

Characteristic functions of path signatures and applications



Ilya Chevyrev
Lincoln College
University of Oxford

A thesis submitted for the degree of
Doctor of Philosophy

Trinity 2015

Abstract

The main object of study in this work is the extension of the classical characteristic function to the setting of path signatures. Our first fundamental result exhibits the following geometric interpretation: the path signature is completely determined by the development of the path into compact Lie groups. This faithful representation of the signature is the primary tool we use to define and study the characteristic function.

Our investigation of the characteristic function can be divided into two parts. First, we employ the characteristic function to study the expected signature of a path as the natural generalisation of the moments of a real random variable. In this direction, we provide a solution to the moment problem, and study analyticity properties of the characteristic function. In particular, we solve the moment problem for signatures arising from families of Gaussian and Markovian rough paths.

Second, we study the characteristic function in relation to the solution map of a rough differential equation. The connection stems from the fact that the signature of a geometric rough path completely determines the path's role as a driving signal. As an application, we demonstrate that the characteristic function can be used to determine weak convergence of flows arising from rough differential equations.

Along the way, we develop tools to study càdlàg processes as rough paths and to determine tightness in p -variation topologies of random walks. As a consequence, we provide a classification of Lévy processes possessing sample paths of finite p -variation and determine a Lévy-Khintchine formula for the characteristic function of the signature of a Lévy process.

Acknowledgements

This work could never have been carried out without the guidance of my advisor Terry Lyons, to whom I am sincerely grateful. His strive for a rigorous, deep, and comprehensive approach in all aspects of mathematics has served as a model for my own mathematical development.

It is also my pleasure to thank my friends and colleagues, and especially my family, who always remained available for support when it was needed.

Contents

1	Introduction	1
2	Universal locally m-convex algebra	5
2.1	Definition and topological properties	5
2.2	Group-like elements	10
2.3	Representations	14
2.3.1	Separation of points	15
2.3.2	Polynomial identities over Lie algebras	16
3	Rough paths and signatures	19
3.1	Rough paths	19
3.1.1	Geometric rough paths	23
3.2	Finite dimensional case	24
3.2.1	Homogeneous norms and the $d_{p\text{-var}}$ metric	25
3.2.2	Hölder reparametrisations	27
3.3	Rough differential equations	31
3.3.1	Lipschitz vector fields	32
3.3.2	Universal limit theorem	32
3.3.3	Non-explosion	35
3.3.4	Linear vector fields	36
3.4	Flows of diffeomorphisms	36
3.4.1	Flows induced by linear vector fields	38
3.5	Euler approximations	43
4	Expected signature	45
4.1	The moment problem	45
4.1.1	Gaussian rough paths	49
4.1.2	Markovian rough paths	51
4.2	Analyticity	54

4.2.1	Markovian rough paths stopped upon exiting a domain	59
4.3	Method of moments	66
5	Random walks in path space and Lévy processes	69
5.1	Iid Arrays and Lévy processes in Lie groups	70
5.1.1	Iid Arrays	70
5.1.2	Lévy processes	71
5.1.3	Convergence towards a Lévy process	72
5.1.4	Scaling functions	74
5.1.5	Convergence towards the generator	77
5.1.6	Maps between Lie groups	79
5.1.7	Approximating walk	80
5.1.8	Infinite p -variation	84
5.2	Random walks in path space	90
5.2.1	Tightness of concatenations	91
5.2.2	Tightness of the random walk	93
5.2.3	Example - random walk connected by a rough path	104
5.3	Path Functions	112
5.3.1	Definition	112
5.3.2	The connecting function on the Skorokhod space	113
5.3.3	Measurability and continuity in the Skorokhod topology	114
5.3.4	p -variation	118
5.3.5	Path functions on $G^N(\mathbb{R}^d)$	119
5.3.6	Globally continuous path function on $G^N(\mathbb{R}^d)$	121
5.4	Lévy processes	123
5.4.1	Finite p -variation of Lévy processes	123
5.4.2	The Lévy-Khintchine formula	125
5.4.3	Convergence towards a Lévy process	128
6	Several open problems	137
	Bibliography	140

Chapter 1

Introduction

A path can be regarded as a representation of an ordered sequence of events. On the one hand, it is a sufficiently versatile and, at least on the surface, simple object, that it has become ubiquitous throughout mathematics. On the other hand, even in the simplest settings, the study of paths can become an extremely deep theory which can be analysed from a wide variety of approaches.

In this work, we will be primarily concerned with paths taking values in a vector space and which play the role of a driving signal in a differential equation. This is of course an extremely diverse area of study in itself. For example, the extension of differential calculus from bounded variation paths to rougher processes has been one of central problems in stochastic analysis over the last century. However, even for a smooth path, determining its effects on a differential equation can be far from straightforward. It thus becomes of great value to find a simple representation of a path which would describe its role as a driving signal.

For sufficiently regular paths, Chen [13] first demonstrated that such a description was given precisely by the sequence of iterated integrals of the path, which has come to be known as the path signature. Chen moreover showed that the set of path signatures possesses a canonical group structure which respects the “concatenation product” of the underlying paths. Many authors have since contributed to our understanding of the signature and it has become one of the central objects of study in the theory of rough paths.

We take a moment to consider the case of a one-dimensional path. The signature of a bounded variation path $\mathbf{x} : [0, T] \mapsto \mathbb{R}$ is given precisely by the sequence of the powers of its increment

$$S(\mathbf{x})_{0,T} := \left(1, \mathbf{x}_T - \mathbf{x}_0, \frac{(\mathbf{x}_T - \mathbf{x}_0)^2}{2!}, \frac{(\mathbf{x}_T - \mathbf{x}_0)^3}{3!}, \dots \right). \quad (1.0.1)$$

We would like to mention three points about this expression. First, for a one-dimensional path, $S(\mathbf{x})_{0,T}$ depends only on the increment of the path $\mathbf{x}_T - \mathbf{x}_0$; this reflects the fact that flowing along a vector field for time t in the first direction and then for time s in the second direction, produces the same effect as first flowing for time s in the second direction and then for time t in the first direction. For multi-dimensional paths, this commutative property no longer holds true, which reflects the fact that when flowing along two *different* vector fields, order becomes important.

Second, observe the factorial decay of the terms in (1.0.1). If one considers a Banach space W and a bounded linear operator $M : W \mapsto W$, it follows that the series

$$M[S(\mathbf{x})_{0,T}] := \sum_{k=0}^{\infty} \frac{M^k(\mathbf{x}_T - \mathbf{x}_0)^k}{k!} \quad (1.0.2)$$

converges rapidly to $\exp(M(\mathbf{x}_T - \mathbf{x}_0))$, which is nothing but the solution of the differential equation

$$d\Phi_t = M(d\mathbf{x}_t) \circ \Phi_t, \quad \Phi_0 = Id_W. \quad (1.0.3)$$

In turn, the corresponding linear map $\Phi_T : W \mapsto W$ is precisely the flow map, sending the starting point $\mathbf{y}_0$ to the endpoint $\mathbf{y}_T$, of the following differential equation

$$d\mathbf{y}_t = M(d\mathbf{x}_t)(\mathbf{y}_t), \quad \mathbf{y}_0 \in W.$$

In contrast to commutativity, the factorial decay of the signature remains true for any bounded variation path $\mathbf{x} : [0, T] \mapsto V$ taking values in a Banach space V . It follows that for a linear map $M : V \mapsto \mathbf{L}(W)$, replacing $(\mathbf{x}_T - \mathbf{x}_0)^k$ by the k -th iterated integral of $\mathbf{x}$ in (1.0.2), and canonically extending M to a linear map defined on the tensor algebra of V , the series $M[S(\mathbf{x})_{0,T}]$ remains a valid expansion for the solution Φ_T .

Furthermore, if one considers a linear map $M : V \mapsto \mathfrak{g}$ taking values in a Lie subalgebra $\mathfrak{g}$ of $\mathbf{L}(\mathbb{R}^e)$, then Φ_T is precisely the Cartan development of $\mathbf{x}$ in the corresponding matrix Lie group G . The important observation here is that any linear map $M : V \mapsto \mathfrak{g}$ canonically induces a representation of the signature group on the Lie group G .

For much rougher paths, the existence of a well-defined signature, as well as its factorial decay, was first shown by Lyons [45] in his introduction of rough paths theory. This decay property of the signature, particularly in connection with linear differential equations, can serve as one of the first stepping stones into the theory of rough paths.

Our final point about (1.0.1), is that if $\mathbf{x}$ is a random variable, then the classical characteristic function of $\mathbf{x}_T - \mathbf{x}_0$ is given by the expectation of

$$\exp(i\lambda(\mathbf{x}_T - \mathbf{x}_0)) = \sum_{k=0}^{\infty} \frac{(i\lambda)^k (\mathbf{x}_T - \mathbf{x}_0)^k}{k!},$$

where λ varies across $\mathbb{R}$. The point of view we emphasise, is that for every $\lambda \in \mathbb{R}$, $i\lambda : x \mapsto i\lambda x$ is a map from $\mathbb{R}$ into $\mathbf{L}(\mathbb{C})$ which takes values in the Lie algebra of the unitary circle group $\mathbb{T}$, and thus induces a unitary representation of the signature group.

In the study of probability theory on algebraic structures, the natural extension of the characteristic function to a random variable X , taking values in a topological group G , is given by $\phi_X(M) := \mathbb{E}[M(X)]$, where M is a unitary representation of G . Under suitable conditions, particularly the existence of sufficiently many unitary representations, ϕ_X is known to uniquely determine the law of X .

Given a linear map $M : V \mapsto \mathfrak{u}$, where $\mathfrak{u}$ is a unitary Lie algebra, it is the remark that M induces a unitary representation of the signature group which serves as the first step into our definition and study of characteristic functions of path signatures.

Layout

We take a moment to briefly summarise the structure of the document. At the beginning of each chapter, we also provide a brief summary of its contents.

In Chapter 2 we introduce a universal topological algebra $E(V)$ associated to a vector space V , in which we embed the signatures of V -valued paths. The primary motivation for a topology on the group of signatures is to make sense of what we mean by a “random signature”. The induced topology, however, exhibits a natural interpretation in terms of differential equations: a sequence of signatures converges if and only if the solution to (1.0.3) converges for every continuous linear map $M : V \mapsto \mathbf{L}(W)$.

In Chapter 3 we recall the fundamental results of rough paths theory. The main link with Chapter 2 is that the signature of a rough path can always be identified with an element of $E(V)$. We highlight the recent work of Boedihardjo, Geng, Lyons and Yang [4] on the classification of rough paths through their signatures, particularly in connection with our work on the characteristic function and weak convergence of solutions to rough differential equations.

Chapter 4 is devoted to the study of the expected value of the signature. The main goal of this chapter is to emphasise the point of view that the expected signature is

the natural generalisation of the moments of a real random variable. In particular, we explore solutions to the moment problem, study analyticity properties of the characteristic function, and show a method of moments for convergence in law of random variables. We demonstrate how signatures arising from classes of Gaussian and Markovian rough paths fit into our framework.

In Chapter 5 we investigate problems concerning weak convergence of rough paths and of rough differential equations. The primary problem we address is to determine conditions under which RDEs driven by concatenations of geometric rough paths converge in law. In passing, we investigate one of the systematic ways to turn càdlàg paths into rough paths, and ultimately apply our results to Lévy processes and stochastic flows.

In Chapter 6 we collect several open problems relating to the current work which we believe to be of interest.

Chapter 2

Universal locally m -convex algebra

The purpose of this chapter is to define the algebra $E(V)$ and collect its basic topological and algebraic properties. The importance of the space $E(V)$ comes from the fact that the signature of a path can always be realised as an element of $E(V)$. This connection will be studied in detail in Chapter 3.

In Section 2.1 we define $E(V)$ as a locally convex space and record some of its basic topological properties. In Section 2.2 we focus on probability measures on $G(V)$, the set of group-like elements of $E(V)$.

In Section 2.3 we study representations of $E(V)$. Our first main result is Theorem 2.3.8 which describes explicitly a family of representations of $E(\mathbb{R}^d)$ which preserves unitary elements and separates the points. An immediate consequence is that one is able to define a meaningful characteristic function for $G(\mathbb{R}^d)$ -valued random variables (Corollary 2.3.12).

For a large part of the later chapters, we will consider only the case $V = \mathbb{R}^d$. However, throughout this chapter we work in the general setting of locally convex spaces as we find this provides the most natural framework to interpret the topology of $E(V)$. We shall make precise whenever finite dimensionality (or further properties) of V are required.

2.1 Definition and topological properties

Unless otherwise stated, we shall assume that all vector spaces are real and all algebras are unital. For topological vector spaces V, W , let $\mathbf{L}(V, W)$ be the space of continuous linear maps from V to W , and denote $\mathbf{L}(V) = \mathbf{L}(V, V)$ and $V' = \mathbf{L}(V, \mathbb{R})$. For terminology and basic properties of topological algebras we refer to [47].

For a topological vector space V , a topological algebra A , and a topology on the tensor algebra $T(V) := \bigoplus_{k \geq 0} V^{\otimes k}$, consider the statement:

$$\text{For all } M \in \mathbf{L}(V, A), \text{ the extension } M : T(V) \mapsto A \text{ is continuous.} \quad (2.1.1)$$

One may then topologize $T(V)$ by requiring that (2.1.1) holds for all topological algebras A of a given category. We consider here the category of locally m -convex algebras whose definition we review.

A family of semi-norms Ψ on a locally convex space V is called *fundamental* if for every semi-norm ξ on V , there exist $\gamma \in \Psi$ and $\varepsilon > 0$ such that $\varepsilon \xi \leq \gamma$ (note that by a semi-norm we always mean a continuous semi-norm).

Definition 2.1.1. A locally m -convex algebra A is an algebra equipped with a locally convex topology for which there exists a fundamental family of defining semi-norms Ψ such that for all $\gamma \in \Psi$

$$\gamma(xy) \leq \gamma(x)\gamma(y) \quad \forall x, y \in A. \quad (2.1.2)$$

When a semi-norm γ satisfies (2.1.2), we say it is *sub-multiplicative*. An locally m -convex algebra A is called *normed* if there exists a single sub-multiplicative norm $\|\cdot\|$ which defines the topology of A .

Definition 2.1.2. Let V be a locally convex space. Let $E_a(V) = T(V)$ equipped with the coarsest topology such that (2.1.1) holds for all locally m -convex algebras A (or equivalently, all normed algebras A). Denote by $E(V)$ the completion of $E_a(V)$.

Thus for any normed algebra A , the set $\text{Hom}(E_a, A)$ of continuous algebra homomorphisms is in bijection with $\mathbf{L}(V, A)$. For any $M \in \mathbf{L}(V, A)$ we shall usually denote by the same letter M the corresponding element in $\text{Hom}(E_a, A)$, but shall write $M_E \in \text{Hom}(E_a, A)$ whenever a clear distinction is needed.

Though in all later chapters we shall assume that V is normed, most results in this chapter are more easily understood for locally convex spaces and so unless stated otherwise, for the remainder of the chapter we let V denote a locally convex space.

In most of our notation, we shall drop the reference to V when it is clear from the context. It holds that E_a and E are locally m -convex algebra ([47] p.14, p.22). While most results in this section are stated for E , it is easy to verify which remain valid for E_a .

This method to obtain a universal topological algebra of a specific category is very natural, and we note that this construction is not new; the same construction

(and essentially Proposition 2.1.4 below) appeared in [15] in relation to cyclic cohomology, while analogous constructions were investigated for locally convex algebras with continuous multiplication in [58] and for commutative locally m -convex algebras (particularly in relation to nuclear spaces) in [18] Section 6.4.

Remark 2.1.3. If we start with V as a general topological vector space, an easy verification shows that we arrive at the same space E_a as when we equip V with the finest locally convex topology coarser than its original.

For semi-norms γ, ξ on locally convex spaces V, W respectively, let $\gamma \otimes \xi$ denote the projective semi-norm on $V \otimes W$. Denote by $V \otimes_\pi W$ the projective tensor product and $V \widehat{\otimes} W$ its completion. For a normed space F , and $M \in \mathbf{L}(V, F)$ denote $\gamma(M) = \sup_{\gamma(v)=1} \|Mv\|$ (possibly infinite).

For any collection of semi-norms Ψ on V , define $\Psi^* = \{n\gamma \mid n \geq 1, \gamma \in \Psi\}$. Define the *projective extension* of a semi-norm γ on V as the semi-norm $\exp(\gamma) = \sum_{k \geq 0} \gamma^{\otimes k}$ on E_a . Remark that $\exp(\gamma)$ is a sub-multiplicative semi-norm on E_a . Moreover for any normed algebra A , $M \in \mathbf{L}(V, A)$, and a semi-norm γ on V such that $\gamma(M) \leq 1$, it holds that $\exp(\gamma)(M_E) \leq 1$. We thus readily obtain the following.

Proposition 2.1.4. *Let Ψ be a family of semi-norms on V . Then Ψ is a fundamental family of semi-norms on V if and only if $\exp(\Psi^*)$ is a fundamental family of semi-norms on E .*

Corollary 2.1.5. *The space E is Hausdorff (resp. metrizable, separable) if and only if V is Hausdorff (resp. metrizable, separable).*

Whenever we speak of a topological space, we shall henceforth always assume it is Hausdorff. The following result identifies E with a subspace of $P(V) := \prod_{k \geq 0} V^{\widehat{\otimes} k}$. For $x \in P$, we write x^k for the projection of x onto $V^{\widehat{\otimes} k}$, so that $x = (x^0, x^1, x^2, \dots)$.

Corollary 2.1.6. *Let Ψ be a fundamental family of semi-norms on V . Then $E = \{x \in P \mid \forall \gamma \in \Psi^*, \sum_{k \geq 0} \gamma^{\otimes k}(x^k) < \infty\}$.*

By noting the identification $P^{\widehat{\otimes} 2} = \prod_{i,j \geq 0} V^{i,j}$, where $V^{i,j} \cong V^{\widehat{\otimes}(i+j)}$, the same considerations show that

$$E^{\widehat{\otimes} 2} = \{x \in P^{\widehat{\otimes} 2} \mid \forall \gamma \in \Psi^*, \sum_{i,j \geq 0} \gamma^{\otimes(i+j)}(x^{i,j}) < \infty\}.$$

Let $\rho^k : E \mapsto V^{\widehat{\otimes} k}$ denote the projection $\rho^k(x) = x^k$ and $\pi^k : E \mapsto \bigoplus_{j=0}^k V^{\widehat{\otimes} j}$ the projection $\pi^k(x) = (x^0, \dots, x^k)$. The following result shall also be useful later and is another consequence of Proposition 2.1.4.

Corollary 2.1.7. *The operators $T^{(n)} := \sum_{k=0}^n \rho^k : E \mapsto E$ converge uniformly on bounded sets to the identity operator on E .*

When V is a normed space, we always equip $V^{\otimes k}$ with the projective norm unless stated otherwise. For an element $x \in P$ define its radius of convergence $R(x)$ as the radius of convergence of the series $\sum_{k \geq 0} \|x^k\| \lambda^k$. Corollary 2.1.6 then implies that $x \in E$ if and only if $R(x) = \infty$.

We now come to a more interesting property of the space E . For a semi-normed space (W, γ) denote the quotient normed space $W_\gamma = (W/\text{Ker}(\gamma), \gamma)$ and $\widehat{W}_\gamma$ its completion. For a locally convex space W and a Banach space A , a map $M \in \mathbf{L}(W, A)$ is called compact (resp. nuclear) if there exists a semi-norm γ on W such that the $\gamma(M) < \infty$ and the induced map $M_\gamma : \widehat{W}_\gamma \mapsto A$ is compact (resp. nuclear). Recall that W is called Schwartz (resp. nuclear) if every $M \in \mathbf{L}(W, A)$ is compact (resp. nuclear) for every Banach space A .

Proposition 2.1.8. *The space E is Schwartz (resp. nuclear) if and only if V is Schwartz (resp. nuclear).*

We note that the case when V is simply Schwartz shall not be used later and is recorded simply for completeness. Moreover nuclearity of E shall only be applied in Section 4.3 to the case $V = \mathbb{R}^d$. However the equivalent statement for $V = \mathbb{R}^d$ uses essentially the same proof and thus we record the result in full generality.

Let Ψ be a fundamental family of sub-multiplicative semi-norms of a locally m -convex algebra F . Equipping Ψ^* with its natural partial order, $(\widehat{F}_\gamma)_{\gamma \in \Psi^*}$ is a projective system of Banach algebras and one obtains a dense topological algebra embedding $F \hookrightarrow \varprojlim_{\gamma \in \Psi^*} \widehat{F}_\gamma$ known as the Arens-Michael decomposition (see [47] Chapter III). As compact (resp. nuclear) operators form an operator ideal, we obtain the following.

Lemma 2.1.9. *Let F be a locally m -convex algebra. Then F is Schwartz (resp. nuclear) if and only if every continuous algebra homomorphism $M : F \mapsto A$ is compact (resp. nuclear) for every Banach algebra A .*

For a normed space V and Banach space W , denote by $\mathcal{N}(V, W)$ the Banach space of nuclear operators from V to W with the nuclear norm $\|\cdot\|_N$.

Lemma 2.1.10. *Let (V, γ) be a normed space and A a Banach algebra. Let $M \in \mathcal{N}(V, A)$ with $\|M\|_N < 1$. Equip $T(V)$ with the norm $\exp(\gamma)$. Then the extension $M_E : T(V) \mapsto A$ is nuclear and $\|M_E\|_N \leq (1 - \|M\|_N)^{-1}$.*

Proof. It holds that product map $M^{\otimes k} : V^{\otimes \pi^k} \mapsto A^{\otimes \pi^k}$ is nuclear with nuclear norm bounded by $\|M\|_N^k$ ([32] Theorem 3.7 - the bound is clear from the proof therein), and the multiplication map $A^{\otimes \pi^k} \mapsto A$ has unit operator norm. It follows that $M^{\otimes k} : V^{\otimes \pi^k} \mapsto A$ is nuclear with nuclear norm at most $\|M\|_N^k$ ([29] p.84). The conclusion follows since $M_E = \sum_{k \geq 0} M^{\otimes k}$ is an absolutely convergent series in $\mathcal{N}(T(V), A)$. $\square$

For a semi-norm γ on V , let $B_\gamma = \{v \in V \mid \gamma(v) < 1\}$, and for a subset $B \subseteq V$, let $\Gamma(B)$ be the absolutely convex hull of B .

Proof of Proposition 2.1.8. The “only if” direction is clear. Let A be a Banach algebra, $M \in \mathbf{L}(V, A)$, and let Ψ be a fundamental family of semi-norms on V . For a semi-norm γ on V , recall that $B_{\gamma^{\otimes k}} = \Gamma(B_\gamma^{\otimes k}) \subset V^{\otimes k}$.

Suppose V is Schwartz. Take $\gamma \in \Psi^*$ such that $M(B_\gamma) \subset A$ is relatively compact and $\gamma(M) < 1$. It follows that $M(B_{\gamma^{\otimes k}})$ is relatively compact in A ([57] Proposition 7.11). Since the unit ball $B_{\exp(\gamma)}$ is given by $\Gamma(\bigcup_{k \geq 0} B_{\gamma^{\otimes k}})$, we obtain that $M(B_{\exp(\gamma)})$ is totally bounded in A . Thus E is Schwartz by Lemma 2.1.9.

Suppose V is moreover nuclear. Take $\gamma \in \Psi^*$ such that the induced map $M_\gamma : V_\gamma \mapsto A$ is nuclear with $\|M_\gamma\|_N < 1$. As $(V^{\otimes \pi^k})_{\gamma^{\otimes k}}$ and $(V_\gamma)^{\otimes \pi^k}$ are isometrically isomorphic ([29] p.38), we have the natural identification $T(V)_{\exp(\gamma)} \cong (T(V_\gamma), \exp(\gamma))$. It follows that $M_E : T(V)_{\exp(\gamma)} \mapsto A$ is nuclear by Lemma 2.1.10. Thus E is nuclear again by Lemma 2.1.9. $\square$

One may also ask when the extension map $\cdot_E : \mathbf{L}(V, A) \mapsto \text{Hom}(E, A)$ is continuous under certain topologies. In the case of the strong topology when V is normed, we obtain a homeomorphism by the following proposition. First, remark that if $\|x_j\| \leq c$ and $\|x_j - y_j\| \leq \varepsilon$ for $x_1, \dots, x_n, y_1, \dots, y_n \in A$, where A is a normed algebra, then

$$\|x_1 \dots x_n - y_1 \dots y_n\| \leq \sum_{j=1}^n \binom{n}{j} \varepsilon^j c^{n-j} = (c + \varepsilon)^n - c^n. \quad (2.1.3)$$

Proposition 2.1.11. *Let V be a normed space and A a Banach algebra. The extension map $\cdot_E : M \mapsto M_E$ from $\mathbf{L}(V, A)$ to $\text{Hom}(E, A)$ is continuous (and thus a homeomorphism) when one equips both sides with the strong topology.*

Proof. Let $(M_j)_{j \geq 1} \rightarrow M$ in $\mathbf{L}(V, A)$. Let γ be a norm on V such that $\gamma(M) \leq 1$ and $\gamma(M_j) \leq 1$ for all $j \geq 1$.

Remark that for any bounded set $B \subset E$ and $\varepsilon > 0$, there exists $k_\varepsilon \geq 1$ such that $\sup_{x \in B} \gamma^{\otimes k}(x^k) \leq \varepsilon^k$ for all $k \geq k_\varepsilon$ (if not, then take a sequence $x_n \in B$ such that

$\gamma^{\otimes n}(x_n^n) > \varepsilon^n$. Then $\exp(c\gamma)(x_n) > c^n \varepsilon^n$ for any $c > 1$ and $n \geq 1$, which implies that $\exp(c\gamma)$ is not bounded on B for some $c > 1$ which is a contradiction).

Remark that every bounded set in $V^{\otimes k}$ is contained in $\overline{\Gamma}(B_1 \otimes \dots \otimes B_k)$ for bounded sets $B_1, \dots, B_k \subset V$, and that the supremum of a convex function on a set is equal to its supremum on the set's convex hull. Together with (2.1.3), this implies that for any fixed n ,

$$\sup_{x \in B} \sum_{0 \leq k \leq n} \|M_i^{\otimes k}(x^k) - M^{\otimes k}(x^k)\| \rightarrow 0.$$

Hence

$$\begin{aligned} \sup_{x \in B} \|M_j(x) - M(x)\| &\leq \sup_{x \in B} \sum_{0 \leq k \leq n} \|M_i^{\otimes k}(x^k) - M^{\otimes k}(x^k)\| \\ &\quad + 2 \sup_{x \in B} \sum_{k > n} \gamma^{\otimes k} x^k \end{aligned}$$

can be made arbitrarily small with sufficiently large n and j . $\square$

Remark 2.1.12. If we assume only that V is locally convex and $M \in \mathbf{L}(V, A)$, where A is a real (resp. complex) Banach algebra, applying the above proposition to the semi-norm $\gamma(x) = \|M(x)\|$ on V implies in particular that the map $\lambda \mapsto (\lambda M)_E$ is continuous from $\mathbb{R}$ (resp. $\mathbb{C}$) to $\text{Hom}(E, A)$, where the latter is equipped with the strong topology.

2.2 Group-like elements

We briefly recall the Hopf algebra structure of $T(V)$. Define the linear map $\Delta : V \mapsto T(V)^{\otimes 2}$ by $\Delta v = 1 \otimes v + v \otimes 1$. By the universal property of $T(V)$, Δ extends to an algebra homomorphism $\Delta : T(V) \mapsto T(V)^{\otimes 2}$.

Now define the linear map $\alpha : V \mapsto V$ by $\alpha(v) = -v$. Then α extends uniquely to an algebra anti-automorphism $\alpha : T(V) \mapsto T(V)$ given by $\alpha(v_1 \dots v_k) = (-1)^k v_k \dots v_1$ for all $v_1 \dots v_k \in V^{\otimes k}$.

One can readily verify that $T(V)$ is a Hopf algebra when equipped with its usual product, coproduct Δ , and antipode α ([50] Proposition 1.10). We shall review certain consequences of the Hopf algebra structure in Section 2.3.

Consider now V a locally convex space. Since $E^{\widehat{\otimes} 2}$ is itself a locally m -convex algebra ([47] p.378), and since $\Delta \in \mathbf{L}(V, E^{\otimes 2})$, the extension $\Delta : E \mapsto E^{\widehat{\otimes} 2}$ is continuous by the universal property of E . Moreover the antipode α extends to a continuous linear map $\alpha : E \mapsto E$. This endows E with an “almost” Hopf algebra

structure (“almost” since E is not mapped to $E^{\otimes 2}$ under the coproduct Δ as for Hopf algebras, but to its completion $E^{\widehat{\otimes 2}}$).

Denote by $U(V) = \{g \in E \mid \alpha(g) = g^{-1}\}$ and $G(V) = \{g \in E \mid \Delta(g) = g \otimes g, g \neq 0\}$ the groups of *unitary* elements and *group-like* elements of E respectively. Note that since multiplication and inversion in E are continuous (and indeed in every locally m -convex algebra, [47] p.5, p.52), U and G are topological groups when endowed with the subspace topology. Moreover, U is closed in E since the map $\phi : x \mapsto (\alpha(x)x, x\alpha(x))$ from E into $E \times E$ is continuous and $U = \phi^{-1}\{(1, 1)\}$. Likewise G is closed in E since $g^0 = 1$ for all $g \in G$ and $G = \psi^{-1}\{0\} \setminus \{0\}$ for the continuous map $\psi : x \mapsto x \otimes x - \Delta(x)$ from E into $E^{\widehat{\otimes 2}}$. Finally, note the inclusion $G \subset U$.

In this section we collect several results concerning measures on G . While these results shall later be applied mostly to the case $V = \mathbb{R}^d$, we find making this assumption does not simplify the proofs, and thus make most statements in full generality.

All measures (resp. random variables) are assumed to be Borel. Denote by $\mathcal{P}(S)$ the space of probability measures on a topological space S endowed with the topology of weak convergence on $C_b(S, \mathbb{C})$.

Recall that for a locally convex space F , an F -valued random variable X is weakly (Gelfand-Pettis) integrable, or that $\mathbb{E}[X]$ exists, if $f(X)$ is integrable for all $f \in F'$ and if there exists $\mathbb{E}[X] := x \in F$ such that $\mathbb{E}[f(X)] = f(x)$. Letting μ be the probability measure associated with X , we denote by $\mu^* = \mathbb{E}[X]$ its barycenter. Unless stated otherwise, we shall always assume that μ is the measure associated to X and that integrals are taken in the weak sense.

Definition 2.2.1. For an E -valued random variable X , we call the sequence

$$\text{ESig}(X) := (\mathbb{E}[X^0], \mathbb{E}[X^1], \dots) \in P = \prod_{k \geq 0} V^{\widehat{\otimes k}}$$

the *expected signature* of X whenever X^k is integrable for all $k \geq 0$.

When V is normed, define $r_1(X)$ as the radius of convergence of the series

$$\sum_{k \geq 0} \mathbb{E}[\|X^k\|] \lambda^k$$

(setting $r_1(X) = 0$ whenever X^k is not norm-integrable for some $k \geq 0$), and $r_2(X)$ as the radius of convergence of the series

$$\sum_{k \geq 0} \|\mathbb{E}[X^k]\| \lambda^k,$$

(setting $r_2(X) = 0$ whenever X^k is not integrable for some $k \geq 0$).

Note that $r_2(X) = R(\text{ESig}(X))$. Remark also that $r_1(X) \leq r_2(X)$ and that Proposition 2.2.4 provides a partial converse when $V = \mathbb{R}^d$ and X is $G(\mathbb{R}^d)$ -valued.

Note that $\text{ESig}(X)$ exists whenever X is integrable as an E -valued random variable. The following proposition now provides a converse when X is G -valued. Recall that we identify E as a subspace of P (Corollary 2.1.6).

Proposition 2.2.2. *Let X be a G -valued random variable. Then X is weakly integrable if and only if $\text{ESig}(X)$ exists and lies in E . In this case $\mathbb{E}[X] = \text{ESig}(X)$.*

In the case that V is normed, note that in order to conclude that a G -valued random variable X is (weakly) integrable (as an E -valued random variable), Proposition 2.2.2 implies that one only needs to check that each projection X^k is (weakly) integrable and that $\|\mathbb{E}[X^k]\|$ decays sufficiently fast as $k \rightarrow \infty$. Remark that this is certainly not true for an arbitrary E -valued random variable.

We observe that for any $f \in E'$, it holds that $f^{\otimes 2} \circ \Delta \in E'$ and $f(g)^2 = f^{\otimes 2}(\Delta g)$ for all $g \in G$. In particular, for all $\mu \in \mathcal{P}(G)$, we have

$$\mu(|f|) \leq \sqrt{\mu(f^2)} = \sqrt{\mu(f^{\otimes 2} \circ \Delta)}. \quad (2.2.1)$$

This simple observation allows for very easy control of a measure through its barycenter. For example, whenever $\mu \in \mathcal{P}(G)$ and $\mathbb{E}[X]$ exists, it follows immediately that for all $f \in E'$, the real random variable $f(X)$ has finite moments of all orders.

