightarrow \mathbb{R}^d$, we shall now investigate iterated integrals of the path against itself. For this, identify \mathbb{R}^d with the homogeneous elements of degree 1. Thus X becomes a smooth curve to the tensor algebra. However, for reasons which will become apparent in a moment, we are more interested in the smooth curve $DX(t) := (0, X'(t), 0, 0, \dots) \in \mathcal{T}^N(\mathbb{R}^d)$. Applying the evolution to DX (viewed as a Lie algebra valued path) and sorting the result of the Volterra series by degree, we obtain

$$\begin{aligned} \text{Evol}(DX)(t) = \\ \left(t \mapsto \left(1, X_{0,t}, \int_0^t \int_0^r dX \otimes dX_r, \int_0^t \int_0^{r_2} \int_0^{r_1} dX \otimes dX \otimes dX, \dots \right) \right), \end{aligned}$$

where we either truncate at $N < \infty$ or take the full tensor series for $N = \infty$. We have now used the evolution of the Lie groups to define a well-known object in the theory of rough paths.

8.4 Definition (Signature of a smooth path) Let $X : [0, 1] \rightarrow \mathbb{R}^d$ be a smooth path. Then we define the N -step signature

$$S_N(X)_{s,t} := \left(1, X_{s,t}, \int_s^t \int_s^{r_1} dX \otimes dX, \int_s^t \int_s^{r_2} \int_s^{r_1} dX \otimes dX \otimes dX, \dots \right). \tag{8.6}$$

If $N = \infty$ we also write the shorter $S(X) := S_{\infty}(X)$ and say that $S(X)$ is the signature of the smooth path X .

As $S_N(X)$ is the evolution of a smooth path it satisfies for every $N \in \mathbb{N} \cup \{\infty\}$ the Lie type differential equation

$$\begin{cases} \frac{d}{dt} S_N(X)_{0,t} = S_N(X)_{0,t} \otimes (dX)_t & t \in [0, 1], \\ S_N(X)_{0,0} = \mathbf{1}. \end{cases} \tag{8.7}$$

Before we continue, it is important to stress that the signature can be defined not only for a smooth path. We shall see later (§8.3) that the signature exists for arbitrary rough paths. The signature of a path has turned out to be an immensely important object in the theory of rough paths and its applications. For example, the signature can be computed in advance for paths of interest and can then be applied in numerical analysis or for machine learning purposes. We refer to Chevyrev and Kormilitzin (2016) for an introduction.

We will generalise (8.7) in Exercise 8.2.4 and show that $S_N(X)_{s,t}$ can be recovered directly from $\text{Evol}(DX)(t) = S_N(X)_{0,t}$ by virtue of Chen’s relation $S_N(X)_{s,t} = S_N(X)_{s,u} \otimes S_N(X)_{u,t}$ for all $s < u < t$. However, our investigation so far also hints at the fact that the unit group of the tensor algebra is much larger than needed and contains many elements which will not turn out to be signatures of smooth paths. For example, the second level of the signature is a matrix (viewed as an element in $\mathbb{R}^d \otimes \mathbb{R}^d$). Its symmetric part is fixed by the shuffle of the level 1 part of the signature with itself (Chen’s relation, cf, §8.4 for more on the shuffle product). Hence the second level uniquely is determined by its antisymmetric part which is in the stochastic theory interpreted as the Levy area. To capture these non-linear constraints, one restricts to a certain Lie subgroup of the unit group which expresses the geometric features of the signature. For this let us first study the restriction of the Lie group exponential.

8.5 Lemma *Let $N \in \mathbb{N} \cup \{\infty\}$ and $d \in \mathbb{N}$ and set $\mathcal{I}_N := (\pi_0^N)^{-1}(0)$. Then \mathcal{I}_N is a Lie algebra ideal in $\mathcal{T}^N(\mathbb{R}^d)$ and the following maps are mutually inverse smooth diffeomorphisms*

$$\begin{aligned} \exp_N : \mathcal{I}_N &\rightarrow \mathbf{1} + \mathcal{I}_N, & v &\mapsto \sum_{0 \leq n \leq N} \frac{X^{\otimes n}}{n!}, \\ \log_N : \mathbf{1} + \mathcal{I}_N &\rightarrow \mathcal{I}_N, & \mathbf{1} + v &\mapsto \sum_{0 \leq n \leq N} (-1)^{n+1} \frac{Y^{\otimes n}}{n}. \end{aligned}$$

Proof Since the Lie bracket is given by the commutator, it is clear that $\mathcal{I}_N = (\pi_0^N)^{-1}(0)$ is a Lie ideal (i.e. $[v, w] \in \mathcal{I}_N$ if either v or w is in \mathcal{I}_N). First, note that since $v \in \mathcal{I}_N$ we have $\pi_0^N(v) = 0$ and thus $v^{\otimes k}$ does not contain contributions by homogeneous elements of degree less than k . In particular, we see that for every degree, the series \exp_N and \log_N reduce to polynomials in v . Thus both mappings are well-defined smooth mappings to $\mathcal{T}^N(\mathbb{R}^d)$. Inserting

and rearranging the formal power series into each other shows that \exp_N and \log_N are mutually inverse, hence diffeomorphisms. \square

Note that $\mathbf{1} + \mathcal{I}_N$ is a subgroup of $\mathcal{T}^N(\mathbb{R}^d)^\times$ and a closed submanifold (as a closed affine subspace of $\mathcal{T}^N(\mathbb{R}^d)$). Thus, in particular, it is a closed Lie subgroup of the unit group. Since the exponential series yields the Lie group exponential of the unit group, we can interpret Lemma 8.5 as the statement that the Lie group exponential of $\mathbf{1} + \mathcal{I}_N$ is a diffeomorphism from the Lie algebra onto the group. However, the Lie group we are after is yet a smaller Lie subgroup of $\mathbf{1} + \mathcal{I}_N$ which nevertheless contains all signatures.

8.6 Definition Let $N \in \mathbb{N} \cup \{\infty\}, d \in \mathbb{N}$. Then we define $\mathfrak{g}^N(\mathbb{R}^d)$ as the smallest closed Lie subalgebra generated by the homogeneous elements of degree 1 in $\mathcal{T}^N(\mathbb{R}^d)$. Explicitly, these algebras are constructed as follows: Set $\mathcal{P}^1(\mathbb{R}^d) := \mathbb{R}^d \subseteq \mathcal{T}^N(\mathbb{R}^d)$ (via the canonical identification). Then define recursively:

$$\begin{aligned} \mathcal{P}^{n+1}(\mathbb{R}^d) &:= \mathcal{P}^n(\mathbb{R}^d) + \text{span}\{[x, y] \mid x \in \mathbb{R}^d, y \in \mathcal{P}^n(\mathbb{R}^d)\}, \\ \mathcal{P}^\infty(\mathbb{R}^d) &:= \{(0, P_1, P_2, \dots) \mid \text{for all } i \in \mathbb{N}, P_i \in (\mathbb{R}^d)^{\otimes i} \cap \mathcal{P}^n(\mathbb{R}^d) \\ &\quad \text{for some } n \in \mathbb{N}\}. \end{aligned}$$

Elements in $\mathcal{P}^n(\mathbb{R}^d)$ are called *Lie polynomials* while elements in $\mathcal{P}^\infty(\mathbb{R}^d)$ are called *Lie series*.⁴ Then we set $\mathfrak{g}^N(\mathbb{R}^d) := \overline{\mathcal{P}^N(\mathbb{R}^d)}$ (where the bar denotes topological closure) and observe that these spaces are closed Lie subalgebras of \mathcal{I}_N for $N \in \mathbb{N} \cup \{\infty\}$.

If we take now the image of $\mathfrak{g}^N(\mathbb{R}^d)$ under the exponential map \exp_N we obtain a closed subset $G^N(\mathbb{R}^d) := \exp_N(\mathfrak{g}^N(\mathbb{R}^d))$ of the unit group. It is non-trivial to see that $G^N(\mathbb{R}^d)$ forms a group under tensor multiplication. The classical proof for this fact employs the Baker–Campbell–Hausdorff series as an essential tool. As this would lead us too far from our objects of interest, we will import this result and investigate just its differentiability.

8.7 Proposition The set $G^N(\mathbb{R}^d)$ is a closed Lie subgroup of the unit group $(\mathcal{T}^N(\mathbb{R}^d))^\times$ for all $N \in \mathbb{N} \cup \{\infty\}$. Moreover, this structure turns it into a (locally) exponential Lie group.

Proof We have seen already that $G^N(\mathbb{R}^d)$ is a closed subset. It is a subgroup by Reutenauer (1993, Corollary 3.3). To see that it is a Lie group, we have to show that it is a submanifold. Exploit that $\mathcal{T}^N(\mathbb{R}^d)$ is locally exponential (see

⁴ The name hails from the custom of writing a Lie series in the form of a formal power series instead of the sequence representation we have chosen. We refer to Reutenauer (1993, Chapter 1) for more information.

