Available online at www.sciencedirect.com

ScienceDirect

Journal of Differential Equations 340 (2022) 151–178

Journal of Differential Equations

www.elsevier.com/locate/jde

Geometric rough paths on infinite dimensional spaces

Erlend Grong [a,*,1], Torstein Nilssen [b], Alexander Schmeding [c]

[a] *University of Bergen, Department of Mathematics, P.O. Box 7803, 5020 Bergen, Norway*
[b] *Department of Mathematics, University of Agder, P.O. Box 422, 4604 Kristiansand, Norway*
[c] *FLU, Nord university, Høgskoleveien 27, 7601 Levanger, Norway*

Received 5 July 2021; revised 6 July 2022; accepted 26 August 2022

Abstract

Similar to ordinary differential equations, rough paths and rough differential equations can be formulated in a Banach space setting. For $\alpha \in (1/3, 1/2)$, we give criteria for when we can approximate Banach space-valued weakly geometric α-rough paths by signatures of curves of bounded variation, given some tuning of the Hölder parameter. We show that these criteria are satisfied for weakly geometric rough paths on Hilbert spaces. As an application, we obtain Wong-Zakai type result for function space valued martingales using the notion of (unbounded) rough drivers.

MSC: 22E65; 53C17; 60H10; 60L20; 60L50

Keywords: Rough paths on Banach spaces; Geometric rough paths; Wong-Zakai result for rough flows; Infinite-dimensional Lie groups; Carnot-Carathéodory geometry

1. Introduction

The theory of rough paths was invented by T. Lyons in his seminal article [33] and provides a fresh look at integration and differential equations driven by rough signals. A rough path consists

* Corresponding author.
E-mail addresses: erlend.grong@uib.no (E. Grong), torstein.nilssen@uia.no (T. Nilssen), alexander.schmeding@nord.no (A. Schmeding).

1 The first author is partially supported by the Trond Mohn Foundation - Grant TMS2021STG02 (GeoProCo).

https://doi.org/10.1016/j.jde.2022.08.034

of a Hölder continuous path in a vector space together with higher level information satisfying certain algebraic and analytical properties. The algebraic identities in turn allow one to conveniently formulate a rough path as a path in nilpotent groups of truncated tensor series, cf. [16] for a detailed account. Similar to the well-known theory of ordinary differential equations, it makes sense to formulate rough paths and rough differential equations with values in a Banach space, [31]. It is expected that the general theory carries over to this infinite-dimensional setting, yet a number of results which are elementary cornerstones of rough path theory are still unknown in the Banach setting.

In [6] the authors introduce the notion of a *rough driver*, which are vector fields with an irregular time-dependence. Rough drivers provide a somewhat generalized description of necessary conditions for the well-posedness of a rough differential equation and the authors use this for the construction of flows generated by these equations. The push-forward of the flow, at least formally, satisfies a (rough) partial differential equation, and this equation is studied rigorously in [4] where the authors introduce the notion of *unbounded rough drives*. This theory was further developed in [13,22,25] in the linear setting (although [13] also tackles the kinetic formulation of conservation laws) as well as nonlinear perturbations in [8,24,23,27,26]. Still, the unbounded rough drivers studied in these papers assume a factorization of time and space in the sense that the vector fields lies in the *algebraic tensor* of the time and space dependence.

Our main motivation for this paper is the observation in [10] that rough drivers can be constructed from rough paths taking values in the space of sufficiently smooth functions, see Section 3.2. Moreover, in [10] the authors needed unbounded rough drivers for which the factorization of time and space was *not* valid, and in particular approximating the unbounded rough driver by smooth drivers. In finite dimensions, sufficient conditions that guarantee the existence of smooth approximations can be easily checked and leads to the so-called *weakly geometric rough paths*. In [10] and ad-hoc method was introduced to tackle the lack of a similar result in infinite dimensions. For other papers dealing with infinite-dimensional rough paths, let us also mention [11,2,7].

In the present paper we address the characterization of weakly geometric rough paths in Banach spaces. Our aims are twofold. Firstly, we describe and develop the infinite-dimensional geometric framework for Banach space-valued rough paths and weakly geometric rough paths. These rough paths take their values in infinite-dimensional groups of truncated tensor products. Some care needs to be taken in this setting, as the tensor product of two Banach spaces will depend on choice of norm on the product. Secondly, we characterize the geometric rough paths that take their values in an Hilbert space and their relationship to weakly geometric rough paths. Our main result is to prove the following well-known relationship for finite dimensional rough paths in an infinite dimensional setting. Recall that for $\alpha \in (1/3, 1/2)$, a geometric α-rough path is an element of the closure in signatures $S^2(x)_{st} = 1 + x_t - x_s + \int_s^t (x_r - x_s) \otimes dx_r$ of curves x_t of bounded variation, while an α-rough path $\mathbf{x}_{st} = 1 + x_{st} + x_{st}^{(2)}$ is called weakly geometric if the symmetric part of $x_{st}^{(2)}$ equals $\frac{1}{2} x_{st} \otimes x_{st}$; a property that holds for all geometric rough paths in particular by an integration by parts argument. Our main result is the following.

Theorem 1.1. *For $\alpha \in (1/3, 1/2)$, let $\mathscr{C}_g^\alpha([0,T], G^2(E))$ and $\mathscr{C}_{wg}^\alpha([0,T], G^2(E))$ denote respectively geometric rough paths and weakly geometric rough paths in a Hilbert space E, defined on the interval $[0,T]$ and relative to the Schatten p-norm, $1 \leq p \leq \infty$ on $E \otimes E$. Then for any $\beta \in (1/3, \alpha)$, we have inclusions*

$$\mathscr{C}_g^\alpha([0,T], G^2(E)) \subset \mathscr{C}_{wg}^\alpha([0,T], G^2(E)) \subset \mathscr{C}_g^\beta([0,T], G^2(E)).$$

We emphasize that this result includes the Hilbert-Schmidt norm, projective tensor norm and injective tensor norm as respectively p equal to 2, 1 and ∞.

The structure of the paper is as follows. In Section 2 we review the infinite-dimensional framework for rough paths with values in Banach spaces. In particular, we discuss good conditions for norms on tensor products and establish that our free nilpotent groups have L^1-regularity in Remark 2.7. We continue with a presentation of Banach space-valued α-rough paths for $\alpha \in (\frac{1}{3}, \frac{1}{2})$ in Section 3. This leads to the three prerequisite assumptions in Theorem 3.3 which states when weakly geometric rough paths can be approximated by signatures of bounded variation path after some tuning of the Hölder parameter. In Section 3.2, we apply Theorem 1.1 to prove Wong-Zakai type results for rough flows; a rough generalization of flows of time-dependent vector fields. This yields a concrete application for rough paths on infinite dimensional space.

The remainder of the paper is dedicated to proving Theorem 1.1 by showing that the criteria of Theorem 3.3 are indeed satisfied in the Hilbert space setting. All of these criteria depends on considering Carnot-Carathéodory geometry or sub-Riemannian geometry of our infinite dimensional groups. Section 4 first establishes the necessary prerequisite results from finite dimensional Hilbert spaces. In particular, we are concerned with formulas and results that are dimension-independent. We then do the proof of Theorem 1.1 in several steps, including a result in Theorem 4.6 where we prove that the Carnot-Carathéodory (CC) metric on the free step 2 nilpotent group generated by a Hilbert space becomes a geodesic distance when restricted to the subset of finite distance from the identity. We emphasize that this is a proper subset as the CC-metric is not Lipschitz-equivalent to the usual homogeneous distance defined by the Banach norms for infinite dimensional spaces. We conclude the proof of Theorem 1.1 in Section 4.5.

2. The infinite-dimensional framework for rough paths

2.1. Tensor products of Banach spaces

If E and F are two Banach spaces, we write $E \otimes_a F$ for their algebraic tensor product. We use the convention that $E^{\otimes_a 0} = \mathbb{R}$. For any $k \geq 0$ we endow the k-fold algebraic tensor product $E^{\otimes_a k}$ with a family of norms $\|\cdot\|_k$ satisfying the following conditions, cf. [5].

1. For every $a \in E^{\otimes_a k}, b \in E^{\otimes_a \ell}$, we have

$$\|a \otimes b\|_{k+\ell} \leq \|a\|_k \cdot \|b\|_\ell .$$

2. For any permutation σ of the integers $1, 2 \dots k$ and for any $x_1, \dots, x_k \in E$,

$$\|x_1 \otimes x_2 \otimes \cdots \otimes x_k\|_k = \|x_{\sigma(1)} \otimes \cdots \otimes x_{\sigma(k)}\|_k .$$

Inductively, for $k, \ell \in \mathbb{N}$ we define the spaces $E^{\otimes k} \otimes E^{\otimes \ell}$ as the completion of $E^{\otimes k} \otimes_a E^{\otimes \ell}$ with respect to the norm $\|\cdot\|_{k+\ell}$. From the inclusions

$$E^{\otimes_a (k+\ell)} \subseteq E^{\otimes k} \otimes_a E^{\otimes \ell} \subseteq E^{\otimes (k+\ell)}$$

it follows that $E^{\otimes k} \otimes E^{\otimes \ell} \cong E^{\otimes (k+\ell)}$ as Banach spaces.

Example 2.1. The projective tensor product of Banach spaces is the completion of the algebraic tensor product with respect to the projective tensor norm

$$\|z\|_\pi := \inf\left\{\sum_{i=1}^n \|x_i\|_E \|y_i\|_F \,:\, z = \sum_{i=1}^n x_i \otimes y_i\right\}.$$

It is well known that the projective tensor norm satisfies properties 1. and 2. since it is a reasonable crossnorm on $E \otimes_a F$ (cf. [40, Section 6]). Similarly, the injective tensor norm, defined by

$$\|z\|_\epsilon = \sup\left\{\left|\sum_{i=1}^n \varphi(x_i)\psi(y_i)\right| \,:\, \varphi \in E^*, \psi \in F^*, \|\varphi\| = \|\psi\| = 1, z = \sum_{i=1}^n x_i \otimes y_i\right\},$$

satisfies 1. and 2. Its completion is the injective tensor product [40, Section 3].

If E is a Hilbert space, then we can identify $E \otimes_a E$ with finite rank operators from E to itself. In this case, the projective and injective norm of $z\colon E \to E$ correspond respectively to the trace norm and the operator norm. Moreover, this identification allows one to identify the projective tensor as the space of nuclear operators $\mathcal{N}(E, E)$ and the injective tensor product as the space of compact operators $\mathcal{K}(E, E)$, see [40, Corollary 4.8 and Corollary 4.13] for details.

2.2. *Algebra of truncated tensor series*

For $N \in \mathbb{N}_0$, we define

$$\mathcal{A}_N := \prod_{k=0}^{N} E^{\otimes k}$$

as the *space of (truncated) formal tensor series of E.*[2] Elements in $\mathcal{A}_N$ will be denoted as sequences $(x^{(k)})_{k\le N}$. A sequence concentrated in the k-th factor $E^{\otimes k}$ is called *homogeneous of degree k*. The set $\mathcal{A}_N$ is an algebra with respect to degree wise addition and the multiplication

$$(x^{(k)})_{k\le N} \cdot (y^{(k)})_{k\le N} := \left(\sum_{n+m=k} x^{(n)} \otimes y^{(m)}\right)_{k\le N}.$$

The algebras $\mathcal{A}_N$ turn out to be Banach algebras. We summarize the relevant results in the following.

Lemma 2.2. *The algebra $\mathcal{A}_N$ is a Banach algebra for $N < \infty$. Moreover, its group of units $\mathcal{A}_N^\times$ is a C^0-regular infinite-dimensional Lie group for any $N \in \mathbb{N}_0$.*

We recall the notion of regularity of a Lie group G. For this it is necessary to endow the occuring function spaces with a topology. This topology is not Banach space topology (but still a completely metrizable topological vector space) and we have to adopt a notion of smoothness

[2] One can also consider the algebra $\mathcal{A}_\infty$ arising as an infinite product, but this algebra is not a Banach space and we have no need for this generality in the present paper (but see the arXiv version of this paper or [41, chapter 8] for more information).

which is called Bastiani calculus. This means that we require the existence and continuity of directional derivatives, see [19,28] for more information. Let now 1 denote the group's identity element and $\mathbf{L}(G)$ its Lie algebra. Then G is called *C^r-regular*, $r \in \mathbb{N}_0 \cup \{\infty\}$, if for each C^r-curve $u\colon [0,1] \to \mathbf{L}(G)$ the initial value problem

$$\dot{\gamma}(t) = \gamma(t) \cdot u(t) \qquad \gamma(0) = 1$$

has a (necessarily unique) C^{r+1}-solution $\mathrm{Evol}(u) := \gamma\colon [0,1] \to G$ and the map

$$\mathrm{evol}\colon C^r([0,1], \mathbf{L}(G)) \to G, \quad u \mapsto \mathrm{Evol}(u)(1)$$

is smooth. A C^∞-regular Lie group G is called *regular (in the sense of Milnor)*. Every Banach Lie group is C^0-regular (cf. [35]). Several important results in infinite-dimensional Lie theory are only available for regular Lie groups, cf. [29].