The main idea behind the proof of Proposition 2.2.2 is that given the existence of $\mathbb{E}[X^k]$ for all $k \geq 0$, we wish to approximate $\mathbb{E}[f(X)]$ by $\sum_{k=0}^n \mathbb{E}[f(X^k)]$. Using the estimate (2.2.1) and the grading of the coproduct Δ , we apply dominated convergence to obtain $\mathbb{E}[f(X)] = \sum_{k \geq 0} \mathbb{E}[f(X^k)]$.

Proof of Proposition 2.2.2. The “only if” direction is clear. Assume that $\text{ESig}(X)$ exists and $\text{ESig}(X) \in E$. As usual, let μ be the measure on G associated to X . We are required to show that f is μ -integrable and that $\mu(f) = \langle f, \text{ESig}(X) \rangle$ for all $f \in E'$.

Recall the projection $\rho^k : E \mapsto V^{\widehat{\otimes} k}$. Treating $V^{\widehat{\otimes} k}$ as a subspace of E , for all $f \in E'$ denote $f^k := f \circ \rho^k \in E'$. Furthermore, we canonically embed $(V^{\widehat{\otimes} k})'$ into E' by $f \mapsto f \circ \rho^k$ for all $f \in (V^{\widehat{\otimes} k})'$. Observe that for all $f \in E'$, by Corollary 2.1.7, $\sum_{k=0}^n f^k$ converges uniformly on bounded sets (and a fortiori pointwise) to f .

Remark that for any $f \in E'$, $f \in (V^{\widehat{\otimes} k})'$ if and only if $f = f^k$. Recall that Δ is a graded linear map from $T(V)$ to $T(V)^{\otimes 2}$. In particular, for all $f_1 \in (V^{\widehat{\otimes} k})'$, $f_2 \in (V^{\widehat{\otimes} m})'$, and $x \in T(V)$, it holds that

$$(f_1 \otimes f_2)\Delta(x) = (f_1 \otimes f_2)\Delta(x^{k+m}). \quad (2.2.2)$$

As $T(V)$ is dense in E , (2.2.2) holds for all $x \in E$, from which it follows that $(f_1 \otimes f_2) \circ \Delta \in (V^{\widehat{\otimes}(k+m)})'$.

Let $f \in E'$ and note that $\mu(f^k) = \langle f^k, \mathbb{E}[X^k] \rangle$ for all $k \geq 0$. Since μ has support on G , it follows from (2.2.1) and (2.2.2) that

$$\mu \left(\sum_{k \geq 0} |f^k| \right) \leq \sum_{k \geq 0} \sqrt{\mu((f^k)^{\otimes 2} \circ \Delta)} = \sum_{k \geq 0} \sqrt{(f^k)^{\otimes 2} \Delta \mathbb{E}[X^{2k}]}. \quad (2.2.3)$$

Without loss of generality, we can assume that $|f(1)| \leq 1$. Let γ be a semi-norm on V such that $\exp(\gamma) \geq |f|$ and ξ a semi-norm on E such that $\xi \geq \exp(\gamma)^{\otimes 2} \circ \Delta$. It follows that $\exp(\gamma) \geq |f^k|$ for all $k \geq 0$, and thus $\xi \geq |(f^k)^{\otimes 2} \circ \Delta|$ for all $k \geq 0$.

Since $\text{ESig}(X) \in E$, it follows from Corollary 2.1.6 that $\sum_{k \geq 0} \sqrt{\xi(\mathbb{E}[X^k])}$ is finite, and hence (2.2.3) is finite. By dominated convergence, we obtain

$$\mu(f) = \lim_{n \rightarrow \infty} \mu \left(\sum_{k=0}^n f^k \right).$$

It then follows that $\mu(f) = \langle f, \text{ESig}(X) \rangle$ as desired since

$$\lim_{n \rightarrow \infty} \mu \left(\sum_{k=0}^n f^k \right) = \lim_{n \rightarrow \infty} \sum_{k=0}^n \langle f^k, \mathbb{E}[X^k] \rangle = \langle f, \text{ESig}(X) \rangle.$$

□

Corollary 2.2.3. *Let V be a normed space and X a G -valued random variable. Then $\mathbb{E}[X] \in E$ exists if and only if $r_2(X) = \infty$, i.e., $\text{ESig}(X)$ exists and has an infinite radius of convergence. In this case $\mathbb{E}[X] = \text{ESig}(X)$.*

We are moreover able to show explicit bounds between $r_1(X)$ and $r_2(X)$ when $V = \mathbb{R}^d$. Suppose first that V is a normed space. Remark that $\|\Delta v\| = 2\|v\|$ for all $v \in V$, from which it follows that $\|\Delta|_{V^{\widehat{\otimes}k}}\| = 2^k$ and thus

$$\|\Delta x^k\| \leq 2^k \|x^k\| \text{ for all } x \in E. \quad (2.2.4)$$

Let $V = \mathbb{R}^d$ equipped with the ℓ^1 norm from its standard basis $e_1, \dots, e_d$, and denote $e_I = e_{i_1} \dots e_{i_k} \in V^{\widehat{\otimes}k}$ for a word $I = i_1 \dots i_k$ in the alphabet $\{1, \dots, d\}$. Then the grading of Δ gives

$$\begin{aligned} \mathbb{E}[\|X^k\|^2] &= \mathbb{E} \left[\left(\sum_{|I|=k} |\langle e_I, X^k \rangle| \right)^2 \right] \leq d^k \mathbb{E} \left[\sum_{|I|=k} \langle e_I, X^k \rangle^2 \right] \\ &= d^k \sum_{|I|=k} e_I^{\otimes 2} \Delta \mathbb{E}[X^{2k}] \\ &\leq d^k \|\Delta \mathbb{E}[X^{2k}]\|, \end{aligned}$$

where the last inequality follows since $(e_I \otimes e_J)_{|I|=|J|=k}$ is an ℓ^1 basis for $V^{\otimes 2k}$. Using (2.2.4) we now obtain the following.

Proposition 2.2.4. *Let X be a $G(\mathbb{R}^d)$ -valued random variable. It follows that*

$$\mathbb{E} \left[\|X^k\|^2 \right] \leq d^k 2^{2k} \|\mathbb{E}[X^{2k}]\|.$$

In particular, $r_1(X) \leq r_2(X) \leq 2\sqrt{d}r_1(X)$.

2.3 Representations

By an algebra of operators, we mean the algebra $\mathbf{L}(H)$ of linear operators $A : H \mapsto H$ for a Hilbert space H , endowed with the product $AB := B \circ A$. Remark that the product may seem in reverse to the usual composition product considered on $\mathbf{L}(H)$, however the two notions are completely equivalent (and can be made identical by considering elements of $\mathbf{L}(H)$ as instead acting on the dual H' via their adjoint operators). We choose this product due to its simple relation with the signature and linear differential equations, as shall be seen in Section 3.4.1.

Recall that for any Hopf algebra, one may define the tensor product and dual of representations via the coproduct and antipode by $M_1 \otimes M_2(x) := (M_1 \otimes M_2)\Delta(x)$ and $M^*(x) := M(\alpha(x))^*$. By virtue of continuity of Δ and α , we observe that the family of continuous representations of E over finite dimensional Hilbert spaces is closed under tensor products and duals.

Definition 2.3.1. Denote by $\mathcal{A}(V)$ the family of finite dimensional representations of E which arise from extensions of all linear maps $M \in \mathbf{L}(V, \mathfrak{u}(H_M))$, where H_M ranges over all finite dimensional complex Hilbert spaces and $\mathfrak{u}(H_M)$ denotes the Lie algebra of the anti-Hermitian operators on H_M . Denote by $\mathcal{C}(V)$ the set of corresponding matrix coefficients, i.e., the set of linear functionals $M_{u,v} \in \mathbf{L}(E, \mathbb{C})$, $M_{u,v}(x) = \langle M(x)u, v \rangle$ for all $M \in \mathcal{A}$ and $u, v \in H_M$.

The family $\mathcal{A}$ possesses the desirable property that it is closed under taking tensor products and duals of representations. Moreover, we see that $\mathcal{A}$ contains exactly those finite dimensional representations of E which preserve involution, i.e., $M(\alpha x) = M(x)^*$ for all $x \in E$. It follows that every $M \in \mathcal{A}$ is a unitary representation of the group U , and thus of G .

Observe that the tensor product $M_1 \otimes M_2$ (of any representations M_1, M_2 of E) coincides on G with the usual group-theoretic tensor product of representations. Moreover, the dual representation M^* of $M \in \mathcal{A}$ can be identified on U with the conjugate representation of M on U . It follows that $\mathcal{C}|_G$ forms a $*$ -subalgebra of $C_b(G, \mathbb{C})$.

Let S be a topological space and F a separating $*$ -subalgebra of $C_b(S, \mathbb{C})$. Recall that for Radon measures μ and ν on S (see [5] Definition 7.1.1), it follows from the Stone-Weierstrass theorem that $\mu = \nu$ if and only if $\mu(f) = \nu(f)$ for all $f \in F$ ([5] Exercise 7.14.79). We now obtain the following from the above discussion.

Lemma 2.3.2. *Assume that $\mathcal{A}$ separates the points of G . Then for tight Borel measures μ, ν on G , $\mu = \nu$ if and only if $\mu(f) = \nu(f)$ for all $f \in \mathcal{C}$, or equivalently, $\mu(M) = \nu(M)$ for all $M \in \mathcal{A}$.*

We show in Theorem 2.3.8 that in fact $\mathcal{A}(\mathbb{R}^d)$ separates the points of $E(\mathbb{R}^d)$.

2.3.1 Separation of points

We investigate conditions under which algebra homomorphisms of E separate points. Though ultimately we apply the theory to the case $V = \mathbb{R}^d$, the arguments used in the general case are exactly the same and we provide them here.

For a Banach algebra A and $M \in \mathbf{L}(V, A)$, let (λM) denote the algebra homomorphism on E induced by $\lambda M \in \mathbf{L}(V, A)$ (λ possibly complex if A is over $\mathbb{C}$). For $\lambda \in \mathbb{R}$, let $\delta_\lambda : E \mapsto E$ denote the dilation operator $\delta_\lambda(x^0, x^1, \dots) = (\lambda^0 x^0, \lambda^1 x^1, \dots)$ (note that $(\lambda M) = M\delta_\lambda$ for $\lambda \in \mathbb{R}$).

Lemma 2.3.3. *Let V be locally convex, A a Banach algebra and $M \in \mathbf{L}(V, A)$. Let $x \in E$ such that $M(x^k) \neq 0$ for some $k \geq 0$. Then there exists $\varepsilon > 0$ sufficiently small such that $(\varepsilon M)(x) \neq 0$.*

Proof. Since $\|M(x)\|$ is a semi-norm on E , $\sum_{k \geq 0} \|M(x^k)\|$ converges by Corollary 2.1.6, from which the conclusion follows. $\square$

Let $\mathbb{F}$ be a field and A an $\mathbb{F}$ -algebra. A polynomial identity over $\mathbb{F}$ on a subset $Q \subseteq A$ is a polynomial in non-commuting indeterminates $x_1, \dots, x_k$, with coefficients in $\mathbb{F}$, which is non-zero (that is, not every coefficient is zero) and which vanishes under all substitutions of variables $x_1, \dots, x_k \in Q$. We refer to Giambruno and Zaicev [28] for further details.

Let V be a vector space with Hamel basis Θ . Then the set of pure tensors $\Theta^{\otimes k} = \{v_1 \dots v_k \mid v_j \in \Theta, 1 \leq j \leq k\}$ is a Hamel basis for $V^{\otimes k}$. Thus for every $x \in V^{\otimes k}$ define

Θ_x as the finite set of vectors in Θ which appear in the representation of x in the basis $\Theta^{\otimes k}$. Define f_x^Θ the canonical formal non-commuting polynomial in indeterminates Θ_x associated with x . As Θ_x is a finite set, the following is a consequence of the Hahn-Banach theorem.

Lemma 2.3.4. *Let V be a locally convex space with Hamel basis Θ , A an algebra which is a topological vector space, and $Q \subseteq A$ a subset. Let $k \geq 0$ and $x \in V^{\otimes k}$. The following two assertions are equivalent.*

- (i) f_x^Θ is not a polynomial identity over $\mathbb{R}$ on Q .
- (ii) There exists a continuous linear map $M : V \mapsto \text{span}(Q)$ such that $M(x)$ is non-zero and $M(v)$ is in Q for all $v \in \Theta_x$.

Remark 2.3.5. If one is not interested in the topological aspects, the same statement holds if one replaces $\mathbb{R}$ by a field $\mathbb{F}$, V by a vector space over $\mathbb{F}$, A by an $\mathbb{F}$ -algebra, and drops the continuity assumption in (ii).

2.3.2 Polynomial identities over Lie algebras

From Lemmas 2.3.3 and 2.3.4, it is clear that to study how representations in $\mathcal{A}(\mathbb{R}^d)$ separate the points of $E(\mathbb{R}^d)$, we must look at polynomial identities in unitary Lie algebras. Let $m \geq 1$ be an integer and denote by $\cdot^s$ the symplectic involution on $M_{2m}(\mathbb{C})$, which we recall is an involution of the first kind (see [27]).

Recall the real Lie algebra $\mathfrak{sp}(m) = \{u \in \mathfrak{u}(\mathbb{C}^{2m}) \mid u^s + u = 0\}$ ($\mathfrak{sp}(m)$ is the Lie algebra of the compact symplectic group $Sp(m)$). A closely related complex Lie subalgebra of $\mathfrak{gl}(\mathbb{C}^{2m})$ is $\mathfrak{sp}(m, \mathbb{C}) = \{u \in M_{2m}(\mathbb{C}) \mid u^s + u = 0\}$. It holds that $\mathfrak{sp}(m, \mathbb{C})$ is the complexification of $\mathfrak{sp}(m)$.

We now illustrate our interest in the Lie algebras $\mathfrak{sp}(m)$ and $\mathfrak{sp}(m, \mathbb{C})$. From the remark that $\mathfrak{sp}(m, \mathbb{C}) = \{u - u^s \mid u \in M_{2m}(\mathbb{C})\}$, we may reformulate a result due to Giambruno and Valenti [27] as follows.

Theorem 2.3.6 ([27] Theorem 6). *Let $m \geq 2$ and $f(x_1, \dots, x_k)$ a polynomial identity over $\mathbb{C}$ on $\mathfrak{sp}(m, \mathbb{C}) \subset M_{2m}(\mathbb{C})$. Then $\deg(f) > 3m$.*

The following is a slight generalization of [28] Theorem 1.3.2 and follows from exactly the same inductive proof.

Lemma 2.3.7. *Let $\mathbb{F}$ be an infinite field, A an $\mathbb{F}$ -algebra and Q a linear subspace of A . If f is a polynomial identity over $\mathbb{F}$ on Q , then every multi-homogeneous component of f is a polynomial identity over $\mathbb{F}$ on Q .*

We remark that every multi-homogeneous polynomial identity over $\mathbb{C}$ (and a fortiori over $\mathbb{R}$) on $\mathfrak{sp}(m) \subset M_{2m}(\mathbb{C})$ is also a polynomial identity over $\mathbb{C}$ on its complexification $\mathfrak{sp}(m, \mathbb{C})$. Thus if f is a polynomial identity over $\mathbb{R}$ on $\mathfrak{sp}(m)$ for $m \geq 2$, then by Theorem 2.3.6 and Lemma 2.3.7, every multi-homogeneous component of f has degree greater than $3m$. Together with Lemmas 2.3.3 and 2.3.4, we have the following result.

Theorem 2.3.8. *Let $x \in E(\mathbb{R}^d)$ such that $x^k \neq 0$ for some $k \geq 0$. Then for any integer $m \geq \max\{2, k/3\}$ there exists $M \in \mathbf{L}(\mathbb{R}^d, \mathfrak{sp}(m))$ such that $M(x) \neq 0$. In particular, $\mathcal{A}(\mathbb{R}^d)$ separates the points of $E(\mathbb{R}^d)$.*

Remark 2.3.9. The necessity that $V = \mathbb{R}^d$ only came into the above argument to ensure that $V^{\otimes k} = V^{\widehat{\otimes} k}$. If one was able to find an analogue of Lemma 2.3.4 for elements $x \in V^{\widehat{\otimes} k}$, or an analogue of Theorem 2.3.6 for appropriate series of polynomials of bounded degree but an unbounded number of indeterminates, then one could readily extend Theorem 2.3.8 to the case when V is infinite dimensional.

Corollary 2.3.10. *The group $U(\mathbb{R}^d)$ is maximally almost periodic.*

Remark 2.3.11. For $d \geq 2$, the topological group $G(\mathbb{R}^d)$ (and thus $U(\mathbb{R}^d)$) is not locally compact. To observe this, let $V = \mathbb{R}^d$ and $L(V)$ be the smallest Lie algebra in $T(V)$ containing V . Since every $\ell \in L(V)$ satisfies $\Delta(\ell) = 1 \otimes \ell + \ell \otimes 1$ ([50] Theorem 1.4), a direct calculation shows that $\exp(\ell) \in G$.

Let $u, v \in V$ be linearly independent elements and $W = \text{span}(u, v)$. Observe that $L(W)$ contains a non-zero element in $W^{\otimes k}$ for every $k \geq 1$. In light of Proposition 2.1.4, for any neighbourhood of zero B of $L(W)$ one can construct a sequence $(\ell_n)_{n \geq 1} \in B$ such that $\gamma(\exp(\ell_i) - \exp(\ell_j)) \geq 1$ for all $i \neq j$ and some semi-norm γ on E . Since $\exp : L(W) \mapsto G$ is continuous ([1] Theorem 3), it follows that no neighbourhood of the identity in G is contained in a sequentially compact set (the same argument more generally applies whenever V is metrizable).

It follows from Corollary 2.1.5 that E is Polish whenever V is metrizable and separable, and thus G , as a closed subset of E , is also Polish. As every probability measures on a Polish space is tight, Lemma 2.3.2 and Theorem 2.3.8 imply the following.

Corollary 2.3.12. *For Borel probability measures μ and ν on $G(\mathbb{R}^d)$, it holds that $\mu = \nu$ if and only if $\mu(f) = \nu(f)$ for all $f \in \mathcal{C}(\mathbb{R}^d)$, or equivalently, $\mu(M) = \nu(M)$ for all $M \in \mathcal{A}(\mathbb{R}^d)$.*

For a Borel probability measure μ on $G(\mathbb{R}^d)$, with associated random variable X , we are thus able to define its characteristic function (or Fourier transform) by $\phi_X = \widehat{\mu} := \mu|_{\mathcal{A}}$, which uniquely characterizes μ .

Chapter 3

Rough paths and signatures

In this chapter we present the fundamental definitions and results from the theory of rough paths.

In Section 3.1 we define the space of p -rough paths $\Omega_p(V)$. We point out that the signature of a rough path can be viewed as an element of the space $E(V)$. We also show that the map from $\Omega_p(V)$ to $E(V)$ sending the rough path to its signature is a continuous map. In Section 3.2 we consider in more detail the finite-dimensional case $V = \mathbb{R}^d$.

In Sections 3.3 and 3.4 we review rough differential equations, their associated solution maps, and the universal limit theorem. Lastly, in Section 3.5 we recall the Euler scheme associated with an RDE.

The results of this chapter are mostly well-known with the exception of Sections 3.2.2 and 3.4.1. These two sections focus on convergence in law of reparametrised rough paths and of RDE solution maps. We in particular highlight the recent work of Boedihardjo, Geng, Lyons and Yang [4] on the classification of rough paths by their signatures in relation to the characteristic function of the previous chapter.

3.1 Rough paths

Let $T > 0$ and define $\Delta_{[0,T]} = \{(s, t) \mid 0 \leq s \leq t \leq T\}$. We first recall the definition of a control function.

Definition 3.1.1. A *control* on $[0, T]$ is a non-negative continuous function $\omega : \Delta_{[0,T]} \mapsto [0, \infty)$ for which

$$\omega(s, t) + \omega(t, u) \leq \omega(s, u)$$

for all $0 \leq s \leq t \leq u \leq T$, and $\omega(t, t) = 0$ for all $t \in [0, T]$.

Let V be a Banach space and $T^n(V) = \bigoplus_{0 \leq k \leq n} V^{\widehat{\otimes} k}$ the truncated tensor algebra. Remark that T^n is equipped with a natural algebra structure (realised as the quotient of $P(V)$ by the ideal $\prod_{k > n} V^{\widehat{\otimes} k}$). In particular, the (multiplicative) identity of T^n is $1_n := (1, 0, \dots, 0)$.

Definition 3.1.2. Let $p \geq 1$. The space of p -rough paths $\Omega_p(V)$ is the collection of all maps $\mathbf{x} : \Delta_{[0,T]} \mapsto T^{[p]}$ satisfying

- (a) $\mathbf{x}_{s,t}^0 = 1$ and $\mathbf{x}_{s,t}\mathbf{x}_{t,u} = \mathbf{x}_{s,u}$ for all $0 \leq s \leq t \leq u \leq T$, and
- (b) for some control ω one has

$$\sup_{1 \leq k \leq [p]} ((k/p)! \beta_p \|\mathbf{x}_{s,t}^k\|)^{p/k} \leq \omega(s, t), \quad \forall (s, t) \in \Delta_{[0,T]}, \quad (3.1.1)$$

where β_p is a constant that only depends on p .

If a map $\mathbf{x} : \Delta_{[0,T]} \mapsto T^{[p]}$ satisfies (a), we say $\mathbf{x}$ is multiplicative, while if $\mathbf{x}$ satisfies (b), we say that the p -variation of $\mathbf{x}$ is controlled by ω .

The map $\mathbf{x}$ may equivalently be viewed as a path $\mathbf{x}_{0,\cdot} : [0, T] \mapsto T^{[p]}, t \mapsto \mathbf{x}_{0,t}$ of finite p -variation, that is,

$$\|\mathbf{x}\|_{p\text{-var};[0,T]} := \sum_{1 \leq k \leq [p]} \sup_{\mathcal{D} \subset [0,T]} \left(\sum_{t_j \in \mathcal{D}} ((k/p)! \beta_p \|\mathbf{x}_{t_j, t_{j+1}}^k\|)^{p/k} \right)^{1/p} < \infty, \quad (3.1.2)$$

where the sup is over all finite partitions $\mathcal{D} = (t_0, \dots, t_k) \subset [0, T]$, for which we define $t_{k+1} := T$ in the corresponding summation. Observe that the path $\mathbf{x}_{0,\cdot}$ completely characterizes $\mathbf{x}$ due to the multiplicative property (a) (noting that $\mathbf{x}_{s,t}$ is always invertible by $\mathbf{x}_{s,t}^0 \equiv 1$ and that $\mathbf{x}_{s,t} = \mathbf{x}_{0,s}^{-1} \mathbf{x}_{0,t}$). Conversely, given a path $\mathbf{x} : [0, T] \mapsto T^{[p]}$ satisfying (3.1.2) and such that $\mathbf{x}^0 \equiv 1$, we may associate to $\mathbf{x}$ an element of Ω_p by setting $\mathbf{x}_{s,t} := \mathbf{x}_s^{-1} \mathbf{x}_t$.

Let $\mathbf{x} \in \Omega_p$ satisfy (3.1.1) for some control ω . A fundamental result of rough paths theory is the following extension theorem, which allows one to uniquely extend $\mathbf{x}$ to a multiplicative map of finite p -variation taking values in T^n for all $n \geq [p]$.

Theorem 3.1.3 ([42] Theorem 3.1.2). *Let $\mathbf{x} \in \Omega_p$. Then for all $n \geq [p]$ there exists a unique map $S_n(\mathbf{x}) : \Delta_{[0,T]} \mapsto T^n$ such that (a) and (b) remain true for the same ω and with $\sup_{1 \leq k \leq [p]}$ replaced by $\sup_{1 \leq k \leq n}$ in (3.1.1).*

Theorem 3.1.3 allows us to embed the space Ω_p into Ω_q for any $q \geq p$. In particular, we shall often refer to elements $\mathbf{x} \in \Omega_p$ which take values in T^n for some $n > \lfloor p \rfloor$, in which case we are always referring to the multiplicative map $S_n(\mathbf{x})$.

An equivalent formulation of Theorem 3.1.3 is that there exists a unique lift to the entire product space $S(\mathbf{x}) : \Delta_{[0,T]} \mapsto P = \prod_{k \geq 0} V^{\widehat{\otimes} k}$ such that (a) and (b) remain true for the same ω and with $\sup_{1 \leq k \leq \lfloor p \rfloor}$ replaced by $\sup_{1 \leq k}$ in (3.1.1).

An immediate consequence of the factorial decay in (3.1.1) is that the lift $S(\mathbf{x})$ takes values in the space E for any $p \geq 1$ (see Corollary 2.1.6).

Remark 3.1.4. While the value of β_p does not affect the definition of the space Ω_p , its existence is crucial to ensure the factorial decay arising from the lift. On this point, we mention the work of Hara and Hino [31] who have resolved a conjecture on the optimal possible value of β_p .

We thus make a canonical extension of the space Ω_p .

Definition 3.1.5. Define the space ΩE_p as the set of maps $\mathbf{x} : \Delta_{[0,T]} \mapsto E$ which satisfy (a) and (b) with $\sup_{1 \leq k \leq \lfloor p \rfloor}$ replaced by $\sup_{1 \leq k}$ in (3.1.1).

It follows that the lift S is a bijective map from Ω_p to ΩE_p , with inverse provided by the $\lfloor p \rfloor$ -th level truncation $\pi^{[p]} : (\mathbf{x}_{s,t}^0, \mathbf{x}_{s,t}^1, \dots) \mapsto (\mathbf{x}_{s,t}^0, \mathbf{x}_{s,t}^1, \dots, \mathbf{x}_{s,t}^{[p]})$. The terms $\mathbf{x}_{s,t}^k$ are often called the iterated integrals of $\mathbf{x}$.

The element $S(\mathbf{x})_{0,T} \in E$ is called the *signature* of a rough path $\mathbf{x} \in \Omega_p$. We denote by $\mathcal{I}_p : \Omega_p \mapsto E$ the signature map sending $\mathbf{x} \mapsto S(\mathbf{x})_{0,T}$.

For $1 \leq p < 2$, $S(\mathbf{x})_{0,T}$ is precisely the sequence of iterated integrals of the path $\mathbf{x}_{0,\cdot} : [0, T] \mapsto V$ taken in the sense of Young.

Remark 3.1.6. The only property of the projective tensor norm used above is that the projective extension provides a sub-multiplicative system of norms. Completely analogous results hold true if one equips $T(V)$ with any system of sub-multiplicative norms and defines E as the completion of $T(V)$ under scalar dilations of these norms. Note that in the case $V = \mathbb{R}^d$, all these systems lead to identical definitions and topologies on the space E .

The lift S moreover exhibits a natural continuity property with respect to the p -variation topology on Ω_p . We first recall that each individual lift S_n exhibits the following continuity property.

Theorem 3.1.7 ([42] Theorem 3.1.3). *Let $a > 0$, ω a control, and $\mathbf{x}, \mathbf{y} \in \Omega_p$ with p -variation controlled by ω . Suppose that*

$$\sup_{1 \leq k \leq \lfloor p \rfloor} ((k/p)! \beta_p a \|\mathbf{x} - \mathbf{y}\|)^{p/k} \leq \omega(s, t), \quad \forall (s, t) \in \Delta_{[0,T]}. \quad (3.1.3)$$

Then for all $n \geq \lfloor p \rfloor$, $S_n(\mathbf{x})$ and $S_n(\mathbf{y})$ satisfy (3.1.3) with $\sup_{1 \leq k \leq \lfloor p \rfloor}$ replaced by $\sup_{1 \leq k \leq n}$.

Now for $\mathbf{x}, (\mathbf{x}(n))_{n \geq 1} \in \Omega_p$, a control ω and a sequence of positive reals $(a_n)_{n \geq 1}$ consider the statement

$$\omega \text{ controls the } p\text{-variation of } \mathbf{x} \text{ and } \mathbf{x}(n) \text{ for all } n \geq 1, \text{ and} \quad (3.1.4)$$

$$\sup_{1 \leq k \leq \lfloor p \rfloor} ((k/p)! \beta_p a_n \|\mathbf{x}(n)_{s,t}^k - \mathbf{x}_{s,t}^k\|)^{p/k} \leq \omega(s, t), \quad \forall (s, t) \in \Delta_{[0, T]}.$$

When (3.1.4) is satisfied for some control ω and a sequence $(a_n)_{n \geq 1}$ such that $a_n \rightarrow \infty$, we say that $\mathbf{x}(n) \rightarrow \mathbf{x}$ in the p -variation topology of Ω_p . One makes the same definition for $\mathbf{x}, (\mathbf{x}(n))_{n \geq 1} \in \Omega E_p$ with $\sup_{1 \leq k \leq \lfloor p \rfloor}$ replaced by $\sup_{1 \leq k}$ in (3.1.4).

The following is an immediate consequence of Theorem 3.1.7.

Proposition 3.1.8. *If $\mathbf{x}, (\mathbf{x}(n))_{n \geq 1} \in \Omega_p$ satisfy (3.1.4) for some ω and $(a_n)_{n \geq 1}$, then $S(\mathbf{x}), (S(\mathbf{x}(n)))_{n \geq 1} \in \Omega E_p$ satisfy (3.1.4) for the same control ω and sequence $(a_n)_{n \geq 1}$.*

In particular, $S : \Omega_p \mapsto \Omega E_p$ is continuous (and thus a homeomorphism) when Ω_p and ΩE_p are equipped with their respective p -variation topologies.

Corollary 3.1.9. *The signature map $\mathcal{I}_p : \Omega_p \mapsto E$ is continuous when Ω_p is equipped with the p -variation topology.*

We also recall another commonly used topology on Ω_p . The (inhomogeneous) metric $\rho_{p\text{-var}}$ on Ω_p is defined by

$$\rho_{p\text{-var}}(\mathbf{x}, \mathbf{y}) = \sup_{1 \leq k \leq \lfloor p \rfloor} \sup_{\mathcal{D} \subset [0, T]} \left(\sum_{t_j \in \mathcal{D}} \|\mathbf{x}_{t_j, t_{j+1}}^k - \mathbf{y}_{t_j, t_{j+1}}^k\|^{p/k} \right)^{k/p}. \quad (3.1.5)$$

The space $(\Omega_p(\mathbb{R}^d), \rho_{p\text{-var}})$ is a complete metric space with a coarser topology than the p -variation topology, but for which convergence of a sequence $\rho_{p\text{-var}}(\mathbf{x}(n), \mathbf{x}) \rightarrow 0$ implies the existence of a subsequence $\mathbf{x}(n_k)$ which converges to $\mathbf{x}$ in the p -variation topology (see [25] Section 8, [42] Proposition 3.3.3, but note the differing notations for homogeneous and inhomogeneous metrics in the two texts; we use the notation of [25] and shall recall the corresponding homogeneous metric $d_{p\text{-var}}$ in Section 3.2).

Using the fact that E is a metric space, we obtain the following refinement of Corollary 3.1.9.

Corollary 3.1.10. *The signature map $\mathcal{I}_p : (\Omega_p, \rho_{p\text{-var}}) \mapsto E$ is continuous.*

3.1.1 Geometric rough paths

We now recall the space $G\Omega_p(V)$ of geometric p -rough paths. We take a moment to highlight the appearance of several different non-equivalent definitions of geometric rough paths in the literature over the last two decades. Lyons [45] originally defined $G\Omega_p(V)$ in the following way: for $n \geq 1$, let $\mathfrak{g}^n(V)$ be the smallest Lie algebra in T^n containing V , and define $G^n(V) = \exp(\mathfrak{g}^n)$. Remark that $\exp : \mathfrak{g}^n \mapsto G^n$ is a polynomial map with a polynomial inverse $\log : G^n \mapsto \mathfrak{g}^n$, and is thus a homeomorphism.

Assume first that V is finite dimensional. Then G^n is a nilpotent Lie group with Lie algebra $\mathfrak{g}^n$ for which $\exp : \mathfrak{g}^n \mapsto G^n$ is a diffeomorphism. For all $\mathbf{x} \in \Omega_1$, it holds that $S_n(\mathbf{x})$ is the development of $\mathbf{x}$ into the space T^n whose derivative (defined for almost all $t \in [0, T]$) lies in $V \subset \mathfrak{g}^n$, and hence $S_n(\mathbf{x})_{s,t} \in G^n$ for all $\mathbf{x} \in \Omega_1$.

When V is an arbitrary Banach space, care needs to be taken as then $\mathfrak{g}^n$ and G^n are no longer closed in T^n due to the incompleteness of $V^{\otimes k}$; to obtain the same conclusion one should replace $\mathfrak{g}^n$ and G^n by their closures and consider G^n as an infinite dimensional Banach-manifold (this point was not addressed in [45]).

Lyons ([45] Definition 2.3.1) originally defined $G\Omega_p$ (at least for V finite dimensional) as the set of all $\mathbf{x} \in \Omega_p$ for which $\mathbf{x}_{s,t} \in G^{\lfloor p \rfloor}$ for all $(s, t) \in \Delta_{[0, T]}$.

In contrast, the monograph [42] avoids considerations of the group $G^{\lfloor p \rfloor}$ (or its closure) by defining $G\Omega_p(V)$ as the closure in Ω_p of Ω_1 under the lift $S_{\lfloor p \rfloor}$, i.e., as the set of $\mathbf{x} \in \Omega_p$ for which there exist bounded variation paths $\mathbf{x}(n) : [0, T] \mapsto V$ such that $S_{\lfloor p \rfloor}(\mathbf{x}(n)) \rightarrow \mathbf{x}$ in the p -variation topology (see [42] Definition 3.3.3, but remark the non-standard use of the term “smooth” as simply possessing finite variation).