Corollary 8.3). Hence there is a 0-neighbourhood $V \subseteq \mathcal{T}^N(\mathbb{R}^d)$ such that \exp_{\otimes} restricts to a diffeomorphism on V . By construction we have $\exp_{\otimes}(\mathfrak{g}^N(\mathbb{R}^d)) = \exp_N(\mathfrak{g}^N(\mathbb{R}^d)) = G^N(\mathbb{R}^d)$, whence $\exp_{\otimes}(V \cap \mathfrak{g}^N(\mathbb{R}^d)) = \exp_{\otimes}(V) \cap G^N(\mathbb{R}^d)$ yields a submanifold chart (φ, V) for $G^N(\mathbb{R}^d)$ around the identity. Exploiting the idea that $\mathcal{T}^N(\mathbb{R}^d)^{\times}$ is a Lie group, we see that a submanifold atlas for $G^N(\mathbb{R}^d)$ is then given by $\varphi_g: V \rightarrow G^N(\mathbb{R}^d), x \mapsto g\varphi(x), g \in G^N(\mathbb{R}^d)$. We conclude that $G^N(\mathbb{R}^d)$ is a closed Lie subgroup of the unit group with Lie algebra $\mathfrak{g}^N(\mathbb{R}^d)$. To see that it is locally exponential, it suffices to notice that \exp_{\otimes} restricts to the Lie group exponential of $G^N(\mathbb{R}^d)$. This is due to the fact that the Lie group exponential is defined via solution to certain differential equations. Given initial values in the closed subspace $\mathfrak{g}^N(\mathbb{R}^d)$, the solutions of the equation in the unit group already stay in $G^N(\mathbb{R}^d)$, whence they solve the differential equation in the subgroup. \square

The algebraic arguments in this section were just cited from the literature as we wished to keep the exposition simple. However, the algebraic structure is the key to understanding the Lie groups at hand, since, as a consequence of Lemma 8.5, the Lie groups are globally diffeomorphic to their Lie algebra via the exponential map. For the next result, we assume familiarity with projective limits (see e.g. Hofmann and Morris, 2007, Chapter 1).

8.8 Proposition *Let $m, n \in \mathbb{N} \cup \{\infty\}$ such that $m \geq n$.*

- (a) *Then the canonical projection $\pi_n^m: \mathcal{T}^m(\mathbb{R}^d) \rightarrow \mathcal{T}^n(\mathbb{R}^d)$ is a morphism of locally convex algebras, which restricts to a Lie group morphism $p_n^m: G^m(\mathbb{R}^d) \rightarrow G^n(\mathbb{R}^d)$.*
- (b) *We obtain a commutative diagram of Lie groups and their associated Lie algebras*

$$\begin{array}{ccccccc}
 \{0\} \times \mathbb{R}^d = \mathfrak{g}^1(\mathbb{R}^d) & \xleftarrow{\mathbf{L}(p_1^2)} & \mathfrak{g}^2(\mathbb{R}^d) & \xleftarrow{\mathbf{L}(p_2^3)} & \mathfrak{g}^3(\mathbb{R}^d) & \xleftarrow{\quad} & \dots \\
 \downarrow \exp_1 & & \downarrow \exp_2 & & \downarrow \exp_3 & & \\
 \{1\} \times \mathbb{R}^d = G^1(\mathbb{R}^d) & \xleftarrow{p_1^2} & G^2(\mathbb{R}^d) & \xleftarrow{p_2^3} & G^3(\mathbb{R}^d) & \xleftarrow{\quad} & \dots,
 \end{array} \tag{8.8}$$

where the upper row is a projective system of locally convex Lie algebras whose limit is $\mathfrak{g}^{\infty}(\mathbb{R}^d)$.

Proof From the definition of the product in the tensor algebra it is clear that the π_n^m are algebra morphisms. As π_n^m is just the projection from a product onto some of its components, it is continuous in the product topology, hence a morphism of locally convex algebras. Note that this entails that π_n^m restricts to a morphism of locally convex Lie algebras $q_n^m: \mathfrak{g}^m(\mathbb{R}^d) \rightarrow \mathfrak{g}^n(\mathbb{R}^d)$ and a

morphism of Lie groups $p_n^m : G^m(\mathbb{R}^d) \rightarrow G^n(\mathbb{R}^d)$ (it is smooth as the restriction of a continuous linear map to closed submanifolds). Since π_n^m is linear, we have $\mathbf{L}(p_n^m) = T_1\pi_n^m|_{\mathfrak{g}^m(\mathbb{R}^d)} = q_n^m$. Thus the naturality of the exponential map (3.8) yields $p_n^m \circ \exp_m = \exp_n \circ \mathbf{L}(p_n^m)$. Summing up, this proves (a) and establishes the commutativity of (8.8).

For part (b) let us note that (a) establishes that the upper row of (8.8) is a projective system of locally convex Lie algebras (as $\mathbf{L}(p_i^j) \circ \mathbf{L}(p_k^i) = \mathbf{L}(p_k^j)$ and these mappings are continuous morphisms of Lie algebras). Consider now $x \in \mathfrak{g}^\infty(\mathbb{R}^d)$. Its projection $\pi_n^\infty(x)$ is contained in $\mathfrak{g}^n(\mathbb{R}^d)$ (this is clear for a Lie series and follows by considering converging sequences for the elements in the closure since the Lie algebras $\mathfrak{g}^n(\mathbb{R}^d)$ are closed). Since π_n^∞ restricts to the Lie algebra morphism $\mathbf{L}(p_n^\infty)$ on $\mathfrak{g}^\infty(\mathbb{R}^d)$, this implies that $\mathfrak{g}^\infty(\mathbb{R}^d)$ is the projective limit of the projective system in the category of Lie algebras. In addition, the product topology on $\mathcal{T}^\infty(\mathbb{R}^d)$ is the projective limit of the locally convex spaces $\mathcal{T}^n(\mathbb{R}^d)$, whence the topology on $\mathfrak{g}^\infty(\mathbb{R}^d)$ is the locally convex projective limit topology induced by the projective system. In conclusion, $\mathfrak{g}^\infty(\mathbb{R}^d)$ is the projective limit in the category of locally convex Lie algebra. □

8.9 Remark (Projective limits of finite-dimensional Lie groups) In Proposition 8.8 we exploited that the projective limit of locally convex Lie algebras can be described as the projective limit of Lie algebras with the (locally convex) projective limit topology. Thanks to the commutativity of (8.8), the lower row also forms a projective system of finite-dimensional Lie groups. Projective limits for Lie groups may not exist (while they always exist in the category of topological groups). Topological groups which are projective limits of finite-dimensional Lie groups are called *pro-Lie groups* (Hofmann and Morris, 2007). Hence (8.8) shows that $G^\infty(\mathbb{R}^d)$ is a pro-Lie group which is simultaneously a Lie group. This situation has been studied in Hofmann and Neeb (2009). It is worth mentioning that groups with these properties inherit a surprising amount of structure from the finite-dimensional Lie groups which were used in their construction.

Summing up, the truncated groups are closely connected to their projective limit $G^\infty(\mathbb{R}^d)$. Moreover, (truncated) signatures of smooth paths extend naturally to the projective limit. In the next section we will review the concept of a rough path, which, in a certain sense, generalises the signature of paths of low regularity. For this, the geometry of the groups $G^N(\mathbb{R}^d)$ will be instrumental.

Exercises

8.2.1 Declare a Hilbert space structure on $\mathcal{T}^N(\mathbb{R}^d)$ for $N \in \mathbb{N}$ by defining the canonical basis elements $e_{i_1} \otimes e_{i_2} \otimes \cdots \otimes e_{i_k}$ to be orthonormal. Show that:

- (a) The norm corresponding to the inner product satisfies $\|v \otimes w\| \leq \|v\| \cdot \|w\|$ and $\|v \otimes w\| = \|w \otimes v\|$. Deduce that $\mathcal{T}^N(\mathbb{R}^d)$ becomes a Banach algebra.
- (b) Identifying the homogeneous elements of degree k with elements \mathbb{R}^{d^k} by sending the canonical bases to each other, the resulting isomorphism is an isometry of Hilbert spaces.

8.2.2 Consider the Step 2 truncated tensor algebra $\mathcal{T}^2(\mathbb{R}^d) = \mathbb{R} \times \mathbb{R}^d \times \mathbb{R}^d \otimes \mathbb{R}^d$.

- (a) Show that the multiplication of the truncated tensor algebra is given by

$$(a, b, c) \cdot (x, y, z) = (ax, ay + xb, az + xc + b \otimes y)$$

$$\text{and } (1, b, c)^{-1} = (1, -b, -c + b \otimes b).$$

Deduce that if $d \neq 1$, the product is not commutative.

- (b) Let $X: [0, 1] \rightarrow \mathbb{R}^d$ be a smooth path and $\mathbb{X}_{s,t} := \int_s^t X_{s,r} \otimes X'_r dr$. Define $\mathbf{X}_{s,t} := (1, X_{s,t}, \mathbb{X}_{s,t}) \in \mathcal{T}^2(\mathbb{R}^d)$. Establish that X satisfies Chen's relation (8.4)

$$\mathbb{X}_{s,t} - \mathbb{X}_{s,u} - \mathbb{X}_{u,t} = X_{s,u} \otimes X_{u,t}, \quad \text{for all } u \in [s, t].$$

Prove then that $\mathbf{X}_{s,t} = \mathbf{X}_{s,u} \otimes \mathbf{X}_{u,t}$ (also called Chen's relation).

8.2.3 Show that for $N \in \mathbb{N}$ the algebra $\mathcal{T}^N(\mathbb{R}^d)$ is a quotient of $\mathcal{T}^\infty(\mathbb{R}^d)$ modulo the algebra ideal $I_N = \{(x_k)_{k \in \mathbb{N}} \in \mathcal{T}^\infty(\mathbb{R}^d) \mid x_1 = \cdots = x_N = 0\}$.