Proof of Lemma 2.2. By construction of the algebra structure we have for elements of degree k and ℓ that $E^{\otimes k} \cdot E^{\otimes \ell} \subseteq E^{\otimes (k+\ell)}$. Hence the choice of tensor norms in section 2.1 shows that $\mathcal{A}_N$ is a Banach algebra, its unit group $\mathcal{A}_N^\times$ is an open subset of $\mathcal{A}_N$. Following [18,21] the submanifold structure turns $\mathcal{A}_N^\times$ into a C^0-regular Banach Lie group. $\square$

Remark 2.3. The unit group $\mathcal{A}_N^\times$ of $\mathcal{A}_N$ is even a real analytic Lie group in the sense that the group operations extend analytically to the complexification.

2.3. *Exponential map*

Define the canonical projection $\pi_0^N\colon \mathcal{A}_N \to \mathbb{R} = \mathcal{A}_0$ and the closed ideal $\mathcal{I}_{\mathcal{A}_N} := \ker \pi_0^N = \prod_{0<k\le N} E^{\otimes k}$. Related to this ideal, we consider the following maps.

Lemma 2.4 *(Exponential and logarithm). The exponential and logarithm series*

$$\exp_N\colon \mathcal{I}_{\mathcal{A}_N} \to 1 + \mathcal{I}_{\mathcal{A}_N}, \qquad X \mapsto \sum_{0\le n\le N} \frac{X^{\otimes n}}{n!},$$

$$\log_N\colon 1 + \mathcal{I}_{\mathcal{A}_N} \to \mathcal{I}_{\mathcal{A}_N}, \qquad 1 + Y \mapsto \sum_{0\le n\le N} (-1)^{n+1} \frac{Y^{\otimes n}}{n},$$

yield mutually inverse real analytic isomorphisms.

Proof. It suffices to note that due to the truncated multiplication, the series are given by polynomials (which are analytic mappings). That they are mutually invers follows via the familiar argument for the exponential and logarithm series $\square$

Remark 2.5. Due to [18, Theorem 5.6] the Lie group exponential of $\mathcal{A}_N^\times$ is

$$\exp_{\mathcal{A}_N}\colon \mathcal{A}_N = \mathbf{L}(\mathcal{A}_N^\times) \to \mathcal{A}_N^\times, \quad x \mapsto \sum_{n\in\mathbb{N}_0} \frac{x^{\otimes n}}{n!}.$$

2.4. Free nilpotent groups

Using the exponential map, we are ready to define the subgroups of $\mathcal{A}_N^\times$ we are interested in. Observe that $\mathcal{A}_N = \mathbf{L}(\mathcal{A}_N^\times)$ is a Lie algebra with respect to the commutator bracket $[x, y] := x \otimes y - y \otimes x$. We define inductively the space $\mathcal{P}_a^n(E)$ of *Lie polynomials* over E of degree $n \in \mathbb{N}$ by $\mathcal{P}_a^1(E) := E$ and

$$\mathcal{P}_a^{n+1}(E) := \mathcal{P}_a^n(E) + \operatorname{span}\{[x, y] \mid x \in \mathcal{P}_a^n(E), y \in E\} \subseteq \mathcal{A}_{n+1}$$

The set of all Lie polynomials or Lie series is a Lie subalgebra of $(\mathcal{A}_N, [\cdot,\cdot])$, [38, Chapter 1.2]. Since $\mathcal{A}_N$ is a topological Lie algebra, we see that also $\mathcal{P}^N(E) := \overline{\mathcal{P}_a^N(E)}$ is a closed Lie subalgebra of $(\mathcal{A}_N, [\cdot,\cdot])$. Due to [38, Theorem 1.4], we have $\mathcal{P}^N(E) \subseteq \mathcal{I}_{\mathcal{A}_N}$. Hence we can apply Lemma 2.4 and [38, Corollary 3.3] to see that the set

$$\mathrm{G^N}(E) := \exp_{\mathcal{A}_N}(\mathcal{P}^N(E)) = \overline{\exp_{\mathcal{A}_N}(\mathcal{P}^N(E))},$$

forms a closed subgroup of $\mathcal{A}_N^\times$. Closed subgroups of Banach Lie groups are in general not Lie subgroups [35, Remark IV.3.17]. So indeed the next proposition is non-trivial.

Proposition 2.6. *The group* $\mathrm{G^N}(E)$ *is a closed submanifold of* $\mathcal{A}_N$ *and this structure turns it into a Banach Lie group. Moreover,* $\mathrm{G^N}(E)$ *is a* C^0*-regular Lie group and the exponential map* $\exp\colon \mathcal{P}^N(E) \to \mathrm{G^N}(E)$ *is a diffeomorphism.*

Observe that the group $\mathrm{G^N}(E)$ is a nilpotent group of step N generated by E.

Proof. The group $\mathrm{G^N}(E)$ is a closed subgroup of the locally exponential Lie group $\mathcal{A}_N^\times$. Due to Remark 2.5, the Lie group exponential of this group is $\exp_{\mathcal{A}_N}$. Define

$$\mathbf{L}^{(N)} := \{x \in \mathcal{I}_{\mathcal{A}_N} \subseteq \mathbf{L}(\mathcal{A}_N^\times) \mid \exp_{\mathcal{A}_N}(\mathbb{R}x) \subseteq \mathrm{G^N}(E)\}.$$

Due to construction of the closed Lie subalgebra $\mathcal{P}^N(E)$, we have $\mathcal{P}^N(E) \subseteq \mathbf{L}^{(N)}$. Conversely as $\mathcal{P}^N(E) \subseteq \mathcal{I}_{\mathcal{A}_N}$ and $\mathrm{G^N}(E) \subseteq 1 + \mathcal{I}_{\mathcal{A}_N}$, we deduce from Lemma 2.4 that also $\mathbf{L}^{(N)} \subseteq \mathcal{P}^N(E)$ holds, hence the two sets coincide. It follows that $\mathrm{G^N}(E)$ is a locally exponential Lie subgroup of $\mathcal{A}_N^\times$ by [35, Theorem IV.3.3].

Since $\mathcal{P}^N(E) \subseteq \mathcal{I}_{\mathcal{A}_N}$, $\mathrm{G^N}(E) \subseteq 1 + \mathcal{I}_{\mathcal{A}_N}$ and the exponential $\exp_{\mathcal{A}_N}$ is a diffeomorphism between those sets (Lemma 2.4), the Lie group exponential induces a diffeomorphism between Lie algebra and Lie group as $\exp = \exp_{\mathcal{A}_N}|_{\mathcal{P}^N(E)}^{\mathrm{G^N}(E)}$ due to [35, Theorem IV.3.3]. As all Banach Lie groups are C^0-regular, so are the $\mathrm{G^N}(E)$, cf. also Remark 2.7 below. □

Remark 2.7. The regularity of the Lie groups $\mathrm{G^N}(E)$ can be strengthened by weakening the requirements on the curves in the Lie algebra. This results in a notion of L^p-regularity [20] for infinite-dimensional Lie groups. One can show that Banach Lie groups such as $\mathrm{G^N}(E)$ are L^1-regular. Furthermore, as in the proof of Proposition 2.6, one sees that the limit $G^\infty(E)$ is L^1-regular. Note that L^1-regularity implies all other known types of measurable regularity for Lie groups.

Example 2.8 *(Step 2).* For the remainder of the paper, we will mostly focus on the special case of $N = 2$. In this case $\mathcal{P}^2(E)$ is the closure in $\mathcal{A}_2$ of sums of elements X, $Y \wedge Z = Y \otimes Z - Z \otimes Y$ with $X, Y, Z \in E$ and Lie brackets

$$[X + \mathbb{X}, Y + \mathbb{Y}] = X \wedge Y, \qquad X, Y \in E, \mathbb{X}, \mathbb{Y} \in \mathcal{P}^2(E) \cap E^{\otimes 2}.$$

3. Applications to infinite dimensional rough paths

3.1. Rough paths and geometric rough paths in Banach space

Let us first recall the notion of a Banach-space valued rough path, see e.g. [7]. The definition of a rough path involves higher level components with values in a completed tensor product.

Definition 3.1. Fix $\alpha \in (\frac{1}{3}, \frac{1}{2})$ and a tensor product completion $E \otimes E$ by a choice of a tensornorm $\|\cdot\|_\otimes$ satisfying the assumptions from Section 2.1. An $(E, \otimes)$-valued *α-rough path* consists of a pair $(x, x^{(2)})$

$$x : [0, T] \to E, \qquad x^{(2)} : [0, T]^2 \to E^{\otimes 2} = E \otimes E$$

where x is an α-Hölder continuous path and $x^{(2)}$ is "twice Hölder continuous", i.e.

$$\|x_t - x_s\| \lesssim |t - s|^\alpha, \qquad \|x^{(2)}_{st}\|_2 \lesssim |t - s|^{2\alpha}. \tag{3.1}$$

In addition, we require

$$x^{(2)}_{st} - x^{(2)}_{su} - x^{(2)}_{ut} = (x_u - x_s) \otimes (x_t - x_u) \tag{3.2}$$

usually called Chen's relation. The set of rough paths equipped with the metric induced by (3.1) is denoted $\mathscr{C}^\alpha([0, T], G^2(E))$.

To be more precise about this distance, we write $\mathbf{x}_{st} = 1 + x_t - x_s + x^{(2)}_{st}$ in $\mathcal{A}_2$, the two step-truncated tensor algebra over E. Chen relation (3.2) can then be rewritten as $\mathbf{x}_{st} = \mathbf{x}_{su}\mathbf{x}_{ut}$. Introduce a metric d on $1 + \mathcal{I}_N = \{\mathbf{x} = 1 + x + x^{(2)} : x \in E, x^{(2)} \in E \otimes E\}$, by

$$|\mathbf{x}| = \max\{\|x\|, \|x\|_\otimes^{1/2}\},$$
$$d(\mathbf{x}, \mathbf{y}) = |\mathbf{x}^{-1} \cdot \mathbf{y}| = |(1 + x + x^{(2)})^{-1} \cdot (1 + y + y^{(2)})|.$$

We then define the distance between two α-rough paths $(s, t) \mapsto \mathbf{x}_{st}, \mathbf{y}_{st}$ on $[0, T]^2$ as

$$d_\alpha(\mathbf{x}, \mathbf{y}) = \sup_{0 \le s < t \le T} \frac{d(\mathbf{x}_{st}, \mathbf{y}_{st})}{|t - s|^\alpha}. \tag{3.3}$$

Rephrasing these properties, we can define $\mathbf{x}_t := \mathbf{x}_{0t} = 1 + x_t + x^{(2)}_{0t} = 1 + x_t + x^{(2)}_t$ and regard $t \mapsto \mathbf{x}_t$ as a α-Hölder continuous path with values in $\mathcal{A}_2$. The relations (3.2) tell us that $\mathbf{x}_{st} = \mathbf{x}_s^{-1}\mathbf{x}_t$ and we have the identification $\mathscr{C}^\alpha([0, T], G^2(E)) \simeq C^\alpha([0, T], 1 + \mathcal{I}_N)$.

If x_t is a smooth path in E, then we can lift it to a rough path $\mathbf{x}_t = 1 + x_t + x_t^{(2)}$, where $x_{st}^{(2)} = \int_s^t (x_r - x_s) \otimes dx_r$. Using integration by parts,

$$\int_s^t (x_r - x_s) \otimes dx_r + \int_s^t dx_r \otimes (x_r - x_s) = (x_t - x_s) \otimes (x_t - x_s), \tag{3.4}$$

that is, the symmetric part of $x_{st}^{(2)}$ is $(x_t - x_s) \otimes (x_t - x_s)$. This algebraic condition is equivalent to $\mathbf{x}_t$ taking values in $G^2(E)$. We note that $\log_2(\mathbf{x}_{st}) = x_t - x_s + \frac{1}{2}\int_s^t (x_r - x_s) \wedge dx_r$.

Definition 3.2 *(Weakly geometric and geometric rough paths).* We say that α-rough path $\mathbf{x}_t$ is *weakly geometric* if it takes values in $G^2(E)$. The set of weakly geometric rough paths can again can be given the structure of a metric space $\mathscr{C}_{wg}^{\alpha}([0,T], G^2(E))$ with the metric d_α as in (3.3) and can be identified with $C^\alpha([0,T], G^2(E))$.

The space of *geometric rough paths* is defined as the closure in the rough path topology of the set canonical lift of smooth paths and is denoted $\mathscr{C}_g^{\alpha}([0,T], G^2(E))$.

Since (3.4) is stable under limits, we get that the set of geometric rough paths can be regarded as a subspace of $C^\alpha([0,T], G^2(E))$. The reversed question, namely if any $\mathbf{x} \in C^\alpha([0,T], G^2(E))$ can be approximated by a sequence of smooth paths is answered positively modulo some tuning of the Hölder parameter α given the following conditions.

We recall the definition of *the Carnot-Caratheodory metric*, which we will often abbreviate as the CC-metric. We define this metric ρ on $G^2(E)$ by $\rho(\mathbf{y}, \mathbf{z}) = \rho(1, \mathbf{y}^{-1} \cdot \mathbf{z})$ and

$$\rho(1, \mathbf{y}) = \inf \left\{ \int_0^T \|\dot{x}_t\|\, dt \; : \; \begin{matrix} x \in C([0,T],E), x_0 = 0,\, x_t \text{ has bounded variation} \\ \mathbf{y} = S^2(x)_T := 1 + x_T + \int_0^T x_t \otimes dx_t \end{matrix} \right\}.$$

Theorem 3.3. *Write*

$$M_{cc} = \{\mathbf{z} \in G^2(E) \, : \, \rho(1, \mathbf{z}) < \infty\},$$

and $C([0,T], M_{cc})$ for the space of continuous curves in M_{cc} with respect to ρ.