The Saint-Flour lecture notes [46] provide the same definition of $G\Omega_p$ as in [42] (see [46] Definition 3.13). Moreover, the space $WG\Omega_p(V)$ of weakly geometric rough paths is defined in [46] as the set of all $\mathbf{x} \in \Omega_p$ for which $\mathbf{x}_{s,t} \in G^{\lfloor p \rfloor}$ for all $(s, t) \in \Delta_{[0, T]}$ (again, at least for finite dimensional V).

A certain disadvantage of the definitions of $G\Omega_p$ found in [45], [42] and [46] is that the space $G\Omega_p$ is non-separable even for finite dimensional V (indeed, the space $\Omega_1(\mathbb{R})$ is not separable under the metric $\rho_{1\text{-var}}$, see [46] p.6, or [25] Example 1.26).

The conflicting definitions are clearly highlighted by Friz and Victoir [21] and the approach taken therein, is to define $G\Omega_p$ as the closure in Ω_p of all smooth paths $C^\infty([0, T], V)$ under the lift $S_{\lfloor p \rfloor}$ (see [21] Definition 7, and [25] Definition 9.15, though note that in [21] the authors claim this is the same definition as that in [42], which we remarked above is not the case). We shall use the definition of Friz and Victoir [21] here.

Definition 3.1.11. For $p \geq 1$, define the space of geometric p -rough paths $G\Omega_p(V)$ as the closure in Ω_p of the image of $C^\infty([0, T], V)$ under the lift $S_{[p]} : \Omega_1 \mapsto \Omega_p$.

Define $R_p(V) = \{S(\mathbf{x})_{0,T} \mid \mathbf{x} \in G\Omega_p\} \subset E$ as the set of signatures of all geometric p -rough paths.

Since a piecewise smooth path can be approximated in the 1-variation topology by smooth paths (applying mollifications at the joining points), an equivalent definition of $G\Omega_p(V)$ is the closure of piecewise smooth paths. It follows that R_p is closed under multiplication in E .

Moreover for all $\mathbf{x} \in G\Omega_p$, the inverse of $S(\mathbf{x})_{0,T}$ is given by $S(\overleftarrow{\mathbf{x}})_{0,T}$, where $\overleftarrow{\mathbf{x}} \in G\Omega_p$ is the time reversal of $\mathbf{x}$ defined by $\overleftarrow{\mathbf{x}}_{s,t} := \mathbf{x}_{T-t, T-s}$. Moreover, $S(\overleftarrow{\mathbf{x}})_{0,T}$ coincides with $\alpha(S(\mathbf{x})_{0,T})$, where α is the antipode of E defined in Section 2.2 (this follows from the same proof as [42] Theorem 3.3.3). Thus R_p is a subgroup of $U = \{g \in E \mid \alpha(g) = g^{-1}\}$.

Equipping R_p with the subspace topology from E , we remark that R_1 is dense in R_p as a consequence of Corollary 3.1.9.

3.2 Finite dimensional case

In this section we consider $V = \mathbb{R}^d$. It follows that $P(\mathbb{R}^d)$ (resp. $E(\mathbb{R}^d)$) can be identified with the algebra of non-commuting formal power series in d indeterminates (resp. with an infinite radius of convergence).

We remark that the coproduct Δ of $E(\mathbb{R}^d)$ is given by a locally finite formula involving the shuffle product ([50] Proposition 1.8) and an element $g \in E(\mathbb{R}^d)$ is in $G(\mathbb{R}^d)$ precisely when $(g^0, g^1, \dots, g^n)$ is in the free n -step nilpotent Lie group $G^n(\mathbb{R}^d)$ for all $n \geq 1$ ([46] Lemma 2.24).

A fundamental result of Chen [12] is that the signature of a smooth path in $\mathbb{R}^d$ is a group-like element of $E(\mathbb{R}^d)$ (see also [46] Section 2.2.5), and thus $R_1(\mathbb{R}^d) \subset G(\mathbb{R}^d)$.

As mentioned in Section 3.1.1, a closely related set to $G\Omega_p(\mathbb{R}^d)$ is the space $WG\Omega_p(\mathbb{R}^d) \subset \Omega_p(\mathbb{R}^d)$ of *weakly* geometric p -rough paths, that is, those p -rough paths $\mathbf{x} \in \Omega_p(\mathbb{R}^d)$ which take values in the free $[p]$ -step nilpotent Lie group, i.e.,

$$(\mathbf{x}_{s,t}^0, \mathbf{x}_{s,t}^1, \dots, \mathbf{x}_{s,t}^{[p]}) \in G^{[p]}(\mathbb{R}^d), \quad \forall (s, t) \in \Delta_{[0,T]}.$$

We note the strict inclusions $G\Omega_p(\mathbb{R}^d) \subsetneq WG\Omega_p(\mathbb{R}^d) \subsetneq G\Omega_{p'}(\mathbb{R}^d)$ for any $p' > p \geq 1$ ([25] Section 8.5), and thus

$$WR_p(\mathbb{R}^d) := \{S(\mathbf{x})_{0,T} \mid \mathbf{x} \in WG\Omega_p(\mathbb{R}^d)\} \subset R_{p'}(\mathbb{R}^d).$$

Since G is closed in E , we obtain the inclusions for all $p \geq 1$

$$R_p(\mathbb{R}^d) \subset WR_p(\mathbb{R}^d) \subset \overline{R_1(\mathbb{R}^d)} \subseteq G(\mathbb{R}^d).$$

Both spaces $G\Omega_p(\mathbb{R}^d)$ and $WG\Omega_p(\mathbb{R}^d)$ shall play an important role in the sequel; $G\Omega_p(\mathbb{R}^d)$ is a Polish space (Theorem 3.2.1) and is thus convenient for the study of random rough paths, while $WG\Omega_p(\mathbb{R}^d)$ is the natural space of paths for which one can define solutions to rough differential equations (Section 3.3).

3.2.1 Homogeneous norms and the $d_{p\text{-var}}$ metric

Recall the so-called inhomogeneous metric $\rho_{p\text{-var}}$ on Ω_p defined in Section 3.1. We now introduce the closely defined homogeneous metric $d_{p\text{-var}}$ on $WG\Omega_p(\mathbb{R}^d)$.

A function $|||\cdot||| : G^n(\mathbb{R}^d) \mapsto [0, \infty)$ is called a homogeneous norm if

1. $|||x||| = 0$ if and only if $x = 1_n$, and
2. $|||\delta_\lambda x||| = |\lambda| |||x|||$ for all $\lambda \in \mathbb{R}$,

where 1_n is the identity of $G^n(\mathbb{R}^d)$ and δ_λ is the dilation operator defined in Section 2.3.1.

A homogeneous norm is called *symmetric* if $|||x^{-1}||| = |||x|||$ and *sub-additive* if $|||xy||| \leq |||x||| + |||y|||$ for all $x, y \in G^n(\mathbb{R}^d)$. For a homogeneous norm $|||\cdot|||$ on $G^n(\mathbb{R}^d)$ define the distance function $d(x, y) := |||x^{-1}y|||$. Note that d is a left-invariant metric whenever $|||\cdot|||$ is symmetric and sub-additive.

Recall that all homogeneous norms on $G^n(\mathbb{R}^d)$ are equivalent ([25] Theorem 7.44) and thus give rise to equivalent distance functions d .

A simple choice for $|||\cdot|||$ is $|||x||| = \max_{1 \leq k \leq n} |x^k|^{1/k}$ (which is neither symmetric nor sub-additive). However the norm most commonly used in the literature (and primarily in [25]) is the Carnot–Carathéodory norm on $G^n(\mathbb{R}^d)$ defined in terms of its sub-Riemannian structure. That is, the Carnot–Carathéodory norm is defined by

$$|||x||| = \inf\{|||\mathbf{x}|||_{1\text{-var};[0,T]} \mid \mathbf{x} \in \Omega_1, S_n(\mathbf{x})_{0,T} = x\}.$$

Moreover, one readily checks that $|||\cdot|||$ is both symmetric and sub-additive ([25] Proposition 7.40), and thus induces a metric d on $G^n(\mathbb{R}^d)$. Unless other stated, we always equip $G^n(\mathbb{R}^d)$ with the Carnot–Carathéodory norm $|||\cdot|||$ and the induced Carnot–Carathéodory metric d .

For a general metric space (E, d) , recall the space $C^{p\text{-var}}([0, T], E)$ of paths of finite p -variation, that is, the continuous paths $\mathbf{x} : [0, T] \mapsto E$ such that

$$\|\mathbf{x}\|_{p\text{-var};[0,T]} := \sup_{\mathcal{D} \subset [0,T]} \left(\sum_{t_j \in \mathcal{D}} d(\mathbf{x}_{t_j}, \mathbf{x}_{t_{j+1}})^p \right)^{1/p} < \infty. \quad (3.2.1)$$

For a set of paths taking values in $G^n(\mathbb{R}^d)$, we shall use the symbol o to denote the subset of paths $\mathbf{x}$ whose starting point is the identity element of $G^n(\mathbb{R}^d)$. For example, $C_o([0, T], G^n(\mathbb{R}^d))$ shall denote the set of all continuous paths $\mathbf{x} : [0, T] \mapsto G^n(\mathbb{R}^d)$ such that $\mathbf{x}_0 = 1_n$.

Remark that by equivalence of homogeneous norms on $G^n(\mathbb{R}^d)$, $\|\mathbf{x}\|_{p\text{-var}}$ is equivalent, up to multiplicative constants, to $\|\mathbf{x}\|_{p\text{-var}}$ defined by (3.1.2). In particular, in light of the discussion following Definition 3.1.2, $WG\Omega_p(\mathbb{R}^d)$ can be identified with the space $C_o^{p\text{-var}}([0, T], G^{\lfloor p \rfloor}(\mathbb{R}^d))$.

On the space $C^{p\text{-var}}([0, T], G^n(\mathbb{R}^d))$ we define the homogeneous p -variation metric by

$$d_{p\text{-var};[0,T]}(\mathbf{x}, \mathbf{y}) = d(\mathbf{x}_0, \mathbf{y}_0) + \sup_{\mathcal{D} \subset [0,T]} \left(\sum_{t_j \in \mathcal{D}} d(\mathbf{x}_{t_j, t_{j+1}}, \mathbf{y}_{t_j, t_{j+1}})^p \right)^{1/p}. \quad (3.2.2)$$

When we consider elements $\mathbf{x}, \mathbf{y} \in C_o^{p\text{-var}}([0, T], G^n(\mathbb{R}^d))$, note that the term $d(\mathbf{x}_0, \mathbf{y}_0)$ disappears. Unless otherwise stated, we always equip $C^{p\text{-var}}([0, T], G^n(\mathbb{R}^d))$ with the metric $d_{p\text{-var}}$.

Recall the definition of $\rho_{p\text{-var}}$ from (3.1.5). Remark that by replacing $\sup_{1 \leq k \leq \lfloor p \rfloor}$ with $\sup_{1 \leq k \leq n}$ generalises the definition of $\rho_{p\text{-var}}$ to a metric on $C_o^{p\text{-var}}([0, T], G^n(\mathbb{R}^d))$.

A fundamental link between $\rho_{p\text{-var}}$ and $d_{p\text{-var}}$ is that both metrics induce the same topology on $C_o^{p\text{-var}}([0, T], G^n(\mathbb{R}^d))$ and give rise to the same notion of bounded sets. In fact, for all $n \geq 1$, the identity map

$$Id : (C_o^{p\text{-var}}([0, T], G^n(\mathbb{R}^d)), d_{p\text{-var}}) \longleftrightarrow (C_o^{p\text{-var}}([0, T], G^n(\mathbb{R}^d)), \rho_{p\text{-var}})$$

is Lipschitz continuous on bounded sets in $\rightarrow$ direction, and $1/\lfloor p \rfloor$ -Hölder continuous on bounded sets in $\leftarrow$ direction ([25] Theorem 8.10).

We define the space

$$C_o^{0,p\text{-var}}([0, T], G^n(\mathbb{R}^d)) := \overline{S_n(C_o^\infty([0, T], \mathbb{R}^d))} \subset C_o^{p\text{-var}}([0, T], G^n(\mathbb{R}^d)).$$

Observe that $G\Omega_p(\mathbb{R}^d)$ is precisely the space $C_o^{0,p\text{-var}}([0, T], G^{\lfloor p \rfloor}(\mathbb{R}^d))$ and we have the following strict inclusions for all $1 \leq p < p'$ ([25] Corollary 8.24)

$$C_o^{0,p\text{-var}}([0, T], G^n(\mathbb{R}^d)) \subsetneq C_o^{p\text{-var}}([0, T], G^n(\mathbb{R}^d)) \subsetneq C_o^{0,p'\text{-var}}([0, T], G^n(\mathbb{R}^d)).$$

The following result summarises the important completeness and separability properties of these spaces.

Theorem 3.2.1 ([25] Theorem 8.13, Proposition 8.25). *For all $p \geq 1$ and integers $n \geq 1$, $C_o^{p\text{-var}}([0, T], G^n(\mathbb{R}^d))$ is a complete non-separable metric space. Moreover $C_o^{0,p\text{-var}}([0, T], G^n(\mathbb{R}^d))$ is a complete separable metric space, and thus Polish.*

3.2.2 Hölder reparametrisations

For a metric space (E, d) and $\alpha \in [0, 1]$, denote by $C^{\alpha\text{-Höl}}([0, T], E)$ the space of α -Hölder paths, that is, the continuous paths $\mathbf{x} : [0, T] \mapsto E$ such that

$$\|\mathbf{x}\|_{\alpha\text{-Höl};[0,T]} := \sup_{0 \leq s < t \leq T} \frac{d(\mathbf{x}_s, \mathbf{x}_t)}{|t - s|^\alpha} < \infty.$$

For $\mathbf{x} \in C^{p\text{-var}}([0, T], E)$, define the continuous non-decreasing function $\lambda_{p;\mathbf{x}} : [0, T] \mapsto [0, T]$ by

$$\lambda_{p;\mathbf{x}}(t) = T \|\mathbf{x}\|_{p\text{-var};[0,t]}^p \|\mathbf{x}\|_{p\text{-var};[0,T]}^{-p}.$$

We drop the reference to p and $\mathbf{x}$ when λ is clear from the context.

Define the Hölder reparametrisation function

$$P_p : C^{p\text{-var}}([0, T], E) \mapsto C^{p\text{-var}}([0, T], E)$$

by $P_p(\mathbf{x})_{\lambda(t)} = \mathbf{x}_t$. Observe that $P_p(\mathbf{x})$ is well-defined since, for all $s, t \in [0, T]$, $\lambda(s) = \lambda(t)$ implies $\mathbf{x}_s = \mathbf{x}_t$ (cf. [25] Proposition 5.14).

Lemma 3.2.2. *Let (E, d) be a metric space and $p \geq 1$. Then P_p maps $C^{p\text{-var}}([0, T], E)$ into $C^{1/p\text{-Höl}}([0, T], E)$. Moreover, for all $\mathbf{x} \in C^{p\text{-var}}([0, T], E)$, the function $t \mapsto \|P_p(\mathbf{x})\|_{p\text{-var};[0,t]}^p$ is linear in $t \in [0, T]$, and*

$$\|P_p(\mathbf{x})\|_{1/p\text{-Höl};[0,T]} \leq \|\mathbf{x}\|_{p\text{-var};[0,T]} T^{-1/p}. \quad (3.2.3)$$

Proof. For $t \in [0, T]$, consider $t' \in [0, T]$ such that $\lambda(t') = t$, and observe that

$$\|P_p(\mathbf{x})\|_{p\text{-var};[0,t]}^p = \lambda(t') \|\mathbf{x}\|_{p\text{-var};[0,T]}^p / T = t \|\mathbf{x}\|_{p\text{-var};[0,T]}^p / T.$$

In particular, it follows that $\|P_p(\mathbf{x})\|_{p\text{-var};[0,t]}^p$ is linear in $t \in [0, T]$.

Furthermore, note that for all $\mathbf{y} \in C^{p\text{-var}}([0, T], E)$ we have

$$d(\mathbf{y}_s, \mathbf{y}_t)^p \leq \|\mathbf{y}\|_{p\text{-var};[0,t]}^p - \|\mathbf{y}\|_{p\text{-var};[0,s]}^p,$$

from which it follows that

$$d(P_p(\mathbf{x})_t, P_p(\mathbf{x})_s)^p \leq (t - s) \|\mathbf{x}\|_{p\text{-var};[0,T]}^p / T.$$

□

The importance of the map P_p comes from the fact that for all $p \geq 1$, $\alpha \in (0, 1]$, and $c, r > 0$, the set

$$\left\{ \mathbf{x} \in C_o^{p\text{-var}}([0, T], G^n(\mathbb{R}^d)) \mid \|\mathbf{x}\|_{p\text{-var};[0,T]} \leq r, \|\mathbf{x}\|_{\alpha\text{-H\"ol};[0,T]} \leq c \right\}$$

is compact in $C_o^{0,p'\text{-var}}([0, T], G^n(\mathbb{R}^d))$ for all $p' > p$, which is a special case of an interpolation estimate and the Arzelà-Ascoli theorem ([25] Lemma 5.12, Proposition 8.17).

As an application, we can show that $WR_p(\mathbb{R}^d)$, the set of signatures of weakly geometric p -rough paths, is a σ -compact subset of $G(\mathbb{R}^d)$.

Proposition 3.2.3. *Let $p \geq 1$. Then $WR_p(\mathbb{R}^d)$ is σ -compact in $G(\mathbb{R}^d)$. In particular, $WR_p(\mathbb{R}^d)$ is a Borel set of $G(\mathbb{R}^d)$.*

Proof. For $r > 0$, consider the sets

$$B_p^r := \{ \mathbf{x} \in WG\Omega_p(\mathbb{R}^d) \mid \|\mathbf{x}\|_{p\text{-var};[0,T]} \leq r \}$$

and

$$C_p^r = \{ \mathbf{y} \in WG\Omega_p(\mathbb{R}^d) \mid \|\mathbf{y}\|_{1/p\text{-H\"ol};[0,T]} \leq rT^{-1/p} \}.$$

Note that $C_p^r \subset B_p^r$, and, due to (3.2.3), $P_p : B_p^r \mapsto C_p^r$. Let $p' > p$ be such that $\lfloor p' \rfloor = \lfloor p \rfloor$. It holds that C_p^r is compact in $(WG\Omega_{p'}(\mathbb{R}^d), d_{p'\text{-var}})$.

Since $\mathcal{I}_{p'} : (WG\Omega_{p'}(\mathbb{R}^d), d_{p'\text{-var}}) \mapsto WR_{p'}(\mathbb{R}^d)$ is continuous by Corollary 3.1.10, and $\mathcal{I}_{p'}(C_p^r) = \mathcal{I}_{p'}(B_p^r)$ by invariance of $S_n(\mathbf{x})_{0,T}$ under reparametrisation of $\mathbf{x}$, it follows that $\mathcal{I}_{p'}(B_p^r)$ is compact. Since $WR_p(\mathbb{R}^d) = \bigcup_{r \geq 1} \mathcal{I}_{p'}(B_p^r)$, it follows that $WR_p(\mathbb{R}^d)$ is σ -compact in $G(\mathbb{R}^d)$. $\square$

The following result will be heavily used in our study of convergence of RDEs and of finiteness of p -variation of limiting processes.

For the remainder of the section, we adopt the shorthand notation $G^N := G^N(\mathbb{R}^d)$ and let all unspecified path spaces be defined on $[0, T]$ and take values in G^N , for example, $C_o^{p\text{-var}}$ shall denote $C_o^{p\text{-var}}([0, T], G^N(\mathbb{R}^d))$.

We equip C_o with the uniform metric d_∞ . Recall that $(C_o^{0,p\text{-var}}, d_{p\text{-var}})$ and (C_o, d_∞) are Polish spaces. Observe that P_p is a measurable function from $C_o^{p\text{-var}}$ to $C_o^{1/p\text{-H\"ol}}$.

Proposition 3.2.4. *Let $p \geq 1$ and $(\mathbf{X}_n)_{n \geq 1}$ a sequence of C_o -valued random variables such that $(\|\mathbf{X}_n\|_{p\text{-var};[0,T]})_{n \geq 1}$ is a tight collection of real random variables. Suppose that $\mathbf{X}_n \xrightarrow{\mathcal{D}} \mathbf{X}$ as C_o -valued random variables.*

Then for every $p' > p$, $\|\mathbf{X}\|_{p'\text{-var};[0,T]} < \infty$ a.s. and $P_{p'}(\mathbf{X}_n) \xrightarrow{\mathcal{D}} P_{p'}(\mathbf{X})$ as $C_o^{0,p'\text{-var}}$ -valued random variables.

For the proof of Proposition 3.2.4, we require the following lemmas.

Lemma 3.2.5. *Let $(\mathbf{X}_n)_{n \geq 1}$ be a sequence of C_o -valued random variables such that $(\|\mathbf{X}_n\|_{p\text{-var};[0,T]})_{n \geq 1}$ is a tight collection of real random variables. Then for every $p' > p$, $(P_{p'}(\mathbf{X}_n))_{n \geq 1}$ is a tight collection of $C_o^{0,p'\text{-var}}$ -valued random variables.*

Proof. Let $p' > p$ and denote $\tilde{\mathbf{X}}_n = P_{p'}(\mathbf{X}_n)$. Observe that

$$\left\| \tilde{\mathbf{X}}_n \right\|_{p\text{-var};[0,T]} = \|\mathbf{X}_n\|_{p\text{-var};[0,T]},$$

and, due to (3.2.3),

$$\left\| \tilde{\mathbf{X}}_n \right\|_{1/p'\text{-Hö};[0,T]} \leq \|\mathbf{X}_n\|_{p'\text{-var};[0,T]} T^{-1/p'} \leq \|\mathbf{X}_n\|_{p\text{-var};[0,T]} T^{-1/p'}.$$

Recall that for all $c, r > 0$,

$$\left\{ \mathbf{x} \in C_o^{p\text{-var}} \mid \|\mathbf{x}\|_{p\text{-var};[0,T]} \leq r, \|\mathbf{x}\|_{1/p'\text{-Hö};[0,T]} \leq c \right\}$$

is a compact subset of $C_o^{0,p'\text{-var}}$. Since $(\|\mathbf{X}_n\|_{p\text{-var};[0,T]})_{n \geq 1}$ is tight, it follows that $(\tilde{\mathbf{X}}_n)_{n \geq 1}$ is a tight collection of $C_o^{0,p'\text{-var}}$ -valued random variables. $\square$

Lemma 3.2.6. *Let $q \geq 1$, $\mathbf{x} \in C_o^{q\text{-var}}$, and $(\mathbf{x}_n)_{n \geq 1}$ a sequence in $C_o^{q\text{-var}}$ such that $t \mapsto \|\mathbf{x}_n\|_{q\text{-var};[0,t]}^q$ is linear for all $n \geq 1$. Suppose that $\lim_{n \rightarrow \infty} d_{q\text{-var}}(\mathbf{x}_n, \mathbf{x}) = 0$. Then $t \mapsto \|\mathbf{x}\|_{q\text{-var};[0,t]}^q$ is linear.*

Proof. Observe that for all $t \in [0, T]$

$$\|\mathbf{x}\|_{q\text{-var};[0,t]}^q = \lim_{n \rightarrow \infty} \|\mathbf{x}_n\|_{q\text{-var};[0,t]}^q = \lim_{n \rightarrow \infty} (t/T) \|\mathbf{x}_n\|_{q\text{-var};[0,T]}^q = (t/T) \|\mathbf{x}\|_{q\text{-var};[0,T]}^q.$$

$\square$

Denote by Λ the set of continuous non-decreasing surjections $\lambda : [0, T] \mapsto [0, T]$.

Lemma 3.2.7. *Let (E, d) be a metric space and $\mathbf{x} \in C([0, T], E)$. Suppose there does not exist a non-empty interval $(s, t) \subseteq [0, T]$ such that $\mathbf{x}_u = \mathbf{x}_s$ for all $u \in (s, t)$. Let $A_{\mathbf{x}} := \{\mathbf{x} \circ \lambda \mid \lambda \in \Lambda\}$. Then $A_{\mathbf{x}}$ is closed in $C([0, T], E)$ under the uniform topology.*

Proof. Let $\lambda_n \in \Lambda$ and $\mathbf{x}_n = \mathbf{x} \circ \lambda_n \in A_{\mathbf{x}}$ such that $\lim_{n \rightarrow \infty} d_{\infty}(\mathbf{x}_n, \mathbf{y}) = 0$ for some $\mathbf{y} \in C([0, T], E)$. The proof of the lemma follows by exactly the same argument as the proof of [3] Proposition 5.3. Indeed, to follow the notation of [3], we set $x := \mathbf{y}$, $x' := \mathbf{x}$, and $\sigma_n := \lambda_n$. Observe that parts (1), (2) and (3) of the proof of [3] Proposition 5.3 only require that x' is not constant on any interval, and that $\sigma_n : [0, T] \mapsto [0, T]$ is a continuous non-decreasing surjection. It follows from the conclusion of the proof of [3] Proposition 5.3, part (3), that there exists a continuous non-decreasing surjection $\sigma : [0, T] \mapsto [0, T]$ such that $x = x' \circ \sigma$. $\square$

Proof of Proposition 3.2.4. Let $p' > p$ and let $\tilde{\mathbf{X}}_n := P_{p'}(\mathbf{X}_n)$. By Lemma 3.2.5, $(\tilde{\mathbf{X}}_n)_{n \geq 1}$ is a tight collection of $C_o^{0,p'-\text{var}}$ -valued random variables. Recall that $C_o^{0,p'-\text{var}}$ is a Polish space (Theorem 3.2.1).

Let $\tilde{\mathbf{X}}$ be a $C_o^{0,p'-\text{var}}$ -valued random variable which is a weak limit point of $\tilde{\mathbf{X}}_n$, so that $\tilde{\mathbf{X}}_{n_k}$ converges in law to $\tilde{\mathbf{X}}$ along a subsequence $(n_k)_{k \geq 1}$.

Consider the pairs $\mathbf{Y}_k := (\mathbf{X}_{n_k}, \tilde{\mathbf{X}}_{n_k})$ as $C_o \times C_o^{0,p'-\text{var}}$ -valued random variables (we keep in mind that $\tilde{\mathbf{X}}_{n_k}$ is a measurable function of $\mathbf{X}_{n_k}$). The collection $(\mathbf{Y}_k)_{k \geq 1}$ is tight, so let $\mathbf{Y}$ be a $C_o \times C_o^{0,p'-\text{var}}$ -valued random variable which is a weak limit point of $\mathbf{Y}_k$ along the subsequence $(k_m)_{m \geq 1}$.

Let $\mathbf{y}^1, \mathbf{y}^2$ denote the projections of $\mathbf{y} \in C_o \times C_o^{0,p'-\text{var}}$ onto C_o and $C_o^{0,p'-\text{var}}$ respectively. Then $\mathbf{Y}^1 \stackrel{\mathcal{D}}{=} \mathbf{X}$ and $\mathbf{Y}^2 \stackrel{\mathcal{D}}{=} \tilde{\mathbf{X}}$. We claim that it suffices to prove

- (1) $\mathbf{Y}^1$ is in $C_o^{p'-\text{var}}$ almost surely, and
- (2) $\mathbf{Y}^2 = P_{p'}(\mathbf{Y}^1)$ almost surely.

Indeed, suppose (1) and (2) hold. Since $\mathbf{Y}^1 \stackrel{\mathcal{D}}{=} \mathbf{X}$, (1) implies that $\mathbf{X}$ takes values in $C_o^{p'-\text{var}}$ almost surely. Moreover, since $\mathbf{Y}^2 \stackrel{\mathcal{D}}{=} \tilde{\mathbf{X}}$, (2) implies that $\tilde{\mathbf{X}} \stackrel{\mathcal{D}}{=} P_{p'}(\mathbf{X})$. Since $\tilde{\mathbf{X}}$ was an arbitrary weak limit point of $\tilde{\mathbf{X}}_n$, it holds that every weak limit point of $\tilde{\mathbf{X}}_n$ is equal in law to $P_{p'}(\mathbf{X})$. Hence $\tilde{\mathbf{X}}_n \xrightarrow{\mathcal{D}} P_{p'}(\mathbf{X})$ as $C_o^{0,p'-\text{var}}$ -valued random variables as desired.

To complete the proof, it thus suffices to show (1) and (2). Since $C_o \times C_o^{0,p'-\text{var}}$ is a Polish space, by the Skorokhod representation theorem, there exist $C_o \times C_o^{0,p'-\text{var}}$ -valued random variables $\tilde{\mathbf{Y}}_m$ and $\tilde{\mathbf{Y}}$ defined on a common probability space $(\Omega, \mathcal{F}, \mathbb{P})$ such that $\tilde{\mathbf{Y}}_m \stackrel{\mathcal{D}}{=} \mathbf{Y}_{k_m}$ and $\tilde{\mathbf{Y}} \stackrel{\mathcal{D}}{=} \mathbf{Y}$ and such that $\lim_{m \rightarrow \infty} \tilde{\mathbf{Y}}_m(\omega) = \mathbf{Y}(\omega)$ for $\mathbb{P}$ -almost all $\omega \in \Omega$. In particular, for $\mathbb{P}$ -almost all ω ,

$$\lim_{m \rightarrow \infty} d_{p'-\text{var}}(\tilde{\mathbf{Y}}_m^2(\omega), \tilde{\mathbf{Y}}^2(\omega)) = 0.$$

Recall that for all $m \geq 1$ and $\omega \in \Omega$, $\left\| \tilde{\mathbf{Y}}_m^2(\omega) \right\|_{p'-\text{var};[0,t]}^{p'}$ is linear in $t \in [0, T]$. It follows from Lemma 3.2.6 that for $\mathbb{P}$ -almost all ω , $\left\| \tilde{\mathbf{Y}}^2(\omega) \right\|_{p'-\text{var};[0,t]}^{p'}$ is also linear in $t \in [0, T]$, from which it follows there does not exist a non-empty interval $(s, t) \subseteq [0, T]$ such that $\tilde{\mathbf{Y}}^2(\omega)_u = \tilde{\mathbf{Y}}^2(\omega)_s$ for all $u \in (s, t)$.

For every $m \geq 1$, $\omega \in \Omega$, and $t \in [0, T]$, define $\lambda_m^\omega(t) := \lambda_{p'; \tilde{\mathbf{Y}}_m^1(\omega)}(t)$ and $\mathbf{Z}_m(\omega)(t) := \tilde{\mathbf{Y}}^2(\omega)(\lambda_m^\omega(t))$. Since $\tilde{\mathbf{Y}}_m^2(\omega) \circ \lambda_m^\omega = \tilde{\mathbf{Y}}_m^1(\omega)$ for all $m \geq 1$ and $\mathbb{P}$ -almost all ω , we have that $\mathbb{P}$ -a.s.

$$d_\infty(\tilde{\mathbf{Y}}_m^2, \tilde{\mathbf{Y}}^2) = d_\infty(\tilde{\mathbf{Y}}_m^1, \mathbf{Z}_m).$$

Since

$$\lim_{m \rightarrow \infty} d_\infty(\tilde{\mathbf{Y}}_m^1(\omega), \tilde{\mathbf{Y}}^1(\omega)) = 0 = \lim_{m \rightarrow \infty} d_\infty(\tilde{\mathbf{Y}}_m^2(\omega), \tilde{\mathbf{Y}}^2(\omega))$$

for $\mathbb{P}$ -almost all ω , we have

$$\lim_{m \rightarrow \infty} d_\infty(\mathbf{Z}_m(\omega), \tilde{\mathbf{Y}}^1(\omega)) = 0 \text{ for } \mathbb{P}\text{-almost all } \omega \in \Omega. \quad (3.2.4)$$

Since for all $\omega \in \Omega$, by definition, $\mathbf{Z}_m(\omega) = \tilde{\mathbf{Y}}^2(\omega) \circ \lambda_m^\omega$ for $\lambda_m^\omega \in \Lambda$, it follows from Lemma 3.2.7 and (3.2.4) that for $\mathbb{P}$ -almost all $\omega \in \Omega$ there exists $\lambda^\omega \in \Lambda$ such that $\tilde{\mathbf{Y}}^1(\omega) = \tilde{\mathbf{Y}}^2(\omega) \circ \lambda^\omega$ (note we make no claims about the measurability of λ^ω with respect to ω). In particular, since $\tilde{\mathbf{Y}}^2$ is a $C_o^{0,p'\text{-var}}$ -valued random variable,

$$\left\| \tilde{\mathbf{Y}}^1(\omega) \right\|_{p'\text{-var};[0,T]} = \left\| \tilde{\mathbf{Y}}^2(\omega) \right\|_{p'\text{-var};[0,T]} < \infty,$$

for $\mathbb{P}$ -almost all ω , which proves (1).