8.2.4 Let $X: [0, 1] \rightarrow \mathbb{R}^d$ be a smooth path, $N \in \mathbb{N} \cup \{\infty\}$ and $DX: [0, 1] \rightarrow \mathcal{T}^N(\mathbb{R}^d)$, $t \mapsto (0, X'_t, 0, \dots)$. Show that

- (a) for fixed s the signature satisfies the differential equation

$$\begin{cases} \frac{d}{dt} S_N(X)_{s,t} = S_N(X)_{s,t} \otimes (DX)_t, & s < t \leq 1, \\ S_N(X)_{s,s} = \mathbf{1} = (1, 0, \dots), \end{cases}$$

- (b) the signature satisfies Chen's relation

$$S_N(X)_{s,t} = S_N(X)_{s,u} \otimes S_N(X)_{u,t}, \quad 0 \leq s \leq u \leq t \leq 1.$$

Hint: It suffices to prove this for every projection of $S_N(X)$ to $(\mathbb{R}^d)^{\otimes k}$ and consider the iterated integral of $dX_{r_1} \otimes \cdots \otimes dX_{r_k}$ over the simplex $\Delta^N = \{s < r_1 < r_2 < \cdots < r_n < t\}$.

- (c) Deduce that $S_N(X)_{s,t} = S_N(X)_{0,s}^{-1} \otimes S_N(X)_{0,t}$, whence the signature can be recovered from the curve $\text{Evol}(DX)$ via the group operations.
- 8.2.5 Let $a \in \mathcal{T}^N(\mathbb{R}^d)$, $N \in \mathbb{N} \cup \{\infty\}$ such that $a_0 := \pi^N(a) \neq 0$. Write $a = a_0(\mathbf{1} + b)$ and prove that the Neumann inverse (8.5) yields $a^{-1} = a_0^{-1}(\mathbf{1} + b)^{-1} = a_0^{-1} \sum_{k=1}^{\infty} (-1)^k b^{\otimes k}$.

8.3 A Rough Introduction to Rough Paths

In this section we will recall the notion of a rough path. The main idea of rough path theory is that paths of much lower regularity than being smooth can be augmented with extra information replacing the iterated integrals we studied in the last section. Indeed the basic idea is to declare the signature to be the object of interest and define signature-like objects in the tensor algebra. While we will present the basic theory of rough paths, we recommend one of the excellent introductions to rough path theory available (see e.g. Friz and Victoir, 2010; Friz and Hairer, 2020) for more in-depth information.

As a starting point, let us formalise the properties observed for the signature in the last section.

8.10 Definition Let $\Delta := \{(s, t) \in [0, 1] \mid 0 \leq s \leq t \leq 1\}$ be the *standard simplex*, $N \in \mathbb{N} \cup \{\infty\}$ and $d \in \mathbb{N}$. We call a map

$$\mathbf{X}: \Delta \rightarrow \mathcal{T}^N(\mathbb{R}^d), \quad \mathbf{X}_{s,t} = (X_{s,t}^0, X_{s,t}^1, X_{s,t}^2, \dots)$$

*multiplicative functional*⁵ of degree N if $X_{s,t}^0 = 1$ for all $(s, t) \in \Delta$ and the map satisfies

$$\text{Chen's relation: } \mathbf{X}_{s,t} = \mathbf{X}_{s,u} \otimes \mathbf{X}_{u,t} \text{ for all } s, u, t \in [0, 1], s \leq u \leq t. \quad (8.9)$$

Note that Chen's relation for the first two (non-trivial) components of \mathbf{X} reduces to (cf. Exercise 8.2.2)

$$X_{s,t}^1 = X_{s,u}^1 + X_{u,t}^1, \quad X_{s,t}^2 = X_{s,u}^2 + X_{u,t}^2 + X_{s,u}^1 \otimes X_{u,t}^1. \quad (8.10)$$

Moreover, Lemma 8.2 (a) shows that every multiplicative functional is invertible in the tensor algebra, whence Chen's relation entails $\mathbf{X}_{s,t} = \mathbf{X}_{0,s}^{-1} \otimes \mathbf{X}_{0,t}$. So instead of a multiplicative functional, we may think of the path $[0, 1] \rightarrow \mathcal{T}^N(\mathbb{R}^d)$, $t \mapsto \mathbf{X}_{0,t}$. Hence the term 'rough path' will make more sense once

⁵ There is a deeper story going on which motivates the term 'multiplicative functional'. We will briefly discuss this in §8.4.

we define it. For the first level, however, this translates to the statement that there is a path $X: [0, 1] \rightarrow \mathbb{R}^d$ with $X_{s,t}^1 = X_t - X_s$.

8.11 Example In the previous section we saw that the signature of a smooth path yields a multiplicative functional. For example, we can augment the zero-path in different ways to generate multiplicative functionals: $\mathbf{X}_{s,t} := (1, 0, (t-s)w)$ is a multiplicative functional of degree 2 for any $w \in \mathbb{R}^d \otimes \mathbb{R}^d$. Note that it coincides with the level 2-signature of the zero path only if w is the zero element. More generally, for any function $F: [0, 1] \rightarrow (\mathbb{R}^d)^{\otimes N}$ the map $\mathbf{X}_{s,t} = (1, 0, \dots, 0, F(t) - F(s)) \in \mathcal{T}^N(\mathbb{R}^d)$ is a multiplicative functional (see Exercise 8.3.1).

Now we need to add an analytic condition to the algebraic objects we just defined. In this case we wish to consider rough signals in the framework of Hölder regular paths.

8.12 Definition Let $X: [0, 1] \rightarrow E$ be a continuous map with values in a Banach space $(E, \|\cdot\|)$. We say X is an α -Hölder path for $0 < \alpha < 1$ if

$$\|X\|_\alpha := \sup_{\substack{t, s \in [0, 1], \\ t \neq s}} \frac{\|X_t - X_s\|}{|t - s|^\alpha} < \infty.$$

Denote by $C^\alpha([0, 1], E)$ the space of all α -Hölder continuous functions.

It is well known (see Exercise 8.3.2) that $C^\alpha([0, 1], E)$ is a Banach space on which $\|\cdot\|_\alpha$ is a continuous seminorm (it annihilates all constant paths).

8.13 Remark Again we will be restricted here to paths on the interval $[0, 1]$. The results carry over to any interval by composing the Hölder paths with a reparametrisation of the interval. Note, however, that the Hölder norm is *not* invariant under reparametrisation. This is one reason why in the rough paths literature one often considers the (more or less) equivalent formulation via p -variation paths for $p > 1$. The p -variation norm is invariant under reparametrisation; see Friz and Victoir (2010).

Due to Young's theorem, it is not possible to augment an α -Hölder path with iterated integrals against itself if $\alpha < 1/2$. The core idea of rough path theory is to augment an α -Hölder path with additional information to make integration against it feasible and circumvent the restrictions of Young integration theory. To this end we augment the concept of a multiplicative functional with a Hölder condition.

8.14 Definition Fix $\alpha \in]0, 1[$ and $N \geq \lfloor 1/\alpha \rfloor$. An \mathbb{R}^d -valued α -rough path consists of a multiplicative functional $\mathbf{X} := (1, X^1, X^2, X^3, \dots, X^N)$ such that the k th component X^k is ‘ k -times α Hölder continuous’, that is,

$$\|X^k_{s,t}\| \lesssim |t - s|^{k\alpha}, \text{ where ‘}\lesssim\text{’ means inequality up to a constant.} \quad (8.11)$$

If, in addition, the α -rough path takes values in $G^N(\mathbb{R}^d)$, we call \mathbf{X} a *weakly geometric rough path*. We define $\text{RP}^\alpha([0, 1], \mathbb{R}^d)$ as the *set of weakly geometric α -rough paths* (we shall see in Theorem 8.15 that it makes sense to drop N from the notation for $\text{RP}^\alpha([0, 1], \mathbb{R}^d)$).

The higher levels of a rough path are not given by increments of a function, whence they do not become constant even if k -times α -Hölder means that the Hölder index satisfies $k\alpha > 1$. Lyons’ original concept of rough path does not require the rough path to take its image in the group $G^N(\mathbb{R}^d)$. However, it has turned out that the general notion is not sufficient to solve non-linear rough differential equations (see e.g. Friz and Victoir, 2010; Friz and Hairer, 2020). For this purpose one should consider weakly geometric rough paths and we will do this in the rest of the section. To ease notation we shall (unless we explicitly say otherwise) only consider weakly geometric rough paths and simply call them rough paths. Let us recall that the cutoff level N in the definition of a rough path is unimportant as long as $N \geq \lfloor 1/\alpha \rfloor$.

8.15 Theorem (Lyons’ lifting theorem (Friz and Victoir, 2010, Theorem 9.5)) *Let $\alpha \in]0, 1[$ and $\lfloor 1/\alpha \rfloor \leq n$. Assume that $\mathbf{X}: \Delta \rightarrow G^n(\mathbb{R}^d)$ is an α -rough path. Then there exists a unique α -rough path $\mathbf{X}^{n+1}: \Delta \rightarrow G^{n+1}(\mathbb{R}^d)$ extending \mathbf{X} , that is, if $\pi_n^{n+1}: \mathcal{T}^{n+1}(\mathbb{R}^d) \rightarrow \mathcal{T}^n(\mathbb{R}^d)$ is the canonical projection, then $\pi_n^{n+1}(\mathbf{X}^{n+1}) = \mathbf{X}$. A rough path extending \mathbf{X} in this way is called a Lyons lift of \mathbf{X} .*

Thus every α -rough path with values in $G^n(\mathbb{R}^d)$ extends to an α -rough path with values in $G^m(\mathbb{R}^d)$, $m \geq n \geq \lfloor 1/\alpha \rfloor$.

As a consequence of Theorem 8.15, every α -rough path can be obtained as a restriction of an α -rough path with values in $G^\infty(\mathbb{R}^d)$, or conversely as an extension of an α -rough path with values in $G^{\lfloor 1/\alpha \rfloor}(\mathbb{R}^d)$. The remarkable fact here is of course that the extension is uniquely determined once information up to level $\lfloor 1/\alpha \rfloor$ is available.