Let $\alpha \in (\frac{1}{3}, \frac{1}{2})$ be given and let $\beta \in (\frac{1}{3}, \alpha)$ be arbitrary. Assume that the following conditions are satisfied.

- *(I) For some $C > 0$ and any $\mathbf{z} \in G^2(E)$, we have $d(1, \mathbf{z}) \leq C\rho(1, \mathbf{z})$.*
- *(II) The metric space (M_{cc}, ρ) is a complete, geodesic space.*
- *(III) The set*

$$C^\alpha([0,T], G^2(E)) \cap C([0,T], M_{cc}),$$

is dense in $C^\alpha([0,T], G^2(E))$ relative to the metric d_β.

Then for any $\mathbf{x} \in C^\alpha([0,T], G^{(2)}(E))$ there exists a sequence of bounded variation paths $x^n : [0,T] \to E$ such that

$$\mathbf{x}^n = S^2(x^n) \to \mathbf{x} \textit{ in } \mathscr{C}^\beta([0,T], E).$$

In particular, we have the inclusions

$$\mathscr{C}_g^\alpha([0,T], G^2(E)) \subset C^\alpha([0,T], G^2(E)) \subset \mathscr{C}_g^\beta([0,T], E).$$

To explain condition (II) in more details, recall that if (M, ρ) is a metric space, then a curve $\gamma\colon [0,T] \to M$ is said to have *constant speed* if $\mathrm{Length}(\gamma|_{[s,t]}) = c|t-s|$ for any $0 \le s \le t \le T$ and some $c \ge 0$. A constant speed curve is a *geodesic* if $\mathrm{Length}(\gamma|_{[s,t]}) = \rho(\gamma(s), \gamma(t)) = |t-s|\rho(\gamma(0), \gamma(T))$. The metric space (M, ρ) is called *geodesic* if any pair of points can be connected by a geodesic.

If E is finite dimensional, the assumptions (I), (II) and (III) hold as ρ and d are then equivalent and we have access to the Hopf-Rinow theorem, see e.g. [16]. If E is a general Hilbert space, the Hopf-Rinow theorem is no longer available [14]. We will also show that the metrics ρ and d will not be equivalent in the infinite dimensional case, yet assumptions (I), (II) and (III) will be satisfied, giving us the result in Theorem 1.1. We will prove this statement in Section 4, finishing the proof in Section 4.5.

Proof of Theorem 3.3. We first consider the case when $\mathbf{x} \in C^\alpha([0,T], G^2(E))\} \cap C([0,T], M_{cc})$. As (M_{cc}, ρ) is a geodesic space, [17, Lemma 5.21] implies that there exists a sequence of truncated signatures $\mathbf{x}^n = S^2(x^n)\colon [0,T] \to M_{cc}$ of bounded variation paths x^n such that

$$\sup_{t\in[0,T]} \rho(\mathbf{x}_t, \mathbf{x}_t^n) \to 0, \qquad \text{for } n \to \infty,$$

and we have the uniform bound $\sup_n d(1, \mathbf{x}_{st}^n) \le C|t-s|^\alpha$. From (I), we conclude that $\mathbf{x}^n$ converges to $\mathbf{x}$ in $C([0,T], G^{(2)}(E))$. To show the stronger convergence in $C^\beta([0,T], G^2(E))$ we perform a classical interpolation argument. Since d is left invariant we see that

$$\begin{aligned} d(\mathbf{x}_{st}^n, \mathbf{x}_{st}) &\le d((\mathbf{x}_s^n)^{-1}\mathbf{x}_t^n, (\mathbf{x}_s)^{-1}\mathbf{x}_t^n) + d((\mathbf{x}_s)^{-1}\mathbf{x}_t^n, (\mathbf{x}_s)^{-1}\mathbf{x}_t) \\ &\le 2 \sup_{t\in[0,T]} d(\mathbf{x}_t^n, \mathbf{x}_t) \le 2C \sup_{t\in[0,T]} \rho(\mathbf{x}_t^n, \mathbf{x}_t), \end{aligned}$$

so that there exists a sequence of real numbers $\varepsilon_n \to 0$ with

$$d(\mathbf{x}_{st}^n, \mathbf{x}_{st}) \le \varepsilon_n.$$

From the construction of $\mathbf{x}^n$ we have $d(1, \mathbf{x}_{st}^n), d(1, \mathbf{x}_{st}) \le C|t-s|^\alpha$. Using the interpolation $\min\{a, b\} \le a^\theta b^{1-\theta}$ for every $a, b \ge 0$ and $\theta \in [0,1]$ we have

$$d(\mathbf{x}_{st}^n, \mathbf{x}_{st}) \le \varepsilon_n \wedge C|t-s|^\alpha \le \varepsilon_n^\theta C^{1-\theta}|t-s|^{\alpha(1-\theta)}$$

and by choosing θ such that $\alpha(1-\theta) = \beta$ we get convergence

$$d_\beta(\mathbf{x}^n, \mathbf{x}) = \sup_{s,t\in[0,T]} \frac{d(\mathbf{x}_{st}^n, \mathbf{x}_{st})}{|t-s|^\beta} \le \varepsilon_n^\theta C^{1-\theta} \to 0, \qquad n \to \infty.$$

Finally, from the density of $C^\alpha([0,T],G^2(E)) \cap C([0,T],M_{cc})$ by (III) it follows that if $\mathbf{x}^m \in C^\alpha([0,T],G^2(E)) \cap C([0,T],M_{cc})$ is a sequence converging to an arbitrary $\mathbf{x} \in C^\alpha([0,T],G^2(E))$ with respect to d_β, and $\mathbf{x}^{n,m}$ is a sequence of truncated signatures of bounded variation curves converging to $\mathbf{x}^m$, then $\mathbf{x}^{m,m}$ converge to $\mathbf{x}$. This completes the proof. □

3.2. Wong-Zakai for stochastic flows

As an application of Theorem 3.3 and Theorem 1.1 we prove an approximation result for martingales with values in a Banach space of sufficiently smooth functions, as systematically explored in [30]. We note that the approximation is in general not of Wong-Zakai type since we are not constructing piecewise linear interpolation of the noise. Rather, existence of the approximation follows from our general result Theorem 1.1. Let $(f_k)_{k=0}^K$ be a collection of time-dependent vector fields $f_k : [0,T] \times \mathbb{R}^d \to \mathbb{R}^d$ of class $C_b^p(\mathbb{R}^d,\mathbb{R}^d)$ in the x-variable for some p to be determined later, and let $(\omega_t)_{t\in[0,T]}$ be a K-dimensional Brownian motion on some filtered probability space $(\Omega,\mathcal{F},\mathbb{P})$. The study of the Stratonovich equation (for notational convenience we write $\omega_t^0 = t$)

$$dy_t = \sum_{k=0}^{K} f_k(t,y_t) \circ d\omega_t^k \tag{3.5}$$

is by now classical. The book [30] stresses the importance of considering the $C_b^p(\mathbb{R}^d,\mathbb{R}^d)$-valued semi-martingale

$$m_t(\xi) := \sum_{k=0}^{K} \int_0^t f_k(r,\xi) d\omega_r^k \tag{3.6}$$

which allows for a one-to-one characterization of stochastic flows (see [30] for precise statement and result). Equation (3.5) is then understood as $dy_t = m_{\circ dt}(y_t)$.

Consider now the tensor product on $C_b^p(\mathbb{R}^d,\mathbb{R}^d)$,

$$(f \otimes g)(\xi,\zeta) := f(\xi)g(\zeta)^T,$$

which allows us to identify $C_b^p(\mathbb{R}^d,\mathbb{R}^d)^{\otimes 2}$ with a subspace of $C_b^p(\mathbb{R}^d \times \mathbb{R}^d,\mathbb{R}^{d\times d})$. Let us define the iterated integral

$$\begin{aligned} m_{st}^{(2)}(\xi,\zeta) &:= \int_s^t (m_r - m_s) \otimes \circ dm_r(\xi,\zeta) \\ &:= \sum_{k,l=0}^{K} \int_s^t \int_s^r f_l(v,\xi) f_k(r,\zeta)^T d\omega_v^l \circ d\omega_r^k, \end{aligned} \tag{3.7}$$

as a $C_b^p(\mathbb{R}^d \times \mathbb{R}^d,\mathbb{R}^{d\times d})$-valued random field. Checking the symmetry condition then boils down to checking (3.4) for this tensor product. We have, for $\mu,\nu \in \{1,\dots,d\}$

$$\begin{aligned}
&m_{st}^{(2),\mu,\nu}(\xi,\zeta)+m_{st}^{(2),\nu,\mu}(\zeta,\xi)\\
&=\int_s^t (m_r^\mu(\xi)-m_s^\mu(\xi))\circ dm_r^\nu(\zeta)+\int_s^t (m_r^\nu(\zeta)-m_s^\nu(\zeta))\circ dm_r^\mu(\xi)\\
&=(m_t^\mu(\xi)-m_s^\mu(\xi))(m_t^\nu(\zeta)-m_s^\nu(\zeta))
\end{aligned}\tag{3.8}$$

by the well-known integration by parts formula for the Stratonovich integral. We note that the particular decomposition of (3.6) and (3.7) in terms of the vector fields f and ω are not important for this property; only the choice of Stratonovich integration in the definition of $m^{(2)}$ plays a role.

The thread of [30] was picked up in the rough path setting in [6] where the authors introduce so-called "rough drivers", which are vector field analogues of rough paths. Rough drivers consist of a family of differential operators $(X_{st},\mathbb{X}_{st})_{0\le s\le t\le T}$ such that X_{st} (respectively $\mathbb{X}_{st}$) are first- (respectively second-) order differential operators for all $0\le s\le t\le T$ and the following Chen's relation holds true;

$$X_{st}=X_{su}+X_{ut},\qquad \mathbb{X}_{st}=\mathbb{X}_{su}+\mathbb{X}_{ut}+X_{su}X_{ut}.$$

A rough driver is called weakly geometric provided the second order derivative operator

$$\mathbb{W}_{st}:=\mathbb{X}_{st}-\frac{1}{2}X_{st}X_{st}$$

is actually a first order derivative operator, i.e. a vector field. Additionally, we require the following regularity:

$$\|X_{st}\|_{\mathfrak{X}^p(\mathbb{R}^d)}\lesssim|t-s|^\alpha,\qquad \|\mathbb{W}_{st}\|_{\mathfrak{X}^{p-1}(\mathbb{R}^d)}\lesssim|t-s|^{2\alpha},$$

where $\mathfrak{X}^p(\mathbb{R}^d)$ denotes the set of vector fields of spatial regularity C^p.

It was noted in [10] that rough drivers can be canonically defined from infinite-dimensional, i.e. $C_b^p(\mathbb{R}^d,\mathbb{R}^d)$, valued rough paths. In fact, the set of C^p-vector fields $\mathfrak{X}^p(\mathbb{R}^d)$ is canonically identified with $C_b^p(\mathbb{R}^d,\mathbb{R}^d)$ via

$$\begin{array}{ccc} C_b^p(\mathbb{R}^d,\mathbb{R}^d) & \to & \mathfrak{X}^p(\mathbb{R}^d)\\ f & \mapsto & f\cdot\nabla=\sum_\mu f^\mu\frac{\partial}{\partial\xi^\mu}. \end{array}$$

Moreover, define by linearity on the algebraic tensor

$$\begin{array}{ccc} C_b^p(\mathbb{R}^d,\mathbb{R}^d)^{\otimes_a 2} & \to & \mathfrak{X}^{p-1}(\mathbb{R}^d)\\ f\otimes g & \mapsto & (f\cdot\nabla(g\cdot\nabla))=\sum_{\mu,\nu} f^\mu\frac{\partial g^\nu}{\partial\xi^\mu}\frac{\partial}{\partial\xi^\nu} \end{array}$$

and denote by $\nabla_2^\otimes$ the extension to $C_b^p(\mathbb{R}^d\times\mathbb{R}^d,\mathbb{R}^{d\times d})$. Moreover, for a matrix a we let $a\nabla^2:=\sum_{\mu,\nu}a^{\mu,\nu}\frac{\partial}{\partial\xi^\mu}\frac{\partial}{\partial\xi^\nu}$. Then, given a rough path $\mathbf{x}\in\mathscr{C}_{wg}^\alpha([0,T],G^2(C_b^p(\mathbb{R}^d,\mathbb{R}^d)))$, if we let

$$X_{st}(\xi):=x_{st}(\xi)\cdot\nabla,\qquad \mathbb{X}_{st}(\xi):=\nabla_2^\otimes x_{st}^{(2)}(\xi,\xi)+x_{st}^{(2)}(\xi,\xi)\nabla^2,\tag{3.9}$$

then $\mathbf{X} := (X, \mathbb{X})$ is a weakly geometric rough driver in the sense of [6].