Moreover for $\mathbb{P}$ -almost all ω , since $t \mapsto \left\| \tilde{\mathbf{Y}}^2(\omega) \right\|_{p'\text{-var};[0,t]}^{p'}$ is linear, it holds that for all $t \in [0, T]$

$$\begin{aligned} \left\| \tilde{\mathbf{Y}}^1(\omega) \right\|_{p'\text{-var};[0,t]}^{p'} &= \left\| \tilde{\mathbf{Y}}^2(\omega) \circ \lambda^\omega \right\|_{p'\text{-var};[0,t]}^{p'} \\ &= \left\| \tilde{\mathbf{Y}}^2(\omega) \right\|_{p'\text{-var};[0,\lambda^\omega(t)]}^{p'} \\ &= (\lambda^\omega(t)/T) \left\| \tilde{\mathbf{Y}}^2(\omega) \right\|_{p'\text{-var};[0,T]}^{p'}, \end{aligned}$$

which is equivalent to the statement that $\lambda^\omega(t) = \lambda_{p'; \tilde{\mathbf{Y}}^1(\omega)}(t)$. Thus $\tilde{\mathbf{Y}}^2(\omega) = P_{p'}(\tilde{\mathbf{Y}}^1(\omega))$ for $\mathbb{P}$ -almost all ω , which proves (2) and completes the proof. $\square$

3.3 Rough differential equations

We now discuss the concept of a rough differential equation (RDE), which forms one of the central topics in the study of rough paths. We shall follow the approach via discrete Euler approximations, which was introduced by Davie [16] and adopted as the main tool in studying RDEs by Friz and Victoir [25]. In doing so, we shall only define and study RDEs in the finite dimensional setting $V = \mathbb{R}^d$.

We note however that the original approach taken by Lyons [45], and presented in [42] and [46], is focused around Picard iterations and allows to give meaning to RDEs where V is a general Banach space. The two contrasting methods may be compared to the difference between the Cauchy-Peano and Picard–Lindelöf existence theorems.

3.3.1 Lipschitz vector fields

We recall now a simplified definition of Lipschitz maps in the sense of E. Stein. We stress that the following definition is less general than the standard definition considered in the literature (e.g., [42] Definition 5.1.1, or [46] Definition 1.21), but it shall cover all cases of interest to us.

Definition 3.3.1. For $\gamma > 0$, normed spaces V and W , and an open set $U \subseteq V$, a map $f : U \mapsto W$ is called γ -Lipschitz on U if f is in $C^{\lfloor \gamma \rfloor}(U, W)$ (i.e., $\lfloor \gamma \rfloor$ -times continuously differentiable on U) and if there exists $K \geq 0$ such that

$$\sup_{0 \leq k \leq \lfloor \gamma \rfloor} \sup_{x \in U} \|D^k f(x)\| \leq K, \text{ and } \|D^{\lfloor \gamma \rfloor} f\|_{(\gamma - \lfloor \gamma \rfloor)\text{-H\"{o}l}; U} \leq K.$$

The smallest $K \geq 0$ for which this holds is denoted by $\|f\|_{\text{Lip}^\gamma(U)}$, the Lip^γ -norm of f on U .

Let $\text{Lip}^\gamma(U, W)$ denote the space of all γ -Lipschitz maps on U , and $\text{Lip}_{\text{loc}}^\gamma(V, W)$ denote the space of all $f : V \mapsto W$ which are in $\text{Lip}^\gamma(U, W)$ for all open bounded subsets $U \subset V$.

Remark 3.3.2. The k -th derivative of f is a map $D^k f : U \mapsto \mathbf{L}(V^{\otimes k}, W)$, so that for $x \in U$, $\|D^k f(x)\|$ depends implicitly on the tensor norm equipped on $V^{\otimes k}$ (where $\mathbf{L}(V^{\otimes k}, W)$ is equipped with the operator norm). When V is finite dimensional, all such norms lead to equivalent definitions of the Lip^γ -norm. However in the infinite dimensional setting, one should first fix an admissible norm on $V^{\otimes k}$, as different norms in general lead to non-equivalent definitions of Lip^γ (see [42] Section 3.1).

Let $F(\mathbb{R}^e, \mathbb{R}^e)$ denote the vector space of functions mapping $\mathbb{R}^e$ to $\mathbb{R}^e$. We call a linear map $f : \mathbb{R}^d \mapsto F(\mathbb{R}^e, \mathbb{R}^e)$, a collection of d vector fields on $\mathbb{R}^e$.

Note that f can be equivalently viewed as a map $f : \mathbb{R}^e \mapsto \mathbf{L}(\mathbb{R}^d, \mathbb{R}^e)$. When $\mathbf{L}(\mathbb{R}^d, \mathbb{R}^e)$ is equipped with the operator norm, we say that f is a collection of Lip^γ vector fields if $f \in \text{Lip}^\gamma(\mathbb{R}^e, \mathbf{L}(\mathbb{R}^d, \mathbb{R}^e))$.

3.3.2 Universal limit theorem

One of the fundamental results in rough paths theory is the *universal limit theorem*, which, under sufficient regularity on the driving vector fields, asserts the existence, uniqueness and continuity (in p -variation topologies) of solutions to the differential equation

$$d\mathbf{y}_t = f(\mathbf{y}_t)d\mathbf{x}_t, \quad \mathbf{y}_0 = y \in \mathbb{R}^e, \tag{3.3.1}$$

where f is a collection of d vector fields on $\mathbb{R}^e$, and $\mathbf{x}$ is an element of $WG\Omega_p(\mathbb{R}^d)$ for some $p \geq 1$.

To give meaning to an RDE, we first consider paths of bounded variation. For $\mathbf{x} \in C^{1\text{-var}}([0, T], \mathbb{R}^d)$ and $\mathbf{z} \in C([0, T], \mathbf{L}(\mathbb{R}^d, \mathbb{R}^e))$, recall the Riemann-Stieltjes integral

$$\int_0^T \mathbf{z}_t d\mathbf{x}_t := \lim_{n \rightarrow \infty} \sum_{t_j \in \mathcal{D}_n} \mathbf{z}_{t_j}(\mathbf{x}_{t_j, t_{j+1}}),$$

where, as usual, we denote $\mathbf{x}_{t_j, t_{j+1}} := \mathbf{x}_{t_{j+1}} - \mathbf{x}_{t_j}$, and where $\mathcal{D}_n = (t_0^n, \dots, t_{k_n}^n)$ is a sequence of partitions of $[0, T]$ such that $\lim_{n \rightarrow \infty} \max_{1 \leq j \leq k_n} |t_j^n - t_{j-1}^n| = 0$. Moreover, recall that $\int_0^T \mathbf{z}_t d\mathbf{x}_t$ is independent of the particular sequence $\mathcal{D}_n$. Furthermore $\int_0^T \mathbf{z}_t d\mathbf{x}_t$ is independent of the starting point of $\mathbf{x}$ in the sense that

$$\int_0^T \mathbf{z}_t d\mathbf{x}_t = \int_0^T \mathbf{z}_t d\tilde{\mathbf{x}}_t,$$

where $\tilde{\mathbf{x}} := \mathbf{x}_t - \mathbf{x}_0$. We shall thus restrict our driving signal $\mathbf{x}$ to have starting point zero.

A solution to (3.3.1) for a driving signal $\mathbf{x} \in C_o^{1\text{-var}}([0, T], \mathbb{R}^d)$, continuous vector fields $f \in C(\mathbb{R}^e, \mathbf{L}(\mathbb{R}^d, \mathbb{R}^e))$, and a starting point $y \in \mathbb{R}^e$, is a path $\mathbf{y} \in C([0, T], \mathbb{R}^e)$ such that $\mathbf{y}_0 = y$ and for all $t \in [0, T]$

$$\mathbf{y}_t = \mathbf{y}_0 + \int_0^t f(\mathbf{y}_s) d\mathbf{x}_s.$$

Recall that the classical Cauchy-Peano and Picard-Lindelöf theorems assert that if f is continuous, then there exists at least one solution to (3.3.1), and if moreover f is Lipschitz continuous, then this solution is unique.

Denote by $\pi_{(f)}(0, y; \mathbf{x}) : [0, T] \mapsto \mathbb{R}^e$ the unique solution to (3.3.1) for Lipschitz continuous f (note that in [25] $\pi_{(f)}(0, y; \mathbf{x})$ refers to the collection of all solutions in the case that f is not necessarily Lipschitz continuous). More generally, for $s \in [0, T]$, let $\pi_{(f)}(s, y; \mathbf{x}) : [s, T] \mapsto \mathbb{R}^e$ denote the solution with time- s initial condition $\mathbf{y}_s = y$. The map

$$\pi_{(f)} : C_o^{1\text{-var}}([0, T], \mathbb{R}^d) \times \mathbb{R}^e \mapsto C([0, T], \mathbb{R}^e), \quad (\mathbf{x}, y) \mapsto \pi_{(f)}(0, y; \mathbf{x}),$$

is called the Itô map.

We summarise the existence, uniqueness and continuity results concerning the Itô map in the following theorem.

Theorem 3.3.3 ([25] Theorems 3.4, 3.18). *Let $y \in \mathbb{R}^e$ and $\mathbf{x} \in C_o^{1\text{-var}}([0, T], \mathbb{R}^d)$. If $f \in C_b(\mathbb{R}^e, \mathbf{L}(\mathbb{R}^d, \mathbb{R}^e))$, then there exists at least one solution to (3.3.1) and all solutions are elements of $C^{1\text{-var}}([0, T], \mathbb{R}^e)$.*

If moreover $f \in \text{Lip}^1(\mathbb{R}^e, \mathbf{L}(\mathbb{R}^d, \mathbb{R}^e))$, then the solution to (3.3.1) is unique, and the Itô map

$$\pi_{(f)} : C_o^{1\text{-var}}([0, T], \mathbb{R}^d) \times \mathbb{R}^e \mapsto C^{1\text{-var}}([0, T], \mathbb{R}^e)$$

is uniformly continuous on every set of the form

$$\{\|\mathbf{x}\|_{1\text{-var};[0,T]} < R\} \times \mathbb{R}^e, \quad R > 0,$$

when $C_o^{1\text{-var}}([0, T], \mathbb{R}^d)$ and $C^{1\text{-var}}([0, T], \mathbb{R}^e)$ are equipped with their respective $d_{1\text{-var}}$ metrics.

We are now ready to state what we mean by a solution to (3.3.1) where $\mathbf{x}$ is a weakly geometric p -rough path.

Definition 3.3.4 (Rough differential equation). Let $f \in C(\mathbb{R}^e, \mathbf{L}(\mathbb{R}^d, \mathbb{R}^e))$ be a collection of vector fields on $\mathbb{R}^e$ and $\mathbf{x} \in WGO_p(\mathbb{R}^d)$. We say that $\mathbf{y} \in C([0, T], \mathbb{R}^e)$ is a solution to (3.3.1) if there exists a sequence $(\mathbf{x}_n)_{n \geq 1}$ in $C_o^{1\text{-var}}([0, T], \mathbb{R}^d)$ such that

$$\lim_{n \rightarrow \infty} d_\infty(S_{[p]}(\mathbf{x}_n), \mathbf{x}) = 0 \quad \text{and} \quad \sup_{n \geq 1} \|S_{[p]}(\mathbf{x}_n)\|_{p\text{-var};[0,T]} < \infty, \quad (3.3.2)$$

and there are solutions $\mathbf{y}_n \in C^{1\text{-var}}([0, T], \mathbb{R}^e)$ to (3.3.1) driven by $\mathbf{x}_n$ with initial condition $\mathbf{y}_n(0) = y$ such that

$$\lim_{n \rightarrow \infty} d_\infty(\mathbf{y}_n, \mathbf{y}) = 0.$$

We now state a version of the universal limit theorem which extends Theorem 3.3.3 to weakly geometric rough paths.

Theorem 3.3.5 (Universal limit theorem, [25] Theorem 10.14, Corollary 10.28). *Let $\gamma > p \geq 1$, $y \in \mathbb{R}^e$, and $\mathbf{x} \in WGO_p(\mathbb{R}^d)$. If $f \in \text{Lip}^{\gamma-1}(\mathbb{R}^e, \mathbf{L}(\mathbb{R}^d, \mathbb{R}^e))$, then there exists at least one solution to (3.3.1) and all solutions are elements of $C^{p\text{-var}}([0, T], \mathbb{R}^e)$.*

If moreover $f \in \text{Lip}^\gamma(\mathbb{R}^e, \mathbf{L}(\mathbb{R}^d, \mathbb{R}^e))$, then the solution to (3.3.1) is unique, and the Itô map

$$\pi_{(f)} : WGO_p(\mathbb{R}^d) \times \mathbb{R}^e \mapsto C^{p\text{-var}}([0, T], \mathbb{R}^e)$$

is uniformly continuous on every set of the form

$$\{\|\mathbf{x}\|_{p\text{-var};[0,T]} < R\} \times \mathbb{R}^e, \quad R > 0,$$

when $WGO_p(\mathbb{R}^d)$ and $C^{p\text{-var}}([0, T], \mathbb{R}^e)$ are equipped with their respective $d_{p\text{-var}}$ metrics.

3.3.3 Non-explosion

In connection with the characteristic function defined in Chapter 2, we shall be interested in RDEs driven along linear vector fields, which do not lie in $\text{Lip}^\gamma(\mathbb{R}^e, \mathbf{L}(\mathbb{R}^d, \mathbb{R}^e))$ for any $\gamma > 0$. We shall see however that RDEs driven along linear vector fields do admit solutions with a bound on their growth. This leads to the following non-explosion definition (which is a variation of [25] Definition 11.1).

Definition 3.3.6 (*p*-non-explosion). Let $p \geq 1$. We say that vector fields $f \in C(\mathbb{R}^e, \mathbf{L}(\mathbb{R}^d, \mathbb{R}^e))$ satisfy the *p*-non-explosion condition if for every $\mathbf{x} \in W\Omega_p(\mathbb{R}^d)$ and $y \in \mathbb{R}^e$ there exists at least one solution $\mathbf{y} \in C([0, T], \mathbb{R}^e)$ to (3.3.1), and for every $c > 0$ there exists $r > 0$ such that if $\|\mathbf{x}\|_{p\text{-var};[0,T]} + \|y\| < c$, then every solution to (3.3.1) satisfies $\|\mathbf{y}\|_{\infty;[0,T]} < r$.

Evidently, by Theorem 3.3.5, all vector fields $f \in \text{Lip}^\gamma(\mathbb{R}^e, \mathbf{L}(\mathbb{R}^d, \mathbb{R}^e))$ for $\gamma > p$ satisfy the *p*-non-explosion condition (in fact, all $f \in \text{Lip}^{\gamma-1}(\mathbb{R}^e, \mathbf{L}(\mathbb{R}^d, \mathbb{R}^e))$ satisfy the *p*-non-explosion condition, see [25] Theorem 10.14).

Proposition 3.3.7. Let $\gamma > p \geq 1$, $f \in \text{Lip}_{\text{loc}}^\gamma(\mathbb{R}^e, \mathbf{L}(\mathbb{R}^d, \mathbb{R}^e))$ satisfy the *p*-non-explosion condition, $y \in \mathbb{R}^e$, and $\mathbf{x} \in W\Omega_p(\mathbb{R}^d)$.

Then there exists a unique solution $\pi_{(f)}(0, y; \mathbf{x})$ to (3.3.1). Moreover $\pi_{(f)}(0, y; \mathbf{x})$ is an element of $C^{p\text{-var}}([0, T], \mathbb{R}^e)$ and the Itô map

$$\pi_{(f)} : W\Omega_p(\mathbb{R}^d) \times \mathbb{R}^e \mapsto C^{p\text{-var}}([0, T], \mathbb{R}^e)$$

is uniformly continuous on every set of the form

$$\{\|\mathbf{x}\|_{p\text{-var};[0,T]} < R_1\} \times \{\|y\| < R_2\}, \quad R_1, R_2 > 0,$$

when $W\Omega_p(\mathbb{R}^d)$ and $C^{p\text{-var}}([0, T], \mathbb{R}^e)$ are equipped with their respective $d_{p\text{-var}}$ metrics.

Proof. Consider a set $A = \{\|\mathbf{x}\|_{p\text{-var};[0,T]} < R_1\} \times \{\|y\| < R_2\}$, for $R_1, R_2 > 0$. By the *p*-non-explosion condition, for all $(\mathbf{x}, y) \in A$, there exists $r > 0$ such that every solution $\mathbf{y} \in C([0, T], \mathbb{R}^e)$ to (3.3.1) (of which there is at least one) satisfies $\|\mathbf{y}\|_{\infty;[0,T]} < r$.

Let $g \in \text{Lip}^\gamma(\mathbb{R}^e, \mathbf{L}(\mathbb{R}^d, \mathbb{R}^e))$ such that $g \equiv f$ on the set $\{\|y\| < r + 1\}$. It follows that for all $(\mathbf{x}, y) \in A$, every solution to (3.3.1) is also a solution to

$$d\mathbf{y}_t = g(\mathbf{y})d\mathbf{x}_t, \quad \mathbf{y}_0 = y \in \mathbb{R}^e. \quad (3.3.3)$$

However by Theorem 3.3.5, (3.3.3) admits a unique solution $\pi_{(g)}(0, y; \mathbf{x})$, which is an element of $C^{p\text{-var}}([0, T], \mathbb{R}^e)$, and $\pi_{(g)} : A \mapsto C^{p\text{-var}}([0, T], \mathbb{R}^e)$ is uniformly continuous, from which the desired result follows. $\square$

3.3.4 Linear vector fields

By linear vector fields we mean a linear map $M \in \mathbf{L}(\mathbb{R}^e, \mathbf{L}(\mathbb{R}^d, \mathbb{R}^e))$, which, as usual, can equivalently be viewed as a map $M \in \mathbf{L}(\mathbb{R}^d, \mathbf{L}(\mathbb{R}^e))$.

Note that every $M \in \mathbf{L}(\mathbb{R}^d, \mathbf{L}(\mathbb{R}^e))$ is in $\text{Lip}_{\text{loc}}^\gamma(\mathbb{R}^e, \mathbf{L}(\mathbb{R}^d, \mathbb{R}^e))$ but that $M \notin \text{Lip}^\gamma(\mathbb{R}^e, \mathbf{L}(\mathbb{R}^d, \mathbb{R}^e))$ (unless $M = 0$) for all $\gamma > 0$, since M is a smooth and (globally) unbounded map.

Theorem 3.3.8 (RDEs along linear vector fields, [25] Theorem 10.53). *Let $p \geq 1$, $M \in \mathbf{L}(\mathbb{R}^e, \mathbf{L}(\mathbb{R}^d, \mathbb{R}^e))$, and $\mathbf{x} \in W\Omega_p(\mathbb{R}^d)$. Then there exists a unique solution in $C([0, T], \mathbb{R}^e)$ to*

$$d\mathbf{y}_t = M(\mathbf{y}_t)d\mathbf{x}_t \quad \mathbf{y}_0 = y \in \mathbb{R}^e. \quad (3.3.4)$$

Moreover there exists a constant C depending only on p and $\|M\|$ such that for all $0 \leq s \leq t \leq T$

$$\|\mathbf{y}_{s,t}\| \leq C(1 + \|y\|) \|\mathbf{x}\|_{p\text{-var};[s,t]} \exp\left(C \|\mathbf{x}\|_{p\text{-var};[0,T]}^p\right). \quad (3.3.5)$$

Remark 3.3.9. A similar result holds for general Lipschitz continuous vector fields f which are locally Lip^γ ([25] Exercise 10.56).

In particular, M satisfies the p -non-explosion condition for all $p \geq 1$, so the uniqueness and continuity results of Proposition 3.3.7 apply to M . As before, we denote the unique solution to (3.3.4) by $\pi_{(M)}(0, y; \mathbf{x}) \in C^{p\text{-var}}([0, T], \mathbb{R}^e)$.

3.4 Flows of diffeomorphisms

Let $\mathbf{x} \in W\Omega_p(\mathbb{R}^d)$ and f a collection of vector fields. Whenever the equation (3.3.1) admits a unique solution for all $y \in \mathbb{R}^e$, we denote by $U_{T \leftarrow 0}^\mathbf{x} : \mathbb{R}^e \mapsto \mathbb{R}^e$ the associated flow map which maps $y = \mathbf{y}_0 \mapsto \mathbf{y}_T$.

Recall that $\overleftarrow{\mathbf{x}}_{s,t} := \mathbf{x}_{T-t, T-s}$ denotes the time-reversal of $\mathbf{x} \in \Omega_p(\mathbb{R}^d)$. For $\mathbf{x} \in C_o^{1\text{-var}}([0, T], \mathbb{R}^d)$, $f \in \text{Lip}^1(\mathbb{R}^e, \mathbf{L}(\mathbb{R}^d, \mathbb{R}^e))$ and $y \in \mathbb{R}^e$, let $\mathbf{y} = \pi_{(f)}(0, y; \mathbf{x})$. One can readily check that $\pi_{(f)}(0, \mathbf{y}_T; \overleftarrow{\mathbf{x}})$, the unique solution to

$$d\mathbf{z}_t = f(\mathbf{z}_t)d\overleftarrow{\mathbf{x}}_t, \quad \mathbf{z}_0 = \mathbf{y}_T,$$

is precisely $\mathbf{z}_t = \mathbf{y}_{T-t}$. It follows that for all $\mathbf{x} \in C_o^{1\text{-var}}([0, T], \mathbb{R}^d)$, the flow map $U_{T \leftarrow 0}^\mathbf{x} : y \mapsto \mathbf{y}_T$ is a bijection of $\mathbb{R}^e$ onto $\mathbb{R}^e$, and

$$U_{T \leftarrow 0}^{\overleftarrow{\mathbf{x}}} = (U_{T \leftarrow 0}^\mathbf{x})^{-1}. \quad (3.4.1)$$

For continuous paths $\mathbf{x} \in C_o([0, t], G^n(\mathbb{R}^d))$ and $\mathbf{y} \in C_o([t, T], G^n(\mathbb{R}^d))$, consider their concatenation $\mathbf{z} := \mathbf{x} * \mathbf{y} \in C_o([0, T], G^n(\mathbb{R}^d))$ defined by $\mathbf{z}_s = \mathbf{x}_s$ and $\mathbf{z}_s = \mathbf{x}_t \mathbf{y}_s$ for $s \in [0, t]$ and $s \in [t, T]$ respectively.

We extend the definition to arbitrary $\mathbf{x}^1, \dots, \mathbf{x}^k \in C_o([0, T], G^n(\mathbb{R}^d))$ by letting

$$\hat{\mathbf{x}}^j \in C_o([(j-1)T/k, jT/k], G^n(\mathbb{R}^d)), \quad \hat{\mathbf{x}}_t^j := \mathbf{x}_{\psi_j(t)}^j,$$

where ψ_j is the unique increasing linear bijection from $[(j-1)T/k, jT/k]$ to $[0, T]$, and defining $\mathbf{x}^1 * \dots * \mathbf{x}^k := (\dots (\hat{\mathbf{x}}^1 * \hat{\mathbf{x}}^2) * \dots) * \hat{\mathbf{x}}^k$.

Observe that for $\mathbf{x}, \mathbf{y} \in C_o^{1\text{-var}}([0, T], \mathbb{R}^d)$, the flow map $U_{T \rightarrow 0}^{\mathbf{x}}$ satisfies the algebraic property

$$U_{T \rightarrow 0}^{\mathbf{x} * \mathbf{y}} = U_{T \rightarrow 0}^{\mathbf{y}} \circ U_{T \rightarrow 0}^{\mathbf{x}}. \quad (3.4.2)$$

Since the image of $\Omega_1(\mathbb{R}^d)$ under the lift $S_{[p]}$ is dense in $WG\Omega_p(\mathbb{R}^d)$, it follows that for all $\mathbf{x} \in WG\Omega_p(\mathbb{R}^d)$ and $f \in \text{Lip}_{\text{loc}}^\gamma(\mathbb{R}^e, \mathbf{L}(\mathbb{R}^d, \mathbb{R}^e))$ which satisfies the p -non-explosion condition, there is a well-defined bijective flow map $U_{T \leftarrow 0}^{\mathbf{x}} : \mathbb{R}^e \mapsto \mathbb{R}^e$. Moreover, the identities (3.4.1) and (3.4.2) remain true for all $\mathbf{x}, \mathbf{y} \in WG\Omega_p(\mathbb{R}^d)$.

Furthermore, by continuity of the Itô map $\pi_{(f)}$ (Theorem 3.3.5), $U_{T \leftarrow 0}^{\mathbf{x}}$ is a homeomorphism. In fact, the uniform continuity of $\pi_{(f)}$ from Theorem 3.3.5 and Proposition 3.3.7 implies that $U_{T \leftarrow 0}^{\mathbf{x}} : \mathbf{x} \mapsto U_{T \rightarrow 0}^{\mathbf{x}}$ is a continuous map from $WG\Omega_p(\mathbb{R}^d)$ to $\text{Homeo}(\mathbb{R}^e)$ when $\text{Homeo}(\mathbb{R}^e)$ is equipped with an appropriate topology. We summarise these properties in the following theorem.

Theorem 3.4.1. *Let $\gamma > p \geq 1$ and $f \in \text{Lip}_{\text{loc}}^\gamma(\mathbb{R}^e, \mathbf{L}(\mathbb{R}^d, \mathbb{R}^e))$ satisfy the p -non-explosion condition.*

Then the associated flow map $U_{T \leftarrow 0}^{\mathbf{x}} : WG\Omega_p(\mathbb{R}^d) \mapsto \text{Homeo}(\mathbb{R}^e)$ is continuous when $\text{Homeo}(\mathbb{R}^e)$ is equipped with the compact-open topology.

If moreover $f \in \text{Lip}^\gamma(\mathbb{R}^e, \mathbf{L}(\mathbb{R}^d, \mathbb{R}^e))$, then $U_{T \leftarrow 0}^{\mathbf{x}}$ is continuous when $\text{Homeo}(\mathbb{R}^e)$ is equipped with the uniform topology.

Remark 3.4.2. By requiring additional regularity on the vector fields f , one can moreover show that $U_{T \leftarrow 0}^{\mathbf{x}}$ is a continuous map from $WG\Omega_p(\mathbb{R}^d)$ to $\text{Diff}^k(\mathbb{R}^e)$, the group of C^k -diffeomorphisms of $\mathbb{R}^e$ (where k depends on the additional regularity of f and $\text{Diff}^k(\mathbb{R}^e)$ is equipped with the C^k -topology, see [25] Section 11.2 and Corollary 11.14). For simplicity, we only consider here the group $\text{Homeo}(\mathbb{R}^e)$.

The following recent result of Boedihardjo et al. [4] establishes a fundamental link between the signature of $\mathbf{x} \in WG\Omega_p(\mathbb{R}^d)$ and the solution map $\pi_{(f)}(0, y; \mathbf{x})$. This result generalises the analogous theorem of Hambly and Lyons [30] for paths

of bounded variation, which in turn is a generalisation of the work of Chen [13] on characterisation of paths through their signatures.

The following theorem is a consequence of [4] Theorem 3 (this formulation was pointed out to the author by H. Boedihardjo in a personal communication).

Theorem 3.4.3 (Boedihardjo et al. [4]). *Let $p \geq 1$ and $\mathbf{x}, \mathbf{y} \in W\Omega_p(\mathbb{R}^d)$. Then $S(\mathbf{x})_{0,T} = S(\mathbf{y})_{0,T}$ if and only if for all $\gamma > p$, integers $e \geq 1$, $y \in \mathbb{R}^e$, and $f \in \text{Lip}_{\text{loc}}^\gamma(\mathbb{R}^e, \mathbf{L}(\mathbb{R}^d, \mathbb{R}^e))$ which satisfies the p -non-explosion condition, it holds that $\pi_{(f)}(0, y; \mathbf{x})_T = \pi_{(f)}(0, y; \mathbf{y})_T$.*

Recall that $WR_p(\mathbb{R}^d)$ denotes the set of group-like elements which arise as the signatures of weakly geometric p -rough paths. Observe that $S(\mathbf{x} * \mathbf{y})_{0,T} = S(\mathbf{x})_{0,T} S(\mathbf{y})_{0,T}$ and $S(\mathbf{x})_{0,T}^{-1} = S(\overleftarrow{\mathbf{x}})_{0,T}$ for all $\mathbf{x}, \mathbf{y} \in W\Omega_p(\mathbb{R}^d)$.

Moreover, observe that the “only if” direction of Theorem 3.4.3 implies that the flow map $U_{T \leftarrow 0} : W\Omega_p(\mathbb{R}^d) \mapsto \text{Homeo}(\mathbb{R}^e)$ factors through the signature. It follows that for all $f \in \text{Lip}_{\text{loc}}^\gamma(\mathbb{R}^e, \mathbf{L}(\mathbb{R}^d, \mathbb{R}^e))$ satisfying the p -non-explosion condition, there exists a well-defined “signature flow map” $U^\cdot : WR_p(\mathbb{R}^d) \mapsto \text{Homeo}(\mathbb{R}^e)$ such that $U^x = U_{T \leftarrow 0}^\mathbf{x}$ whenever $x = S(\mathbf{x})_{0,T}$.

We now obtain the following result.

Corollary 3.4.4. *Let $\gamma > p \geq 1$ and $f \in \text{Lip}^\gamma(\mathbb{R}^e, \mathbf{L}(\mathbb{R}^d, \mathbb{R}^e))$ satisfy the p -non-explosion condition. Then the flow map $U^\cdot : WR_p(\mathbb{R}^d) \mapsto \text{Homeo}(\mathbb{R}^e)$ is group homomorphism.*

3.4.1 Flows induced by linear vector fields

In this section, we shall provide a concrete description of the flow map $U^\cdot$ for linear vector fields $M \in \mathbf{L}(\mathbb{R}^e, \mathbf{L}(\mathbb{R}^d, \mathbb{R}^e))$.

Recall that every linear map $M \in \mathbf{L}(\mathbb{R}^e, \mathbf{L}(\mathbb{R}^d, \mathbb{R}^e))$ may equivalently be viewed as a map $M \in \mathbf{L}(\mathbb{R}^d, \mathbf{L}(\mathbb{R}^e))$, where $\mathbf{L}(\mathbb{R}^e)$ is the space of linear operators on $\mathbb{R}^e$.

Observe that every linear vector field $x \in \mathbf{L}(\mathbb{R}^e)$ induces a linear vector field $L_x \in \mathbf{L}(\mathbf{L}(\mathbb{R}^e))$ by composition $L_x : z \mapsto x \circ z$.

Consider now a collection of linear vector fields $M \in \mathbf{L}(\mathbb{R}^d, \mathbf{L}(\mathbb{R}^e))$. It follows that M induces a collection of linear vector fields $L_M \in \mathbf{L}(\mathbb{R}^d, \mathbf{L}(\mathbf{L}(\mathbb{R}^e)))$ given by $L_M(x) = L_{M(x)}$.

For $\mathbf{x} \in W\Omega_p(\mathbb{R}^d)$, we can therefore consider the unique solution $\mathbf{z}_t$ to the RDE

$$d\mathbf{z}_t = L_M(\mathbf{z}_t)d\mathbf{x}_t, \quad \mathbf{z}_0 = Id_{\mathbb{R}^e} \in \mathbf{L}(\mathbb{R}^e). \quad (3.4.3)$$

The solution to (3.4.3) will play a central role in connection with the signature of $\mathbf{x}$. For all $M \in \mathbf{L}(\mathbb{R}^d, \mathbf{L}(\mathbb{R}^e))$, define the map

$$M : WG\Omega_p(\mathbb{R}^d) \mapsto \mathbf{L}(\mathbb{R}^e), \quad M(\mathbf{x}) = \mathbf{z}_T,$$

where $\mathbf{z}$ is the solution of (3.4.3).

Defining $\mathbf{y}_t := \mathbf{z}_t(y)$ for some $y \in \mathbb{R}^e$, one readily sees that $\mathbf{y}$ is the solution to

$$d\mathbf{y}_t = M(\mathbf{y}_t)d\mathbf{x}_t, \quad \mathbf{y}_0 = y \in \mathbb{R}^e.$$

In particular, it follows that $M(\mathbf{x})$ is precisely the flow map $U_{T \leftarrow 0}^{\mathbf{x}} : \mathbb{R}^e \mapsto \mathbb{R}^e$ associated with the vector fields $M \in \mathbf{L}(\mathbb{R}^e, \mathbf{L}(\mathbb{R}^d, \mathbb{R}^e))$.

As a special case of the map $U_{T \leftarrow 0}^{\mathbf{x}} : WG\Omega_p(\mathbb{R}^d) \mapsto \text{Homeo}(\mathbb{R}^e)$, we remark that $M(\overleftarrow{\mathbf{x}}) = M(\mathbf{x})^{-1}$ and $M(\mathbf{x} * \mathbf{y}) = M(\mathbf{y}) \circ M(\mathbf{x}) = M(\mathbf{x})M(\mathbf{y})$, where we recall that $\mathbf{L}(\mathbb{R}^e)$ is equipped with the product $AB := B \circ A$. Furthermore, by continuity of the Itô map $\pi_{(L_M)}$ (with respect to the driving signal $\mathbf{x}$), we have that M is a continuous map from $WG\Omega_p(\mathbb{R}^d)$ into $\mathbf{L}(\mathbb{R}^e)$.

The following proposition establishes the connection between $M(\mathbf{x})$, the signature $S(\mathbf{x})_{0,T} \in G(\mathbb{R}^d)$ of $\mathbf{x}$, and the algebra homomorphism $M \in \text{Hom}(E(\mathbb{R}^d), \mathbf{L}(\mathbb{R}^e))$ which was studied in Chapter 2.