8.16 Example We have seen that for a smooth path X and $N \in \mathbb{N} \cup \{\infty\}$, the signature $S_N(X)$ is a multiplicative functional. More generally, if we start with an α -Hölder path for $\alpha > 1/2$, we can compute its signature using iterated Young integrals. It is then a consequence of Young’s inequality (Young, 1936)

that the resulting multiplicative functional is indeed an α -rough path (see Friz and Hairer, 2020, Section 4, for a detailed discussion).

For paths of lower regularity, the idea is that the components X^k of a rough path \mathbf{X} replace the (in general, ill-defined) iterated integrals of the path giving the first level increments. Due to its importance, let us stress it here again explicitly: If $\mathbf{X} = (1, X_{s,t}^1, X_{s,t}^2)$ is an α -rough path for $\alpha \in]1/3, 1/2[$ and $X_{s,t}^1 = X_t - X_s$, one should interpret the second level as

$$\int_s^t X_{u,s} dX_u := X_{s,t}^2, \quad 0 \leq s \leq t \leq 1,$$

where the left-hand side is defined via the right-hand and not the other way around! The levels of an α -rough path thus encode information similar to the iterated integrals in the signature of a smooth path.

It is not obvious at all whether a given α -Hölder path with values in \mathbb{R}^d can be enhanced to yield an α -rough path. Thanks to a result by Lyons and Victoire, this can be achieved in many cases through an abstract extension result. If $1/\alpha \notin \mathbb{N}$, every α -Hölder path can be enhanced (non-uniquely) to an α -Hölder rough path (this is the Lyons–Victoire lifting theorem; Lyons and Victoir, 2007, Theorem 1). As an example of the non-uniqueness of the extension let us mention the following example for rough paths arising from Brownian motion.

8.17 Example (Brownian motion as a rough path) Brownian motion models the (random) movement of a particle in a fluid. It can be modelled as a stochastic process $B: \Omega \times [0, 1] \rightarrow \mathbb{R}^d$, with independent Gaussian increments and continuous sample paths. Here $(\Omega, \mathcal{F}, \mathbb{P})$ is a probability space. However, we refer to the stochastic literature for explanations and more details. One can show that for partitions P of $[0, 1]$, the Riemann sum

$$\int_0^1 B_r \otimes d^\theta B_r = \lim_{|P| \rightarrow 0} \sum_{P=(t_i)} B_{t_i + \theta(t_{i+1} - t_i)} \otimes B_{t_i, t_{i+1}}$$

converges in $L^2(\Omega, \mathbb{R}^{d \times d})$ for mesh size $|P|$ converging to 0 and $\theta \in [0, 1]$. Contrary to usual Riemann–Lebesgue integration theory, the integral $\int B_r \otimes d^\theta B_r$ is *not* independent of θ . The choice $\theta = 0$ leaves the martingale structure invariant and is referred to as the Itô integral, whereas $\theta = \frac{1}{2}$ is compatible with regular calculus rules and is referred to as the Stratonovich integral. It is well known that B is an α -Hölder path for every $\alpha \in]1/3, 1/2[$ and if we write $B^{\text{Itô}}$ and B^{Strat} for the Itô and the Stratonovich integrals, we obtain two α -rough paths $(1, B, B^{\text{Itô}})$ and $(1, B, B^{\text{Strat}})$. Note that the rough path obtained via the Itô integral does not take its values in $G^2(\mathbb{R}^d)$ (i.e. it is not a weakly geometric rough path), while the one obtained from Stratonovich

integration takes values in $G^2(\mathbb{R}^d)$ and is weakly geometric. We refer to Friz and Hairer (2020, Chapter 3) for a detailed discussion of these examples.

The reader may wonder now in what sense a rough path is a path, that is, can we interpret it as a Hölder continuous path $[0, 1] \rightarrow G^N(\mathbb{R}^d)$? For this we have to consider the metric geometry of the groups $G^N(\mathbb{R}^d)$ for $N < \infty$. However, as a first step to make sense of the following constructions, we need the following result (whose full proof we postpone to Remark 8.25).

8.18 Lemma *Let $\mathbf{X}: \Delta \rightarrow G^N(\mathbb{R}^d), \mathbf{X} = (1, X^1, X^2, \dots)$ be an α -rough path. Taking the pointwise inverse, the map $\mathbf{X}^{-1}: \Delta \rightarrow G^N(\mathbb{R}^d), t \mapsto (\mathbf{X}_t)^{-1} = (1, X_t^{-1}, X_t^{-2}, \dots)$ is graded $k\alpha$ -Hölder, that is, $\|X^{-k}\|_{k\alpha} < \infty$ for all $k \in \mathbb{N}$.*

Proof We will only prove the case where $N = 2$ and $d = 2$. Without more techniques from §8.4 the general case turns out to be quite involved. Note that for $N = 2$, thanks to Exercise 8.2.2(a), we have $\mathbf{X}^{-1} = (1, -X^1, -X^2 + X^1 \otimes X^1)$. Thus the degree 1 component is α -Hölder since X^1 is α -Hölder. Immediately, we see that the degree 2 component might not be 2α -Hölder, as $X^1 \otimes X^1$ is a product of α -Hölder functions. To circumvent this, we express \mathbf{X} with the help of the standard basis e_1, e_2 of \mathbb{R}^2 as $\mathbf{X} = (1, x_1e_1 + x_2e_2, \sum_{1 \leq i, j \leq 2} y_{ij}e_i \otimes e_j)$. Working out the constraints imposed by $G^2(\mathbb{R}^2) = \exp_2(g^2(\mathbb{R}^2))$ (the reader should check this!), we find that

$$x_1x_2 = y_{12} + y_{21}, \quad x_1^2 = 2y_{11}, \quad x_2^2 = 2y_{22}.$$

Plugging this into the inversion formula we find for the degree 2 component,

$$\begin{aligned} X^{-2} &= - \sum_{1 \leq i, j \leq 2} y_{ij}e_i \otimes e_j + x_1^2e_1 \otimes e_1 + x_1x_2e_1 \otimes e_2 + x_1x_2e_2 \otimes e_1 + x_2^2e_2 \otimes e_2 \\ &= y_{11}e_1 \otimes e_1 + y_{22}e_2 \otimes e_2 + y_{12}e_2 \otimes e_1 + y_{21}e_1 \otimes e_2. \end{aligned}$$

Thus the second component is 2α -Hölder as it contains only contributions from the 2α -Hölder maps y_{ij} comprising the second level of \mathbf{X} . This happens in general, as the non-linear constraints imposed by $G^N(\mathbb{R}^d)$ allow us to rewrite the k th degree component of \mathbf{X}^{-1} as a sum of the functions comprising the k th degree component of \mathbf{X} (whence they preserve the Hölder condition). \square

The discussion of the $N = 2, d = 2$ case reveals that Lemma 8.18 will, in general, be false if the rough path is not weakly geometric. We now define a suitable metric on the group $G^N(\mathbb{R}^d)$ for which every α -rough path will correspond to a Hölder function.

8.19 Definition Consider for $N \in \mathbb{N}$ the subset $\mathcal{T}_1^N(\mathbb{R}^d) = \{x \in \mathcal{T}^N(\mathbb{R}^d) \mid \pi_0^N(x) = 1\}$. Let $x = (1, x^1, x^2, \dots, x^N) \in \mathcal{T}_1^N(\mathbb{R}^d)$; then we define

$$\|x\| := \max_{k=1,\dots,N} \{(k! \|x^k\|)^{1/k}\} + \max_{k=1,\dots,N} \{(k! \|(x^{-1})^k\|)^{1/k}\}, \tag{8.12}$$

where $\|\cdot\|$ is the norm we have chosen on $\mathcal{T}^N(\mathbb{R}^d)$. Then we set $\rho_N(x, y) := |x^{-1} \otimes y|$. Identifying $\mathcal{T}_1^N(\mathbb{R}^d)$ with the vector space \mathcal{I}_N (by subtracting $\mathbf{1}$), it is easy to see that $|\cdot|$ becomes a norm and ρ_N induces a left-invariant, symmetric and subadditive metric on $G^N(\mathbb{R}^d)$ (see Exercise 8.3.4 for the details).

The metric ρ_N turns the group $G^N(\mathbb{R}^d)$ into a homogeneous group in the sense of Folland and Stein (1982).⁶ Now if \mathbf{X} is an α -rough path (with values in $G^N(\mathbb{R}^d)$), then the path $x_t := \mathbf{X}_{0,t}$ satisfies

$$\|x\|_{\alpha}^{\rho_N} := \sup_{\substack{t \neq s \\ s, t \in [0,1]}} \frac{\rho_N(x_s, x_t)}{|t - s|^{\alpha}} = \sup_{\substack{t \neq s \\ s, t \in [0,1]}} \frac{\rho_N(\mathbf{1}, \mathbf{X}_{s,t})}{|t - s|^{\alpha}} < \infty. \tag{8.13}$$

In other words, an α -rough path is an α -Hölder continuous path with values in metric space $(G^N(\mathbb{R}^d), \rho_N)$. This statement hinges on the rough path taking values in $G^N(\mathbb{R}^d)$ (i.e. the path being weakly geometric) as we need Hölder continuity of the pointwise inverse \mathbf{X}^{-1} .

Indeed due to Chen’s relation this is an equivalent point of view, up to forgetting the starting point of the path in \mathbb{R}^d . Unfortunately this point of view is limited to the truncated groups, as the metric (8.12) does not extend to $G^{\infty}(\mathbb{R}^d)$.