In [6] the authors prove Wong-Zakai approximations of $dy_t = m_{\circ dt}(y_t)$ by using linear interpolation of the Banach-space martingale m, showing that the corresponding iterated integral converges to $m^{(2)}$ in the appropriate sense and using continuity of the Itô-Lyons map, see [6] for details. As such, the Wong-Zakai approximation is constructed by hand. Our proposition below is also proved using continuity of the Itô-Lyons map and the continuity of the mapping $\mathbf{x} \mapsto \mathbf{X}$. However, let us emphasize that we do not construct the Wong-Zakai approximation by hand, but instead use Theorem 1.1 which guarantees the existence of a smooth approximation of the rough driver.

Theorem 3.4. *Let $\mathbf{x} \in \mathscr{C}^{\alpha}_{wg}([0,T], G^2(H^k(\mathbb{R}^d, \mathbb{R}^d)))$ for $k > \frac{d}{2} + p + 1$ for some $p \geq 3$ and suppose y solves $dy_t = \mathbf{X}_{dt}(y_t)$ where $\mathbf{X}_t = (X_t, \mathbb{X}_t)$ is the rough driver built from $\mathbf{x}$. Then there exists a sequence of functions $x^n : [0,T] \times \mathbb{R}^d \to \mathbb{R}^d$ of bounded variation of t such that the solution y^n of*

$$\dot{y}^n_t = x^n_t(y^n_t)$$

converges to y in $C^{\beta}([0,T], C^p(\mathbb{R}^d, \mathbb{R}^d))$ for any $\beta \in (\frac{1}{3}, \alpha)$.

Proof. Since $\mathbf{x}$ is weakly geometric we have

$$x^{(2),\mu,\nu}_{st}(\xi,\zeta) + x^{(2),\nu,\mu}_{st}(\zeta,\xi) = (x^{\mu}_t(\xi) - x^{\mu}_s(\xi))(x^{\nu}_t(\zeta) - x^{\nu}_s(\zeta)) \tag{3.10}$$

for all $\mu, \nu \in \{1, \dots, d\}$ which gives $x^{(2)}_{st}\nabla^2 = \frac{1}{2}(x_t - x_s)(x_t - x_s)^T \nabla^2$. It follows that

$$\mathbb{X}_{st}(\xi) - \frac{1}{2} X_{st}(X_{st})(\xi) = \nabla^{\otimes}_2 \left(x^{(2)}_{st} - \frac{1}{2} x_{st} \otimes x_{st} \right)(\xi, \xi)$$

is actually a vector field (so it is a weakly geometric rough driver in the sense of [6]). From Theorem 1.1 we get can approximate the infinite dimensional rough path $(x, x^{(2)})$ by a sequence of smooth paths. The result now follows from [6, Theorem 2.6] since the embedding $H^k(\mathbb{R}^d, \mathbb{R}^d) \subset C^p_b(\mathbb{R}^d, \mathbb{R}^d)$ is continuous. $\square$

Notice that we use the Sobolev embedding in the above proof to put ourselves in a Hilbert-space setting and it is the reason for requiring the high spatial regularity, $k > \frac{d}{2} + p + 1$.

4. Geometric rough paths on Hilbert spaces

4.1. Free nilpotent groups of step 2 in finite dimensions

We will first consider Carnot-Carathéodory geometry for finite dimensional, with the aim of proving result that are valid in the infinite-dimensional setting as well.

Let E be a finite dimensional inner product space and use the notation $X^* = \langle X, \cdot \rangle$ for any $X \in E$. In the notation of Section 2.4, define a Lie algebra $\mathfrak{g}(E) = \mathcal{P}^2(E)$. By Example 2.8, we can identify $\mathfrak{g}(E)$ with $E \oplus \wedge^2 E$ equipped with a Lie bracket structure

$$[X + \mathbb{X}, Y + \mathbb{Y}] = X \wedge Y, \qquad X, Y \in E, \mathbb{X}, \mathbb{Y} \in \wedge^2 E. \tag{4.1}$$

We identify $\wedge^2 E$ with the space of skew-symmetric endomorphisms $\mathfrak{so}(E)$ by writing

$$X \wedge Y = X^* \otimes Y - Y^* \otimes X. \tag{4.2}$$

Consider the corresponding simply connected Lie group $G^2(E)$. For the rest of this section, we will use the fact that $\exp : \mathfrak{g}(E) \to G^2(E)$ is a diffeomorphism to identify these as spaces. Using group exponential coordinates $G^2(E)$ is then the space $E \oplus \mathfrak{so}(E)$ with multiplication

$$(x + x^{(2)}) \cdot (y + y^{(2)}) = x + y + x^{(2)} + y^{(2)} + \frac{1}{2} x \wedge y, \tag{4.3}$$

$x, y \in E$, $x^{(2)}, y^{(2)} \in \mathfrak{so}(E)$. With this identification, the identity is 0 and inverses are given by $(x + x^{(2)})^{-1} = -x - x^{(2)}$. Recall Lemma 2.4 for relating the presentation of $G^2(E)$ as a subset of $\mathcal{A}_2$ and its representation in exponential coordinates.

An absolutely continuous curve $\mathbf{\Gamma}(t)$ in $G^2(E)$ with an L^1-derivative is called *horizontal* if for almost every t,

$$\mathbf{\Gamma}(t)^{-1} \cdot \dot{\mathbf{\Gamma}}(t) \in E.$$

In other words, if we write $\mathbf{\Gamma}(t) = \gamma(t) + \gamma^{(2)}(t)$ with $\gamma(t) \in E$ and $\gamma^{(2)}(t) \in \wedge^2 E$, then for some L^1-function $u(t) \in E$, we have

$$\dot{\gamma}(t) = u(t), \qquad \dot{\gamma}^{(2)}(t) = \frac{1}{2} \gamma(t) \wedge u(t).$$

Since E is a generating subspace of $\mathfrak{g}(E)$, it follows from the Chow-Rashevskiĭ Theorem [9,37] that any pair of points in $G^2(E)$ can be connected by a horizontal curve. For any pair of points in $\mathbf{x}, \mathbf{y} \in G^2(E)$, define the Carnot-Carathéodory metric (CC-metric) by

$$\rho(\mathbf{x}, \mathbf{y}) = \left\{ \int_0^1 \|\mathbf{\Gamma}(t)^{-1} \cdot \dot{\mathbf{\Gamma}}(t)\|_E \, dt \; : \; \begin{array}{c} \mathbf{\Gamma} : [0,1] \to G^2(E) \text{ horizontal}, \\ \mathbf{\Gamma}(0) = \mathbf{x}, \mathbf{\Gamma}(1) = \mathbf{y} \end{array} \right\}.$$

Note that if $\mathbf{\Gamma}(t)$ is horizontal, then so is $\mathbf{x} \cdot \mathbf{\Gamma}(t)$. It follows that the distance ρ is left invariant.

From e.g. [1, Section 7.3], length minimizers of ρ are all on the form,

$$\gamma(t) = x_0 + \int_0^t e^{s\Lambda} u_0 \, ds, \qquad \gamma^{(2)}(t) = x_0^{(2)} + \frac{1}{2} \int_0^t \gamma(s) \wedge e^{s\Lambda} u_0 ds, \tag{4.4}$$

for some constant element $\Lambda \in \mathfrak{so}(H)$ and $u_0 \in E$.

Example 4.1 *(Heisenberg group).* When E is two-dimensional, the group $G^2(E)$ is known as the Heisenberg group. For any choice of orthogonal frame X, Y, define $Z = \frac{1}{2}(X - iY)$. This means that we can represent any element $\mathbf{y} = aX + bY + cX \wedge Y$ as

$$\mathbf{y} = (a + ib)Z + (a - ib)\bar{Z} + cX \wedge Y.$$

We will use a similar notation in the rest of the paper.

If $\Lambda = \lambda X \wedge Y$, $u_0 = u_0 Z + \bar{u}_0 \bar{Z}$, $\gamma(t) = z(t)Z + \bar{z}(t)\bar{Z}$ and $\gamma^{(2)}(t) = \sigma(t) X \wedge Y$ with $z(t), u_0 \in \mathbb{C}$ and $\sigma(t), \lambda \in \mathbb{R}$, then (4.4) becomes

$$z(t) = z_0 + \int_0^t e^{i\lambda s} u_0 \, ds = z_0 + \frac{e^{i\lambda t} - 1}{i\lambda} u_0 = z_0 + \frac{2\sin(\lambda t/2)}{\lambda} e^{i\lambda/2t} u_0,$$

$$\begin{aligned} \sigma(t) &= \sigma_0 + \frac{1}{2} \int_0^t \operatorname{Im}(\bar{z}(s) e^{i\lambda s} u_0) \, ds \\ &= \sigma_0 + \frac{2\sin(\lambda t/2)}{\lambda} \operatorname{Im}\left(e^{i\lambda/2t} \bar{z}_0 u_0\right) - \frac{1}{2} \frac{|u_0|^2}{\lambda} \left(t - \frac{\sin(\lambda t)}{\lambda}\right), \end{aligned}$$

where we interpret $\frac{\sin(\lambda t)}{\lambda}$ as t if $\lambda = 0$. If the initial point is the identity 0, we have

$$z(t) = \frac{2\sin(\lambda t/2)}{\lambda} e^{it\lambda/2} u_0, \qquad \sigma(t) = -\frac{|u_0|^2}{2\lambda} \left(t - \frac{\sin(\lambda t)}{\lambda}\right).$$

If the above geodesic is defined on the interval [0, 1], then it has length $|u_0|$. In particular, we observe the following.

(a) A minimizing geodesic defined on [0, 1] from 0 to $zZ + \bar{z}\bar{Z}$ is given by the choice $\lambda = 0$. It follows that

$$\rho(0, zZ + \bar{z}\bar{Z}) = |z|.$$

(b) A minimizing geodesic defined on [0, 1] from 0 to $\sigma X \wedge Y$ is given by the choice $\lambda = \pm 2\pi$ depending on the sign of σ. Hence, we have that

$$|\sigma| = \frac{\rho(0, \sigma X \wedge Y)^2}{4\pi}.$$

(c) Note that since

$$|z(1)| = |z| = \left| \int_0^1 u(t) dt \right| \leq \int_0^1 |u(t)| dt = |u_0|,$$

we have $|z| \leq \rho(0, zZ + \bar{z}\bar{Z} + \sigma X \wedge Y)$. It then also follows that

$$\begin{aligned} &2\sqrt{\pi} |\sigma|^{1/2} = \rho(0, \sigma X \wedge Y) \\ &\leq \rho(0, -zZ + \bar{z}\bar{Z}) + \rho(-zZ - \bar{z}\bar{Z}, \sigma X \wedge Y) \\ &= |z| + \rho(0, zZ + \bar{z}\bar{Z} + \sigma X \wedge Y) \leq 2\rho(0, zZ + \bar{z}\bar{Z} + \sigma X \wedge Y). \end{aligned}$$

Using this fact along with the upper bound from the triangle inequality and left invariance, we have

$$\max\{|z|, \sqrt{\pi}|\sigma|^{1/2}\} \leq \rho(0, zZ + \bar{z}\bar{Z} + \sigma X \wedge Y) \leq |z| + 2\sqrt{\pi}|\sigma|^{1/2}. \tag{4.5}$$

We want to generalize the inequality (4.5) to free nilpotent groups of step 2 of arbitrary dimensions, but without having any dimension-dependent constants. The inequality can be concluded from formulas of the CC-distance to the vertical space in [39, Appendix A], but we include some more details here for the sake of completion and for applications to infinite dimensional vector spaces in Section 4.2.

Consider the case of a Hilbert space E of arbitrary finite dimension $n \geq 2$. We want to introduce a class of norms and quasi-norms on $\mathfrak{so}(E)$. Any element $\mathbb{X} \in \mathfrak{so}(E)$ will have non-zero eigenvalues

$$\{\pm i\sigma_1, \dots, \pm i\sigma_k\} \tag{4.6}$$

or some $k \geq 0$. We order them in such a way that

$$\sigma_1 \geq \cdots \geq \sigma_k > 0.$$

These are also the singular values of $\mathbb{X}$ as $|\mathbb{X}| = \sqrt{-\mathbb{X}^2}$ has exactly these non-zero eigenvalues, with each σ_j appearing twice. Define a sequence $\sigma(\mathbb{X}) = (\sigma_j)_{j=1}^{\infty}$ of non-negative numbers such that $\sigma_j = 0$ for $j > k$. For $0 < p \leq \infty$, we define

$$\|\mathbb{X}\|_{\mathrm{Sch}^p} = 2^{1/p}\|\sigma(\mathbb{X})\|_{\ell^p}.$$

For $p \geq 1$, these are norms called the Schatten p-norms [34, 16]. We will also introduce the following map

$$\|\mathbb{X}\|_{cc} = \|\sigma(\mathbb{X})\|_{\ell^1(\mathbb{R};\mathbb{N})} = \sum_{j=1}^{\infty} j\sigma_j.$$

It is simple to see that $\|\cdot\|_{cc}$ is not a norm when $\dim E > 2$. However, we will show that it is a quasi-norm. Recall that a quasi-norm is a map satisfying the norm axioms except the triangle inequality which is assumed in the form $\|x + y\| \leq K(\|x\| + \|y\|)$ for some $K \geq 1$, [12, Section I.9]. From the definition of $\|\cdot\|_{cc}$, we note that

$$\frac{1}{2}\|\mathbb{X}\|_{\mathrm{Sch}^1} \leq \|\mathbb{X}\|_{cc} \leq \frac{1}{4}\|\mathbb{X}\|_{\mathrm{Sch}^{1/2}}. \tag{4.7}$$

The latter follows from the fact that for any $k > 0$, $\sqrt{a + kb} \leq \sqrt{a} + \sqrt{b}$ if $b \geq 0$ and $a \geq \frac{(k-1)^2}{4}b$. Hence

$$\sqrt{\sigma_1 + \cdots + k\sigma_k} \leq \sqrt{\sigma_1 + \cdots + (k-1)\sigma_{k-1}} + \sqrt{\sigma_k},$$

since $\sigma_1 + \cdots + (k-1)\sigma_{k-1} \geq \frac{k(k-1)}{2}\sigma_k$.