Proposition 3.4.5. *Let $\mathbf{x} \in WG\Omega_p(\mathbb{R}^d)$ and $M \in \mathbf{L}(\mathbb{R}^d, \mathbf{L}(\mathbb{R}^e))$. Then $M(\mathbf{x}) = M(S(\mathbf{x})_{0,T})$.*

Proof. Consider first $\mathbf{x} \in \Omega_1(\mathbb{R}^d)$. Then for all $k \geq 0$, the k -th term of the signature $S(\mathbf{x})_{0,t}^k$ (up to time $t \in [0, T]$) is given by the k -th iterated integral

$$S(\mathbf{x})_{0,t}^k = \int_{0 < t_1 < \dots < t_k < t} d\mathbf{x}_{t_1} \otimes \dots \otimes d\mathbf{x}_{t_k} \in (\mathbb{R}^d)^{\otimes k}.$$

Defining $\mathbf{z}_t := M(S(\mathbf{x})_{0,t})$ (as an absolutely convergent series), it is straightforward to verify that $\mathbf{z}_t$ is indeed the solution to (3.4.3) (cf. [46] Section 2.1).

For general $\mathbf{x} \in WG\Omega_p(\mathbb{R}^d)$, let $\mathbf{x}_n \in \Omega_1(\mathbb{R}^d)$ satisfy (3.3.2) so that the corresponding solutions $\mathbf{z}_n$ to $\mathbf{z}$ in the uniform metric d_∞ . In particular, we have $\lim_{n \rightarrow \infty} M(\mathbf{x}_n) = M(\mathbf{x})$.

Moreover (3.3.2) implies that $\lim_{n \rightarrow \infty} d_{p' \cdot \text{var}; [0, T]}(S_{[p]}(\mathbf{x}_n), \mathbf{x}) = 0$ for all $p' > p$, so by continuity of the signature map $\mathcal{I}_{p'} : \Omega_{p'}(\mathbb{R}^d) \mapsto E(\mathbb{R}^d)$ (Corollary 3.1.10),

$$\lim_{n \rightarrow \infty} S(\mathbf{x}_n)_{0,T} = S(\mathbf{x})_{0,T}.$$

As $M : E(\mathbb{R}^d) \mapsto \mathbf{L}(\mathbb{R}^e)$ is continuous, it follows that

$$\lim_{n \rightarrow \infty} M(S(\mathbf{x}_n)_{0,T}) = M(S(\mathbf{x})_{0,T}).$$

However we have just seen that $M(\mathbf{x}_n) = M(S(\mathbf{x}_n)_{0,T})$, so that indeed $M(\mathbf{x}) = M(S(\mathbf{x})_{0,T})$. $\square$

Recall from Theorem 2.3.8 that the collection of algebra homomorphisms $M \in \text{Hom}(E(\mathbb{R}^d), \mathbf{L}(H))$ arising from all linear maps $M \in \mathbf{L}(\mathbb{R}^d, \mathbf{u}(H))$, where H ranges over all finite dimensional Hilbert spaces, separates the points of $E(\mathbb{R}^d)$. Moreover, recall that M sends $U(\mathbb{R}^d)$, the group of unitary elements of $E(\mathbb{R}^d)$ (which contains $G(\mathbb{R}^d)$), into $\mathcal{U}(H)$, the group of unitary operators on H .

The following is now a consequence of Proposition 3.4.5 and Theorem 2.3.8.

Corollary 3.4.6. *Let $\mathbf{x}, \mathbf{y} \in W\Omega_p(\mathbb{R}^d)$. Then $S(\mathbf{x})_{0,T} = S(\mathbf{y})_{0,T}$ if and only if $M(\mathbf{x}) = M(\mathbf{y})$ for all $M \in \mathbf{L}(\mathbb{R}^d, \mathbf{u}(H))$, where H ranges over all finite dimensional Hilbert spaces. Moreover, every such M is a continuous map from $W\Omega_p(\mathbb{R}^d)$ into $\mathcal{U}(H)$.*

We now demonstrate an application of the characteristic functions introduced in Chapter 2 and the result of Boedihardjo et al. [4] (Theorem 3.4.3).

Theorem 3.4.7. *Let $\gamma > p \geq 1$ and $f \in \text{Lip}_{\text{loc}}^\gamma(\mathbb{R}^e, \mathbf{L}(\mathbb{R}^d, \mathbb{R}^e))$ satisfy the p -non-explosion condition (resp. $f \in \text{Lip}^\gamma(\mathbb{R}^e, \mathbf{L}(\mathbb{R}^d, \mathbb{R}^e))$). Let $U_{T \leftarrow 0} : W\Omega_p(\mathbb{R}^d) \mapsto \text{Homeo}(\mathbb{R}^d)$ be the corresponding flow map.*

Let $\mathbf{X}$ and $\mathbf{Y}$ be $G\Omega_p(\mathbb{R}^d)$ -valued random variables such that $\mathbb{E}[M(\mathbf{X})] = \mathbb{E}[M(\mathbf{Y})]$ for all $M \in \mathbf{L}(\mathbb{R}^d, \mathbf{u})$.

Then $U_{T \leftarrow 0}^{\mathbf{X}} \stackrel{\mathcal{D}}{=} U_{T \leftarrow 0}^{\mathbf{Y}}$ as $\text{Homeo}(\mathbb{R}^d)$ -valued random variables, where $\text{Homeo}(\mathbb{R}^d)$ is equipped with the compact-open (resp. uniform) topology.

The two hard elements of the proof have already been established, namely that the signature $S(\mathbf{x})_{0,T}$ completely determines the homeomorphism $U_{T \leftarrow 0}^{\mathbf{x}}$ (Theorem 3.4.3, Boedihardjo et al. [4]), and that the collection of maps $M(\mathbf{X})$ completely determines the signature (Corollary 3.4.6). One simply needs to address several measure-theoretic questions to establish the desired probabilistic statement, which is done in the proof below.

Proof of Theorem 3.4.7. Denote by μ and ν the probability measures associated to $\mathbf{X}$ and $\mathbf{Y}$ on $G\Omega_p(\mathbb{R}^d)$ respectively. Recall the continuous signature map $\mathcal{I}_p : G\Omega_p(\mathbb{R}^d) \mapsto G(\mathbb{R}^d)$ (Corollary 3.1.10) and the set $R_p(\mathbb{R}^d) := \mathcal{I}(G\Omega_p(\mathbb{R}^d))$

For the remainder of the proof, equip $R_p(\mathbb{R}^d)$ with the finest topology for which $\mathcal{I}_p : G\Omega_p \mapsto R_p(\mathbb{R}^d)$ is continuous. It follows in particular that the induced flow map $U : R_p(\mathbb{R}^d) \mapsto \text{Homeo}(\mathbb{R}^e)$ of Corollary 3.4.4 is continuous (under either of the stated assumptions on the vector fields f).

Let us write $\mathcal{I}_p^G : G\Omega_p(\mathbb{R}^d) \mapsto G(\mathbb{R}^d)$ and $\mathcal{I}_p^R : G\Omega_p(\mathbb{R}^d) \mapsto R_p(\mathbb{R}^d)$ (note that $\mathcal{I}_p^G$ and $\mathcal{I}_p^R$ are equal as functions, but we use the superscript to distinguish the two different topologies of the image space).

Let $\mu_R := \mu \circ (\mathcal{I}_p^R)^{-1}$ and $\nu_R := \nu \circ (\mathcal{I}_p^R)^{-1}$ be the pushforward Borel probability measures on $R_p(\mathbb{R}^d)$. It holds that the measure associated to $U_{T \leftarrow 0}^{\mathbf{X}}$ and $U_{T \leftarrow 0}^{\mathbf{Y}}$ is the pushforward of μ_R and ν_R respectively by the continuous map U . It thus suffices to show that $\mu_R = \nu_R$.

Since $G\Omega_p(\mathbb{R}^d)$ is a Polish space, it holds that μ is a Radon measure (in the sense of [5] Definition 7.1.1). Since μ_R is the pushforward of μ by the continuous map $\mathcal{I}_p^R : G\Omega_p(\mathbb{R}^d) \mapsto R_p(\mathbb{R}^d)$, it readily follows that μ_R is also Radon. Similarly, ν_R is Radon.

Let F be the set of all functions $h : R_p(\mathbb{R}^d) \mapsto \mathbb{R}$ such that there exists $h_0 \in C_b(G(\mathbb{R}^d), \mathbb{R})$ for which $h_0|_{R_p(\mathbb{R}^d)} = h$. Observe that F is a subalgebra of $C_b(R_p(\mathbb{R}^d), \mathbb{R})$ which separates the points.

Let $\mu_G := \mu \circ (\mathcal{I}_p^G)^{-1}$ and $\nu_G := \nu \circ (\mathcal{I}_p^G)^{-1}$ be the pushforward Borel probability measures on $G(\mathbb{R}^d)$. Observe that $\mathbb{E}[M(\mathbf{X})] = \mathbb{E}[M(\mathbf{Y})]$ for all $M \in \mathbf{L}(\mathbb{R}^d, \mathbf{u})$ is equivalent to the statement that $\mu_G(M) = \nu_G(M)$ for all $M \in \mathcal{A}$. It follows by Corollary 2.3.12 that $\mu_G = \nu_G$.

However, for every $h \in F$ and $h_0 \in C_b(G(\mathbb{R}^d), \mathbb{R})$ such that $h_0|_{R_p(\mathbb{R}^d)} = h$, it holds that $h \circ \mathcal{I}_p^R \equiv h_0 \circ \mathcal{I}_p^G$, and thus

$$\mu_R(h) = \mu(h \circ \mathcal{I}_p^R) = \mu(h_0 \circ \mathcal{I}_p^G) = \mu_G(h_0).$$

Similarly $\nu_R(h) = \nu_G(h_0)$. Since $\mu_G = \nu_G$, it follows that $\mu_R(h) = \nu_R(h)$ for all $h \in F$.

Since μ_R and ν_R are Radon measures on $R_p(\mathbb{R}^d)$, and F is a separating subalgebra of $C_b(R_p(\mathbb{R}^d), \mathbb{R})$, it follows by the Stone-Weierstrass theorem that $\mu_R = \nu_R$ ([5] Exercise 7.14.79). $\square$

We can now state the following consequence on weak convergence of flows associated with RDEs. Recall that $\text{Homeo}(\mathbb{R}^e)$ equipped with the compact-open topology is a Polish space ([17] p.51).

Corollary 3.4.8. *Let $\gamma > p \geq 1$ and $f \in \text{Lip}_{\text{loc}}^\gamma(\mathbb{R}^e, \mathbf{L}(\mathbb{R}^d, \mathbb{R}^e))$ satisfy the p -non-explosion condition. Let $U_{T \leftarrow 0} : WG\Omega_p(\mathbb{R}^d) \mapsto \text{Homeo}(\mathbb{R}^e)$ be the associated flow map. Equip $\text{Homeo}(\mathbb{R}^e)$ with the compact-open topology.*

Let $(\mathbf{X}_n)_{n \geq 1}$ be $WG\Omega_p(\mathbb{R}^d)$ -valued random variables such that $(\|\mathbf{X}_n\|_{p\text{-var};[0,T]})_{n \geq 1}$ is a tight collection of real random variables.

(1) *Suppose that $\lim_{n \rightarrow \infty} \mathbb{E}[M(\mathbf{X}_n)]$ exists for all $M \in \mathbf{L}(\mathbb{R}^d, \mathbf{u}(H))$. Then the $\text{Homeo}(\mathbb{R}^e)$ -valued random variables $U_{T \leftarrow 0}^{\mathbf{X}_n}$ converge in law.*

(2) *Suppose moreover that for some $1 \leq q < \gamma$, there exists a $WG\Omega_q(\mathbb{R}^d)$ -valued random variable $\mathbf{X}$ such that $\lim_{n \rightarrow \infty} \mathbb{E}[M(\mathbf{X}_n)] = \mathbb{E}[M(\mathbf{X})]$ for all $M \in \mathbf{L}(\mathbb{R}^d, \mathbf{u}(H))$. Then $U_{T \leftarrow 0}^{\mathbf{X}_n} \xrightarrow{\mathcal{D}} U_{T \leftarrow 0}^{\mathbf{X}}$.*

Proof. Let $p < p' < \gamma \wedge (\lfloor p \rfloor + 1)$. Denote $\tilde{\mathbf{X}}_n = P_{p'}(\mathbf{X}_n)$. By Lemma 3.2.5, $(\tilde{\mathbf{X}}_n)_{n \geq 1}$ is a tight collection of $G\Omega_{p'}(\mathbb{R}^d)$ -valued random variables.

By invariance of $U_{T \leftarrow 0}^{\mathbf{X}}$ under reparametrisation of $\mathbf{x} \in G\Omega_{p'}(\mathbb{R}^d)$, it holds that $U_{T \leftarrow 0}^{\mathbf{X}_n} \stackrel{\mathcal{D}}{=} U_{T \leftarrow 0}^{\tilde{\mathbf{X}}_n}$ for all $n \geq 1$. Moreover, since $U_{T \leftarrow 0} : G\Omega_{p'}(\mathbb{R}^d) \mapsto \text{Homeo}(\mathbb{R}^e)$ is continuous, it follows that $(U_{T \leftarrow 0}^{\mathbf{X}_n})_{n \geq 1}$ is a tight collection of $\text{Homeo}(\mathbb{R}^e)$ -valued random variables.

Furthermore, one readily sees that every weak limit point of $(U_{T \leftarrow 0}^{\mathbf{X}_n})_{n \geq 1}$ is the pushforward by $U_{T \leftarrow 0}$ of some $G\Omega_{p'}(\mathbb{R}^d)$ -valued random variable $\tilde{\mathbf{X}}$ which is a weak limit point of $(\tilde{\mathbf{X}}_n)_{n \geq 1}$.

(1) For two weak limit points $\tilde{\mathbf{X}}, \tilde{\mathbf{Y}}$ of $(\tilde{\mathbf{X}}_n)_{n \geq 1}$, it holds that

$$\mathbb{E}[M(\tilde{\mathbf{X}})] = \mathbb{E}[M(\tilde{\mathbf{Y}})]$$

for all $M \in \mathbf{L}(\mathbb{R}^d, \mathbf{u}(H))$. It follows by Theorem 3.4.7 that $U_{T \leftarrow 0}^{\tilde{\mathbf{X}}} \stackrel{\mathcal{D}}{=} U_{T \leftarrow 0}^{\tilde{\mathbf{Y}}}$, and thus all weak limit points of $(U_{T \leftarrow 0}^{\mathbf{X}_n})_{n \geq 1}$ are equal in law.

(2) It now holds that, for every weak limit point $\tilde{\mathbf{X}}$ of $(\tilde{\mathbf{X}}_n)_{n \geq 1}$,

$$\mathbb{E}[M(\tilde{\mathbf{X}})] = \mathbb{E}[M(\mathbf{X})]$$

for all $M \in \mathbf{L}(\mathbb{R}^d, \mathbf{u}(H))$. Treating $\tilde{\mathbf{X}}$ and $\mathbf{X}$ as $G\Omega_{q'}(\mathbb{R}^d)$ -valued random variables for some $p \vee q < q' < \gamma$ (replacing them by their lifts to $G^{\lfloor q' \rfloor}(\mathbb{R}^d)$ -valued random variables if either $p < \lfloor q' \rfloor$ or $q < \lfloor q' \rfloor$), it follows again by Theorem 3.4.7 that $U_{T \leftarrow 0}^{\tilde{\mathbf{X}}} \stackrel{\mathcal{D}}{=} U_{T \leftarrow 0}^{\mathbf{X}}$, so indeed $U_{T \leftarrow 0}^{\mathbf{X}_n} \xrightarrow{\mathcal{D}} U_{T \leftarrow 0}^{\mathbf{X}}$. $\square$

3.5 Euler approximations

We end the chapter with a result concerning the Euler approximation of an RDE. Fix a basis $e_1, \dots, e_d$ of $\mathbb{R}^d$ and for $x \in T^N(\mathbb{R}^d)$, let $x^{k;i_1, \dots, i_k}$ denote the coefficient of $e_1 \otimes \dots \otimes e_{i_k}$ in x under the basis $\{e_{i_1} \otimes \dots \otimes e_{i_k} \mid k \in \{0, \dots, N\}, i_1, \dots, i_k \in \{1, \dots, d\}\}$ of $T^N(\mathbb{R}^d)$.

For $y \in \mathbb{R}^e$, $x = (1, x^1, \dots, x^N) \in G^N(\mathbb{R}^d)$, and $N-1$ times continuously differentiable vector fields $f = (f_1, \dots, f_d)$, $f_i \in C^{N-1}(\mathbb{R}^e, \mathbb{R}^e)$, define the Euler approximation

$$\mathcal{E}_{(f)}(y, x) := \sum_{k=1}^N \sum_{i_1, \dots, i_k \in \{1, \dots, d\}} x^{k;i_1, \dots, i_k} (f_{i_1} \dots f_{i_k} Id_{\mathbb{R}^e})(y),$$

where we treat f_i as a first order differential operator.

We now recall the Euler approximation of the solution of an RDE.

Theorem 3.5.1 ([25] Corollary 10.15). *Let $\gamma > p \geq 1$. Then there exists a constant $C = C(p, \gamma) > 0$ such that for all $f \in \text{Lip}^\gamma(\mathbb{R}^e, \mathbf{L}(\mathbb{R}^d, \mathbb{R}^e))$, $y \in \mathbb{R}^e$, and $\mathbf{x} \in WG\Omega_p(\mathbb{R}^d)$*

$$\|\pi_{(f)}(s, y; \mathbf{x})_{s,t} - \mathcal{E}_{(f)}(y, S_{\lfloor \gamma \rfloor}(\mathbf{x})_{s,t})\| \leq C \left(\|f\|_{\text{Lip}^{\gamma-1}} \|\mathbf{x}\|_{p\text{-var};[s,t]} \right)^\gamma.$$

For linear vector fields $M_1, \dots, M_k \in \mathbf{L}(\mathbb{R}^e)$, note that

$$(M_1 \dots M_k Id_{\mathbb{R}^e})(y) = M_k \circ \dots \circ M_1(y),$$

where we treat M_i as a differential operator on the left side (cf. [25] Exercise 10.50). In particular, for the vector fields $L_{M_i} \in \mathbf{L}(\mathbf{L}(\mathbb{R}^e))$, $L_{M_i} : z \mapsto M_i \circ z$, it holds that

$$(L_{M_1} \dots L_{M_k} Id_{\mathbf{L}(\mathbb{R}^e)})(z) = L_{M_k} \circ \dots \circ L_{M_1}(z) = M_k \circ \dots \circ M_1 \circ z.$$

In particular, for vector fields $M = (M_1, \dots, M_d) \in \mathbf{L}(\mathbb{R}^d, \mathbf{L}(\mathbb{R}^e))$, setting $z = Id_{\mathbb{R}^e}$ and recalling the product $AB = B \circ A$ on $\mathbf{L}(\mathbb{R}^e)$, we have

$$\mathcal{E}_{(L_M)}(Id_{\mathbb{R}^e}, x) = \sum_{k=1}^N \sum_{i_1, \dots, i_k \in \{1, \dots, d\}} x^{k;i_1, \dots, i_k} M_{i_1} \dots M_{i_k}.$$

Extending $M \in \mathbf{L}(\mathbb{R}^d, \mathbf{L}(\mathbb{R}^e))$ canonically to a linear map $M \in \mathbf{L}(T^N(\mathbb{R}^d), \mathbf{L}(\mathbb{R}^e))$ by $M(e_{i_1} \otimes \dots \otimes e_{i_k}) = M(e_{i_1}) \dots M(e_{i_k})$ (but which we note is not an algebra homomorphism), this relation is written simply as

$$\mathcal{E}_{(L_M)}(Id_{\mathbb{R}^e}, x) = M(x) - Id_{\mathbb{R}^e}. \quad (3.5.1)$$

Lemma 3.5.2. *Let $M \in \mathbf{L}(\mathbb{R}^d, \mathbf{L}(\mathbb{R}^e))$ and $1 \leq p < \gamma$. Then there exist $r, C > 0$ such that for all $\mathbf{x} \in W\Omega_p(\mathbb{R}^d)$ with $\|\mathbf{x}\|_{p\text{-var};[0,T]} < r$, we have*

$$\|M(\mathbf{x}) - M(S_{\lfloor \gamma \rfloor}(\mathbf{x})_{0,T})\| \leq C \|\mathbf{x}\|_{p\text{-var};[0,T]}^\gamma.$$

Proof. Fix a bounded neighbourhood $W \subset \mathbf{L}(\mathbb{R}^e)$ of $Id_{\mathbb{R}^e}$ and consider vector fields $f \in \text{Lip}^\gamma(\mathbf{L}(\mathbb{R}^e), \mathbf{L}(\mathbb{R}^d, \mathbf{L}(\mathbb{R}^e)))$ such that $f \equiv L_M$ on W .

By Theorem 3.5.1, there exists $C_2 = C_2(p, \gamma) > 0$ such that for all $\mathbf{x} \in W\Omega_p(\mathbb{R}^d)$

$$\|(\pi_{(f)}(0, Id_{\mathbb{R}^e}; \mathbf{x})_T - Id_{\mathbb{R}^e}) - \mathcal{E}_{(f)}(Id_{\mathbb{R}^e}, S_{\lfloor \gamma \rfloor}(\mathbf{x})_{0,T})\| \leq C_2 \left(\|f\|_{\text{Lip}^{\gamma-1}} \|\mathbf{x}\|_{p\text{-var};[0,T]} \right)^\gamma.$$

We can now choose $r > 0$ sufficiently small such that $\|\mathbf{x}\|_{p\text{-var};[0,T]} < r$ implies

$$\pi_{(f)}(0, Id_{\mathbb{R}^e}; \mathbf{x})_T = \pi_{(L_M)}(0, Id_{\mathbb{R}^e}; \mathbf{x})_T$$

and

$$\mathcal{E}_{(f)}(Id_{\mathbb{R}^e}, S_{\lfloor \gamma \rfloor}(\mathbf{x})_{0,T}) = \mathcal{E}_{(L_M)}(Id_{\mathbb{R}^e}, S_{\lfloor \gamma \rfloor}(\mathbf{x})_{0,T}).$$

The desired result now follows from (3.5.1). □

Chapter 4

Expected signature

Our main focus in this chapter is the expected signature of $G(\mathbb{R}^d)$ -valued random variables.

In Section 4.1, we study the analogue of the moment problem, that is, conditions under which the law of a $G(\mathbb{R}^d)$ -valued random variable is uniquely determined by its expected signature. Our primary tool is the characteristic function defined at the end of Section 2.3. Proposition 4.1.1 provides a general criterion under which the law of a $G(\mathbb{R}^d)$ -valued random variable is uniquely determined by its expected signature. In turn, Theorem 4.1.3 provides a method to verify this criterion without explicit knowledge of the expected signature itself. We demonstrate applications of these results to Lévy processes studied in [24] and to families of Gaussian and Markovian rough paths studied in [10] and [11].

In Section 4.2 we study analyticity properties of the characteristic function. The main result is Theorem 4.2.5 (and its Corollaries 4.2.9 and 4.2.10), which provides a criterion to establish analyticity of the characteristic function and solve the moment problem within a restricted family of $G(\mathbb{R}^d)$ -valued random variables. We demonstrate an application to Markovian rough paths stopped upon exiting a domain.

In Section 4.3 we demonstrate a method of moments for weak convergence of $G(\mathbb{R}^d)$ -valued random variables (Theorem 4.3.5).

4.1 The moment problem

When V is a normed space, recall from Corollary 2.2.3 that if X is a G -valued random variable such that $\text{ESig}(X)$ exists and has an infinite radius of convergence, then $\mathbb{E}[X]$ exists as an element of E and is equal to $\text{ESig}(X)$. Thus $\mathbb{E}[f(X)]$ is completely determined by $\text{ESig}(X)$ for all $f \in E'$, and in particular for all $M \in \mathcal{A}$.

The following is now a consequence of the uniqueness of probability measures from Corollary 2.3.12.

Proposition 4.1.1. *Let X and Y be $G(\mathbb{R}^d)$ -valued random variables such that*

$$\text{ESig}(X) = \text{ESig}(Y).$$

Suppose furthermore that $\text{ESig}(X) \in E$, i.e., $\text{ESig}(X)$ has an infinite radius of convergence. Then $X \stackrel{\mathcal{D}}{=} Y$.

Recall from Corollary 3.1.10 that the signature map $\mathcal{I}_p : (\Omega_p, \rho_{p\text{-var}}) \mapsto E$ is continuous. It follows that the signature $S(\mathbf{X})_{0,T}$ of any Ω_p -valued (resp. $G\Omega_p$ -valued) random variable $\mathbf{X}$ is a well-defined (Borel) E -valued (resp. U -valued, or $G(\mathbb{R}^d)$ -valued in case $V = \mathbb{R}^d$) random variable.

Example 4.1.2. We apply Proposition 4.1.1 to the Lévy–Khintchine formula established in [24]. Recall that every Lévy process in $\mathbb{R}^d$ admits a lift to a $G\Omega_p(\mathbb{R}^d)$ -valued random variable $\mathbf{X}$ for any $p > 2$ by adding appropriate adjustments for jumps (see [59] Section 2, and [24] Section 9.1, Corollary 44; cf. Sections 5.3 and 5.4 in the following chapter). Concretely, the lift is first constructed as a càdlàg process in $G^2(\mathbb{R}^d)$ as the solution of a stochastic differential equation (in the sense of Marcus) driven by the Lévy process, and then mapped to an element of $G\Omega_p(\mathbb{R}^d)$ by linear traversals at the jumps.

Let (a, b, K) denote the triplet of a Lévy process in $\mathbb{R}^d$, $\mathbf{X}$ the associated $G\Omega_p(\mathbb{R}^d)$ -valued random variable, and $X = S(\mathbf{X})_{0,T}$ its signature. It follows from [24] Theorem 43 that $\text{ESig}(X)$ exists (as an element of $P(\mathbb{R}^d) = \prod_{k \geq 0} (\mathbb{R}^d)^{\otimes k}$) whenever the Lévy measure K has finite moments of all orders. Furthermore, $\text{ESig}(X) \in E$ exactly when

$$\int_{\mathbb{R}^d} (e^{\lambda \|y\|} - 1 - \lambda \mathbf{1}\{\|y\| \leq 1\} \|y\|) K(dy) < \infty \quad \text{for all } \lambda > 0. \quad (4.1.1)$$

It follows by Proposition 4.1.1 that whenever (4.1.1) is satisfied, the expected signature of $S(\mathbf{X})_{0,T}$ uniquely determines the law of $S(\mathbf{X})_{0,T}$ as a $G(\mathbb{R}^d)$ -valued random variable.

Recall the radius of convergence $r_1(X)$ from Definition 2.2.1. Theorem 4.1.3 below provides sufficient conditions to ensure that $r_1(X) > 0$ or $r_1(X) = \infty$ without explicit knowledge of $\text{ESig}(X)$.

For a subset $B \subseteq A$ of an algebra A and $n \geq 1$, define $B^n = \{x_1 \dots x_n \mid x_1, \dots, x_n \in B\}$. For an element $x \in A$, define $B(x) = \inf\{n \geq 1 \mid x \in B^n\}$ (taking $B(x) = \infty$ if $x \notin B^n$ for all $n \geq 1$).

Note that for a topological algebra A with (jointly) continuous multiplication, an A -valued random variable X , and a (Borel) measurable set $B \subset A$, $B(X)$ is a well-defined random variable in $\{1, 2, \dots\} \cup \{\infty\}$.

Theorem 4.1.3. *Let V be a normed space and X an E -valued random variable. Suppose there exists a bounded, measurable set $B \subset E$ such that $B(X)$ has an exponential tail, i.e., $\mathbb{E}[e^{\lambda B(X)}] < \infty$ for some $\lambda > 0$. Then $r_1(X) > 0$. If moreover $\mathbb{E}[e^{\lambda B(X)}] < \infty$ for all $\lambda > 0$, then $r_1(X) = \infty$.*

Proof. Equip E with the projective extension of the norm on V . For any $r > 0$ and $\lambda > 0$ such that $\sup_{x \in B} \|\delta_r(x)\| < e^\lambda$, it holds that

$$\sum_{k \geq 0} r^k \mathbb{E}[\|X^k\|] = \mathbb{E}[\|\delta_r(X)\|] \leq \mathbb{E}[e^{\lambda B(X)}], \quad (4.1.2)$$

where the inequality follows from the fact that $\delta_r(X) = \delta_r(X_1) \dots \delta_r(X_{B(X)})$ for some $X_1, \dots, X_{B(X)} \in B$.

Suppose first that $\mathbb{E}[e^{\lambda B(X)}] < \infty$ for all $\lambda > 0$. Then since B is bounded, for all $r > 0$ there exists $\lambda > 0$ sufficiently large such that $\sup_{x \in B} \|\delta_r(x)\| < e^\lambda$. Then (4.1.2) implies that $r_1(X) \geq r$. As r was arbitrary, it follows that $r_1(X) = \infty$.

Suppose now that $\mathbb{E}[e^{\lambda B(X)}] < \infty$ for some $\lambda > 0$. By Proposition 2.1.11, the functions δ_r converge strongly (that is, uniformly on bounded sets) to δ_0 as $r \rightarrow 0$ and, in particular, uniformly on B . Thus there exists $r > 0$ such that $\sup_{x \in B} \|\delta_r(x)\| < e^\lambda$. Then (4.1.2) implies that $r_1(X) \geq r > 0$ as desired. $\square$

We demonstrate how to apply Theorem 4.1.3 to random variables arising from signatures of geometric rough paths whose expected signature is not directly known.

Let V be a Banach space and $p \geq 1$. We note that for any $\mathbf{x} \in \Omega_p$, $\omega_{\mathbf{x}}(s, t) := \|\mathbf{x}\|_{p\text{-var};[s,t]}^p$ defines a control for which (3.1.1) is satisfied. Thus for all $k \geq 0$, the lift $S(\mathbf{x}) : \Delta_{[0,T]} \mapsto E$ satisfies

$$\|S(\mathbf{x})_{0,T}^k\| \leq \frac{\omega(0, T)^{k/p}}{\beta_p(k/p)!}.$$

We hence define

$$K_p = \left\{ x \in E \mid \sup_{k \geq 0} \beta_p(k/p)! \|x^k\| \leq 1 \right\}$$

and observe that $S(\mathbf{x})_{0,T} \in K_p$ for every $\mathbf{x} \in \Omega_p$ with $|||\mathbf{x}|||_{p\text{-var};[0,T]} \leq 1$. Observe furthermore that K_p is bounded (due to the factorial term $(k/p)!$) and measurable in E .

For $\mathbf{x} \in \Omega_p$, define $k_p(\mathbf{x}) = K_p(S(\mathbf{x})_{0,T})$, i.e., the minimum positive integer k for which there exist $x_1, \dots, x_k \in K_p$ such that $S(\mathbf{x})_{0,T} = x_1 \dots x_k$.

We briefly recall the construction of the greedy sequence and function $N_{\kappa,[0,T],p}(\mathbf{x})$ introduced by Cass, Litterer and Lyons [10]. For $\kappa > 0$ define the sequence of times $\tau_0 = 0$,

$$\tau_{j+1} = \inf\{t > \tau_j \mid \omega_{\mathbf{x}}(\tau_j, t) \geq \kappa\} \wedge T,$$

so that $\omega_{\mathbf{x}}(\tau_j, \tau_{j+1}) = \kappa$ for all $0 \leq j < N = N_{\kappa,[0,T],p}(\mathbf{x}) := \sup\{j \geq 0 \mid \tau_j < T\}$ and $\omega_{\mathbf{x}}(\tau_N, \tau_{N+1}) \leq \kappa$ (see [10] Definition 4.7, [23] p.158). Note that $k_p(\mathbf{x}) \leq N_{1,[0,T],p}(\mathbf{x}) + 1$.

Remark 4.1.4. For any $p, q \geq 1$ and $\mathbf{x} \in \Omega_q$, note that the signature $S(\mathbf{x})_{0,T}$ exists and so $k_p(\mathbf{x})$ is meaningfully defined. Moreover, in case $q \leq p$, $\mathbf{x}$ can canonically be viewed as an element of Ω_p by its lift $S_{[p]}\mathbf{x} \in \Omega_p$, and we have $S(\mathbf{x})_{0,T} = S(S_{[p]}\mathbf{x})_{0,T}$.

However, if $q < [p]$ and $N_{1,[0,T],p}(\mathbf{x})$ and the greedy sequence $(\tau_j)_{j=1}^{\infty}$ are defined in terms of $\mathbf{x}$ (not its lift $S_{[p]}\mathbf{x}$), then $N_{1,[0,T],p}(\mathbf{x})$ does not yield a deterministic bound on $k_p(\mathbf{x})$ since the individual signatures $S(\mathbf{x})_{\tau_j, \tau_{j+1}}$ will in general fail to be elements of K_p .

To obtain a bound on $k_p(\mathbf{x})$, one needs to consider $N_{1,[0,T],p}(S_{[p]}\mathbf{x})$ and $(\tau_j)_{j=1}^{\infty}$ defined in terms of $S_{[p]}\mathbf{x} \in \Omega_p$. Then $S(\mathbf{x})_{\tau_j, \tau_{j+1}} = S(S_{[p]}\mathbf{x})_{\tau_j, \tau_{j+1}} \in K_p$ for all $j = 0, 1, \dots$, and so $k_p(\mathbf{x}) \leq N_{1,[0,T],p}(S_{[p]}\mathbf{x}) + 1$.

This point shall be revisited in Remark 4.2.8 and Remark 4.2.17 of Section 4.2.1.