8.20 Remark In Remark 8.9 we saw that $G^{\infty}(\mathbb{R}^d)$ is the projective limit (as a Lie group) of the truncated groups $G^N(\mathbb{R}^d)$. Moreover, this Lie group is modelled on the Fréchet space $\mathfrak{g}^{\infty}(\mathbb{R}^d)$. So, in view of the characterisation of rough paths as Hölder continuous paths in the truncated groups, the question is of course whether we can use the metric on $\mathfrak{g}^{\infty}(\mathbb{R}^d)$ to construct a suitable metric dist which allows us to cast α -rough path as an α -Hölder path with values in the metric space $(G^{\infty}(\mathbb{R}^d), \text{dist})$.

This problem was considered in Le Donne and Züst (2021) starting from the *Carnot–Carathéodory metric* on $G^N(\mathbb{R}^d)$, $N < \infty$. It is defined as $d_{CC}^N(y, z) := d_{CC}^N(\mathbf{1}, y^{-1} \otimes z)$ and

$$d_{CC}^N(\mathbf{1}, y) := \inf \left\{ \int_0^1 \|X'_t\| dt \mid X \in C([0, 1], \mathbb{R}^d), \begin{array}{l} X_0=0, X_t \text{ has bounded variation} \\ \text{and } y=S_N(X)_{0,1} \end{array} \right\},$$

where the signature of the bounded variation path is defined as in (8.6) using (iterated) Riemann–Stieltjes integrals. We will abbreviate this metric as the

⁶ The groups $G^N(\mathbb{R}^d)$ possess many strong structural properties which are exploited in rough path theory. They are homogeneous groups and connected and simply connected nilpotent Lie groups (i.e. Carnot groups, which are studied in sub-Riemannian geometry; see Le Donne and Züst (2021)). The details are beyond the scope of this chapter and we refer the reader to the literature.

CC-metric and note that it is not straightforward to prove that it is a metric. Moreover, one can show that the CC-metric is equivalent to the metric ρ_N (see Friz and Victoir, 2010, 7.5.4 for details on the CC-metric). Now one can argue that the metric d_∞ should correspond to the metric on the projective limit of the system $(G^N(\mathbb{R}^d), \mathbf{1}, d_{CC}^N)$ in the category of (pointed) metric spaces (with morphisms given by submetrics). However, one of the main results of Le Donne and Züst (2021) is the somewhat surprising insight that the limiting object $(G_\infty, \mathbf{1}, d_\infty)$ cannot be a topological group in the topology induced by d_∞ (let us note that $G_\infty \neq G^\infty(\mathbb{R}^d)$).

Hence there seems to be no straightforward way to construct a metric on the projective limit which simultaneously captures the desired convergence and geometry, is left-invariant and turns the projective limit into a Lie group.

Summing up, rough paths can naturally be identified as maps with values in the infinite-dimensional Lie group $G^\infty(\mathbb{R}^d)$. This group is closely connected to the truncated groups $G^N(\mathbb{R}^d)$ and inherits many properties from the projective system of these groups. However, there are important geometric properties connected to the sub-Riemannian geometry of the groups $G^N(\mathbb{R}^d)$, which have no counterpart in the infinite-dimensional group. This leaves us at an uncomfortable situation: If it is enough to work with the finite-dimensional groups, why bother with the more complicated situation of $G^\infty(\mathbb{R}^d)$?

One reason to care about the infinite-dimensional group is that it hosts all rough paths regardless of their regularity (recall that an α -rough path can only be defined on the group $G^N(\mathbb{R}^d)$, where $N \geq \lfloor 1/\alpha \rfloor$). It would be interesting to understand the geometry and manifold structure of the set of all elements which can be reached by α -rough paths (the construction of tangent spaces to α -rough paths for $\alpha \in]1/3, 1/2[$ in Qian and Tudor (2011) can be understood as a step in this direction). Another reason might be that one could be interested in different flavours of rough paths such as Gubinelli's branched rough paths (Gubinelli, 2010). The properties of branched rough paths require a different geometry and one has to replace the groups $G^N(\mathbb{R}^d)$. We shall describe the general construction in the next section (but mention that for the branched rough paths the construction actually yields isomorphic groups due to a deep algebraic result).

Exercises

- 8.3.1 Let $N \in \mathbb{N}$ and $\mathbf{X}, \mathbf{Y}: \Delta \rightarrow \mathcal{T}^N(\mathbb{R}^d)$ be multiplicative functionals which agree up to degree $N - 1$ (i.e. their projections onto the components up to m th level are equal for all $m \leq N$). Then prove that

- (a) The difference function $F_{s,t} = X_{s,t}^N - Y_{s,t}^N \in (\mathbb{R}^d)^{\otimes N}$ is additive in the sense that for all $s \leq u \leq t$ one has $F_{s,t} = F_{s,u} + F_{u,t}$.
- (b) If $F: \Delta \rightarrow (\mathbb{R}^d)^{\otimes N}$ is an additive function (in the above sense), then $(s,t) \mapsto \mathbf{X}_{s,t} + F_{s,t}$ is also an additive functional (where addition is addition in the tensor algebra).

8.3.2 We consider the α -Hölder space $C^\alpha([0, 1], E)$ for $n \in \mathbb{N}$, $0 < \alpha < 1$ for $(E, \|\cdot\|)$ a Banach space. Show that the map

$$\|x\|_\alpha := \sup_{\substack{t,s \in [0,1], \\ t \neq s}} \frac{\|x_t - x_s\|}{|t - s|^\alpha}$$

is a semi-norm on $C^\alpha([0, 1], E)$ which is not a norm. Then deduce that

$$\|x\|_{C^\alpha} := \sup_{t \in [0,1]} \|x_t\| + \|x\|_\alpha$$

is a norm, turning $C^\alpha([0, 1], E)$ into a Banach space.

8.3.3 Assume that a path $x: [0, 1] \rightarrow E$ with values in a Banach space is α -Hölder in the sense of Definition 8.12. If $\alpha > 1$, show that x is differentiable and constant.

8.3.4 Consider for $N \in \mathbb{N}$ the subset $\mathcal{T}_1^N(\mathbb{R}^d) = \{x \in \mathcal{T}^N(\mathbb{R}^d) \mid \pi_0(x) = 1\}$ (i.e. all elements in the tensor algebra whose zeroth degree term is 1) (recall that these are invertible!) set $\rho_N(x, y) := |x^{-1} \otimes y|$ where, as in (8.12), we have

$$\begin{aligned} |(1, x^1, x^2, \dots, x^n)| \\ = \max_{k=1, \dots, N} \{(k! \|x^k\|)^{1/k}\} + \max_{k=1, \dots, N} \{(k! \|((x^{-1})^k)\|)^{1/k}\}. \end{aligned}$$

- (a) Show that $|\cdot|$ is subadditive, that is, $|x \otimes x'| \leq |x| + |x'|$.
- (b) Given $\lambda \in \mathbb{R}$, define the dilation $\delta_\lambda: \mathcal{T}_1^N(\mathbb{R}^d) \rightarrow \mathcal{T}_1^N(\mathbb{R}^d)$, $(x^k)_{0 \leq k \leq N} \mapsto (\lambda^k x^k)_{0 \leq k \leq N}$. Show that $|\delta_\lambda(x)| = |\lambda| |x|$.
- (c) Show that ρ_N is a left-invariant metric on $G^N(\mathbb{R}^d)$, that is, $\rho_N(b \otimes x, b \otimes y) = \rho_N(x, y)$.
- (d) Let $N \geq \lfloor 1/\alpha \rfloor$. Prove that a multiplicative functional $X: \Delta \rightarrow G^N(\mathbb{R}^d)$ is an α -rough path if and only if (8.13) is satisfied.
Hint: You will need Lemma 8.18 and can use (without a proof) that the k th component of \mathbf{X}^{-1} arises by applying a linear map to the k th component of \mathbf{X} .

8.4 Rough Paths and the Shuffle Algebra

In this section we will broaden the scope of the investigation. The reason for this is twofold. On one hand, it allows us to give some further motivation to the question why infinite-dimensional differential geometry is of relevance in rough path theory. On the other hand, there is a pleasing framework which generalises the construction of the groups $G^\infty(\mathbb{R}^d)$ and leads to a whole class of infinite-dimensional Lie groups which have recently been found in a variety of mathematical contexts.

To start, consider again a smooth path $X: [0, 1] \rightarrow \mathbb{R}^d$ with components $X^i, i=1, \dots, d$. Now the iterated integrals comprising the signature $S_N(X)$ satisfy several algebraic conditions. For example, for $S_2(X)_{s,t} = (1, X_{s,t}, \int_s^t X_{s,r} dX)$ the product rule of ordinary calculus yields

$$X_{s,t}^i \cdot X_{s,t}^j = \int_s^t X_{s,r}^i dX^j + \int_s^t X_{s,r}^j dX^i = \int_s^t \int_s^r dX^i dX^j + \int_s^t \int_s^r dX^j dX^i. \tag{8.14}$$

There are similar identities for higher-order iterated integrals and this is ultimately responsible for the non-linear set $G^N(\mathbb{R}^d)$ being the correct state space of a rough path. The question is of course how these identities and the associated combinatorics can be conveniently expressed. These questions lead to the so-called shuffle product.