Define a homogeneous norm

$$|||x + x^{(2)}||| = \max\left\{ \|x\|_E, \sqrt{\pi}\|x^{(2)}\|_{cc}^{1/2} \right\}.$$

We then have the following result.

Theorem 4.2. *Let E be an arbitrary finite dimensional Hilbert space. If ρ is the Carnot-Carathéodory distance on $G^2(E)$, then*

$$|||x + x^{(2)}||| \leq \rho(0, x + x^{(2)}) \leq 3|||x + x^{(2)}|||.$$

We emphasize that the above inequality holds independent of dimension. If we allow constants depending on dimension, then any homogeneous gauge will be equivalent, see e.g. [32, Proposition 10].

Proof. The minimal geodesic from 0 to $x \in E$ is just a straight line in E, and hence

$$\|x\|_E = \rho(0, x).$$

We will show that we also have

$$\rho(0, x^{(2)}) = 2\sqrt{\pi}\|x^{(2)}\|_{cc}^{1/2}, \qquad x^{(2)} \in \mathfrak{so}(E). \tag{4.8}$$

The result then follows from similar steps as in Example 4.1.

We will use the geodesic equations in (4.4). Consider a general solution $\mathbf{\Gamma}(t) = \gamma(t) + \gamma^{(2)}(t)$ on $G^2(E)$ with $\mathbf{\Gamma}(0) = 0$ and $\mathbf{\Gamma}(1) = x^{(2)}$. Consider arbitrary initial values $\Lambda \neq 0$ and $u_0 \neq 0$ for the geodesic equation as in (4.4). Choose an orthonormal basis $X_1, \dots, X_k, Y_1, \dots, Y_k, T_1, \dots, T_{n-k}$ such that we can write

$$\Lambda = \sum_{j=1}^{k} \lambda_j X_j \wedge Y_j, \qquad \lambda_j > 0.$$

Introduce again complex notation $Z_j = \frac{1}{2}(X_j - iY_j)$ and write

$$u_0 = \sum_{j=1}^{k} w_j Z_j + \sum_{j=1}^{k} \bar{w}_j \bar{Z}_j + \sum_{j=1}^{n-k} c_j T_j, \qquad w_j \in \mathbb{C}, c_j \in \mathbb{R}.$$

We will then have

$$u(t) = \sum_{j=1}^{k} e^{i\lambda_j t} w_j Z_j + \sum_{j=1}^{k} e^{-i\lambda_j t} \bar{w}_j \bar{Z}_j + \sum_{j=1}^{n-k} c_j T_j, \qquad w_j \in \mathbb{C}, c_j \in \mathbb{R}.$$

We make the following simplifications. If $w_j = 0$, then the value of λ_j has no effect on $u(t)$. We may hence set it to zero and reduce the value of k. Without any loss of generality, we can hence assume that every w_j is non-zero. Next, if we have $\lambda_j = \lambda_l$ for some $1 \leq j, l \leq k$ then $e^{i\lambda_j t} w_j Z_j + e^{i\lambda_k t} w_k Z_k = e^{i\lambda_j t}(w_j Z_j + w_k Z_k) =: e^{i\lambda_j t} \frac{w_{jl}}{2}(X_{jl} - iY_{jl})$ for some orthonormal pair of vectors X_{jl}, Y_{jl}. Hence we again obtain the same $u(t)$ if we replace $\lambda_j X_j \wedge Y_j + \lambda_l X_l \wedge Y_l$ with $\lambda_j X_{jl} \wedge Y_{jl}$. By repeating such replacements, we may assume that all values of $\lambda_1, \dots, \lambda_k$ are different.

If $\mathbf{\Gamma}(t) = \gamma(t) + \gamma^{(2)}(t)$ is the corresponding geodesic, then

$$\gamma(t) = \sum_{j=1}^{k} \frac{2\sin(\lambda_j t/2)}{\lambda_j} e^{i\lambda_j t/2} w_j Z_j + \sum_{j=1}^{k} \frac{2\sin(\lambda_j t/2)}{\lambda_j} e^{-i\lambda_j t/2} \bar{w}_j \bar{Z}_j + \sum_{j=1}^{n-k} t c_j T_j.$$

From the condition $\gamma(1) = 0$, it follows that $c_1, \dots, c_{n-k}$ all vanish for every $1 \leq j \leq n-k$. Furthermore, since we assume that $w_j \neq 0$, it follows that $\lambda_j = 2\pi n_j$ for some positive integers n_j.

Computing $x^{(2)}$ and using that the integers $n_1, \dots, n_k$ are all different, we obtain

$$x^{(2)} = \frac{1}{4\pi} \sum_{j=1}^{k} \operatorname{Im}\left(\frac{|w_j|^2}{i n_j} \int_0^1 (1 - e^{2i\pi n_j t}) dt \right) X_j \wedge Y_j = -\frac{1}{4\pi} \sum_{j=1}^{k} \frac{|w_j|^2}{n_j} X_j \wedge Y_j.$$

It follows that the endpoint $x^{(2)}$ has $2k$ non-zero eigenvalues $\{\pm i\sigma_1, \dots, \pm i\sigma_k\}$ with

$$\sigma_j = \frac{1}{4 n_j \pi} |w_j|^2.$$

In other words, any local length minimizer $\mathbf{\Gamma}(t)$ from 0 to the point $x^{(2)}$ has length

$$\text{Length}(\mathbf{\Gamma})^2 = \sum_{j=1}^{k} |w_j|^2 = \sum_{j=1}^{k} 4\pi n_j \sigma_j.$$

In order to obtain the minimal value, we use $n_j = l$ if σ_j is the l-th largest eigenvalue. The result follows. □

Using the identity (4.8) we also obtain the following result.

Corollary 4.3. $\| \cdot \|_{cc}$ *is a quasi-norm on* $\mathfrak{so}(E)$, *even a* $1/2$*-norm [12, Section I.9], in that it satisfies*

$$\|\mathbb{X} + \mathbb{Y}\|_{cc}^{1/2} \leq \|\mathbb{X}\|_{cc}^{1/2} + \|\mathbb{Y}\|_{cc}^{1/2}, \qquad \|\mathbb{X} + \mathbb{Y}\|_{cc} \leq 2(\|\mathbb{X}\|_{cc} + \|\mathbb{Y}\|_{cc}).$$

4.2. *Free nilpotent groups on step 2 from infinite dimensional Hilbert spaces*

Let E be a real Hilbert space. We will not assume that E is finite dimensional or even separable, but our result and notation from the previous section will still be essential. We choose and fix a tensor norm $\| \cdot \|_\otimes$ on the algebraic tensor product $E \otimes_a E$ which is assumed to satisfy properties 1. and 2. from Section 2.1. Moreover, we assume that $\| \cdot \|_\otimes$ lies (pointwise) between the injective and projective tensor norms (cf. e.g. [40]). As mentioned in Example 2.1, we can identity $E \otimes_a E$ with finite rank operators, and we can consider $E \otimes E$ as the closure of finite rank operators with respect to $\| \cdot \|_\otimes$.

In describing $\mathfrak{g}(E) = \mathcal{P}^2(E) = E \oplus \wedge^2 E$, through (4.2) we identify the algebraic wedge product $\wedge_a^2 E$ with the space of all finite rank skew-symmetric operators, which we denote by $\mathfrak{so}_a(E)$. We identify $\mathfrak{g}(E)$ with $E \oplus \mathfrak{so}_\otimes(E)$ where $\mathfrak{so}_\otimes(E)$ are the skew-symmetric operators on E that are in the closure of $\mathfrak{so}_a(E)$ with respect to $\| \cdot \|_\otimes$ and with brackets as in (4.1). If we give $\mathfrak{g}(E)$ a norm

$$\|x + x^{(2)}\|_{\mathfrak{g}(E)} = \max \left\{ \|x\|_E, \|x^{(2)}\|_\otimes \right\},$$

then it has the structure of a Banach Lie algebra.

For any compact skew-symmetric map $\mathbb{X} : E \to E$, define a sequence $\sigma(\mathbb{X}) = (\sigma_j)_{j=1}^\infty$ such that $|\mathbb{X}| = \sqrt{-\mathbb{X}^2}$ has eigenvalues in non-increasing order $\sigma_1 = \sigma_1 \geq \sigma_2 = \sigma_2 \geq \cdots$. For $p \in (0, \infty]$, let $\mathfrak{so}_p(E)$ denote the space of compact skew-symmetric operators $\mathbb{X}$ with finite Schatten p-norm $\|\mathbb{X}\|_{\mathrm{Sch}^p} = 2^{1/p} \|\sigma(X)\|_{\ell^p}$. As $p = \infty$ and $p = 1$ correspond to respectively the injective norm and the projective norm, we have

$$\mathfrak{so}_1(E) \subseteq \mathfrak{so}_\otimes(E) \subseteq \mathfrak{so}_\infty(E).$$

Introduce the space $\mathfrak{so}_{cc}(E)$ as the subspace of $\mathfrak{so}_\infty(E)$ of elements $\mathbb{X}$ such that $\|\mathbb{X}\|_{cc} := \sum_{j=1}^\infty j\sigma_j$ is finite. Since all compact operators are limits of finite rank operators ([34, Corollary 16.4]), all the previously mentioned inequalities from Section 4.1 still hold. In particular, we get that $\| \cdot \|_{cc}$ is a quasi-norm and that the inequality (4.7) holds for any $\mathbb{X} \in \mathfrak{so}_a(E)$. It then follows that

$$\mathfrak{so}_{1/2}(E) \subseteq \mathfrak{so}_{cc}(E) \subseteq \mathfrak{so}_1(E).$$

We emphasize here that for $\mathfrak{so}_1(E)$, the completion of $\mathfrak{so}_a(E)$ is with respect to the norm $\| \cdot \|_1$, while for $\mathfrak{so}_{1/2}(E)$ and $\mathfrak{so}_{cc}(E)$, we are considering the completion with respect to the respective induced distances $(\mathbb{X}, \mathbb{Y}) \mapsto \|\mathbb{X} - \mathbb{Y}\|_{1/2}^{1/2}$ and $(\mathbb{X}, \mathbb{Y}) \mapsto \|\mathbb{X} - \mathbb{Y}\|_{cc}^{1/2}$.

The group $G^2(E)$ corresponding to $\mathfrak{g}(E)$ can be considered in exponential coordinates as the set $\mathfrak{g}(E)$ with group operation as in (4.3). We define the distance d on $G^2(E)$ by $d(\mathbf{x}, \mathbf{y}) = \|\mathbf{x}^{-1}\mathbf{y}\|_{\mathfrak{g}(E)}$. Let $t \mapsto u(t)$ be any function in $L^1([0, 1], E)$, and let $\mathbf{\Gamma}_u$ be the solution of

$$\mathbf{\Gamma}_u(t)^{-1} \cdot \dot{\mathbf{\Gamma}}_u(t) = u(t), \qquad \mathbf{\Gamma}_u(0) = 0.$$

Recall that 0 is the identity, since we are using exponential coordinates. This curve always exists from the L^1-regularity property of the Banach Lie group $G^2(E)$ (see [20] and also Remark 2.7). For any $\mathbf{x}, \mathbf{y} \in G^2(E)$, we define $\rho(\mathbf{x}, \mathbf{y}) \in [0, \infty]$ by

$$\begin{aligned} \rho(\mathbf{x}, \mathbf{y}) &= \rho(0, \mathbf{x}^{-1} \cdot \mathbf{y}), \\ \rho(0, \mathbf{x}) &= \inf \left\{ \|u\|_{L^1} : u \in L^1([0, 1], E), \mathbf{\Gamma}_u(1) = \mathbf{x} \right\}. \end{aligned}$$

4.3. Properties of projections

Let F be a closed subspace of E. Let $\mathrm{Pr}_F : E \to F$ be the corresponding orthonormal projection. We then write the map $\mathrm{pr}_F : G^2(E) \to G^2(F)$ for the corresponding map

$$\mathrm{pr}_F(x + x^{(2)}) = \mathrm{Pr}_F + (\mathrm{Pr}_F\, x^{(2)}\, \mathrm{Pr}_F), \qquad x \in E, x^{(2)} \in \mathfrak{so}_\otimes(E).$$

We then emphasize the following properties.

Lemma 4.4.

(a) pr_F *is a group homomorphism from* $G^2(E)$ *and* $G^2(F)$.
(b) Let ρ_F *denote the Carnot-Carathéodory distance defined on* $G^2(F)$. *For any* $\mathbf{x}, \mathbf{y} \in G^2(E)$, *we have*

$$\rho_F(\mathrm{pr}_F\,\mathbf{x}, \mathrm{pr}_F\,\mathbf{y}) = \rho(\mathrm{pr}_F\,\mathbf{x}, \mathrm{pr}_F\,\mathbf{y}) \le \rho(\mathbf{x}, \mathbf{y}).$$

In particular, if there is a geodesic from $\mathbf{x}$ *to* $\mathbf{y}$ *in* F *with respect to* ρ_F, *then this is also the geodesic in* $G^2(E)$ *with respect* ρ.