Let $\mathcal{K}_p(V)$ be the family of Ω_p -valued random variables $\mathbf{X}$ such that $\mathbb{E}[e^{\lambda k_p(\mathbf{X})}] < \infty$ for all $\lambda > 0$. The following is a consequence of Theorem 4.1.3.

Corollary 4.1.5. *Let V be a Banach space, $p \geq 1$ and $\mathbf{X} \in \mathcal{K}_p(V)$. Then the expected signature $\text{ESig}[S(\mathbf{X})_{0,T}]$ has an infinite radius of convergence.*

Applying Proposition 4.1.1, we obtain the following result which allows us to answer the moment problem for signatures arising from a wide range of random geometric rough paths.

Corollary 4.1.6. *Let $p \geq 1$ and $\mathbf{X}$ a $G(\mathbb{R}^d)$ -valued random variable in $\mathcal{K}_p(\mathbb{R}^d)$. Then the law of its signature $S(\mathbf{X})_{0,T}$ is the unique probability measure on $G(\mathbb{R}^d)$ whose expected signature is $\text{ESig}[S(\mathbf{X})_{0,T}]$.*

We now demonstrate two important examples of $G\Omega_p(\mathbb{R}^d)$ -valued random variables in $\mathcal{K}_p(\mathbb{R}^d)$. Remark that a non-negative random variable Z satisfies $\mathbb{E}[e^{\lambda Z}] < \infty$ for all $\lambda > 0$ whenever Z^θ has a Gaussian tail for some $\theta > 1/2$, i.e., $\mathbb{P}[Z^\theta > z] \leq C^{-1}e^{-Cz^2}$ for all $z > 0$ and a constant $C > 0$. In both of the following sections $[0, T]$ is a fixed time interval.

4.1.1 Gaussian rough paths

We recall first the definition of Gaussian processes and their lifts to geometric rough paths. A Gaussian process $X = (X^1, \dots, X^d)$ in $\mathbb{R}^d$ defined on $[0, T]$ is a stochastic process for which $(X_{t_1}^{i_1}, \dots, X_{t_k}^{i_k})$ is a multivariate Gaussian random variable for all finite collection of times $t_1, \dots, t_k \in [0, T]$ and indexes $i_1, \dots, i_k \in \{1, \dots, d\}$. The law of X is completely determined by the couple (μ, R) , where $\mu : [0, T] \mapsto \mathbb{R}^d$ is the mean function

$$\mu(t)_i := \mathbb{E}[X_t^i],$$

and $R : [0, T]^2 \mapsto (\mathbb{R}^d)^2$ is the covariance function

$$R \binom{s}{t}_{i,j} := \mathbb{E}[(X_s^i - \mu(s)_i)(X_t^j - \mu(t)_j)].$$

The process X has independent components precisely when $R_{i,j} \equiv 0$ for all $i \neq j$. Moreover X is said to be centred if $\mu \equiv 0$.

For a function $R : [0, T]^2 \mapsto \mathbb{R}^d$, denote

$$R \binom{s, t}{u, v} := R \binom{s}{u} + R \binom{t}{v} - R \binom{s}{v} - R \binom{t}{u}.$$

For $\rho \geq 1$, the 2D ρ -variation of $R : [0, T]^2 \mapsto \mathbb{R}^d$ is defined by

$$\|R\|_{\rho\text{-var}; [0, T]^2} = \sup_{\mathcal{D}, \mathcal{D}' \subset [0, T]} \left(\sum_{t_i \in \mathcal{D}, t'_i \in \mathcal{D}'} \left| R \binom{t_i, t_{i+1}}{t'_i, t'_{i+1}} \right|^\rho \right)^{1/\rho},$$

where, as usual, the supremum is taken over all finite partitions $\mathcal{D}, \mathcal{D}'$ of $[0, T]$.

We recall that every centred continuous Gaussian process in $\mathbb{R}^d$ with independent components and covariance function of finite 2D ρ -variation for some $\rho \in [1, 2)$ admits a natural lift to a $G\Omega_p(\mathbb{R}^d)$ -valued random variable $\mathbf{X}$ for any $p > 2\rho$ ([25] Theorem 15.33). The lift may be realised concretely as the almost sure limit in p -variation topology of the piecewise linear interpolation (or appropriate mollifications) of the Gaussian process.

We remark moreover that the covariance function of Brownian motion has finite 2D 1-variation (which corresponds to its lift as a $G\Omega_p(\mathbb{R}^d)$ -valued random variable for any $p > 2$). Hence the range $\rho \in [1, 2)$ allows one to deal with Gaussian processes of far worse regularity than Brownian motion.

We recall now several results from [10] on estimates of the function $N_{1,[0,T],p}$ (see [10] Corollary 5.5, Theorem 6.3, and [23] Theorem 11.13).

Theorem 4.1.7 ([10]). *Suppose that X is a centred Gaussian process in $\mathbb{R}^d$ with independent components whose covariance function has finite 2D ρ -variation for some $\rho \in [1, 3/2)$. Let $\mathbf{X}$ be its natural lift to a $G\Omega_p(\mathbb{R}^d)$ -valued random variable for some $2\rho < p < 3$. Then $N_{1,[0,T],p}(\mathbf{X})^{1/\rho}$ has a Gaussian tail.*

Note that the covariance function of fractional Brownian motion B^H in $\mathbb{R}^d$ with Hurst parameter $H \in (0, 1)$ and independent components is

$$R\left(\begin{smallmatrix} s \\ t \end{smallmatrix}\right)_{i,j} = \delta_{i,j} \frac{1}{2} (t^{2H} + s^{2H} + |t - s|^{2H}).$$

It readily follows that R has finite 2D ρ -variation for any $\rho \geq 1/(2H)$. Thus for all $H > 1/4$, B^H admits a natural lift to a $G\Omega_p(\mathbb{R}^d)$ -valued random variable $\mathbf{X}$ for any $p > H^{-1}$. Observe that Theorem 4.1.7 applies directly to $\mathbf{X}$ only in the case $H > 1/3$. However, one is able to exploit directly the structure of the Cameron-Martin space of B^H to obtain the following sharper result.

Theorem 4.1.8 ([10]). *Let B^H be a fractional Brownian motion in $\mathbb{R}^d$ with independent components and $H > 1/4$. Let $\mathbf{X}$ be its natural lift to a $G\Omega_p(\mathbb{R}^d)$ -valued random variable for some $p > H^{-1}$. Then $N_{1,[0,T],p}(\mathbf{X})^{1/2+1/p}$ has a Gaussian tail.*

Since $k_p(\mathbf{X}) \leq N_{1,[0,T],p}(\mathbf{X}) + 1$ as remarked before, it follows that $\mathbf{X}$ is a $G\Omega_p(\mathbb{R}^d)$ -valued random variable in $\mathcal{K}_p(\mathbb{R}^d)$ in both of the above cases. Applying Corollary 4.1.6, we obtain the following solution to the moment problem for signatures arising from Gaussian rough paths.

Corollary 4.1.9. *Let $\mathbf{X}$ be the $G\Omega_p(\mathbb{R}^d)$ -valued random variable in Theorem 4.1.7 or Theorem 4.1.8. Then the expected signature of $S(\mathbf{X})_{0,T}$ uniquely determines the law of $S(\mathbf{X})_{0,T}$ as a $G(\mathbb{R}^d)$ -valued random variable.*

4.1.2 Markovian rough paths

We recall first the construction of Markovian rough paths in the sense of Friz and Victoir [22], [25]. One of the main motivations behind the construction is to extend the applicability of rough differential equations to Markov diffusions well outside the scope of semi-martingales (see Remark 4.1.12). We review the construction in some detail as we shall further study Markovian rough paths in Section 4.2.1. All the material in this section, unless otherwise stated, is taken from [25] Chapter 16.

For $n \geq 1$, denote by $\mathfrak{g}^n := \mathfrak{g}^n(\mathbb{R}^d)$ the Lie algebra of $G^n := G^n(\mathbb{R}^d)$, which we realise as the smallest Lie subalgebra of $T^n(\mathbb{R}^d)$ containing $\mathbb{R}^d$.

For $\Lambda \geq 1$, denote by $\Xi^{n,d}(\Lambda)$ the space of all measurable maps $a : G^n \mapsto \mathbf{L}(\mathbb{R}^d)$ such that $a(x)$ is symmetric and

$$\Lambda^{-1}|\xi|^2 \leq \langle \xi, a(x)\xi \rangle \leq \Lambda|\xi|^2 \quad (4.1.3)$$

for all $x \in G^n$ and $\xi \in \mathbb{R}^d$.

Let $u_1, \dots, u_d$ denote a basis for $\mathbb{R}^d$ viewed as a subspace of $\mathfrak{g}^n$. By definition, it holds that $u_1, \dots, u_d$ generate the Lie algebra $\mathfrak{g}^n$, and we identify each u_i with the corresponding left-invariant vector field on G^n .

Let $C_c^\infty(G^n)$ denote the space of smooth compactly supported functions $f : G^n \mapsto \mathbb{R}$. For $a \in \Xi^{n,d}(\Lambda)$, define the carré-du-champ operator $\Gamma^a : C_c^\infty(G^n) \times C_c^\infty(G^n) \mapsto C_c^\infty(G^n)$

$$\Gamma^a(f, g) := \sum_{i,j=1}^d a_{i,j} u_i f u_j g.$$

Note that by symmetry of a , Γ^a is a symmetric bilinear map.

Denote by m the Haar measure on G^n (which we recall is unique up to a positive multiplying constant, is both left- and right-invariant, and coincides with the push-forward of the Lebesgue measure on $\mathfrak{g}$ by the exponential map $\exp : \mathfrak{g}^n \mapsto G^n$, [25] Proposition 16.40). Define the corresponding symmetric non-negative definite bilinear form $\mathcal{E}^a : C_c^\infty(G^n) \times C_c^\infty(G^n) \mapsto \mathbb{R}$

$$\mathcal{E}^a(f, g) := \int_{G^n} \Gamma^a(f, g) dm.$$

One can check that for a sequence f_n in $C_c^\infty(G^n)$ for which $f_n \rightarrow 0$ in $L^2(m)$ and $\mathcal{E}^a(f_n - f_m, f_n - f_m) \rightarrow 0$ as $n, m \rightarrow \infty$, it holds that $\mathcal{E}^a(f_n, f_n) \rightarrow 0$. It follows that $\mathcal{E}^a$ is closable with $C_c^\infty(G^n)$ as a core and domain $\mathcal{D}(\mathcal{E}^a)$ defined as the closure of $C_c^\infty(G^n)$ in $L^2(m)$ under the norm

$$f \mapsto \mathcal{E}^a(f, f) + \langle f, f \rangle_{L^2(m)}.$$

Moreover $\mathcal{E}^a$ satisfies the Markovian property, that is, for all $f \in \mathcal{D}(\mathcal{E}^a)$, it holds that $g := (0 \vee f) \wedge 1$ is in $\mathcal{D}(\mathcal{E}^a)$ and $\mathcal{E}^a(g, g) \leq \mathcal{E}^a(f, f)$. Hence $\mathcal{E}^a$ defines a regular symmetric Dirichlet form on $L^2(m)$. Furthermore, since the u_i are first order differential operators, $\mathcal{E}^a$ is strongly local ($\mathcal{E}^a(f, g) = 0$ whenever f is constant on a neighbourhood of the support of g).

Following the above construction, one may now define a Markov process $\mathbf{X}^a$ on G^n which is symmetric with respect to m from the general fact that every regular symmetric Dirichlet form gives rise to a symmetric Markov process ([26] Chapter 7 Theorem 7.2.1). Moreover, since $\mathcal{E}^a$ is strongly local, the process $\mathbf{X}^a$ is a diffusion.

However to study properties of $\mathbf{X}^a$, particularly sample path regularity crucial to rough path theory, we shall study (and may equivalently define) $\mathbf{X}^a$ using the heat kernel associated to $\mathcal{E}^a$, whose main properties we now recall.

Recall first that for a non-positive self-adjoint operator $\mathcal{L}$ defined on a dense subspace $\mathcal{D}(\mathcal{L})$ of a Hilbert space H , using spectral analysis one may define the non-negative operator $\sqrt{-\mathcal{L}}$ on $\mathcal{D}(\sqrt{-\mathcal{L}}) \supset \mathcal{D}(\mathcal{L})$ for which the associated (non-negative definite) symmetric bilinear form $\mathcal{E}(f, g) := \langle \sqrt{-\mathcal{L}}f, \sqrt{-\mathcal{L}}g \rangle$ is closed. Conversely, every closed, densely defined, symmetric, non-negative definite bilinear form $\mathcal{E}$ on H arises in such a way from a non-positive self-adjoint operator $\mathcal{L}$. Furthermore, again by spectral analysis, one may define a strongly continuous contractive semi-group of self-adjoint operators $\mathcal{P}_t = e^{-t\mathcal{L}} : H \mapsto H$.

It follows that associated to the Dirichlet form $\mathcal{E}^a$, there is a non-positive self-adjoint operator $\mathcal{L}^a$ defined on a dense subspace $\mathcal{D}(\mathcal{L}^a) \subset \mathcal{D}(\mathcal{E}^a)$ in $L^2(m)$, and a semi-group $\mathcal{P}_t^a : L^2(m) \mapsto L^2(m)$.

Observe that the uniform sub-ellipticity of a in (4.1.3) implies that $\mathcal{E}^a$ is quasi-isometric to $\mathcal{E}^I$ in the sense that

$$\Lambda^{-1}\mathcal{E}^I(f, f) \leq \mathcal{E}^a(f, f) \leq \Lambda\mathcal{E}^I(f, f),$$

where $\mathcal{E}^I$ is the Dirichlet form associated with $\Gamma^I(f, g) := \sum_{i,j=1}^d u_i f u_j g$. Quasi-isometry between $\mathcal{E}^I$ and $\mathcal{E}^a$ allows us to carry over many properties of $\mathcal{P}_t^I$ to $\mathcal{P}_t^a$, in particular an ultracontractivity property $\mathcal{P}_t^a : L^1 \mapsto L^\infty$ (with operator norm $\|\mathcal{P}_t^a\|_{L^1 \mapsto L^\infty}$ decreasing in t) from which one can establish the existence of the heat kernel $p_t^a : G^n \times G^n \mapsto [0, \infty)$ for which

$$(\mathcal{P}_t^a f)(x) = \int_{G^n} f(y) p_t(x, y) m(dy)$$

for all $f \in L^2(m)$. We remark for each fixed $x \in G^n$, $(t, y) \mapsto p_t^a(x, y)$ is a weak solution of $\partial_t p = \mathcal{L}^a p_t(x, \cdot)$.

Finally, due to the Chapman-Kolmogorov equations satisfied by p_t^a , one may realise $\mathbf{X}^a$ as the Markov process whose finite dimensional distributions are completely determined by p_t^a ([25] Proposition E.15).

The quasi-isometry between $\mathcal{E}^a$ and $\mathcal{E}^I$ further allows us to compare the intrinsic distance, and doubling and Poincaré constants associated to $\mathcal{E}^a$ to those of $\mathcal{E}^I$, which in turn leads to upper (and lower) bounds on $p_t^a(x, y)$ (as well as Hölder regularity properties by the de Giorgi–Moser–Nash theorem). The upper bounds on $p_t^a(x, y)$ then imply that the sample paths of $\mathbf{X}^a$ have almost surely finite p -variation for any $p > 2$ ([25] Corollary 16.12).

Remark 4.1.10. The process $\mathbf{X}^I$ associated to $\mathcal{E}^I$ is readily seen to be the natural lift of Brownian motion in $\mathbb{R}^d$ given by its Stratonovich iterated integrals. The regularity properties of $\mathbf{X}^a$ may thus be interpreted as a statement about propagation of sample path regularity of Markov diffusions through quasi-isometry.

Remark 4.1.11. Due to the sample path regularity of $\mathbf{X}^a$, we can associate to every realisation $\mathbf{X}^a : [0, T] \mapsto G^n$ an element of $G\Omega_p(\mathbb{R}^d)$ which we denote by the same letter $\mathbf{X}^a$, as discussed in Section 3.1 following Definition 3.1.2. Note however that the law of $\mathbf{X}^a$ depends in general on the starting point $x := \mathbf{X}_0^a \in G^n$, so we shall use the notation $\mathbf{X}^{a,x}$ whenever we wish to emphasise the dependence on x .

Remark 4.1.12. Taking $n = 1$ above and $a \in \Xi^{1,d}(\Lambda)$, one obtains a Markov diffusion in $G^1(\mathbb{R}^d) = \mathbb{R}^d$ of a.s. finite p -variation for $p > 2$. We then define its natural lift to $G^2(\mathbb{R}^d)$ (to obtain a $G\Omega_p(\mathbb{R}^d)$ -valued random variable) as the process $\mathbf{X}^{a \circ \rho^1}$, where we recall $\rho^1 : T^2(\mathbb{R}^d) \mapsto \mathbb{R}^d$ is the natural projection onto the first level $\mathbb{R}^d$, and remark that $a \circ \rho^1 \in \Xi^{2,d}(\Lambda)$.

We note that given $a \in \Xi^{1,d}(\Lambda)$ sufficiently regular, for example in C^2 , one may realise $\mathbf{X}^a$ as the solution of an (Itô) stochastic differential equation, and is thus a semi-martingale (see [25] Section 16.1). However in the case that a is only measurable, $\mathbf{X}^a$ will not in general be a semi-martingale. Lifting the process $\mathbf{X}^a$ to a rough path thus allows us to define rough differential equations for Markov processes well outside the range of semi-martingales.

Furthermore, the general case $n \geq 2$ corresponds to diffusions in $\mathbb{R}^d$ of finite p -variation, $p > 2$, whose lifts to $G\Omega_p$ -valued random variables are given by the construction above and which are Markovian with respect to their first n iterated integrals.

We now recall a recent result of Cass and Ogrodnik [11].

Theorem 4.1.13 ([11] Theorem 5.3). *Let $\Lambda \geq 1$ and $a \in \Xi^{n,d}(\Lambda)$. Let $p > 2$ and $\mathbf{X}$ the $G\Omega_p(\mathbb{R}^d)$ -valued random variable associated to the Markov process $\mathbf{X}^{a,x}$. Then $N_{1,[0,T],p}(\mathbf{X})^{1-1/p}$ has a Gaussian tail.*

It follows that $\mathbf{X}$ is a $G\Omega_p(\mathbb{R}^d)$ -valued random variable in $\mathcal{K}_p(\mathbb{R}^d)$ for every $p > 2$. Applying Corollary 4.1.6, we obtain the following solution to the moment problem for signatures arising from Markovian rough paths.

Corollary 4.1.14. *Let $\mathbf{X}$ be the $G\Omega_p(\mathbb{R}^d)$ -valued random variable in Theorem 4.1.13. Then the expected signature of $S(\mathbf{X})_{0,T}$ uniquely determines the law of $S(\mathbf{X})_{0,T}$ as a $G(\mathbb{R}^d)$ -valued random variable.*

4.2 Analyticity

In this section we investigate conditions under which the characteristic function is analytic. We apply these results to situations where the expected signature does not necessarily have an infinite radius of convergence.

Definition 4.2.1. Let X be an E -valued random variable, H a finite dimensional Hilbert space, and $M \in \mathbf{L}(V, \mathbf{L}(H))$. For $\lambda \in \mathbb{C}$, define $\phi_{X,M}(\lambda) = \mathbb{E}[(\lambda M)(X)]$ whenever $\|(\lambda M)(X)\|$ is integrable.

The above definition of $\phi_{X,M}$ does not introduce any new concept to the previously defined ϕ_X and simply makes the results in this section easier to state.

Recall that for a real random variable X , if $\mathbb{E}[|e^{\lambda X}|] < \infty$ for all $\lambda \in (-\varepsilon, \varepsilon)$, then $\phi_X(\lambda) = \mathbb{E}[e^{i\lambda X}]$ is well-defined and analytic on the strip $|\operatorname{Im}(z)| < \varepsilon$. This property is known as the propagation of regularity (and similar results hold for C^{2k} regularity of ϕ_X on $\mathbb{R}$, see, e.g., [41]).

We start by showing that the analogue of this property is not in general true for G -valued random variables whenever $\dim(V) \geq 2$. The propagation of regularity for real (or equivalently $G(\mathbb{R})$ -valued) random variables relies crucially on commutativity in $E(\mathbb{R})$, and we show how the lack of commutativity prevents the same phenomenon from occurring when $\dim(V) \geq 2$. Recall the radius of convergence $r_1(X)$ from Definition 2.2.1.

Example 4.2.2. Let V be a normed space with $\dim(V) \geq 2$. We construct a G -valued random variable X such that

- (1) $r_1(X) > 0$, thus in particular, $\operatorname{ESig}(X)$ exists and $\phi_{X,M}$ is analytic in a neighbourhood of zero for all $M \in \mathcal{A}$,

- (2) there exists $M \in \mathbf{L}(V, \mathfrak{u}(\mathbb{C}^2))$ such that the set of $\lambda \in \mathbb{C}$ for which $|\lambda| > 1$ and $\mathbb{E} [||(\lambda M)(X)||] = \infty$ forms a dense subset of $\{z \in \mathbb{C} \mid |z| > 1\}$, and
- (3) $\phi_{X,M}$ is nowhere differentiable on $(1, \infty)$.

Let e_1, e_2 be fixed linearly independent vectors in V of unit length, s a non-negative real random variable, and N a random variable in $\mathbb{N} = \{0, 1, 2, \dots\}$ independent of s . Define the G -valued random variable $X = \exp(sf_N)$, where $f_N = [e_1, [\dots, [e_1, e_2] \dots]]$ with e_1 appearing N times.

Suppose there exists $r_0 > 0$ such that $\mathbb{E}[e^{rs}] < \infty$ for all $0 \leq r < r_0$. We claim this implies (1). Indeed, remark that $||f_n|| \leq 2^n$, and thus $||\delta_r X|| \leq \exp(2^N r^N s)$. Denote $p_n = \mathbb{P}[N = n]$. It follows that for $r > 0$ sufficiently small

$$\mathbb{E} [||\delta_r X||] \leq \sum_{n \geq 0} p_n \mathbb{E} [\exp(2^n r^n s)] < \infty,$$

which implies $r_1(X) > 0$ as claimed.

Let $\mathfrak{su}(2)$ be the special unitary Lie algebra of dimension 3 with the standard basis u_1, u_2, u_3 satisfying $[u_1, u_2] = u_3$, $[u_2, u_3] = u_1$, $[u_3, u_1] = u_2$. Let $M : V \mapsto \mathfrak{su}(2)$ defined by $e_i \mapsto u_i$ for $i = 1, 2$ and arbitrary otherwise.

Suppose moreover that $\mathbb{E}[e^{rs}] = \infty$ for all $r > r_0$ and that N has unbounded support. We claim this implies (2). Indeed, let $v_n = M(f_n)$ (thus $v_0 = u_2, v_1 = u_3, v_2 = -u_2, v_3 = -u_3, v_4 = v_0, \dots$). Denote $\lambda = re^{i\theta}$ for $r, \theta \geq 0$, so that $(\lambda M)(X) = \exp(sr^N e^{iN\theta} v_N)$. We obtain

$$\begin{aligned} \mathbb{E} [||(\lambda M)(X)||] &= \sum_{n \geq 1} p_n \mathbb{E} [||\exp(sr^n e^{in\theta} v_n)||] \\ &= \sum_{n \geq 1} p_n \mathbb{E} \left[\exp \left(\frac{1}{2} |sr^n \sin(n\theta)| \right) \right], \end{aligned}$$

where the last equality follows since iv_n is Hermitian with eigenvalues $\pm \frac{1}{2}$.

Let D be any open subset of $\{z \in \mathbb{C} \mid |z| > 1\}$. We observe that there exists $n \geq 1$ sufficiently large and $r > 1, \theta > 0$, such that $p_n > 0, re^{i\theta} \in D$ and $\mathbb{E} [\exp(\frac{1}{2} |sr^n \sin(n\theta)|)] = \infty$. Thus (2) holds as claimed.

Finally, we make specific choices for s and N to obtain (3). Observe by Fubini's theorem that for all $r \geq 0$

$$\mathbb{E} [(rM)(X)] = \sum_{n \geq 1} p_n \mathbb{E} [\exp(sr^n v_n)].$$

Suppose $p_n > 0$ only if $n = 4m$ for some integer m . Since

$$\exp(tu_3) = \begin{pmatrix} e^{it/2} & 0 \\ 0 & e^{-it/2} \end{pmatrix},$$

it follows that

$$\mathbb{E}[(rM)(X)] = \sum_{n \geq 0} p_{4n} \begin{pmatrix} \exp(ir^{4n}s/2) & 0 \\ 0 & \exp(-ir^{4n}s/2) \end{pmatrix}.$$

It follows that $\phi_{X,M}$ has the same regularity at $r > 0$ as $\sum_{n \geq 0} p_{4n} \phi_s(r^{4n}/2)$, where ϕ_s is the characteristic function of s .

It is now easy to find ϕ_s and p_n such that the above series defines a nowhere differentiable function on $(1, \infty)$. For example, let $\phi_s(\lambda) = (1 - q)(1 - qe^{i\lambda})^{-1}$ for any $0 < q < 1$, i.e., s is geometrically distributed with parameter $1 - q$, and let p_n decay faster than any geometric sequence, i.e., for any $\alpha \in (0, 1)$ there exists n_α such that $p_n < \alpha^n$ for all $n \geq n_\alpha$. The statement of (3) then follows by Dini's general construction of a nowhere differentiable function ([37] p.24).

Remark 4.2.3. The random variable X constructed above is the exponential of a Lie polynomial of degree N . Thus when $V = \mathbb{R}^d$, X is the signature of a random weakly geometric N -rough path ([25] Exercise 9.17), and thus of a random geometric p -rough path for $p > N$. Moreover, as the decay of $\|X^k\|$ is exactly of the order $(k/N)!^{-1}$, there does not exist a fixed $p \geq 1$ such that X is almost surely the signature of a random geometric p -rough path.

One can however approximate each sample of X by the signature $S(\mathbf{X})_{0,T}$ of a bounded variation path $\mathbf{X}_{0,\cdot} : [0, T] \mapsto \mathbb{R}^d$ in such a way that (1) and (2) in Example 4.2.2 still hold for the $G(\mathbb{R}^d)$ -valued random variable $S(\mathbf{X})_{0,T}$ with the change that the stated λ in (2) will be dense in the annulus $\{z \in \mathbb{C} \mid 1 < |z| < R\}$ for any fixed $R > 1$ (where the random variable $\mathbf{X}$ depends on R).

Definition 4.2.4. Let V be a normed space. Denote by $\Phi(V)$ the set of G -valued random variables X which satisfy

(P1) $r_1(X) > 0$, and

(P2) $\phi_{X,M}$ is (weakly) analytic on $\mathbb{R}$ for all $M \in \mathcal{A}$.

The importance of the set Φ is that when $V = \mathbb{R}^d$ and $X, Y \in \Phi(\mathbb{R}^d)$ such that $\text{ESig}(X) = \text{ESig}(Y)$, we have $X \stackrel{\mathcal{D}}{=} Y$. To observe this, remark that for V normed and X an E -valued random variable with $r_1(X) =: \varepsilon > 0$, it follows by dominated

convergence that $\mathbb{E}[M(X)] = \sum_{k \geq 0} M^{\otimes k} \mathbb{E}[X^k]$ whenever $\|M\| < \varepsilon$. Hence for all $M \in \mathcal{A}$, $\phi_{X,M}(\lambda)$ is completely determined by $\text{ESig}(X)$ whenever $|\lambda|$ is sufficiently small, and the claim follows by Corollary 2.3.12.

Theorem 4.2.5 is the main result of this section and provides a criterion to ensure that $X \in \Phi$.

Theorem 4.2.5. *Let V be a normed space and X a U -valued random variable. Suppose there exists a bounded, measurable set $B \subset U$ such that $B(X)$ has an exponential tail, i.e., $\mathbb{E}[e^{\lambda B(X)}] < \infty$ for some $\lambda > 0$. Then (P1) and (P2) hold for X .*

Proof of Theorem 4.2.5, (P1). This follows immediately from Theorem 4.1.3. $\square$

For the proof of (P2), we require the following two lemmas.

Lemma 4.2.6. *Let F be a topological algebra, A a normed algebra, $M \in \text{Hom}(F, A)$, $B \subset F$ a bounded measurable set, and let $c = \sup_{x \in B} \|M(x)\|$. Let X be an F -valued random variable such that $\|M(X)\|$ and $(c + \varepsilon)^{B(X)}$ are integrable for some $\varepsilon > 0$.*

Then $\sup_{M' \in \mathcal{U}} \|M'(X)\|$ is an integrable random variable, where

$$\mathcal{U} = \{M' \in \text{Hom}(F, A) \mid \sup_{x \in B} \|M(x) - M'(x)\| < \varepsilon\}$$

is an open subset of $\text{Hom}(F, A)$ (under the strong topology).

Proof. By definition of the strong topology, $\sup_{x \in B} \|\cdot(x)\|$ is a semi-norm on $\mathbf{L}(F, A)$ and so $\mathcal{U}$ is indeed an open subset.

Moreover $x \mapsto \sup_{M' \in \mathcal{U}} \|M'(x)\|$ is the supremum of a family of continuous functions, thus lower semi-continuous, and thus measurable. The claim now follows by a direct application of (2.1.3). $\square$

For a bounded complex domain $D \subset \mathbb{C}$, denote by $H_b(\overline{D})$ the space of continuous functions on $\overline{D}$ which are analytic on D . Recall that $H_b(\overline{D})$ equipped with the uniform norm is a separable Banach space.

Lemma 4.2.7. *Let V be a normed space, A a separable Banach algebra and $M \in \mathbf{L}(V, A)$. Let X be an E -valued random variable. Assume that for a bounded domain $D \subset \mathbb{C}$, $\sup_{\lambda \in \overline{D}} \|(\lambda M)(X)\|$ is an integrable random variable. Then for every $f \in A'$, the map $\lambda \mapsto \mathbb{E}[\langle f, (\lambda M)(X) \rangle]$ is in $H_b(\overline{D})$.*

Proof. Let $f \in A'$. For $x \in E$ consider the map $\phi_{M,f}(x) : \lambda \mapsto \langle f, (\lambda M)(x) \rangle$, which is an entire function on $\mathbb{C}$.

We claim that the corresponding linear map $\phi_{M,f} : E \mapsto H(\mathbb{C})$, where $H(\mathbb{C})$ is the space of entire functions on $\mathbb{C}$, is bounded when we equip $H(\mathbb{C})$ with the compact-open topology. Indeed, since $\lambda \mapsto (\lambda M)$ is a continuous map from $\mathbb{C}$ into $\text{Hom}(E, A)$ by Proposition 2.1.11, the collection of maps $(\lambda M)_{\lambda \in K}$ is strongly bounded in $\text{Hom}(E, A)$ for any bounded set $K \subset \mathbb{C}$. Thus for every bounded set $L \subset E$, it holds that

$$\sup_{x \in L} \sup_{\lambda \in K} \|(\lambda M)(x)\| < \infty.$$

In particular, this implies that $\phi_{M,f}(L)$ is a bounded subset of $H(\mathbb{C})$ for every bounded set $L \subset E$ as claimed.

Since E is a Fréchet space (hence bornological), it follows moreover that $\phi_{M,f} : E \mapsto H(\mathbb{C})$ is continuous. Hence $\phi_{M,f}(X)|_{\overline{D}}$ is a norm-integrable $H_b(\overline{D})$ -valued random variable and thus possesses a barycenter $h \in H_b(\overline{D})$.

Let $\lambda \in \overline{D}$. Since the evaluation map $\langle \cdot, \lambda \rangle : x \mapsto x(\lambda)$ is in the continuous dual of $H_b(\overline{D})$, it follows that

$$h(\lambda) = \mathbb{E}[\langle \phi_{M,f}(X), \lambda \rangle] = \mathbb{E}[\langle f, (\lambda M)(X) \rangle] = \langle f, \mathbb{E}[(\lambda M)(X)] \rangle,$$

where the last equality follows since $(\lambda M)(X)$ is a norm-integrable A -valued random variable and is thus weakly integrable by the separability of A . As h is in $H_b(\overline{D})$, the conclusion follows. $\square$

Proof of Theorem 4.2.5, (P2). Let $M \in \mathcal{A}$. Since $\|M(g)\| = 1$ for all $g \in U$, one obtains from Proposition 2.1.11 and Lemma 4.2.6 that there exists a domain D containing $1 \in \mathbb{C}$ such that $\sup_{\lambda \in \overline{D}} \|(\lambda M)(X)\|$ is an integrable random variable. The conclusion now follows by applying Lemma 4.2.7. $\square$

In light of the discussion following Theorem 4.1.3, we define

$$N_p := K_p \cap U = \left\{ g \in U \mid \sup_{k \geq 0} \beta_p(k/p)! \|g^k\| \leq 1 \right\}.$$

Observe that $S(\mathbf{x})_{0,T} \in N_p$ for every $\mathbf{x} \in G\Omega_p$ with $\|\mathbf{x}\|_{p\text{-var};[0,T]} \leq 1$. As with K_p , N_p is bounded and measurable in E .