8.21 Definition (Shuffle algebra (Reutenauer, 1993)) Consider the set $\mathcal{A} = \{1, 2, \dots, d\}$, which in this context is called an *alphabet* and its elements *letters*. By concatenation we can construct words from the letters and the set \mathcal{A}^* of all words including the empty word \emptyset . We construct now the shuffle algebra $\text{Sh}(\mathcal{A}) = \mathbb{R}\mathcal{A}^*$ as the vector space generated by the words over \mathcal{A} . Note that, as a locally convex space, $\text{Sh}(\mathcal{A})$ is isomorphic to the direct sum of countably many copies of \mathbb{R} ; see Example A.36. Moreover, \mathcal{A}^* and thus also $\text{Sh}(\mathcal{A})$ is graded by word length, that is, if $w = a_1 \cdots a_n \in \mathcal{A}^*$ for letters a_i , we set $|w| = n$ and say that an element of $\text{Sh}(\mathcal{A})$ is homogeneous of degree $n \in \mathbb{N}_0$ if it is a linear combination of words of length n . For the algebra structure let $a, b \in \mathcal{A}$ and $u, w \in \mathcal{A}^*$. Then the *shuffle product* is defined recursively by

$$\emptyset \sqcup w = w \sqcup \emptyset = w, \quad (au) \sqcup (bw) := a(u \sqcup (bw)) + b((au) \sqcup w).$$

We will see in Exercise 8.4.1 that it extends to a continuous product on $\text{Sh}(\mathcal{A})$, whence $\text{Sh}(\mathcal{A})$ becomes a unital locally convex algebra, called the *shuffle algebra*.⁷

⁷ The shuffle algebra is actually a graded bialgebra. The coproduct is given by deconcatenation of words

$$\delta: \text{Sh}(\mathcal{A}) \rightarrow \text{Sh}(\mathcal{A}) \otimes \text{Sh}(\mathcal{A}),$$

It turns out that the shuffle product encodes the combinatorics keeping track of the non-linear identities in the signature. For this we define for a smooth path $X: [0, 1] \rightarrow \mathbb{R}^d$ with (truncated) signature $S_N(X), N \in \mathbb{N} \cup \{\infty\}$, the evaluation

$$\langle S_N(X), w \rangle := \int_s^t \int_s^{t_{n-1}} \dots \int_s^{t_1} dX^{a_1} dX^{a_2} \dots dX^{a_m}$$

for $w = a_1 \dots a_n \in \mathcal{A}^*, |w| \leq N$. With this notation in place, one can show (see Exercise 8.4.2) that the chain rule of ordinary calculus implies

$$\langle S_N(X)_{s,t}, w \rangle \langle S_N(X)_{s,t}, u \rangle = \langle S_N(X)_{s,t}, w \sqcup u \rangle, \quad \text{for all } w, u \in \mathcal{A}^*. \quad (8.15)$$

Thus in view of (8.15), we see that the chain rule induces some non-linear constraints to the signature. Indeed, these constraints are responsible for the signature to take its values in $G^N(\mathbb{R}^d)$ instead of the full tensor algebra. The role these identities play will become clearer once we change our point of view slightly. For this we identify the k th level of the signature $S_\infty(X)$ as a mapping into $(\mathbb{R}^d)^{\otimes k} \cong \mathbb{R}^{d^k}$ whose components are given by the maps $\Delta \rightarrow \langle S_\infty(X), w \rangle$ for all $w \in \mathcal{A}^*$ with $|w| = k$. Hence for every pair $(s, t) \in \Delta$, we can identify the signature with a functional

$$\langle S_\infty(X)_{s,t}, \cdot \rangle: \text{Sh}(\mathcal{A}) \rightarrow \mathbb{R}, \quad \mathcal{A}^* \ni w \mapsto \langle S_\infty(X)_{s,t}, w \rangle.$$

In Exercise 8.4.2 we will show that $\langle S_\infty(X)_{s,t}, \cdot \rangle$ is continuous, that is, it takes its values in the continuous dual $(\text{Sh}(\mathcal{A}))' \cong \prod_{w \in \mathcal{A}^*} \mathbb{R}$ (this follows from Exercise B.1.2 as the dual of a locally convex direct sum of copies of \mathbb{R} is the direct product; Meise and Vogt, 1997, Proposition 24.3). From this identification of the dual, one deduces at once that the signature gives rise to a continuous map $\langle S_\infty(X), \cdot \rangle: \Delta \rightarrow (\text{Sh}(\mathcal{A}))'$. Elements in the image of $\langle S_\infty(X)_{s,t}, \cdot \rangle$ map the algebra product to the multiplication in \mathbb{R} by (8.15), whence they are algebra morphisms.

Algebra morphisms from the shuffle algebra are called *characters* of the algebra $\text{Sh}(\mathcal{A})$. We have already mentioned that the shuffle algebra carries more structure and is indeed a Hopf algebra. A *Hopf algebra* A is an algebra which

$$\delta(a_1 a_2 \dots a_n) = \sum_{i=0}^n a_1 a_2 \dots a_i \otimes a_{i+1} \dots a_{i+2} \dots a_n, \quad a_1, \dots, a_n \in \mathcal{A}.$$

As the subspace of homogeneous elements of degree 0 is 1-dimensional, the graded bialgebra $\text{Sh}(\mathcal{A})$ becomes a Hopf algebra (see Manchon, 2008 for an introduction). This means that $\text{Sh}(\mathcal{A})$ admits an antipode, that is, a mapping $S: \text{Sh}(\mathcal{A}) \rightarrow \text{Sh}(\mathcal{A})$ connecting the algebra and coalgebra structures and given by

$$S(a_1 a_2 \dots a_n) = (-1)^n a_n a_{n-1} \dots a_1 \quad \text{for all } a_1, \dots, a_n \in \mathcal{A}.$$

is simultaneously a coalgebra $A \rightarrow A \otimes A$ such that the algebra and coalgebra structures are connected via the so-called antipode $S: A \rightarrow A$. We refer to Manchon (2008), for example, for the full definition and many examples. Now, for a Hopf algebra such as the shuffle algebra, the characters form a group $\mathcal{G}(\text{Sh}(\mathcal{A}), \mathbb{R})$ under the convolution product

$$\varphi \star \psi(w) := \varphi \otimes \psi(\delta(w)), \text{ where } \delta \text{ is the deconcatenation coproduct.}$$

Note that we exploit here $\mathbb{R} \otimes \mathbb{R} \cong \mathbb{R}$. Inversion in the character group is given by precomposition with the antipode, that is, $\varphi^{-1} = \varphi \circ S$ (see e.g. Bogfjellmo et al., 2016, Lemma 2.3). Summing up, we have just proved the following.

8.22 Lemma *For a smooth path $X: [0, 1] \rightarrow \mathbb{R}^d$, the signature induces a continuous map*

$$\langle S_\infty(X), \cdot \rangle: \Delta \rightarrow \mathcal{G}(\text{Sh}(\mathcal{A}), \mathbb{R}) \subseteq \text{Sh}(\mathcal{A})', \quad (s, t) \mapsto \langle S_\infty(X)_{s,t}, \cdot \rangle$$

with values in the character group of the Hopf algebra $\text{Sh}(\mathcal{A})$.

More generally, if we consider the k th level of an element

$$\mathbf{X} = (1, X^1, X^2, \dots) \in G^\infty(\mathbb{R}^d)$$

we see that $X^k \in (\mathbb{R}^d)^{\otimes k}$ can be written as

$$X^k = \sum_{w \in \mathcal{A}^*, w=a_1 \dots a_k} \alpha_{\mathbf{X}}(w) e_{a_1} \otimes \dots \otimes e_{a_k},$$

where e_{a_ℓ} is the standard basis vector labelled by $a_\ell \in \mathcal{A} = \{1, \dots, d\}$ and $\alpha_{\mathbf{X}}(w)$ some real coefficient. With some work on the algebra, one can prove the following.

8.23 Lemma (Ree, 1958, Theorem 2.6) *Let $\mathcal{A} = \{1, \dots, d\}$ and for $\mathbf{X} = (1, X^1, X^2, \dots)$ we write $X^k = \sum_{w \in \mathcal{A}^*, w=a_1 \dots a_k} \alpha_{\mathbf{X}}(w) e_{a_1} \otimes \dots \otimes e_{a_k}$. Then the following map is a well-defined group isomorphism:*

$$\Psi: G^\infty(\mathbb{R}^d) \rightarrow \mathcal{G}(\text{Sh}(\mathcal{A}), \mathbb{R}), \quad \mathbf{X} = (1, X^1, X^2, \dots) \mapsto \psi_{\mathbf{X}},$$

$$\text{where } \psi_{\mathbf{X}} \text{ is defined on } \mathcal{A}^* \text{ as } \psi_{\mathbf{X}}(w) = \alpha_{\mathbf{X}}(w) \text{ for all } w \in \mathcal{A}^*.$$

$$(8.16)$$

Moreover, a similar statement holds for the character group of the truncated shuffle algebra and the groups $G^N(\mathbb{R}^d), N \in \mathbb{N}$.

Now turning to the topological side, $G^\infty(\mathbb{R}^d)$ is a Lie group with respect to the subspace topology induced by the tensor algebra. As a locally convex space, the tensor algebra is isomorphic to a countable product of copies of the reals. We can also topologise the character group $\mathcal{G}(\text{Sh}(\mathcal{A}), \mathbb{R}) \subseteq \text{Sh}(\mathcal{A})' = \prod_{w \in \mathcal{A}^*} \mathbb{R}$ with the subspace topology. This topology even turns the character

group into an infinite-dimensional Lie group (Bogfjellmo et al., 2016, Theorem A). Thus we have the following additional statement.

8.24 Proposition *The group isomorphism (8.16) is an isomorphism of Lie groups. Thus there is a bijection between α -rough path \mathbf{X} with values in $G^\infty(\mathbb{R}^d)$ and continuous maps*

$$\Psi_{\mathbf{X}} : \Delta \rightarrow \mathcal{G}(\text{Sh}(\mathcal{A}), \mathbb{R}), \quad (t, s) \mapsto \psi_{\mathbf{X}_{s,t}}$$

such that for every $w \in \mathcal{A}^$, the component path $\Delta \rightarrow \mathbb{R}, (t, s) \mapsto \psi_{\mathbf{X}_{s,t}}(w)$ satisfies the ‘ $k\alpha$ -Hölder condition’*

$$\sup_{\substack{(s,t) \in \Delta \\ s \neq t}} \frac{|\psi_{\mathbf{X}_{s,t}}(w)|}{|t - s|^{k\alpha}} < \infty. \tag{8.17}$$

Proof In Exercise 8.4.3 you will show that (8.16) is the restriction of an isomorphism of locally convex spaces, hence an isomorphism of topological groups. Now by Proposition 8.7 the group $G^\infty(\mathbb{R}^d)$ is a Lie group and a submanifold of the tensor algebra. From Bogfjellmo et al. (2016, Theorem B) we know that also $\mathcal{G}(\text{Sh}(\mathcal{A}), \mathbb{R})$ is a Lie group and a submanifold of the dual space of the shuffle algebra. Hence we deduce from Lemma 1.39 that (8.16) is already smooth and thus an isomorphism of Lie groups.