Proof. (a) follows from the definition of the definition of the group operation. Using (a), we only need to prove that $\rho(0, \mathrm{pr}_F\,\mathbf{x}) \le \rho(0, \mathbf{x})$ to prove (b). We observe that if $\mathbf{\Gamma}(t)$ is a horizontal curve from 0 to $\mathbf{x}$, then $\mathrm{pr}_F\,\mathbf{\Gamma}(t)$ is a horizontal curve of less or equal length with endpoint $\mathrm{pr}_F\,\mathbf{x}$. $\square$

Lemma 4.5. *If* $\|\cdot\|_\otimes = \|\cdot\|_p$ *is the Schatten* p*-norm, the following properties hold.*

(a) For any closed subspace F *of* E *and* $\mathbf{x}, \mathbf{y} \in G^2(E)$, *we have*

$$d(\mathrm{pr}_F\,\mathbf{x}, \mathrm{pr}_F\,\mathbf{y}) \le d(\mathbf{x}, \mathbf{y}).$$

(b) For any $\mathbf{x} = x + x^{(2)}$, *there is a sequence of finite dimensional subspaces* $F_1 \subseteq F_2 \subseteq \cdots$ *such that*

$$x \in F_n \textit{ for any } n, \qquad \lim_{n\to 0} d(\mathbf{x}, \mathrm{pr}_{F_n}\,\mathbf{x}) = 0.$$

Proof. (a) Again it is sufficient to prove that $d(0, \mathrm{pr}_F\,\mathbf{x}) \le d(0, \mathbf{x})$ for any $\mathbf{x} = x + x^{(2)} \in G^2(E)$. We see that $\|\mathrm{Pr}_F\, x\|_E \le \|x\|_E$ and furthermore,

$$\|\mathrm{Pr}_F\, x^{(2)}\, \mathrm{Pr}_F\|_{\mathrm{Sch}^p} \le \|x^{(2)}\|_{\mathrm{Sch}^p}$$

since

$$\begin{aligned}\sigma_{j+1}(\mathrm{Pr}_F\, x^{(2)}\, \mathrm{Pr}_F) &= \max_{\mathrm{rank}(\tilde{E})=2j+1}\ \min_{y\in\tilde{E}\,\|y\|_E=1} \|\mathrm{Pr}_F x^{(2)} \mathrm{Pr}_F\, y\|_E \\ &\le \max_{\mathrm{rank}(\tilde{E})=2j+1}\ \min_{y\in\tilde{E}\,\|y\|_E=1} \|x^{(2)} y\|_E = \sigma_{j+1}(x^{(2)}).\end{aligned}$$

We remark that we have here used the skew-symmetry of $\mathrm{Pr}_F\, x^{(2)}\, \mathrm{Pr}_F$ and the relationship between the singular values and the eigenvalues given in (4.6).

(b) Write $\mathbf{x} = x + x^{(2)}$ and give the singular value decomposition

$$x^{(2)} = \sum_{j=1}^{\infty} \sigma_j X_j \wedge Y_j, \qquad \sigma_1 \geq \sigma_2 \geq \cdots, \tag{4.9}$$

with $X_1, Y_1, X_2, Y_2, \ldots$ all orthogonal unit vector fields. We define

$$\tilde{F}_n = \mathrm{span}\{X_j, Y_j \,:\, j = 1, \ldots, n\}, \qquad F_n = \mathrm{span}\{x, \tilde{F}_n\}.$$

Then by left invariance

$$d(\mathbf{x}, \mathrm{pr}_{F_n} \mathbf{x}) = d(0, \mathrm{pr}_{F_n^\perp} x^{(2)}) \overset{(a)}{\leq} d(0, \mathrm{pr}_{\tilde{F}_n^\perp} x^{(2)}) = 2^{1/p} \left(\sum_{j=n+1}^{\infty} \sigma_j (x^{(2)})^p \right)^{1/p}$$

which converge to zero by definition. □

4.4. Geodesic completeness

One of the main steps in completing Theorem 1.1 will be to establish that ρ makes a subset into a geodesic space.

Theorem 4.6. *Let E be a Hilbert space and define $G^2(E)$ relative to a tensor norm $\|\cdot\|_\otimes$ satisfying 1. and 2. from Section 2.1 and bounded from below by the injective tensor product $\|\cdot\|_{\mathrm{Sch}^\infty}$. If we define $G^2(E) := \exp(E \oplus \mathfrak{so}_{cc}(E))$, then*

$$G^2_{cc}(E) = \{\mathbf{x} \in G^2(E) \,:\, \rho(0, \mathbf{x}) < \infty\}.$$

Furthermore, the metric space $(G^2_{cc}(E), \rho)$ is a complete, geodesic space and if we define

$$|||x + x^{(2)}||| = \max\left\{ \|x\|_E, \sqrt{\pi}\, \|x^{(2)}\|_{cc}^{1/2} \right\},$$

then

$$|||\mathbf{x}||| \leq \rho(0, \mathbf{x}) \leq 3|||\mathbf{x}|||. \tag{4.10}$$

We will do the proof of this theorem in two parts. In the first part, we will show that $G^2_{cc}(E)$ is indeed exactly the set with finite ρ-distance and that the inequality (4.10) holds. In the second part, we show that it is a geodesic space.

Proof of Theorem 4.6, Part I. We will begin by introducing the following notation, which we will use in both parts of the proof. Recall the definition of $\mathrm{pr}_F : G^2(E) \to G^2(F) \subseteq G^2(E)$ for some closed subspace F from Section 4.3. Write $\mathrm{pr}_{F,\perp} = \mathrm{pr}_{F^\perp}$ and write a projection operator

$$\mathrm{pr}_{F\wedge F^\perp}(\mathbf{x}) = \mathrm{Pr}_F x^{(2)}\mathrm{Pr}_{F^\perp} + \mathrm{Pr}_{F^\perp} x^{(2)}\mathrm{Pr}_F = \mathbf{x} - \mathrm{pr}_F\,\mathbf{x} - \mathrm{pr}_{F,\perp}\,\mathbf{x}, \qquad \mathbf{x} = x + x^{(2)}.$$

We have already shown the result for finite dimensional spaces, so we assume that E is infinite dimensional.

Step 1: The CC-distance is finite on algebraic elements. Let $G_a(E) = \exp(E \oplus \mathfrak{so}_a(E)$ and consider an arbitrary element $\mathbf{x} = x + x^{(2)} \in G_a^2(E)$ with $x^{(2)} = \sum_{j=1}^n \sigma_j X_j \wedge Y_j$ being the singular value decomposition as in (4.9). Define the finite dimensional subspace $F = \mathrm{span}\{x, X_1, Y_1, \dots, X_n, Y_n\}$. We then observe that since $\mathbf{x} \in G^2(F)$, $\rho(0, \mathbf{x}) < \infty$ and there is a minimizing geodesic from 0 to $\mathbf{x}$. Also, any element in $G_a^2(E)$ satisfies the inequality (4.10).

Step 2: Vertical elements. Consider an element $\mathbf{x} = x^{(2)} \in \mathfrak{so}_{cc}(E)$ with $\sigma(x^{(2)}) = (\sigma_j)$. Let $x^{(2)} = \sum_{j=1}^\infty \sigma_j X_j \wedge Y_j$ be the singular value decomposition and define $Z_j = \frac{1}{2}(X_j - Y_j)$. Consider the curve

$$u(t) = 2\sqrt{\pi}\sum_{j=1}^{\infty}(j\sigma_j)^{1/2}(e^{-2\pi j t}Z_j + e^{2\pi j t}\bar{Z}_j).$$

We see that $\|u(t)\|_E = \|u\|_{L^1} = 2\sqrt{\pi\|x^{(2)}\|_{cc}}$. Furthermore, if F_n is the span of $X_1, Y_1, \dots, X_n$, Y_n, then by the proof of Theorem 4.2, it follows that $\mathrm{pr}_{F_n}\boldsymbol{\Gamma}_u$ is a minimizing geodesic from 0 to $\mathbf{x}^n := \sum_{j=1}^n \sigma_j X_j \wedge Y_j$. Since $\mathrm{pr}_{F_n} u$ converges to u in $L^1([0,1], E)$ and $\mathbf{x}^n$ converges to $\mathbf{x}$ in the norm $\|\cdot\|_{\mathfrak{g}(E)}$, it follows that $\boldsymbol{\Gamma}_u$ is a minimizing geodesic from 0 to $\mathbf{x}$, and in particular,

$$\rho(0, \mathbf{x}) = \mathrm{Length}(\boldsymbol{\Gamma}_u) = 2\sqrt{\pi}\|x^{(2)}\|_{cc}^{1/2}.$$

Step 3: The CC-distance is exactly finite on $G_{cc}^2(E)$. For any element $\mathbf{x} = x + x^{(2)} \in G_{cc}^2(E)$, we can construct a horizontal curve $\boldsymbol{\Gamma}$ from 0 to $\mathbf{x}$ by a concatenation of the straight line from 0 to x with a minimizing geodesic from 0 to $x^{(2)}$ left translated by x. The result is that

$$\rho(0, \mathbf{x}) \le \mathrm{Length}(\boldsymbol{\Gamma}) = \|x\| + 2\sqrt{\pi}\|x^{(2)}\|_{cc}^{1/2} \le 3|||\mathbf{x}||| < \infty.$$

Conversely if $\mathbf{x} \in G^2(E)$ and $|||\mathbf{x}||| = \infty$, then using (4.10) and any sequence $\mathbf{x}^n$ in $G_a^2(E)$ converging to $\mathbf{x}$ in $\|\cdot\|_{\mathfrak{g}(E)}$, we see that $\rho(0, \mathbf{x}) = \infty$. $G_{cc}^2(E)$ is complete with the distance ρ as it is complete with respect to $|||\cdot|||$ by definition. $\square$

In order for us to complete Part II of the proof of Theorem 4.6 and show that $(G_{cc}(E), \rho)$ is a geodesic space, we will need the following lemma.

Lemma 4.7. *Let $\mathbf{x} = x + x^{(2)} \in G_{cc}^2(E)$ be a fixed arbitrary element with singular value decomposition $x^{(2)} = \sum_{j=1}^\infty \sigma_j X_j \wedge Y_j$ as in (4.9). Define subspaces $F_0 \subseteq F_1 \subseteq F_2 \subseteq \cdots$ by*

$$F_0 = \mathrm{span}\{x\}, \qquad F_{n+1} = \mathrm{span}(F_n \cup \{X_{n+1}, Y_{n+1}\}). \tag{4.11}$$

For any $n \ge 0$, define $\mathrm{pr}_n = \mathrm{pr}_{F_n}$, $\mathrm{pr}_{n,\perp} = \mathrm{pr}_{F_n^\perp}$ *and* $\mathrm{pr}_{n,\wedge} = \mathrm{pr}_{F_n \wedge F_n^\perp}$

(a) The set

$$K(\mathbf{x}) = \left\{ \mathbf{y} \in G^2_{cc}(E) : \begin{array}{c} \text{For any } n \geq 0, \\ \rho(0, \mathrm{pr}_n \mathbf{y}) \leq \sqrt{2\rho(0,\mathbf{x})\rho(0,\mathrm{pr}_n \mathbf{x})} \\ \rho(0, \mathrm{pr}_{n,\perp} \mathbf{y}) \leq \sqrt{2\rho(0,\mathbf{x})\rho(0,\mathrm{pr}_{n,\perp} \mathbf{x})} \\ \| \mathrm{pr}_{n,\wedge} \mathbf{y}\|_\otimes \leq 4\sqrt{2\rho(0,\mathbf{x})^3 \rho(0,\mathrm{pr}_{n,\perp} \mathbf{x})} \end{array} \right\}, \tag{4.12}$$

is relatively compact in $G^2(E)$.

(b) Any minimizing geodesic from 0 *to* $\mathbf{x}$ *is contained in* $K(\mathbf{x})$.

Proof. To simplify notation in the proof, we write $\rho(\mathbf{x}) := \rho(0, \mathbf{x})$.