For $\mathbf{x} \in G\Omega_p$, define $n_p(\mathbf{x}) = N_p(S(\mathbf{x})_{0,T})$, i.e., the minimum positive integer n for which there exist $g_1, \dots, g_n \in N_p$ such that $S(\mathbf{x})_{0,T} = g_1 \dots g_n$. Recall the functions k_p and $N_{1,[0,T],p}$ from Section 4.1 and note that $k_p(\mathbf{x}) \leq n_p(\mathbf{x}) \leq N_{1,[0,T],p}(\mathbf{x}) + 1$.

Remark 4.2.8. As in Remark 4.1.4, we mention again that for $1 \leq q \leq p$, every $\mathbf{x} \in G\Omega_q$ is canonically defined as an element of $G\Omega_p$ via its lift $S_{[p]}\mathbf{x} \in G\Omega_p$. However one cannot bound $n_p(\mathbf{x})$ in terms $N_{1,[0,T],p}(\mathbf{x})$ computed directly in terms of $\mathbf{x}$; instead one has $n_p(\mathbf{x}) \leq N_{1,[0,T],p}(S_{[p]}\mathbf{x}) + 1$.

Let $\mathcal{N}_p(V)$ be the family of $G\Omega_p$ -valued random variables $\mathbf{X}$ such that $n_p(\mathbf{X})$ has an exponential tail. Note that if $\mathbf{X} \in \mathcal{K}_p$ and is $G\Omega_p$ -valued, then $\mathbf{X} \in \mathcal{N}_p$. The following is a consequence of Theorem 4.2.5.

Corollary 4.2.9. *Let V be a Banach space. Then for all $p \geq 1$ and $\mathbf{X} \in \mathcal{N}_p$, the signature $S(\mathbf{X})_{0,T}$ is a U -valued random variable satisfying (P1) and (P2).*

In the finite dimensional setting, we obtain a result analogous to Corollary 4.1.6 but with weaker assumptions and a weaker conclusion.

Corollary 4.2.10. *Let $p \geq 1$ and $\mathbf{X} \in \mathcal{N}_p(\mathbb{R}^d)$. Then $S(\mathbf{X})_{0,T} \in \Phi(\mathbb{R}^d)$. In particular, the expected signature of $S(\mathbf{X})_{0,T}$ uniquely determines the law of $S(\mathbf{X})_{0,T}$ amongst all $G(\mathbb{R}^d)$ -valued random variables in $\Phi(\mathbb{R}^d)$.*

Remark 4.2.11. For a random variable $X \in \Phi(\mathbb{R}^d)$, we cannot exclude the possibility that there exists a $G(\mathbb{R}^d)$ -valued random variable Y (which might arise as the signature of a geometric rough path) such that $Y \notin \Phi(\mathbb{R}^d)$ and $\text{ESig}(X) = \text{ESig}(Y)$. Whether this is possible currently remains unknown.

However, note that Corollaries 4.2.9 and 4.2.10 apply for all $p \geq 1$. Thus for any $p, q \geq 1$ and random geometric rough paths $\mathbf{X} \in \mathcal{N}_p(\mathbb{R}^d)$ and $\mathbf{Y} \in G\Omega_q(\mathbb{R}^d)$, if $S(\mathbf{X})_{0,T}$ and $S(\mathbf{Y})_{0,T}$ are not equal in law (as $G(\mathbb{R}^d)$ -valued random variables) and $\text{ESig}(S(\mathbf{X})_{0,T}) = \text{ESig}(S(\mathbf{Y})_{0,T})$, then the lift $S_{[q']}\mathbf{Y}$ cannot be in $\mathcal{N}_{q'}(\mathbb{R}^d)$ for any $q' \geq q$.

4.2.1 Markovian rough paths stopped upon exiting a domain

We continue with the notation introduced in Section 4.1.2. Recall the result of Cass and Ogrodnik [11] that $N_{1,[0,T],p}(\mathbf{X}^{a,x})^{1-1/p}$ has a Gaussian tail for any $p > 2$ (Theorem 4.1.13).

In this section we shall replace the fixed time $T > 0$ by the first exit time of $\mathbf{X}^{a,x}$ from a suitable set. In particular, we shall show in this case that $N_{\kappa,[0,T],p}(\mathbf{X}^{a,x})$ has an exponential tail and that this result is asymptotically sharp.

Throughout this section we fix $\Lambda \geq 1$ and $n \geq 1$. We shall find it notationally convenient to identify G^n with its Lie algebra $\mathfrak{g} := \mathfrak{g}^n(\mathbb{R}^d)$ via the exponential map

$\exp : \mathfrak{g} \mapsto G^n$, which we recall is a global diffeomorphism (this identification is in fact used throughout [22], [25] and [11]). For the remainder of the section, we shall thus treat $\mathfrak{g}$ as a Lie group.

Equip $\mathfrak{g}$ with the Carnot–Carathéodory norm $\|\cdot\|$ induced by its identification with G^n (see Section 3.2) and the corresponding left-invariant metric $d(x, y) := \|x^{-1}y\|$ (we stress that $\|\cdot\|$ is not a norm on $\mathfrak{g}$ under its vector space structure and is certainly different from the projective extension norm usually considered on $T(\mathbb{R}^d)$).

For a continuous path $\mathbf{h} : [s, t] \mapsto \mathfrak{g}$, define the Sobolev norm

$$\|\mathbf{h}\|_{W^{1,2};[s,t]} := \left(\sup_{\mathcal{D} \subset [s,t]} \sum_{t_j \in \mathcal{D}} \frac{\|\mathbf{h}_{t_j, t_{j+1}}\|^2}{|t_{j+1} - t_j|} \right)^{1/2}.$$

For $x \in \mathfrak{g}$, we define the Sobolev path space $W_x^{1,2}([s, t], \mathfrak{g})$ as the set of all $\mathbf{h} : [s, t] \mapsto \mathfrak{g}$ with $\|\mathbf{h}\|_{W^{1,2};[s,t]} < \infty$ and starting point $\mathbf{h}_s = x$.

For $\alpha \in [0, 1]$, define the α -Hölder norm

$$\|\mathbf{h}\|_{\alpha\text{-Hö};[s,t]} := \sup_{s \leq u < v \leq t} \frac{\|\mathbf{h}_{u,v}\|}{|v - u|^\alpha}.$$

For $x \in \mathfrak{g}$, recall the Hölder path space $C_x^{\alpha\text{-Hö}}([s, t], \mathfrak{g})$ as the set of all $\mathbf{h} : [s, t] \mapsto \mathfrak{g}$ with $\|\mathbf{h}\|_{\alpha\text{-Hö};[s,t]} < \infty$ and starting point $\mathbf{h}_s = x$. On $C_x^{\alpha\text{-Hö}}([s, t], \mathfrak{g})$ we introduce the metric

$$d_{\alpha\text{-Hö};[s,t]}(\mathbf{x}, \mathbf{y}) := \sup_{s \leq u < v \leq t} \frac{d(\mathbf{x}_{u,v}, \mathbf{y}_{u,v})}{|v - u|^\alpha}.$$

Recall the p -variation norm $\|\cdot\|_{p\text{-var};[s,t]}$ defined by (3.2.1). Remark that for any continuous path $\mathbf{h} : [s, t] \mapsto \mathfrak{g}$ and $p \geq 1$

$$\|\mathbf{h}\|_{p\text{-var};[s,t]} \leq (t - s)^{1/p} \|\mathbf{h}\|_{1/p\text{-Hö};[s,t]},$$

and for $\alpha \in [0, 1/2]$

$$\|\mathbf{h}\|_{\alpha\text{-Hö};[s,t]} \leq (t - s)^{1/2-\alpha} \|\mathbf{h}\|_{W^{1,2};[s,t]}.$$

Similarly, recall the p -variation metric $d_{p\text{-var};[s,t]}$ defined by (3.2.2). Remark that for all $p \geq 1$ and continuous paths $\mathbf{x}, \mathbf{y} : [s, t] \mapsto \mathfrak{g}$ with $\mathbf{x}_0 = \mathbf{y}_0$,

$$d_{p\text{-var};[s,t]}(\mathbf{x}, \mathbf{y}) \leq (t - s)^{1/p} d_{1/p\text{-Hö};[s,t]}(\mathbf{x}, \mathbf{y}).$$

Our first result is a slight extension of the support theorem [25] Theorem 16.33 of $\mathbf{X}^{a,x}$ in the Hölder topology. For $\theta > 0$ define the ball

$$W_{\theta;x} := \{\mathbf{h} \in W_x^{1,2}([0, 1], \mathfrak{g}) \mid \|\mathbf{h}\|_{W^{1,2};[0,1]} < \theta\}.$$

Lemma 4.2.12. *For any $\alpha \in [0, 1/4)$, $\theta > 0$ and $c > 0$, there exists $\delta > 0$ such that*

$$\mathbb{P}^{a,x} [d_{\alpha\text{-H\"{o}l};[0,1]}(\mathbf{X}, \mathbf{h}) < c] > \delta$$

for all $a \in \Xi^{n,d}(\Lambda)$, starting points $x \in \mathfrak{g}$, and $\mathbf{h} \in W_{\theta;x}$.

The proof is essentially the same as that in [25] and we defer it to the end of the section. Recall now the greedy sequence $(\tau_j)_{j=1}^\infty$ associated with $N_{\kappa,[0,T],p}(\mathbf{X}^{a,x})$. For ease of notation, we shall not stop τ_j at T for $j > N := N_{\kappa,[0,T],p}(\mathbf{X}^{a,x})$ (i.e., we do not necessarily have $\tau_{N+1} = T$). Note this causes no confusion since $\mathbf{X}_t^{a,x}$ is defined for all times $t \geq 0$ as a diffusion on $\mathfrak{g}$.

Consider first $\mathbf{X}^{a,x} : [0, 1] \mapsto \mathfrak{g}$. Taking $\mathbf{h} \equiv x$ the trivial path, Lemma 4.2.12 implies that for any $p > 4$ and $\kappa > 0$, there exists $\delta > 0$ such that

$$\inf_{x \in \mathfrak{g}} \mathbb{P}^{a,x} \left[\|\mathbf{X}\|_{1/p\text{-H\"{o}l};[0,1]} < \kappa \right] \geq \delta.$$

It follows that

$$\inf_{x \in \mathfrak{g}} \mathbb{P}^{a,x} [\tau_1 > 1] \geq \inf_{x \in \mathfrak{g}} \mathbb{P}^{a,x} \left[\|\mathbf{X}\|_{1/p\text{-H\"{o}l};[0,1]} < \kappa \right] \geq \delta,$$

so by the (strong) Markov property of $\mathbf{X}^{a,x}$ and properties of conditional expectation

$$\mathbb{P}^{a,x} [N_{\kappa,[0,1],p}(\mathbf{X}) \geq k] = \mathbb{P}^{a,x} [\tau_k < 1] \leq (1 - \delta)^k.$$

That is, $N_{\kappa,[0,1],p}(\mathbf{X}^{a,x})$ has an exponential tail (moreover δ does not depend on $x \in \mathfrak{g}$ or $a \in \Xi^{n,d}(\Lambda)$).

While this argument yields a strictly weaker asymptotic bound than that of Cass and Ogrodnik [11] (Theorem 4.1.13), the advantage is that by choosing appropriate $\mathbf{h}$ in Lemma 4.2.12, a very similar argument gives upper and lower bounds on the tail of $N_{\kappa,[0,T],p}(\mathbf{X}^{a,x})$, where T is now the first exit time of $\mathbf{X}^{a,x}$ from a suitable open set. We first show the lower bound.

For any $r > 0$ and $x \in \mathfrak{g}$, define $B_r(x) = \{y \in \mathfrak{g} \mid d(x, y) \leq r\}$.

Proposition 4.2.13. *Let $p > 4$, $\kappa, r > 0$. Define $T = \inf\{t > 0 \mid \mathbf{X}_t^{a,x} \notin B_r(x)\}$ the first exit time of $\mathbf{X}^{a,x}$ from $B_r(x)$. Then there exists $\delta > 0$ such that*

$$\mathbb{P}^{a,x} [N_{\kappa,[0,T],p}(\mathbf{X}) \geq k] \geq \delta^k$$

for all $a \in \Xi^{n,d}(\Lambda)$ and $x \in \mathfrak{g}$.

Proof. Let $\theta > 0$ sufficiently large such for all $x \in \mathfrak{g}$ and $y \in B_{r/2}(x)$ there exists $\mathbf{h}^y \in W_{\theta;y}$ such that $\mathbf{h}_1^y = x$, $\mathbf{h}_t^y \in B_{r/2}(x)$ for all $t \in [0, 1]$, and $\|\mathbf{h}^y\|_{p\text{-var};[0,1]} > \kappa + r/2$ (for example, take $t \mapsto \mathbf{h}_t^y$ as a geodesic from y to x on $[0, 1/2]$ and then $\mathbf{h}_t^y = x\mathbf{h}_{2t-1}$ for $t \in [1/2, 1]$ for a fixed $\mathbf{h} \in W_0^{1,2}([0, 1], \mathfrak{g})$ with $\mathbf{h}_1 = 0$, $\mathbf{h}_t \in B_{r/2}(0)$ for all $t \in [0, 1]$, and $\|\mathbf{h}\|_{p\text{-var};[0,1]} > \kappa + r/2$).

Since $\|\mathbf{X}^{a,y}\|_{p\text{-var};[0,1]} \geq \|\mathbf{h}^y\|_{p\text{-var};[0,1]} - d_{1/p\text{-H\"ol};[0,1]}(\mathbf{X}^{a,y}, \mathbf{h}^y)$, we have for all $x \in \mathfrak{g}$, $y \in B_{r/2}(x)$ and $a \in \Xi^{n,d}(\Lambda)$

$$\begin{aligned} \mathbb{P}^{a,y} \left[\mathbf{X}_t \in B_r(x) \text{ for all } t \in [0, 1], \|\mathbf{X}\|_{p\text{-var};[0,1]} > \kappa, \mathbf{X}_1 \in B_{r/2}(x) \right] \\ \geq \mathbb{P}^{a,y} \left[d_{1/p\text{-H\"ol}}(\mathbf{X}, \mathbf{h}^y) < r/2 \right]. \end{aligned}$$

Applying Lemma 4.2.12 with $c = r/2$ and $\alpha = 1/p$, along with the (weak) Markov property and conditional expectation, concludes the proof. $\square$

Remark 4.2.14. Note that Proposition 4.2.13 deals only with the random variable $N_{\kappa,[0,T],p}(\mathbf{X}^{a,x})$ and does not provide a lower bound on the tail of $n_p(\mathbf{X}_{[0,T]}^{a,x})$. In particular, one cannot conclude that $\text{ESig}[S(\mathbf{X}^{a,x})_{0,T}]$ does not have an infinite radius of convergence.

We now show an upper bound on the tail of $N_{\kappa,[0,T],p}(\mathbf{X}^{a,x})$ which will imply that $S(\mathbf{X}^{a,x})_{0,T} \in \Phi(\mathbb{R}^d)$ (see however Remark 4.2.17). For a subset $D \subset \mathfrak{g}$, consider the following property:

$$\text{There exist } r, c > 0 \text{ such that } \sup_{h \in B_r(0)} \inf_{y \in D} d(xh, y) > c \text{ for all } x \in D. \quad (4.2.1)$$

Note that every subset $D \subset \mathfrak{g}$ which is bounded under the metric d satisfies (4.2.1).

Remark 4.2.15. For $1 \leq k \leq n$, let $\pi^k : \mathfrak{g}^n(\mathbb{R}^d) \mapsto \mathfrak{g}^k(\mathbb{R}^d)$ denote the canonical projection. Then whenever the image $\pi^k(D)$ satisfies (4.2.1) for some $1 \leq k \leq n$ (for the respective metric on $\mathfrak{g}^k(\mathbb{R}^d)$), then so does D (with a different choice of r, c).

Indeed, on the one hand $d(x, y) \geq d(\pi^k(x), \pi^k(y))$ for all $x, y \in \mathfrak{g}^n(\mathbb{R}^d)$. On the other hand, for every $r > 0$, there exists $R > 0$ such that $B_r(0) \subset \pi^k(B_R(0)) \subset \mathfrak{g}^k(\mathbb{R}^d)$. The conclusion readily follows since π^k is a group homomorphism.

Proposition 4.2.16. *Let $p > 4$, $\kappa > 0$, and $D \subset \mathfrak{g}$ be an open set satisfying (4.2.1) for some $r, c > 0$. Define $T = \inf\{t > 0 \mid \mathbf{X}_t^{a,x} \notin D\}$ the first exit time of $\mathbf{X}^{a,x}$ from D . Then there exists $\delta > 0$ such that*

$$\mathbb{P}^{a,x} \left[N_{\kappa,[0,T],p}(\mathbf{X}) \geq k \right] \leq (1 - \delta)^k$$

for all $a \in \Xi^{n,d}(\Lambda)$ and $x \in D$.

Proof. Let $\theta > 0$ be sufficiently large such that for every $h \in B_r(0)$ there exists $\mathbf{h} \in W_{\theta;0}$ such that $\mathbf{h}_1 = h$. Note that it suffices to prove the statement for any fixed $\kappa > 0$. In particular, we may assume that $\kappa > \theta + c$.

It follows that to every point $x \in D$, we can assign $h^x \in B_r(x)$ and $\mathbf{h}^x \in W_{\theta;x}$ such that $\inf_{y \in D} d(h^x, y) > c$ and $\mathbf{h}_1^x = h^x$. Then for all $a \in \Xi^{n,d}(\Lambda)$ and $x \in D$

$$\begin{aligned} \mathbb{P}^{a,x} [\tau_1 > 1 \geq T] &\geq \mathbb{P}^{a,x} \left[\|\mathbf{X}\|_{p\text{-var};[0,1]} < \theta + c, \mathbf{X}_1 \notin D \right] \\ &\geq \mathbb{P}^{a,x} \left[\|\mathbf{X}\|_{1/p\text{-Hö};[0,1]} < \theta + c, d(\mathbf{X}_1, h^x) < c \right] \\ &\geq \mathbb{P}^{a,x} \left[d_{1/p\text{-Hö};[0,1]}(\mathbf{X}, \mathbf{h}^x) < c \right]. \end{aligned}$$

Applying Lemma 4.2.12 with $\alpha = 1/p$, along with the (strong) Markov property and conditional expectation, concludes the proof. $\square$

Remark 4.2.17. The diffusion $\mathbf{X}^{a,x}$ is constructed on the space $\mathfrak{g}^n = \mathfrak{g}^n(\mathbb{R}^d)$ (or equivalently on $G^n(\mathbb{R}^d)$), and Proposition 4.2.16 gives an exponential bound on the tail of $N_{1,[0,T],p}(\mathbf{X}^{a,x})$ computed in terms of $\mathbf{X}^{a,x}$ for any $p > 4$. Fixing $4 < p < 5$, Corollary 4.2.10 thus implies that $S(\mathbf{X}^{a,x})_{0,T} \in \Phi(\mathbb{R}^d)$ for $n \geq 4$.

One could extend this to the case $n = 2$ or 3 (recall for $n = 1$ we consider the diffusion $\mathbf{X}^{a \circ \pi^1, x}$ on $\mathfrak{g}^2$) if the analogue of Lemma 4.2.12 were true for all $\alpha \in [0, 1/2)$. However such a support theorem is currently unknown.

Nonetheless, in light of Remarks 4.1.4 and 4.2.8, for $n = 2$ or 3 we can still show that $S(\mathbf{X}^{a,x})_{0,T} \in \Phi(\mathbb{R}^d)$ by showing that $N_{1,[0,T],p}(S_4 \mathbf{X}^{a,x})$ has an exponential tail.

To show this, note we can apply Proposition 4.2.16 to the diffusion $\mathbf{X}^{a \circ \pi^n, y}$ on $\mathfrak{g}^4$ and the open set $(\pi^n)^{-1}(D) \subset \mathfrak{g}^4$ (which indeed satisfies (4.2.1) due to Remark 4.2.15). We thus obtain that $N_{1,[0,\tilde{T}],p}(\mathbf{X}^{a \circ \pi^n, y})$ has an exponential tail, where $\tilde{T}$ is the first exit time of $\pi^n \mathbf{X}^{a \circ \pi^n, y}$ from D .

To conclude that $N_{1,[0,T],p}(S_4 \mathbf{X}^{a,x})$ has an exponential tail, it suffices to show that $\mathbf{Y}^y := y * S_4 \mathbf{X}_{0,\cdot}^{a, \pi^n y}$ is equal in law to $\mathbf{X}^{a \circ \pi^n, y}$ for all $y \in \mathfrak{g}^4$ as processes on $\mathfrak{g}^4$ ($*$ denoting group multiplication in $\mathfrak{g}^4$). This follows by a similar argument as [22] Section 6: observe that the Markov process $\mathbf{Y}_t^y$ is the solution of an RDE with starting point $y \in \mathfrak{g}^4$ and driven by $\pi^2(\mathbf{X}_t^{a, \pi^n y})$ (which is non-Markov in general) along the (unbounded) canonical left-invariant vector fields $u_1, \dots, u_d$ on $\mathfrak{g}^4$. Denoting by $\mathcal{P}_t$ the semi-group on $C_b(\mathfrak{g}^4)$ of $\mathbf{Y}_t^y$, it suffices to show that

$$\lim_{t \rightarrow 0} \langle t^{-1}(f - \mathcal{P}_t f), g \rangle_{L^2(\mathfrak{g}^4)} = \mathcal{E}^{a \circ \pi^n}(f, g)$$

for all $f, g \in C_c^\infty(\mathfrak{g}^4)$.

Consider $f, g \in C_c^\infty(\mathfrak{g}^4)$ with support in a ball $B_R(0) \subset \mathfrak{g}^4$ and fix smooth vector fields u_i^R which agree with u_i on $B_{2R}(0)$ and have compact support. Let $\mathbf{Y}_t^{R,y}$ denote the RDE driven by $\pi_2(\mathbf{X}_t^{a,\pi^n(y)})$ along u_i^R starting at $\mathbf{Y}_0^{R,y} = y$. For all $y \in B_R(0)$ and $t \in [0, 1]$, we have $\mathbf{Y}_t^{R,y} = \mathbf{Y}_t^y$ whenever $\mathbf{Y}_s^y \in B_{2R}(0)$ for all $s \in [0, t]$. The probability that $\mathbf{Y}_s^y$ leaves $B_{2R}(0)$ in $[0, t]$ is bounded above by $C^{-1} \exp(-Ct^{-2/p})$ for any $2 < p < 3$ and some $C = C(R, p)$ (which follows from Fernique estimates on $\|\mathbf{X}^{a,x}\|_{1/p\text{-Hölder};[0,1]}$).

Defining $\mathcal{P}_t^R f(y) := \mathbb{E} \left[f(\mathbf{Y}_t^{R,y}) \right]$, it follows readily that

$$\lim_{t \rightarrow 0} \langle t^{-1}(f - \mathcal{P}_t f), g \rangle = \lim_{t \rightarrow 0} \langle t^{-1}(f - \mathcal{P}_t^R f), g \rangle.$$

Finally, the latter limit is now seen to equal $\mathcal{E}^{a \circ \pi^n}(f, g)$ following [22] Lemmas 26, 27 and the proof of Proposition 28 (note that one readily extends Lemma 27 to diffusions on $\mathfrak{g}^n$ for $n > 2$, cf. [25] Proposition 16.20).

Following Remark 4.2.17, we may apply Proposition 4.2.16 to conclude that for every $q > 4$, the $G\Omega_q(\mathbb{R}^d)$ -valued random variable $S_{n \vee [q]}(\mathbf{X}^{a,x})$ is in $\mathcal{N}_q(\mathbb{R}^d)$. Applying Corollary 4.2.10, we obtain the following partial solution to the moment problem for signatures arising from Markovian rough paths stopped upon exiting a domain.

Corollary 4.2.18. *Let $n \geq 1$, $\Lambda \geq 1$, $a \in \Xi^{n,d}(\Lambda)$ and $x \in \mathfrak{g}$. Let $D \subset \mathfrak{g}$ be an open set satisfying (4.2.1) and $T = \inf\{t > 0 \mid \mathbf{X}_t^{a,x} \notin D\}$. Let $p > 2$ and $\mathbf{X}$ denote the $G\Omega_p(\mathbb{R}^d)$ -valued random variable associated to $\mathbf{X}^{a,x} : [0, T] \mapsto \mathfrak{g}$.*

Then the expected signature of $S(\mathbf{X})_{0,T}$ uniquely determines the law of $S(\mathbf{X})_{0,T}$ amongst all $G(\mathbb{R}^d)$ -valued random variables in $\Phi(\mathbb{R}^d)$

We conclude this section with the proof of Lemma 4.2.12.

Proof of Lemma 4.2.12. We mimic the proofs of [25] Lemma 16.32 and Theorem 16.33 while keeping track of constants.

For $\alpha \in [0, 1/4)$, $\mathbf{h} \in W_x^{1,2}([0, 1], \mathfrak{g})$ and $\varepsilon > 0$ define the set

$$B_{\varepsilon;\alpha}^{\mathbf{h}} = \{\mathbf{x} \in C_x^{\alpha\text{-Hölder}}([0, 1], \mathfrak{g}) \mid \|\mathbf{x}\|_{\alpha\text{-Hölder}} \leq 2\|\mathbf{h}\|_{\alpha\text{-Hölder}} + 1, d_\infty(\mathbf{x}, \mathbf{h}) \leq \varepsilon\}.$$

We claim that for all $\alpha \in [0, 1/4)$ and $\varepsilon > 0$, there exists $\delta > 0$ such that

$$\mathbb{P}^{a,x} [\mathbf{X} \in B_{\varepsilon;\alpha}^{\mathbf{h}}] > \delta$$

for all $a \in \Xi^{n,d}(\Lambda)$, $x \in \mathfrak{g}$, and $\mathbf{h} \in W_{\theta;x}$.

Indeed, we follow the proof [25] Lemma 16.32 (we also mention here that, directly as stated, [25] Lemma 16.32 contains the minor error that it fails to hold for the

trivial path $\|\mathbf{h}\|_{\alpha\text{-H}\ddot{\text{ö}}\text{l}} = 0$; this is readily fixed by modifying the definition of their $B_\varepsilon^{\mathbf{h}}$ to our definition above; moreover the proof of [25] Theorem 16.33 then goes through unchanged).

Using Step 1 of the proof of [25] Lemma 16.32, we obtain that for any $\beta \in (\alpha, 1/2)$, $\mathbb{P}^{a,x} [\mathbf{X} \in B_{\varepsilon;\alpha}^{\mathbf{h}}] \geq \Delta_1 - \Delta_2$, where $\Delta_1 = \mathbb{P}^{a,x} [d_\infty(\mathbf{X}, \mathbf{h}) \leq \varepsilon]$ and

$$\Delta_2 = \mathbb{P}^{a,x} \left[\|\mathbf{X}\|_{\beta\text{-H}\ddot{\text{ö}}\text{l}} > (\|\mathbf{h}\|_{\alpha\text{-H}\ddot{\text{ö}}\text{l}} + 1)^{\beta/\alpha} (2\varepsilon)^{1-\beta/\alpha} \right].$$

The claim will follow once we show that $\Delta_2/\Delta_1 \rightarrow 0$ as $\varepsilon \rightarrow 0$ uniformly over $a \in \Xi^{n,d}(\Lambda)$, $x \in \mathfrak{g}$, and $\mathbf{h} \in W_{\theta;x}$.

By [25] Theorem E.21, we have $\log(\Delta_1) \geq -c_1 \varepsilon^{-2}$ where $c_1 = C(1 + \|\mathbf{h}\|_{W^{1,2}})^2$ and C is a constant depending only on the doubling and Poincaré constants of $\mathcal{E}^a$, which in turn depend only on Λ, n and d ([25] Proposition 16.5 and Theorem E.8).

On the other hand, the Fernique estimate in [25] Corollary 16.12 implies that

$$\log(\Delta_2) \leq -c_2 (\|\mathbf{h}\|_{\alpha\text{-H}\ddot{\text{ö}}\text{l}} + 1)^{2\beta/\alpha} \varepsilon^{2-2\beta/\alpha} \leq -c_2 \varepsilon^{2-2\beta/\alpha},$$

where c_2 depends only on β and Λ . So for fixed $\alpha \in [0, 1/4)$, choose any $\beta \in (2\alpha, 1/2)$. Since $2-2\beta/\alpha < -2$, we see $\Delta_2/\Delta_1 \rightarrow 0$ as $\varepsilon \rightarrow 0$ uniformly over the desired variables, which proves the claim.

To conclude, we follow the proof of [25] Theorem 16.33. By the d_0/d_∞ estimate on $\mathfrak{g}$ ([25] Proposition 8.15),

$$d_0(\mathbf{x}, \mathbf{h}) \leq C d_\infty(\mathbf{x}, \mathbf{h}) + C d_\infty(\mathbf{x}, \mathbf{h})^{1/n} (\|\mathbf{x}\|_\infty + \|\mathbf{h}\|_\infty)^{1-1/n},$$

where $C = C(n, d)$, and so by interpolation ([25] Lemma 8.16) we have for all $x \in \mathfrak{g}$, $\mathbf{x}, \mathbf{h} \in C_x^{\alpha\text{-H}\ddot{\text{ö}}\text{l}}([0, 1], \mathfrak{g})$ and $0 \leq \alpha' < \alpha < 1/4$ that

$$\begin{aligned} d_{\alpha'\text{-H}\ddot{\text{ö}}\text{l}}(\mathbf{x}, \mathbf{h}) &\leq (\|\mathbf{x}\|_{\alpha\text{-H}\ddot{\text{ö}}\text{l}} + \|\mathbf{h}\|_{\alpha\text{-H}\ddot{\text{ö}}\text{l}})^{\alpha'/\alpha} d_0(\mathbf{x}, \mathbf{h})^{1-\alpha'/\alpha} \\ &\leq C^{1-\alpha'/\alpha} (\|\mathbf{x}\|_{\alpha\text{-H}\ddot{\text{ö}}\text{l}} + \|\mathbf{h}\|_{\alpha\text{-H}\ddot{\text{ö}}\text{l}})^{\alpha'/\alpha} \\ &\quad [d_\infty(\mathbf{x}, \mathbf{h}) + d_\infty(\mathbf{x}, \mathbf{h})^{1/n} (\|\mathbf{x}\|_\infty + \|\mathbf{h}\|_\infty)^{1-1/n}]^{1-\alpha'/\alpha}. \end{aligned}$$

Since $\|\mathbf{h}\|_\infty \leq \|\mathbf{h}\|_{\alpha\text{-H}\ddot{\text{ö}}\text{l}} \leq \|\mathbf{h}\|_{W^{1,2}}$, it follows for all $\mathbf{h} \in W_{\theta;x}$ and $\mathbf{x} \in B_{\varepsilon;\alpha}^{\mathbf{h}}$ that

$$d_{\alpha'\text{-H}\ddot{\text{ö}}\text{l}}(\mathbf{x}, \mathbf{h}) \leq c_3 (\varepsilon + \varepsilon^{1/n})^{1-\alpha'/\alpha},$$

where c_3 depends only on n, d, α, α' and θ .

Choosing $\varepsilon > 0$ so that $c_3 (\varepsilon + \varepsilon^{1/n})^{1-\alpha'/\alpha} < c$, it follows that there exists $\delta > 0$ such that

$$\mathbb{P}^{a,x} [d_{\alpha'\text{-H}\ddot{\text{ö}}\text{l}}(\mathbf{X}, \mathbf{h}) < c] \geq \mathbb{P}^{a,x} [\mathbf{X} \in B_{\varepsilon;\alpha}^{\mathbf{h}} < c] > \delta$$

for all $a \in \Xi^{n,d}(\Lambda)$, $x \in \mathfrak{g}$, and $\mathbf{h} \in W_{\theta;x}$, which concludes the proof. $\square$

4.3 Method of moments

We conclude this chapter with a result analogous to the method of moments for weak convergence of $G(\mathbb{R}^d)$ -valued random variables. We work first with a slightly general notion of coproduct spaces as this is the only structure of E which we require.

Definition 4.3.1. A *coproduct space* (F, Δ) is a locally convex space F and a continuous linear map $\Delta : F \mapsto F^{\widehat{\otimes} 2}$, with the property that $G(F) := \{g \in F \mid \Delta(g) = g \otimes g, g \neq 0\}$ is closed in F . Let $\mathcal{P}_G(F)$ be the set of (weakly) integrable probability measures on $G(F)$.

The condition that $G(F)$ is closed in F will only arise in Lemma 4.3.4 to ensure that $G(F)$ is Polish whenever F is. Remark that (2.2.1) remains true for any coproduct space (F, Δ) , $f \in F'$, and $\mu \in \mathcal{P}(F)$ with support on $G(F)$. Recall that μ^* denotes the barycentre of a measure μ on F (whenever it exists).

Lemma 4.3.2. Let (F, Δ) be a nuclear coproduct space and γ a semi-norm on F . There exists a semi-norm ξ on F such that $\mu(\gamma) \leq \sqrt{\xi(\mu^*)}$ for all $\mu \in \mathcal{P}_G(F)$.

Proof. Let ζ be a semi-norm on F such that the canonical map $\widehat{F}_\zeta \mapsto \widehat{F}_\gamma$ is nuclear. Increasing ζ by a scalar multiple if necessary, it follows that there exist $(f_n)_{n \geq 1} \in F'$ such that $\sum_{n \geq 1} \zeta(f_n) \leq 1$ and $\gamma \leq \sum_{n \geq 1} |f_n|$. The conclusion then follows from (2.2.1) for any semi-norm ξ on F such that $\xi \geq (\zeta^{\otimes 2}) \circ \Delta$. $\square$

Lemma 4.3.3. Let (F, Δ) be a Fréchet nuclear coproduct space. Let $R \subseteq \mathcal{P}_G(F)$ be a family of probability measures on $G(F)$ such that $(\mu^*)_{\mu \in R}$ is bounded. Then R is uniformly tight.