Every α -rough path \mathbf{X} with values in $G^\infty(\mathbb{R}^d)$ is a continuous $G^\infty(\mathbb{R}^d)$ -valued map (since its components are continuous in the product topology $G^\infty(\mathbb{R}^d) \subseteq \mathcal{T}^\infty(\mathbb{R}^d)$) and the k th level satisfies a $k\alpha$ -Hölder condition. Hence $\Psi_{\mathbf{X}} := \Psi \circ \mathbf{X} : \Delta \rightarrow \mathcal{G}(\text{Sh}(\mathcal{A}), \mathbb{R})$ is continuous. Up to changing the Hölder constant, we can change the norm on $(\mathbb{R}^d)^{\otimes k} \cong \mathbb{R}^{d^k}$ to the maximum norm. This shows that the components $\Psi_{\mathbf{X}}(w)$ inherit any Hölder condition the k th level of \mathbf{X} satisfies, that is, (8.17) holds. Vice versa, if (8.17) holds for every $w \in \mathcal{A}^*$ with $|w| = k$, then the k th level of \mathbf{X} satisfies a $k\alpha$ -Hölder condition. This establishes the claimed bijection. □

Proposition 8.24 yields a nice interpretation of a (weakly geometric) rough path as a ‘multiplicative functional’ on the shuffle algebra. Of course this was not intended when T. Lyons originally coined the term ‘multiplicative functional’.

8.25 Remark We are finally able to give a complete and easy proof of Lemma 8.18: In view of Proposition 8.24, a rough path \mathbf{X} takes its values in $\mathcal{G}(\text{Sh}(\mathcal{A}), \mathbb{R})$ and the group structure is isomorphic to the one of $G^\infty(\mathbb{R}^d)$. Hence the pointwise inverse \mathbf{X}^{-1} can be computed via the group structure. Now it is a well-known fact about characters of a Hopf algebra that for a character ϕ , its inverse in the character group is given by $\phi \circ S$, where S is the antipode

of the Hopf algebra (Bogfjellmo et al., 2016, Lemma 2.3). Thus in the case of the shuffle Hopf algebra, this yields for an arbitrary linear combination of the words of length k the formula

$$S\left(\sum_{i_1, \dots, i_k \in \{1, \dots, d\}} a_{(i_1, \dots, i_k)} i_1 \cdots i_k\right) = \sum_{i_1, \dots, i_k \in \{1, \dots, d\}} (-1)^k a_{(i_1, \dots, i_k)} i_k \cdots i_1.$$

Now the k th component of \mathbf{X} corresponds to a linear combination of basis element $e_{i_1} \otimes e_{i_2} \otimes \cdots \otimes e_{i_k}$. The isomorphism $\mathcal{G}(\text{Sh}(A), \mathbb{R}) \cong G^\infty(\mathbb{R}^d)$ identifies $i_1 i_2 \dots i_k$ with $e_{i_1} \otimes \cdots \otimes e_{i_k}$, whence the inversion formula shows the k th component of the pointwise inverse \mathbf{X}^{-1} arises by permuting coefficients (and multiplying by $(-1)^k$). Hence the k th component of \mathbf{X}^{-1} is $k\alpha$ -Hölder if the k th component of \mathbf{X} is so.

Exercises

8.4.1 Let $\mathcal{A} = \{1, \dots, d\}$ be a finite alphabet and \mathcal{A}^* the monoid of words generated by \mathcal{A} . Recall that $\text{Sh}(\mathcal{A}) = \mathbb{R}\mathcal{A}^*$ is the vector space generated by \mathcal{A}^* . Show that:

- (a) for all words $(u \sqcup v) \sqcup w = u \sqcup (v \sqcup w)$, $u \sqcup v = v \sqcup u$ and $|u \sqcup v| = |u| + |v|$ hold;
- (b) one can bilinearly extend the shuffle product to $\sqcup: \text{Sh}(\mathcal{A}) \times \text{Sh}(\mathcal{A}) \rightarrow \text{Sh}(\mathcal{A})$;
- (c) the box-topology from Example A.36 makes $\sqcup: \text{Sh}(\mathcal{A}) \times \text{Sh}(\mathcal{A}) \rightarrow \text{Sh}(\mathcal{A})$ continuous and $(\text{Sh}(\mathcal{A}), \sqcup)$ a locally convex algebra.

Hint: By Exercise B.1.2 it suffices to show that for every pair of words v, w the shuffle induces a continuous map $\mathbb{R} \times \mathbb{R} \rightarrow \text{Sh}(\mathcal{A})$, $(a, b) \mapsto (av) \sqcup (bw)$.

8.4.2 Let $X: [0, 1] \rightarrow \mathbb{R}^d$ be a smooth path with signature $S_N(X)$ (for some $N \in \mathbb{N} \cup \{\infty\}$). Let $w = a_1 \cdots a_n \in \mathcal{A}^*$. Define $\langle S_N(X)_{s,t}, w \rangle = \int_s^t dX^{a_1} dX^{a_2} \cdots dX^{a_n}$. Show that:

- (a) (8.15) holds, that is, $\langle S_N(X)_{s,t}, w \rangle \langle S_N(X)_{s,t}, u \rangle = \langle S_N(X)_{s,t}, w \sqcup u \rangle$, for all $w, u \in \mathcal{A}^*$. *Hint:* Use induction by word length and the chain rule. For $|w| = 1$, the claim becomes (8.14).
- (b) For every pair $(s, t) \in \Delta$, the linear map $\langle S_\infty(X)_{s,t}, \cdot \rangle: \text{Sh}(\mathcal{A}) \rightarrow \mathbb{R}$ is continuous, where we endow the shuffle algebra with the box-topology from Example A.36.

(c) The signature gives rise to a continuous map $\langle S_\infty(X), \cdot \rangle: \Delta \rightarrow (\text{Sh}(\mathcal{A}))' \cong \prod_{w \in \mathcal{A}^*} \mathbb{R}$. *Hint:* The identification is given by $\text{Sh}(\mathcal{A}) \ni f \mapsto (f(w))_{w \in \mathcal{A}^*} \in \prod_{w \in \mathcal{A}^*} \mathbb{R}$.

8.4.3 Show that there is a canonical isomorphism $\text{Sh}(\mathcal{A})' = \prod_{n \in \mathbb{N}} \mathbb{R} \cong \mathcal{T}^\infty(\mathbb{R}^d)$ of locally convex spaces if $\mathcal{A} = \{1, \dots, d\}$. Deduce that this restricts to (8.16), whence (8.16) is a morphism of topological groups if we endow $G^\infty(\mathbb{R}^d)$ with its Lie group topology and $\mathcal{G}(\text{Sh}(\mathcal{A}), \mathbb{R})$ with the subspace topology induced by embedding it in $\text{Sh}(\mathcal{A})'$.

8.5 The Grand Geometric Picture (Rough Paths and Beyond)

In the last section we have seen that rough paths can be understood as certain continuous paths with values in the character group of the shuffle Hopf algebra. We will now extend the focus slightly to different types of rough paths. For this, recall that the signature of a smooth path had to satisfy certain algebraic identities which connected it to the shuffle algebra and its character group. For smooth paths this was a consequence of the chain rule of ordinary calculus. However, there are also many interesting objects, which do not satisfy a classical chain rule. One example for this is the Itô-integral from stochastic calculus (nevertheless we constructed the rough path Itô-lift of Brownian motion in Example 8.17). However, this motivates the idea of relaxing the requirement of (8.15) and recording iterated integrals of a path $X: [0, 1] \rightarrow \mathbb{R}^d$ of the form

$$\int_s^t \left(\int_s^r dX^i \right) \left(\int_s^r dX^j \right) dX^k. \tag{8.18}$$

For the rough paths we have considered so long, the iterated integral (8.18) (respectively the corresponding combination of levels of the rough path) simplifies via the shuffle identity to an iterated integral (resp. level of the rough path) we have already recorded. Relaxing the requirement (8.15) necessitates that we record objects corresponding to (8.18) as additional information. Following this approach, one arrives at Gubinelli’s concept of a *branched rough path* (Gubinelli, 2010). Branched rough paths satisfy different algebraic identities than the rough paths we have been considering (which are called *weakly geometric rough paths* in the literature (Friz and Hairer, 2020)). By dropping (8.15), branched rough paths will not take values in the character group of the shuffle algebra. However, this can easily be remedied by replacing the shuffle algebra with another Hopf algebra. This leads us to the following concept.