(a) Define $F_\infty = \mathrm{span}\{x, X_1, Y_1, X_2, Y_2, \dots\}$. From the definition of $K(\mathbf{x})$ it follows that $\mathrm{pr}_{F_\infty} \mathbf{y} = 0$ for any $\mathbf{y} \in K(\mathbf{x})$. Considering the limit of pr_n, we also have that

$$\|\mathbf{y}\|_{\mathfrak{g}(E)} \leq \rho(\mathbf{y}) \leq \sqrt{2}\rho(\mathbf{x})$$

so $K(\mathbf{x})$ is bounded in both $G^2_{cc}(E)$ and $G^2(E)$. Recall (e.g. from [15, Theorem 4.3.29]) that for a complete metric space, a set is relatively compact if and only if it is totally bounded, i.e. for every $\varepsilon > 0$ there is a finite set of balls of radius $\varepsilon > 0$ covering the set. Let $B(\mathbf{z}, r)$ be the ball of radius r centered at $\mathbf{z} \in \mathfrak{g}(E)$ with respect to the $\|\cdot\|_{\mathfrak{g}(E)}$-norm. We observe that for any $\mathbf{y} \in K(\mathbf{x})$, we have

$$\begin{aligned} \|\mathrm{pr}_n \mathbf{y}\|_{\mathfrak{g}(E)} &\leq \rho(\mathrm{pr}_n \mathbf{y}) \leq \sqrt{2}\rho(\mathbf{x}), \\ \|\mathrm{pr}_{n,\perp} \mathbf{y}\|_{\mathfrak{g}(E)} &\leq \rho(\mathrm{pr}_{n,\perp} \mathbf{y}) \leq \sqrt{2\rho(\mathbf{x})\rho(\mathrm{pr}_{n,\perp} \mathbf{x})}, \\ \|\mathrm{pr}_{n,\wedge} \mathbf{y}\|_{\mathfrak{g}(E)} &= \frac{1}{2}\|\mathrm{pr}_{n,\wedge} \mathbf{y}_3\|_\otimes \leq 4\sqrt{2\rho(\mathbf{x})^3\rho(\mathrm{pr}_{n,\perp} \mathbf{x})}. \end{aligned}$$

Hence, for a given $\varepsilon > 0$, we can choose n sufficiently large such that

$$\max\left\{\sqrt{2\rho(\mathbf{x})\rho(\mathrm{pr}_{n,\perp} \mathbf{x})}, 2\sqrt{2\rho(\mathbf{x})^3\rho(\mathrm{pr}_{n,\perp} \mathbf{x})}\right\} \leq \frac{\varepsilon}{3}.$$

Choose a finite set of points $\mathbf{z}_1, \dots, \mathbf{z}_N$ such that $\cup_{j=1}^N B(\mathbf{z}_j, \frac{\varepsilon}{3})$ covers the relatively compact set $F_n \cap B(0, \rho(\mathbf{x}))$. By our choice of n, we then have that $\cup_{j=1}^N B(\mathbf{z}_j, \varepsilon)$ covers all of $K(\mathbf{x})$.

(b) We observe first that since $\mathrm{pr}_{n,\perp} \mathbf{x}$ is in the center of $G^2(E)$, then by left invariance

$$\rho(\mathbf{x}) = \rho((\mathrm{pr}_n \mathbf{x}) \cdot (\mathrm{pr}_{n,\perp} \mathbf{x})) \leq \rho(\mathrm{pr}_n \mathbf{x}) + \rho(\mathrm{pr}_{n,\perp} \mathbf{x})$$

Let $\mathbf{\Gamma} = \mathbf{\Gamma}_u = \gamma + \gamma^{(2)} : [0, 1] \to G^2(E)$ be any minimizing geodesic with left logarithmic derivative u and write $u_n = \mathrm{pr}_n u$ and $u_{\perp,n} = \mathrm{pr}_{n,\perp} u$. Since u is a minimizing geodesic, then by reparametrization, we may assume that

$$\|u(t)\|_E = \sqrt{\|u_n(t)\|_E^2 + \|u_{n,\perp}(t)\|_E^2} = \rho(\mathbf{x}), \qquad \text{and note} \tag{4.13}$$

$$\rho(\mathrm{pr}_{n,\perp}\,\mathbf{x}) \leq \mathrm{Length}(\mathrm{pr}_{n,\perp}\,\mathbf{\Gamma}) = \int_0^1 \|u_{n,\perp}(t)\|_E\,dt \leq \rho(\mathbf{x}). \tag{4.14}$$

This leads to the following sequence of inequalities

$$\begin{aligned}
&\rho(\mathbf{x})\rho(\mathrm{pr}_n\,\mathbf{x}) \geq \rho(\mathbf{x})\left(\rho(\mathbf{x}) - \rho(\mathrm{pr}_{n,\perp}\,\mathbf{x})\right) \\
&\overset{(4.13)+(4.14)}{\geq} \rho(\mathbf{x}) \int_0^1 \left(\sqrt{\|u_n(t)\|_E^2 + \|u_{n,\perp}(t)\|_E^2} - \|u_{n,\perp}(t)\|_E\right) dt \\
&\overset{(4.13)}{=} \rho(\mathbf{x}) \int_0^1 \left(\frac{\|u_n(t)\|_E^2}{\sqrt{\|u_n(t)\|_E^2 + \|u_{n,\perp}(t)\|_E^2} + \|u_{n,\perp}(t)\|_E} \right) dt \\
&\overset{(4.14)}{\geq} \frac{1}{2}\int_0^1 \|u_n(t)\|_E^2\,dt \overset{\text{Jensen}}{\geq} \frac{1}{2}\left(\int_0^1 \|u_n(t)\|_E\,dt\right)^2 = \frac{1}{2}\mathrm{Length}(\mathrm{pr}_n\,\mathbf{\Gamma})^2.
\end{aligned}$$

It follows that any point $\mathbf{y}$ on the curve $\mathbf{\Gamma}$ will have $\rho(\mathrm{pr}_n\,\mathbf{y}) \leq \sqrt{2\rho(\mathbf{x})\rho(\mathrm{pr}_n\,\mathbf{x})}$. By a similar calculation, we have that $\rho(\mathrm{pr}_{n,\perp}\,\mathbf{y}) \leq \sqrt{2\rho(\mathbf{x})\rho(\mathrm{pr}_{n,\perp}\,\mathbf{x})}$.

We also see that $\mathrm{pr}_{n,\wedge}\,\mathbf{\Gamma}(t) = \mathrm{pr}_{n,\wedge}\,\gamma^{(2)}(t)$ with

$$\mathrm{pr}_{n,\wedge}\,\gamma^{(2)}(t) = \frac{1}{2}\int_0^t ((\mathrm{pr}_n\,\gamma(s)) \wedge u_{n,\perp}(s) + (\mathrm{pr}_{n,\perp}\,\gamma(s)) \wedge u_n(s))ds.$$

We note that $\int_0^t \|u_n(s)\|\,dt \leq \mathrm{Length}(\mathrm{pr}_n\,\mathbf{\Gamma})$, while $\|\mathrm{pr}_n\,\gamma(t)\| \leq \rho(\mathrm{pr}_n\,\gamma(t)) \leq \mathrm{Length}(\mathrm{pr}_n\,\mathbf{\Gamma})$. Since we have similar relation applying pr_n. We finally use that

$$\|\mathrm{pr}_\wedge\,\gamma^{(2)}(t)\|_\otimes \leq \mathrm{Length}(\mathrm{pr}_n\,\mathbf{\Gamma}(t))\,\mathrm{Length}(\mathrm{pr}_{n,\perp}\,\mathbf{\Gamma}(t)) \leq \sqrt{2\rho(\mathbf{x})^3\rho(\mathrm{pr}_{n,\perp}\,\mathbf{x})}.$$

Hence the geodesic satisfies the pointwise bounds from the definition of $K(\mathbf{x})$ and the result follows. □

Lemma 4.8. *For any $r > 0$ and $\mathbf{x}_0 \in G^2(E)$, we have that the set*

$$\bar{B}_\rho(\mathbf{x}_0, r) = \{\mathbf{x} : \rho(\mathbf{x}_0, \mathbf{x}) \leq r\}$$

is closed in $G^2(E)$, that is, with respect to the metric d.

Proof. By left invariance, we only consider $\mathbf{x}_0 = 0$. Assume that $\mathbf{y}^n = y^n + y^{n,(2)}$ is a sequence contained in $\bar{B}_\rho(0, r)$ converging in $G^2(E)$ to some element $\mathbf{y} = y + y^{(2)}$. We then have that

$$\|y - y^n\|_E \to 0, \qquad \|y^{(2)} - y^{n,(2)}\|_\otimes \to 0.$$

In particular, we will have $\|y^{(2)} - y^{n,(2)}\|_{\mathrm{Sch}^\infty} \to 0$ implying that if $\sigma(y^{(2)}) = (\sigma_j)$ and $\sigma(y^{n,(2)}) = (\sigma_j^n)$, then $\sigma_j^n \to \sigma_j$. It follows that

$$|||\mathbf{y}||| \le \|y\|_E + \sqrt{\pi}\sum_{j=1}^{\infty} j\sigma_j = \lim_{n\to\infty}\|y^n\|_E + \sqrt{\pi}\lim_{m\to\infty}\left(\lim_{n\to\infty}\sum_{j=1}^{m} j\sigma_j^n\right) \le 2r.$$

Since $|||\mathbf{y}||| < \infty$, we can conclude the following.

If $y^{(2)} = \sum_{j=1}^\infty \sigma_j X_j \wedge Y_j$ is defined with all vectors orthonormal, we define $F_m = \operatorname{span}\{y, X_1, Y_1, \dots, X_m, Y_m\}$. Then

$$\lim_{m\to\infty}\rho(\mathrm{pr}_{F_m}\mathbf{y}, \mathbf{y}) \le 3\lim_{m\to\infty}|||\mathbf{y} - \mathrm{pr}_{F_m}\mathbf{y}||| = 0.$$

Using that $\rho(0, \mathrm{pr}_{F_m}\mathbf{y}) \le \rho(0,\mathbf{y}) \le \rho(0, \mathrm{pr}_{F_m}\mathbf{y}) + \rho(\mathrm{pr}_{F_m}\mathbf{y}, \mathbf{y})$, it follows that

$$\lim_{m\to\infty}\rho(0, \mathrm{pr}_{F_m}\mathbf{y}) = \rho(0,\mathbf{y}).$$

Furthermore, since

$$\|\mathrm{pr}_{F_m}(\mathbf{y} - \mathbf{y}^n)\|_{\mathrm{Sch}^\infty} \le \|\mathbf{y} - \mathbf{y}^n\|_{\mathrm{Sch}^\infty} \le \|\mathbf{y} - \mathbf{y}^n\|_\otimes,$$

we have that $\lim_{n\to\infty}\|\mathrm{pr}_{F_m}(\mathbf{y} - \mathbf{y}^n)\|_{\mathrm{Sch}^\infty} = 0$. Since all left invariant homogeneous norms are equivalent on a finite dimensional space [32, Proposition 10], we have convergence $\lim_{n\to\infty}\rho(\mathrm{pr}_{F_m}\mathbf{y}^n, \mathrm{pr}_{F_m}\mathbf{y}) \to 0$ for any fixed m. Finally

$$\rho(0,\mathbf{y}) = \lim_{m\to\infty}\rho(0, \mathrm{pr}_{F_m}\mathbf{y}) = \lim_{m\to\infty}\lim_{n\to\infty}\rho(0, \mathrm{pr}_{F_m}\mathbf{y}^n) \le r,$$

so $\mathbf{y} \in \bar{B}_\rho(0, r)$. $\square$

Proof of Theorem 4.6, Part II. We are now ready to complete the proof.

Step 5: Every point has a midpoint. Let $\mathbf{x} = x + x^{(2)} = x + \sum_{j=1}^\infty \sigma_j X_j \wedge Y_j \in G^2_{cc}(E)$ be arbitrary and define F_n as in (4.11). If we write $\mathbf{x}^n = \mathrm{pr}_{F_n}\mathbf{x}$, then by the definition in (4.12), we have that $K(\mathbf{x}^n) \subseteq K(\mathbf{x})$. Since $\mathbf{x}^n \in G^2_a(E)$, there exists a length minimizing geodesic $\mathbf{\Gamma}^n$ from 0 to $\mathbf{x}^n$, which we know is in $K(\mathbf{x})$ by Lemma 4.7.

Let $\mathbf{s}^n$ denote the midpoint of each geodesic $\mathbf{\Gamma}^n$. This satisfies

$$\rho(0, \mathbf{s}^n) = \rho(\mathbf{x}^n, \mathbf{s}^n) = \frac{1}{2}\rho(0, \mathbf{x}^n) \le \frac{1}{2}\rho(0,\mathbf{x}) := r.$$

Write $\delta_m = \rho(\mathbf{x}^m, \mathbf{x})$, and define balls

$$\bar{B}_0 = \{\mathbf{y} \in G^2(E) : \rho(0,\mathbf{y}) \le r\},$$
$$\bar{B}_m = \{\mathbf{y} \in G^2(E) : \rho(\mathbf{x},\mathbf{y}) \le r + \delta_m\}.$$

By the definition of $\mathbf{x}^n$, we have $\mathbf{s}^n \in \bar{B}_0 \cap \bar{B}_m$ for any $n \geq m$ with $\delta_m \to 0$.

Since every $\mathbf{s}^m$ is contained in $K(\mathbf{x})$, by compactness, there is a subsequence $\mathbf{s}^{n_k}$ converging to a point $\mathbf{s}$ in $G^2(E)$. This element hence has to be contained in $\bar{B}_0 \cap \bar{B}_m$ for any $m \geq 1$ by Lemma 4.8. It follows that

$$\rho(0, \mathbf{s}) = \rho(\mathbf{x}, \mathbf{s}) = \frac{1}{2}\rho(0, \mathbf{x}),$$

i.e., $\mathbf{s}$ is a midpoint of $\mathbf{x}$. Since $(G^2_{cc}(E), \rho)$ is a complete length space, it follows from left-invariance of the metric together with [3, Theorem 2.4.16] that existence of such midpoint for any element is equivalent to the space being a geodesic space. This completes the proof. □

4.5. Proof of Theorem 1.1

We now come to the proof of our main result. Namely, if E is a Hilbert space and we define α-weak geometric rough path relative to the tensor norm $\|\cdot\|_{\mathrm{Sch}^p}$, $1 \leq p \leq \infty$ on the tensor product, then for $\beta \in (1/3, \alpha)$

$$\mathscr{C}^\alpha_g([0, T], G^2(E)) \subset \mathscr{C}^\alpha_{wg}([0, T], G^2(E)) \subset \mathscr{C}^\beta_g([0, T], G^2(E)).$$

We can prove this by showing that the conditions (I), (II) and (III) in Theorem 3.3 are satisfied. By Theorem 4.6 it follows that (I) and (II) are satisfied for Hilbert spaces. Hence, we only need to prove that condition (III) holds.