8.26 Definition Let \mathcal{H} be a graded and connected Hopf algebra⁸ and let $(\mathcal{G}(\mathcal{H}, \mathbb{R}), \star)$ be its character group. Fix a basis B of the vector space \mathcal{H} consisting of homogeneous elements. Then an α -rough path over \mathcal{H} is a map

$$\mathbf{X}: \Delta \rightarrow \mathcal{G}(\mathcal{H}, \mathbb{R})$$

which satisfies Chen’s relation

$$\mathbf{X}_{s,u} \star \mathbf{X}_{u,t} = \mathbf{X}_{s,t}, \quad s \leq u \leq t \in [0, 1],$$

and for every $h \in B$ we have the graded Hölder condition $|\langle \mathbf{X}_{s,t}, h \rangle| \lesssim |t - s|^{\alpha|h|}$ (where $\langle \cdot, h \rangle$ is evaluation in h). If $\dim \mathcal{H}_1 = d$ and $B \cap \mathcal{H}_1 = \{e_1, \dots, e_d\}$, an \mathcal{H} -rough path \mathbf{X} lifts $x: [0, 1] \rightarrow \mathbb{R}^d, x(t) = (x_1(t), \dots, x_d(t))$ if $\langle \mathbf{X}_{s,t}, e_i \rangle = x_i(t) - x_i(s), i = 1, \dots, d$ for all $s, t \in \Delta$.

Since the Hopf algebra \mathcal{H} is graded, one can also define truncated versions of the rough paths by looking at the character group of truncated Hopf algebra. Note that the choice of basis B is part of the definition of an \mathcal{H} -rough path and the graded Hölder condition will depend on the choice of grading and the base.

8.27 Example The concept of an α -rough path over \mathcal{H} is quite flexible and we illustrate this with the following list of different types of rough paths appearing in the literature:

- For $\mathcal{H} = \text{Sh}(\mathcal{A})$ and $B = \mathcal{A}^*$, we recover the notion of a (weakly) geometric rough path as discussed in §§8.3 and 8.4.
- If \mathcal{H} is the Butcher–Connes–Kreimer Hopf algebra of (decorated) rooted trees (Bogfjellmo et al., 2016, Example 4.6) and B the base of all forests, one obtains branched rough paths (Gubinelli, 2010).
- For the Munthe-Kaas–Wright Hopf algebra and B the base consisting of planar forests, one obtains the notion of a planarly branched rough path. These objects generalise the concept of a branched rough path to homogeneous spaces. See Curry et al. (2020) for more information.

While switching the Hopf algebra allows us to treat different concepts of rough paths, the general geometric theory stays the same for the different types of rough paths which can be treated in this generalised setting. For example, there is a generalised version of the Lyons–Victoire lifting theorem (see Tapia and Zambotti, 2020, Theorem 3.4) and the character groups can be endowed with an infinite-dimensional Lie group structure for which every rough path becomes a continuous map (see Bogfjellmo et al., 2016, Theorem A). Thus the

⁸ Graded and connected means that $\mathcal{H} = \bigoplus_{n \in \mathbb{N}_0} \mathcal{H}_n$ as vector spaces with $\mathcal{H}_0 \cong \mathbb{R}$ and the algebra and coalgebra structures are compatible with the grading; see Bogfjellmo et al. (2016). Recall that an element $x \in \mathcal{H}$ is called homogeneous of degree $|x| = n$ if $x \in \mathcal{H}_n$.

theory for weakly geometric rough paths from §8.4 carries over to the generalised concept of rough paths over \mathcal{H} . However, there are also new geometric features which appear in this generalised setting. For example, the following result hints at interesting geometry:

8.28 Proposition (Tapia and Zambotti, 2020, Theorem 1.2) *For every $\alpha \in]0, 1[$ with $\alpha^{-1} \notin \mathbb{N}$ the α -branched rough paths form a homogeneous space under the action of a vector space of α -Hölder functions.*

The term ‘homogeneous space’ in Proposition 8.28 means that there is a transitive and free group action $G \times X \rightarrow X$ of a vector space G of Hölder continuous functions on the space of branched rough paths X . The canonical manifold structures on the spaces and groups should turn the branched rough paths into an infinite-dimensional homogeneous space in the sense of differential geometry (see Example 3.55); however, this question was, to the best of my knowledge, not yet investigated. While there is a lot more which could be said about rough paths and their interplay with (finite- and infinite-dimensional) geometry we will not go into more detail here. Instead, let us briefly point out a general theme underlying the idea to identify rough paths as paths into the character group of a suitable Hopf algebra.

8.29 (Elements in the character group as formal power series) Assume that we have a graded Hopf algebra $\mathcal{H} = \bigoplus_{n \in \mathbb{N}_0} \mathcal{H}_n$ with B a vector space base of \mathcal{H} consisting only of homogeneous elements. Instead of thinking of an element in the (continuous) dual space $\mathcal{H}' = \prod_{n \in \mathbb{N}_0} \mathcal{H}'_n$ as a linear map, we can identify it with its values on the base B and write it as a *formal power series*

$$\psi = \sum_{w \in B} \langle \psi, w \rangle w, \quad \text{with } \langle \psi, w \rangle := \psi(w). \tag{8.19}$$

Now elements in the character group $\mathcal{G}(\mathcal{H}, \mathbb{R})$ are algebra morphisms. Hence it suffices to record their values on a set C which generates the algebra \mathcal{H} in the sense that \mathcal{H} is isomorphic (as an algebra) to the (not necessarily commutative) polynomial algebra $\mathbb{R}\langle C \rangle$. If we have chosen a set C generating \mathcal{H} in this sense, it suffices to record, instead of the formal power series (8.19), the series

$$\psi = \sum_{w \in C} \langle \psi, w \rangle w, \quad \text{if } \psi \in \mathcal{G}(\mathcal{H}, \mathbb{R}). \tag{8.20}$$

8.30 Example For the shuffle algebra $\text{Sh}(\mathcal{A})$, we can represent a functional as the formal power series

$$\psi = \sum_{w \in \mathcal{A}^*} \langle \psi, w \rangle w.$$

These series are also called *word series* and they are studied in relation to dynamical systems and numerical integration (Murua and Sanz-Serna, 2017).

Note that the shuffle algebra can be interpreted as a commutative polynomial algebra generated by the Lyndon words (Reutenauer, 1993, Chapter 5).

While 8.29 seems to be only a notational trick to write a functional as a series, this idea leads to a rich source of examples. The idea one should have when looking at the series expansions (8.19) and (8.20) is that these series represent some kind of Taylor expansion.⁹ It is then the role of the Hopf algebra \mathcal{H} to describe the combinatorics for the objects which define the Taylor expansion. Groups of (formal) power series modelled as characters on suitable Hopf algebras arise in a variety of contexts. Beyond rough paths here one should mention the following applications: Hairer’s regularity structures for stochastic partial differential equations (Friz and Hairer, 2020, Chapter 13), word series in numerical analysis (Murua and Sanz-Serna, 2017), Chen–Fliess series in control theory (Gray et al., 2022) and the Connes–Kreimer approach to the renormalisation of quantum field theories (Manchon, 2008). In these applications, the differential geometry of the character group is of interest. As a concrete example, recall the connection to numerical integrators.

8.31 Example Consider the (time-independent) ordinary differential equation

$$y'(t) = F(y), \text{ where } F: \mathbb{R}^d \rightarrow \mathbb{R}^d \text{ is a vector field.} \tag{8.21}$$

Assume that we are trying to compute a power series solution to (8.21). Then we compute derivatives of y via the chain rule:

$$y' = F(y), y'' = dF(y; F(y)), y''' = d^2F(y; F(y), F(y)) + dF(y; dF(y; F(y))), \dots$$

Fixing a starting point y_0 we only need iterated differentials of F and information on where derivatives were inserted in the arguments of the differential. The combinatorics can be handled by encoding the information via rooted trees, that is, as finite graphs with a special node called the root (displayed as the nethermost node in the following):



Formally, we write \emptyset for the empty tree. Every rooted tree τ can be written recursively as $\tau = [\tau_1, \dots, \tau_m]$, where the τ_i are trees whose roots are grafted to a common new root. For example, $[\bullet, \bullet] = \mathcal{V}$. Then the iterated differentials of F can recursively be encoded via the *elementary differentials* defined as $E_F(\bullet, y) := F(y)$, and for a rooted tree τ we set

⁹ For rough paths this is most easily seen in the context of controlled rough paths; see Friz and Hairer (2020, Section 4.6).

$$E_F(\tau)(y) := d^n F(y; E_F(\tau_1, y), \dots, E_F(\tau_n, y)) \text{ for } \tau = [\tau_1, \dots, \tau_n].$$

This leads to a formal power series, called a *B-series*

$$B_F(\psi, y, h) := y + \sum_{\tau \text{ rooted tree}} \frac{h^{|\tau|}}{\sigma(\tau)} \psi(\tau) E_F(\tau, y), \tag{8.22}$$

where $h \in \mathbb{R}, a(\tau) \in \mathbb{R}, |\tau|$ is the number of nodes in the tree τ and $\sigma(\tau)$ is a certain symmetry factor. B-series model certain numerical solutions to (8.21), where the parameter h is interpreted as the step size of the numerical method. One identifies the B-series (8.22) with its coefficients $\{\psi(\tau)\}_{\tau \text{ rooted tree}}$. As the rooted trees generate the Butcher–Connes–Kreimer Hopf algebra, ψ extends to a character of this Hopf algebra. The character group of the Butcher–Connes–Kreimer Hopf algebra is in numerical analysis known as the *Butcher group* (Bogfjellmo and Schmeding, 2017). Its elements encode numerical integration schemes such as Runge–Kutta methods and the group product models composition of schemes.

From the perspective of numerical analysis, the Butcher group is a convenient tool as its algebraic and differential geometric structure is of interest in the analysis of numerical integrators. Unfortunately, it is also a very large group with many elements which do not correspond to any (locally) convergent integration scheme. This leads to subgroups which admit a stronger topology turning them into Lie groups while still containing all elements of interest (Dahmen and Schmeding, 2020). In the context of rough paths, a similar problem is the question of whether there is a differentiable structure on the subgroup generated by all α -rough paths for some fixed $\alpha \in]0, 1[$.