Recall the results of Lemma 4.5. Let $\mathbf{x} = x + x^{(2)} \in C^\alpha([0, T], G^2(E))$ be an arbitrary weakly geometric α-rough path. For any fixed t, define a sequence of increasing finite subspaces $\{F_{t,n}\}_{n=1}^\infty$, such that

$$x_t \in F_{t,n} \qquad d(\mathbf{x}_t, \mathrm{pr}_{F_{t,n}} \mathbf{x}_t) = d\left(0, \mathrm{pr}_{F_{t,n}^\perp} x^{(2)}\right) \leq \frac{1}{n}.$$

Consider a partition $\Pi = \{t_0 = 0 < t_1 < t_2 < \cdots < t_k = T\}$ of the interval $[0, T]$. Write

$$F_{\Pi,n} = \mathrm{span}\{F_{t,n} : t \in \Pi\}.$$

Define $\mathbf{x}^{\Pi,n}_t = \mathrm{pr}_{F_{\Pi,n}} \mathbf{x}_t$. Since ρ and d are equivalent on the finite dimensional $F_{\Pi,n}$, it follows that $\mathbf{x}^{\Pi,n}_t$ is a continuous function in $G^2_{cc}(E)$ with respect to ρ.

Since $\mathbf{x}_t$ is uniformly continuous, we can for each $r > 0$ find a number o_r such that

$$o_r = \mathrm{Osc}(\mathbf{x}_t; r) = \sup_{0 \leq s < t \leq T t - s \leq r} d(\mathbf{x}_s, \mathbf{x}_t),$$

with o_r approaching 0 as $r \to 0$. We now see that for every $t \in [t_i, t_{i+1}]$,

$$\begin{aligned}
d(\mathbf{x}^{\Pi,n}_t, \mathbf{x}_t) &\leq d(\mathbf{x}^{\Pi,n}_t, \mathbf{x}^{\Pi,n}_{t_i}) + d(\mathbf{x}^{\Pi,n}_{t_i}, \mathbf{x}_{t_i}) + d(\mathbf{x}_{t_i}, \mathbf{x}_t) \\
&\leq 2d(\mathbf{x}_t, \mathbf{x}_{t_i}) + d(0, \mathrm{pr}_{F^\perp_{\Pi,n}} x^{(2)}_{t_i}) \\
&\leq 2d(\mathbf{x}_t, \mathbf{x}_{t_i}) + d(0, \mathrm{pr}_{F^\perp_{t_i,n}} x^{(2)}_{t_i}) \leq 2o_{|\Pi|} + \frac{1}{n}
\end{aligned}$$

Defining $\mathbf{x}^n_t = \mathbf{x}^{n,\Pi_n}_t$ where Π_n is a partition with $|\Pi_n| = \frac{1}{n}$, we have that $d(\mathbf{x}^n_t, \mathbf{x}_t)$ converges uniformly to 0. Note that by Lemma 4.5 (a), the sequence $\mathbf{x}^n_t$ consist of α-Hölder curves, so by using this property of $\mathbf{x}^n$ and $\mathbf{x}$ and an interpolation argument as in the proof of Theorem 3.3, we obtain that $d_\beta(\mathbf{x}, \mathbf{x}^n) \to 0$ for any $\beta \in (\frac{1}{3}, \alpha)$. This completes the proof.

Remark 4.9 *(Other cross-norms).* As one can see from the proof in Section 4.5, what is needed for our result is the properties of Lemma 4.5 and Theorem 4.6. Hence, for any norm on the tensor product which satisfy these two results, the result in Theorem 1.1 holds.

4.6. Generalizing the result to Banach spaces

One of the central tools in our proof for geometric rough paths when E is a Hilbert space, is that we can use orthogonal projections $\mathrm{Pr}_F : E \to F$, which all shorten lengths and hence have norm 1. Such contractive projections are in general rare in Banach spaces as we have the following characterization from [36, Theorem 3.1].

Theorem 4.10. *For a Banach space E with $\dim E \geq 3$, the following statements are equivalent:*

(i) E is isometrically isomorphic to a Hilbert space,
(ii) every 2-dimensional subspace of E is the range of a projection of norm 1,
(iii) every subspace of E is the range of a projection of norm 1.

Data availability

No data was used for the research described in the article.

References

[1] A. Agrachev, D. Barilari, U. Boscain, A Comprehensive Introduction to Sub-Riemannian Geometry, Cambridge Studies in Advanced Mathematics, vol. 181, Cambridge University Press, Cambridge, 2020, From the Hamiltonian viewpoint, with an appendix by Igor Zelenko.
[2] I. Bailleul, Flows driven by Banach space-valued rough paths, in: Séminaire de Probabilités XLVI, in: Lecture Notes in Math., vol. 2123, Springer, Cham, 2014, pp. 195–205.
[3] D. Burago, Y. Burago, S. Ivanov, A Course in Metric Geometry, Graduate Studies in Mathematics, vol. 33, American Mathematical Society, Providence, RI, 2001.
[4] I. Bailleul, M. Gubinelli, Unbounded rough drivers, Ann. Fac. Sci. Toulouse Math. 26 (4) (2017) 795–830, https://doi.org/10.5802/afst.1553, https://hal.archives-ouvertes.fr/hal-01278756.
[5] H. Boedihardjo, X. Geng, T. Lyons, D. Yang, Note on the signatures of rough paths in a Banach space 2015, https://arxiv.org/pdf/1510.04172.pdf.
[6] I. Bailleul, S. Riedel, Rough flows, J. Math. Soc. Jpn. 71 (3) (2019) 915–978, https://doi.org/10.2969/jmsj/80108010.
[7] T. Cass, B.K. Driver, N. Lim, C. Litterer, On the integration of weakly geometric rough paths, J. Math. Soc. Jpn. 68 (4) (2016) 1505–1524, https://doi.org/10.2969/jmsj/06841505.
[8] D. Crisan, D.D. Holm, J.-M. Leahy, T. Nilssen, Variational principles for fluid dynamics on rough paths, arXiv:2004.07829, 2020.
[9] W.-L. Chow, Über Systeme von Linearen Partiellen Differentialgleichungen Erster Ordnung, Math. Ann. 117 (1939) 98–105.
[10] M. Coghi, T. Nilssen, Rough nonlocal diffusions, https://arxiv.org/abs/1905.07270, 2019.
[11] S. Dereich, Rough paths analysis of general Banach space-valued Wiener processes, J. Funct. Anal. 258 (9) (2010) 2910–2936, https://doi.org/10.1016/j.jfa.2010.01.018.

[12] A. Defant, K. Floret, Tensor Norms and Operator Ideals, North-Holland Mathematics Studies, vol. 176, North-Holland Publishing Co., Amsterdam, 1993.

[13] A. Deya, M. Gubinelli, M. Hofmanová, S. Tindel, A priori estimates for rough PDEs with application to rough conservation laws, J. Funct. Anal. 276 (12) (2019) 3577–3645.

[14] I. Ekeland, The Hopf-Rinow theorem in infinite dimension, J. Differ. Geom. 13 (2) (1978) 287–301, http://projecteuclid.org/euclid.jdg/1214434494.

[15] R. Engelking, General Topology, PWN—Polish Scientific Publishers, Warsaw, 1977, Translated from the Polish by the author, Monografie Matematyczne, Tom 60. [Mathematical Monographs, Vol. 60].

[16] P. Friz, N. Victoir, A note on the notion of geometric rough paths, Probab. Theory Relat. Fields 136 (3) (2006) 395–416, https://doi.org/10.1007/s00440-005-0487-7.

[17] P.K. Friz, N.B. Victoir, Multidimensional Stochastic Processes as Rough Paths: Theory and Applications, Cambridge Studies in Advanced Mathematics, Cambridge University Press, 2010.

[18] H. Glöckner, Algebras whose groups of units are Lie groups, Stud. Math. 153 (2) (2002) 147–177, https://doi.org/10.4064/sm153-2-4.

[19] H. Glöckner, Lie groups of measurable mappings, Can. J. Math. 55 (5) (2003) 969–999, https://doi.org/10.4153/CJM-2003-039-9.

[20] H. Glöckner, Measurable regularity properties of infinite-dimensional Lie groups, https://arxiv.org/pdf/1601.02568.pdf, 2015.

[21] H. Glöckner, K.-H. Neeb, When unit groups of continuous inverse algebras are regular Lie groups, Stud. Math. 211 (2) (2012) 95–109, https://doi.org/10.4064/sm211-2-1.

[22] A. Hocquet, M. Hofmanová, An energy method for rough partial differential equations, J. Differ. Equ. 265 (4) (2018) 1407–1466.

[23] M. Hofmanova, J.-M. Leahy, T. Nilssen, On a rough perturbation of the Navier-Stokes system and its vorticity formulation, Ann. Appl. Probab. 31 (2) (2021) 736–777, https://doi.org/10.1214/20-AAP1603, arXiv:1902.09348.

[24] M. Hofmanová, J.-M. Leahy, T. Nilssen, On the Navier-Stokes equation perturbed by rough transport noise, J. Evol. Equ. 19 (1) (2019) 203–247, https://doi.org/10.1007/s00028-018-0473-z.

[25] A. Hocquet, T. Nilssen, An Itô formula for rough partial differential equations. Application to the maximum principle, Potential Anal. (2020), https://doi.org/10.1007/s11118-020-09830-y.

[26] A. Hocquet, T. Nilssen, W. Stannat, Generalized Burgers equation with rough transport noise, Stoch. Process. Appl. 130 (4) (2020) 2159–2184, https://doi.org/10.1016/j.spa.2019.06.014, http://www.sciencedirect.com/science/article/pii/S030441491830721X.

[27] A. Hocquet, Quasilinear rough partial differential equations with transport noise, arXiv:1808.09867, 2018.

[28] H.H. Keller, Differential Calculus in Locally Convex Spaces, Lecture Notes in Mathematics, vol. 417, Springer-Verlag, Berlin-New York, 1974.

[29] A. Kriegl, P.W. Michor, Regular infinite-dimensional Lie groups, J. Lie Theory 7 (1) (1997) 61–99.

[30] H. Kunita, Stochastic Flows and Stochastic Differential Equations, Vol. 24, Cambridge University Press, 1997.

[31] T.J. Lyons, M. Caruana, T. Lévy, Differential Equations Driven by Rough Paths, Lecture Notes in Mathematics, vol. 1908, Springer, Berlin, 2007, Lectures from the 34th Summer School on Probability Theory held in Saint-Flour, July 6–24, 2004, with an introduction concerning the Summer School by Jean Picard.

[32] A. Lejay, Yet another introduction to rough paths, in: Séminaire de Probabilités XLII, in: Lecture Notes in Math., vol. 1979, Springer, Berlin, 2009, pp. 1–101.

[33] T.J. Lyons, Differential equations driven by rough signals, Rev. Mat. Iberoam. 14 (2) (1998) 215–310, https://doi.org/10.4171/RMI/240.

[34] R. Meise, D. Vogt, Introduction to Functional Analysis, Oxford Graduate Texts in Mathematics, vol. 2, The Clarendon Press, Oxford University Press, New York, 1997, Translated from the German by M.S. Ramanujan and revised by the authors.

[35] K.-H. Neeb, Towards a Lie theory of locally convex groups, Jpn. J. Math. 1 (2) (2006) 291–468, https://doi.org/10.1007/s11537-006-0606-y.

[36] B. Randrianantoanina, Norm-One Projections in Banach Spaces, vol. 5, 2001, pp. 35–95, International Conference on Mathematical Analysis and its Applications (Kaohsiung, 2000).

[37] P.K. Rashevskiĭ, On the connectability of two arbitrary points of a totally nonholonomic space by an admissible curve, Uchen. Zap. Mosk. Ped. Inst. Ser. Fiz.-Mat. Nauk 3 (2) (1938) 83–94.

[38] C. Reutenauer, Free Lie Algebras, London Mathematical Society Monographs. New Series, vol. 7, The Clarendon Press, Oxford University Press, New York, 1993, Oxford Science Publications.

[39] L. Rizzi, U. Serres, On the cut locus of free, step two Carnot groups, Proc. Am. Math. Soc. 145 (12) (2017) 5341–5357, https://doi.org/10.1090/proc/13658.

[40] R.A. Ryan, Introduction to Tensor Products of Banach Spaces, Springer Monographs in Mathematics, Springer-Verlag London, Ltd., London, 2002.
[41] A. Schmeding, An Introduction to Infinite-Dimensional Differential Geometry, Cambridge Studies in Advanced Mathematics, Cambridge University Press, 2022, in press, arXiv:2112.08114 [math.DG].