Proof. Let $(\gamma_n)_{n \geq 1}$ be a defining non-decreasing sequence of semi-norms on F . It follows from Lemma 4.3.2 that there exists a sequence of semi-norms $(\xi_n)_{n \geq 1}$ on F such that $\mu(\gamma_n) \leq \sqrt{\xi_n(\mu^*)}$ for all $\mu \in \mathcal{P}_G(F)$. Since $(\mu^*)_{\mu \in R}$ is bounded, $\sup_{\mu \in R} \xi_n(\mu^*) < \infty$ for every $n \geq 1$.

Let $B_n = \{x \in F \mid \gamma_n(x) < 1\}$. For any sequence of positive reals $(\lambda_n)_{n \geq 1}$, the set $K := \bigcap_{n \geq 1} \lambda_n B_n$ is bounded in H and thus relatively compact ([57] p.520). For all $\mu \in \mathcal{P}_G(F)$ we have that

$$\mu(K^c) \leq \sum_{n \geq 1} \mu(\{x \mid \gamma_n(x) \geq \lambda_n\}) \leq \sum_{n \geq 1} \lambda_n^{-1} \mu(\gamma_n) \leq \sum_{n \geq 1} \lambda_n^{-1} \sqrt{\xi_n(\mu^*)}.$$

Taking λ_n sufficiently large, it follows that $\sup_{\mu \in R} \mu(K^c)$ can be made arbitrarily small. $\square$

Lemma 4.3.4. *Let (F, Δ) be a Fréchet nuclear coproduct space and let $(\mu_n)_{n \geq 1}$ be a sequence of measures in $\mathcal{P}_G(F)$ such that $\mu_n^* \rightarrow x$ weakly for some $x \in F$. Then there exists $\mu \in \mathcal{P}_G(F)$ and a subsequence $(n(k))_{k \geq 1}$ such that $\mu_{n(k)} \xrightarrow{\mathcal{D}} \mu$ and $x = \mu^*$.*

Proof. Recall that a Fréchet Montel space (thus in particular a Fréchet nuclear space) is always separable ([52] p.195), and hence Polish. As a closed subset of F , $G(F)$ is also Polish.

The sequence $(\mu_n^*)_{n \geq 1}$ is bounded ([51] Theorem 3.18) thus there exists a convergent subsequence $\mu_{n(k)} \rightarrow \mu$ for some probability measure μ on $G(F)$ by Lemma 4.3.3.

Let $f \in F'$. Since $\sup_{n \geq 1} \mu_n(f^2) = \sup_{n \geq 1} (f^{\otimes 2})(\Delta \mu_n^*) < \infty$, the sequence of image measures $(\mu_n f^{-1})_{n \geq 1}$ on $\mathbb{R}$ is uniformly integrable. It follows that f is μ -integrable and $f(\mu_n^*) = \mu_n(f) \rightarrow \mu(f)$ ([5] Lemma 8.4.3). Thus $x = \mu^*$ and $\mu \in \mathcal{P}_G(F)$. $\square$

Recall that E is Fréchet and nuclear whenever V is (Corollary 2.1.5 and Proposition 2.1.8). The following is now a consequence of Lemma 4.3.4 and Proposition 4.1.1.

Theorem 4.3.5. *Let $(X_n)_{n \geq 1}$ be a sequence of $G(\mathbb{R}^d)$ -valued random variables such that $\mathbb{E}[X_n] \in E(\mathbb{R}^d)$ exists (i.e., $r_2(X_n) = \infty$) for all $n \geq 1$. Suppose that $\mathbb{E}[X_n]$ converges to some $x \in E(\mathbb{R}^d)$ in the weak topology of $E(\mathbb{R}^d)$. Then there exists a unique in law integrable $G(\mathbb{R}^d)$ -valued random variable X such that $X_n \xrightarrow{\mathcal{D}} X$ and $x = \mathbb{E}[X]$.*

Remark 4.3.6. We remark that $F := P(\mathbb{R}) = \prod_{k \geq 0} (\mathbb{R})^{\otimes k}$ is also a Fréchet nuclear coproduct space under the product topology. Moreover the exponential map $\exp : \mathbb{R} \mapsto G(\mathbb{R}) = G(F)$ is a homeomorphism. One may then directly apply Lemma 4.3.4 to obtain a proof of the classical method of moments for real random variables: if μ_n are probability measures on $\mathbb{R}$ with finite moments $(m_n(j))_{j \geq 1}$ such that $\lim_{n \rightarrow \infty} m_n(j) = m(j)$ for every $j \geq 1$, then $(m(j))_{j \geq 1}$ are the moments of a probability measure μ on $\mathbb{R}$ for which $\mu_{n(k)} \xrightarrow{\mathcal{D}} \mu$ along a subsequence $(n(k))_{k \geq 1}$ (if μ is moment-determined then in fact $\mu_n \xrightarrow{\mathcal{D}} \mu$).

Chapter 5

Random walks in path space and Lévy processes

This chapter discusses several related topics concerning random rough paths with stationary and independent increments, and their connection to weak convergence of random walks.

In Section 5.1, we collect several results concerning Lévy processes in a general Lie group, and develop tools which will be applied to the nilpotent Lie group $G^N(\mathbb{R}^d)$. The primary tool of the section is Theorem 5.1.1, which is a special case of a theorem due to Feinsilver [19], and provides necessary and sufficient condition for a random walk to converge to a Lévy process in the Skorokhod topology.

In Section 5.2, the main question we address is the following: given a sequence $(\mathbf{X}^n)_{n \geq 1}$ of random walks in $G^N(\mathbb{R}^d)$, and $p \geq 1$, when is the collection of real random variables $(\|\mathbf{X}^n\|_{p\text{-var};[0,1]})_{n \geq 1}$ tight? We provide a sufficient condition in Theorem 5.2.3, and apply the result to show weak convergence in rough path topologies of “random walks in path space” (concatenations of geometric rough paths).

Section 5.3 introduces the concept of a path function. This concept stems from the ideas of Friz and Shekhar [24] and Williams [59] on a systematic way to turn càdlàg paths into rough paths. However our results are essentially new and single out the important elements in the construction of a continuous (or rough) path from a càdlàg path.

Section 5.4 focuses on Lévy processes as rough paths, and constitutes the main contribution of this chapter. The three main results are the following:

- For any given $p > 1$, an (almost) complete characterisation of the generators of Lévy processes in $G^N(\mathbb{R}^d)$ possessing sample paths of a.s. finite p -variation.

- A Lévy-Khintchine formula for the characteristic function (in the sense of Section 2.3) of the signature of the rough path constructed from a Lévy process by applying a path function.
- A sufficient condition for a random walk connected by a path function to converge in law to a Lévy process in rough path topologies.

We apply the last of these results to weak convergence of stochastic flows in several examples. Notably, we provide a significant generalisation of a result of Kunita [39] and of a related result of Breuillard, Friz and Huesmann [7].

5.1 Iid Arrays and Lévy processes in Lie groups

All Lie groups considered are assumed to be second-countable, which in particular implies they are Polish spaces ([35] p.59).

Throughout the section, we fix a Lie group G with Lie algebra $\mathfrak{g}$, and an open neighbourhood $U \subset G$ of the identity $1_G \in G$, such that U has compact closure and $\exp : \mathfrak{g} \mapsto G$ is a diffeomorphism from a neighbourhood of zero in $\mathfrak{g}$ onto U .

Let $u_1, \dots, u_m$ be a basis for $\mathfrak{g}$. We equip $\mathfrak{g}$ with the inner product under which $u_1, \dots, u_m$ is an orthonormal basis. When x is an element of a normed space, we denote its norm by $|x|$.

Let $\xi_i \in C_c^\infty(G, \mathbb{R})$ be smooth functions of compact support such that $\log(x) = \sum_{i=1}^m \xi_i(x) u_i$ for all $x \in U$ (that is, ξ_i provide exponential coordinates of the first kind on U). We denote $\xi : G \mapsto \mathfrak{g}, \xi(x) = \sum_{i=1}^m \xi_i(x) u_i$. For an element $y \in \mathfrak{g}$ we write $y = \sum_{i=1}^m y^i u_i$.

5.1.1 Iid Arrays

For an arbitrary set S , an S -valued array of variables is a sequence of a finite collection of elements $(x_{n1}, \dots, x_{nn})_{n \geq 1}, x_{nj} \in S$. If each X_{nj} is a random variable, we call the corresponding array *iid* when, for all $n \geq 1$, $X_{n1}, \dots, X_{nn}$ are iid.

For an array $x_{n1}, \dots, x_{nn}$ in a Lie group G , we denote by $\mathbf{x}^n : [0, 1] \mapsto G$ the *associated walk* defined by $\mathbf{x}_t^n = 1_G$ for all $t \in [0, n^{-1})$, and $\mathbf{x}_t^n = x_{n1} \dots x_{n[\!tn\!]}$ for all $t \in [n^{-1}, 1]$. In the case of an array of random variables X_{nj} , we call $\mathbf{X}^n$ the *associated random walk*.

Let X_{nj} be an iid array of G -valued random variables with associated probability measures F_n . We call X_{nj} infinitesimal if for every neighbourhood V of 1_G ,

$\lim_{n \rightarrow \infty} \mathbb{P}[X_{n1} \notin V] = 0$. Unless otherwise stated, every iid array we consider on a Lie group shall be infinitesimal.

For all $n \geq 1$, define

$$B_n := \mathbb{E}[\xi(X_{n1})] \in \mathfrak{g},$$

and for all $i, j \in \{1, \dots, m\}$

$$A_n^{i,j} := \mathbb{E}[\xi_i(X_{n1})\xi_j(X_{n1})].$$

Note that since X_{nj} is infinitesimal, it holds that $\lim_{n \rightarrow \infty} B_n = 0$ and $\lim_{n \rightarrow \infty} A_n^{i,j} = 0$.

5.1.2 Lévy processes

For a Polish space E , recall the Skorokhod space $D([0, T], E)$ of càdlàg functions $\mathbf{x} : [0, T] \mapsto E$. We refer to [2] Section 12 for the basic properties of $D([0, T], E)$. We recall in particular that $D([0, T], E)$ is a Polish space.

As usual, we shall use the symbol o to denote spaces of paths whose starting point is the identity. For example $D_o([0, T], G)$ denotes the subset of all $\mathbf{x} \in D([0, T], G)$ such that $\mathbf{x}_0 = 1_G$.

Recall that a (left) Lévy process in G is a $D([0, T], G)$ -valued random variable $\mathbf{X}$ such that

1. the (right) increments of $\mathbf{X}$ are independent, i.e., for all $0 \leq t_1 < \dots < t_k \leq T$, $\mathbf{X}_{t_j}^{-1} \mathbf{X}_{t_{j+1}}$, $j \in \{1, \dots, k\}$, are independent,
2. $\mathbf{X}$ is continuous in probability, i.e., $\mathbf{X}_s^{-1} \mathbf{X}_t \xrightarrow{\mathcal{D}} 1_G$ as $s \rightarrow t$ for all $t \in [0, T]$, and
3. the increments of $\mathbf{X}$ are (right) stationary, i.e., for all $0 \leq s \leq t \leq T$, $\mathbf{X}_s^{-1} \mathbf{X}_t \stackrel{\mathcal{D}}{=} \mathbf{X}_0^{-1} \mathbf{X}_{t-s}$.

We refer to Liao [40] for further details, or Kunita [39] for a concise summary.

Unless otherwise stated, every Lévy process $\mathbf{X}$ which we consider will take values in $D_o([0, T], G)$, i.e., $\mathbf{X}_0 = 1_G$ a.s..

Recall that a Lévy measure Π on G is a σ -finite (Borel) measure such that

$$\Pi(G \setminus U) + \int_U |\xi(x)|^2 \Pi(dx) < \infty$$

(note that any choice of U and ξ with the given properties leads to the same definition of a Lévy measure).

A *Lévy triplet* (or simply *triplet*) is a collection (A, B, Π) of an $m \times m$ covariance matrix $(A_{i,j})_{i,j=1}^m$, a collection of reals $(B_i)_{i=1}^m$, and a Lévy measure Π on G . We equivalently treat B as an element of $\mathfrak{g}$, given canonically by $B = \sum_{i=1}^m B_i u_i$.

Denote by $C^2(G)$ the space of twice continuously differentiable functions $f : G \mapsto \mathbb{R}$ such that $f, u_i f, u_i u_j f \in C_b(G)$ for all $i, j = 1, \dots, m$.

A classical theorem of Hunt [33] asserts that for every Lévy process $\mathbf{X}$ in G , there exists a unique triplet (A, B, Π) such that the generator of $\mathbf{X}$ is given for all $f \in C^2(G)$ and $x \in G$ by

$$\begin{aligned} Lf(x) &:= \lim_{t \rightarrow 0} \frac{\mathbb{E}[f(x\mathbf{X}_t) - f(x)]}{t} \\ &= \sum_{i=1}^m B_i(u_i f)(x) + \frac{1}{2} \sum_{i,j=1}^m A_{i,j}(u_i u_j f)(x) \\ &\quad + \int_G \left[f(xy) - f(x) - \sum_{i=1}^m \xi_i(y)(u_i f)(x) \right] \Pi(dy). \end{aligned}$$

Conversely, every triplet gives rise to a unique Lévy process.

Note that the values of $A_{i,j}$ and B_i in the triplet (A, B, Π) associated to a Lévy process $\mathbf{X}$ depend on our choice of basis $u_1, \dots, u_m$. Moreover, the values B_i depend on our choice of local coordinates ξ . When we wish to emphasise the role of ξ and $u_1, \dots, u_m$, we shall say that (A, B, Π) is the triplet of $\mathbf{X}$ with respect to local coordinates $\xi_1, \dots, \xi_m$.

5.1.3 Convergence towards a Lévy process

In [19], Feinsilver established necessary and sufficient conditions under which the random walk $\mathbf{X}^n$ associated to a general infinitesimal array of independent G -valued random variables X_{nj} (not necessarily identically distributed) converges in law (as a $D_o([0, T], E)$ -valued random variable) to a specified Markov process $\mathbf{X}$. Feinsilver's approach is based upon the characterisation of a Markov process by its associated martingales. This method was initially developed by Stroock and Varadhan [53], [54], and was successively employed in [55] to solve the analogous problem in the case that the limiting Markov process $\mathbf{X}$ is a diffusion.

For the case of iid arrays, similar results to those of Feinsilver were obtained by Kunita [39]. We note however that only sufficient (and not necessary) conditions were established in [39] for convergence in law to hold. Furthermore, Kunita linked his results to convergence in law of flows of diffeomorphisms (to so-called Lévy flows) along

collections of vector fields. We shall revisit, and, in a certain direction, substantially generalise the results of [39] concerning convergence in law of flows in Example 5.4.9.

We remark moreover that Feinsilver [19] (as well as Kunita [55], and Stroock and Varadhan [39]) work with a general system of local coordinates ξ which differentiate as exponential coordinates at the identity. We have chosen, purely for convenience, to work precisely with exponential coordinates.

We shall use the following theorem as the characterisation of random walks which converge to a Lévy process in G . Recall the notation introduced in Section 5.1.1 regarding an iid array X_{nj} on G .

Theorem 5.1.1 (Feinsilver [19]). *Let X_{nj} be an iid array of G -valued random variables and $\mathbf{X}^n$ the associated random walk. Let $\mathbf{X}$ be a Lévy process in G with triplet (A, B, Π) .*

Then $\mathbf{X}^n \xrightarrow{\mathcal{D}} \mathbf{X}$ as $D_o([0, 1], G)$ -valued random variables if and only if

- (1) $\lim_{n \rightarrow \infty} nF_n(f) = \Pi(f)$ for every $f \in C_b(G)$ which is identically zero on a neighbourhood of 1_G ,
- (2) $\lim_{n \rightarrow \infty} nB_n = B$, and
- (3) for all $i, j \in \{1, \dots, m\}$,

$$\lim_{n \rightarrow \infty} nA_n^{i,j} = A_{i,j} + \int_G \xi_i(x) \xi_j(x) \Pi(dx).$$

We emphasise that Theorem 5.1.1 is only a special case of the main results in [19]. In fact, the results of Feinsilver hold in much greater generality, and contain significantly deeper consequences. In particular, Feinsilver was able to give a complete characterisation of $D_o([0, T], G)$ -valued random variables having independent (not necessarily stationary) increments, vastly generalising the aforementioned result of Hunt.

While the implication is elementary, we explain in detail how the results in [19] imply Theorem 5.1.1, as this may seem unclear at first. Note the following two lemmas, whose proofs are elementary and we omit.

Lemma 5.1.2. *Let $(B_n)_{n \geq 1}$ be a sequence in $\mathfrak{g}$ such that $\lim_{n \rightarrow \infty} B_n = 0$, and let $B \in \mathfrak{g}$. Then $\exp(B_n)^{\lfloor tn \rfloor} \rightarrow \exp(tB)$ uniformly in $t \in [0, 1]$ as $n \rightarrow \infty$, if and only if $\lim_{n \rightarrow \infty} nB_n = B$.*

Lemma 5.1.3. *Suppose that (1) and (2) in Theorem 5.1.1 hold. Let V be a neighbourhood of 1_G which is a continuity set of Π . Then (3) in Theorem 5.1.1 holds if and only if for all $i, j \in \{1, \dots, m\}$*

$$\lim_{n \rightarrow \infty} n\mathbb{E}[(\xi_i(X_{n1}) - B_n^i)(\xi_j(X_{n1}) - B_n^j)\mathbf{1}\{X_{n1} \in V\}] = A_{i,j} + \int_V \xi_i(x)\xi_j(x)\Pi(dx). \quad (5.1.1)$$

Proof of Theorem 5.1.1. Observe that $\mathbf{X}$ is the unique $D_o([0, 1], G)$ -valued random variable determined by the martingale property with parameters $B(t) := tB$, $A := tA$ and $M(t, dx) := t\Pi(dx)$, where we employ the definitions and notation of [19] Martingale Characterization Theorem, p.80–81.

Defining the “mean” function $m(t) := \exp(tB)$ (treating B canonically as an element of $\mathfrak{g}$), and the “centred” process $\mathbf{Z}_t := \mathbf{X}_t m(t)^{-1}$, observe that by [19] Uniqueness of Parameters Theorem, 2., p.81., $\mathbf{Z}_t$ satisfies the martingale property of [19] Theorem, 3., p.82–83. It follows that $\mathbf{X}_t$ is the unique process having representation with parameters (m, A, M) (again employing the notation of [19] Theorem, 5., p.83).

It thus follows from [19] Corollary, p.83, that $\mathbf{X}^n \xrightarrow{\mathcal{D}} \mathbf{X}$ as $D_o([0, 1], G)$ -valued random variables if and only if (1), (2’), and (3’) hold, where

$$(2') \quad \exp(B_n)^{\lfloor tn \rfloor} \rightarrow \exp(tB) \text{ uniformly in } t \in [0, 1] \text{ as } n \rightarrow \infty;$$

$$(3') \quad \text{for some neighbourhood } V \text{ of } 1_G \text{ which is a continuity set of } \Pi, (5.1.1) \text{ holds for all } i, j \in \{1, \dots, m\}.$$

However (2’) and (2) are equivalent by Lemma 5.1.2, and (3’) and (3) are equivalent whenever (1) and (2) hold by Lemma 5.1.3, from which the conclusion follows. $\square$

5.1.4 Scaling functions

For a subset $W \subseteq G$ for which 1_G is an accumulation point, a normed space E , and functions $f, h : W \mapsto E$ such that $h(x) \neq 0$ for all $x \neq 1_G$, we shall use throughout the chapter little- o and big- O notation to mean that $f = o(h)$ if

$$\lim_{x \rightarrow 1_G} \frac{|f(x)|}{|h(x)|} = 0,$$

and $f = O(h)$ if

$$\limsup_{x \rightarrow 1_G} \frac{|f(x)|}{|h(x)|} < \infty.$$

If 1_G is not an accumulation point of W , then we say that $f = o(h)$ and $f = O(h)$ always hold for any pair of functions $f, h : W \mapsto E$.

Definition 5.1.4 (Scaling function). A continuous bounded function $\theta : G \rightarrow \mathbb{R}$ is called a *scaling function* if

- (i) $\theta(1_G) = 0$,
- (ii) $\theta(x) > 0$ for all $x \neq 1_G$,
- (iii) $|\xi|^2 = O(\theta)$, and
- (iv) there exists $c > 0$ such that $\theta > c$ outside a neighbourhood of 1_G with compact closure.

Let $X_{n1}, \dots, X_{nn}$ be an array of G -valued random variables. We say that θ *scales the array* X_{nj} if

$$\sup_{n \geq 1} \sum_{j=1}^n \mathbb{E} [\theta(X_{nj})] < \infty.$$

Definition 5.1.5 (Taylor expansion). Consider a scaling function θ on G , a subset $W \subseteq G$ containing 1_G , a normed space E , and a function $f : W \mapsto E$. Let $(b_i)_{i=1}^m$ and $(a_{i,j})_{i,j=1}^m$ be elements in E , and $h : G \mapsto E$ a function such that

$$f(x) = f(1_G) + \sum_{i=1}^m b_i \xi_i(x) + \frac{1}{2} \sum_{i,j=1}^m a_{i,j} \xi_i(x) \xi_j(x) + h(x) \quad (5.1.2)$$

for all $x \in W$. We call (5.1.2) a *Taylor expansion* of f on W of order θ if $h = o(\theta)$.

When W and θ are clear from the context, we simply call (5.1.2) a Taylor expansion of f . Observe that the existence of a Taylor expansion is entirely a local property of f about 1_G . Correspondingly, if f admits a Taylor expansion of order θ , and $\psi : G \mapsto \mathbb{R}$ is any function such that $\psi|_V = \theta|_V$ for a neighbourhood V of 1_G , we shall equivalently state that f admits a Taylor expansion of order ψ . Note that every $f \in C^2(G)$ admits a Taylor expansion of order θ .

Note that in general a Taylor expansion is not unique, even when $W = G$. In particular, if 1_G is not an accumulation point of W , then, by definition, any choice of $(b_i)_{i=1}^m$ and $(a_{i,j})_{i,j=1}^m$ provide a Taylor expansion of f , for which one can choose h identically zero on a neighbourhood of 1_G .

Furthermore, the possible coefficients a, b in (5.1.2) depend on the choice of basis $u_1, \dots, u_m$ and the local coordinates $\xi_1, \dots, \xi_m$. However observe that the definition of a scaling function and the existence of a Taylor expansion does not depend on this choice.

We record a simple consequence of the convergence in (1) of Theorem 5.1.1 which we shall use later.

Lemma 5.1.6. *Let X_{nj} be an iid array scaled by θ such that (1) of Theorem 5.1.1 holds for some Borel measure Π on G . Then*

$$\int_G \theta(x) \Pi(dx) \leq \limsup_{n \rightarrow \infty} n \mathbb{E} [\theta(X_{nj})] < \infty.$$

Proof. Let $(U_i)_{i \geq 1}$ be a decreasing sequence of neighbourhoods which shrinks to the identity in G . Let $(\theta_i)_{i \geq 1}$ be an increasing sequence of continuous functions such that θ_i agrees with θ on U_i^c and is identically zero on a neighbourhood of 1_G . Then (1) implies that for all $i \geq 1$

$$\lim_{n \rightarrow \infty} n \mathbb{E} [\theta_i(X_{n1})] = \int_G \theta_i(x) \Pi(dx).$$

Moreover, by the monotone convergence theorem,

$$\int_G \theta_i(x) \Pi(dx) \nearrow_{i \rightarrow \infty} \int_G \theta(x) \Pi(dx),$$

and for every $n \geq 1$

$$n \mathbb{E} [\theta_i(X_{n1})] \nearrow_{i \rightarrow \infty} n \mathbb{E} [\theta(X_{n1})],$$

from which the conclusion follows. $\square$

The prototypical scaling function is given by the following: let c be sufficiently small such that $W := \{\exp(y) \mid y \in \mathfrak{g}, |y| \leq c\}$ is contained in U . Then

$$\theta(x) := \mathbf{1}\{x \in U\}(c^2 \wedge |\xi(x)|^2) + c^2 \mathbf{1}\{x \notin U\} \quad (5.1.3)$$

is a scaling function. The following result explains the condition $|\xi|^2 = O(\theta)$ in Definition 5.1.4 by ensuring that every iid array, for which the corresponding random walk converges, is scaled by θ in (5.1.3).

Proposition 5.1.7. *Let X_{nj} be an iid array in G and $\mathbf{X}^n$ the associated random walk. Suppose that $\mathbf{X}^n$ converges in law to a Lévy process in G . Then θ defined by (5.1.3) scales the array X_{nj} .*

Proof. From (1) of Theorem 5.1.1, and the fact that $\Pi(f) < \infty$ for any $f \in C_b(G)$ which is identically zero on a neighbourhood of 1_G , it follows that

$$\sup_{n \geq 1} n \mathbb{E} [\mathbf{1}\{X_{n1} \notin W\}] < \infty.$$

Furthermore, from (3) of Theorem 5.1.1, it follows that

$$\limsup_{n \rightarrow \infty} n \mathbb{E} [|\xi(X_{n1})|^2 \mathbf{1}\{X_{n1} \in W\}] \leq \lim_{n \rightarrow \infty} n \mathbb{E} [|\xi(X_{n1})|^2] < \infty,$$

which completes the proof. $\square$

5.1.5 Convergence towards the generator

The main result of this section is Proposition 5.1.8 which provides a condition under which the limit $\lim_{n \rightarrow \infty} n\mathbb{E}[f(X_{n1}) - f(1_G)]$ exists for a function $f : G \mapsto \mathbb{R}$ and an iid array X_{nj} in G .

Let $W \subseteq G$ be a closed subset containing 1_G , and θ a scaling function on G . For a collection of functions $R \subset C_b(W)$, we say that R is *uniformly controlled by θ* if

(U1) for every compact $K \subseteq W$, R is equicontinuous on K ,

(U2) $\sup_{f \in R} \sup_{x \in W} |f(x)| < \infty$,

and every $f \in R$ admits a Taylor expansion of order θ

$$f = f(1_G) + \sum_{i=1}^m b_i^f \xi_i + \frac{1}{2} \sum_{i,j=1}^m a_{i,j}^f \xi_i \xi_j + h^f \quad (5.1.4)$$

such that

(U3) if 1_G is an accumulation point of W , then $\lim_{x \rightarrow 1_G} \sup_{f \in R} |h^f(x)/\theta(x)| = 0$, and

(U4) $\sup_{f \in R} |b_i^f| < \infty$ and $\sup_{f \in R} |a_{i,j}^f| < \infty$ for all $i, j \in \{1, \dots, m\}$.

In particular, observe that for any $f \in C_b(W)$ which admits a Taylor expansion of order θ , the singleton set $\{f\}$ is uniformly controlled by θ .

For a Lévy triplet (A, B, Π) and f with Taylor expansion (5.1.4), define the quantity

$$L^f := \sum_{i=1}^m B_i b_i^f + \frac{1}{2} \sum_{i,j=1}^m A_{i,j} a_{i,j}^f + \int_W \left[f(x) - f(1_G) - \sum_{i=1}^m b_i^f \xi_i(x) \right] \Pi(dx).$$

Note that in general L^f depends on the Taylor expansion of f .

Proposition 5.1.8. *Let X_{nj} be an iid array in G and $\mathbf{X}^n$ the associated random walk. Suppose that X_{nj} is scaled by a scaling function θ , and that $\mathbf{X}^n$ converges in law to a Lévy process $\mathbf{X}$ in G with triplet (A, B, Π) . Let $W \subseteq G$ be a closed subset such that $X_{n1} \in W$ a.s. for all $n \geq 1$, and $\text{supp}(\Pi) \subseteq W$.*

Then for every $R \subset C_b(W)$ which is uniformly controlled by θ , it holds that

$$\lim_{n \rightarrow \infty} \sup_{f \in R} |n\mathbb{E}[f(X_{n1}) - f(1_G)] - L^f| = 0,$$

where L^f does not depend on the Taylor expansion of f .

Remark 5.1.9. Remark that if $W \subseteq G$ is a closed subset and X_{nj} is an infinitesimal array in G such that $X_{nj} \in W$ a.s. for all $n \geq 1$, then $1_G \in W$. In particular, $f(1_G)$ in the above proposition is well-defined.

The following lemma first considers the case that the coefficients of the Taylor expansion of every $f \in R$ are zero.

Lemma 5.1.10. *Use the notation from Proposition 5.1.8. Let $R \subset C_b(W)$ satisfy (U1), (U2), and*

(U3') if 1_G is an accumulation points of W , then $\lim_{x \rightarrow 1_G} \sup_{f \in R} |f(x)/\theta(x)| = 0$.

Then

$$\lim_{n \rightarrow \infty} \sup_{f \in R} \left| n\mathbb{E}[f(X_{n1})] - \int_W f(x)\Pi(dx) \right| = 0.$$

Proof. As $\Pi(U^c) < \infty$, by (1) of Theorem 5.1.1, there exists a sufficiently large compact set $K \subseteq W$ such $\Pi(W \setminus K)$ and $\sup_n nF_n(W \setminus K)$ are sufficiently small so that, due to (U2), $\sup_{f \in R} \int_{W \setminus K} |f|d\Pi$ and $\sup_{n \geq 1, f \in R} \int_{W \setminus K} |f|ndF_n$ are arbitrarily small.

It thus remains to show

$$\lim_{n \rightarrow \infty} \sup_{f \in R} \left| n\mathbb{E}[f(X_{n1})\mathbf{1}\{X_{n1} \in K\}] - \int_K f(x)\Pi(dx) \right| = 0. \quad (5.1.5)$$

Consider $K_0 := K \setminus \{1_G\}$ as a locally compact space and equip $C_c(K_0)$ and $C_0(K_0)$ with the uniform norm. For a measure μ on G , denote by μ^{K_0} the restriction of μ to K_0 . From (1) of Theorem 5.1.1, we have that $nF_n^{K_0}(f) \rightarrow \Pi^{K_0}(f)$ for all $f \in C_c(K_0)$.

On G , define the measures $\nu_n(dx) := \theta(x)nF_n(dx)$. Note that the statement that θ scales the array X_{nj} is equivalent to $\sup_{n \geq 1} \nu_n(G) < \infty$. Hence, by the Riesz–Markov–Kakutani theorem, $\nu_n^{K_0}$ are linear functionals on $C_0(K_0)$ bounded in norm, and thus equicontinuous in the (topological) dual of $C_0(K_0)$. Since $C_c(K_0)$ is dense in $C_0(K_0)$ and $\nu_n^{K_0} \rightarrow \theta\Pi^{K_0}$ pointwise on $C_c(K_0)$, it follows that $\nu_n^{K_0} \rightarrow \theta\Pi^{K_0}$ uniformly on compact subsets of $C_0(K_0)$ ([57] Proposition 32.5).

Observe that (U1), (U2), and (U3') imply that the set $\{(f/\theta)|_{K_0} \mid f \in R\}$ is uniformly bounded, equicontinuous, and vanishes uniformly at infinity, thus is compact in $C_0(K_0)$ by the Arzelà–Ascoli theorem. Thus $\nu_n^{K_0} \rightarrow \theta\Pi^{K_0}$ uniformly on $\{(f/\theta)|_{K_0} \mid f \in R\}$, from which (5.1.5) follows. $\square$

Proof of Proposition 5.1.8. By Theorem 5.1.1, we have

$$\lim_{n \rightarrow \infty} n\mathbb{E}[\xi_i(X_{n1})] = B_i \quad (5.1.6)$$

and

$$\lim_{n \rightarrow \infty} n\mathbb{E} [\xi_i(X_{n1})\xi_j(X_{n1})] = A_{i,j} + \int_W \xi_i(x)\xi_j(x)\Pi(dx). \quad (5.1.7)$$

Note that (U3) implies that (U3') holds for the collection of functions $(h^f|_W)_{f \in R}$, so by Lemma 5.1.10

$$\lim_{n \rightarrow \infty} \sup_{f \in R} \left| n\mathbb{E} [h^f(X_{n1})] - \int_W h^f(x)\Pi(dx) \right| = 0.$$

Since (U4) implies b_i^f and $a_{i,j}^f$ are uniformly bounded for $f \in R$, combining (5.1.6) and (5.1.7) with the definition of h^f from the Taylor expansion (5.1.4) yields the conclusion. $\square$

5.1.6 Maps between Lie groups

Consider an iid array X_{nj} in G such that the associated random walk $\mathbf{X}^n$ converges in law to a Lévy process $\mathbf{X}$ in G with Lévy measure Π . Let $W \subseteq G$ be a closed subset such that $X_{n1} \in W$ a.s. for all $n \geq 1$ (which in particular implies $1_G \in W$) and $\text{supp}(\Pi) \subseteq W$.

Let H be another Lie group with Lie algebra $\mathfrak{h}$. Let $v_1, \dots, v_k$ be a basis for $\mathfrak{h}$ and $\sigma_i \in C_c^\infty(H)$, $i \in \{1, \dots, k\}$, be functions such that $\sigma := \sum_{i=1}^k v_i \sigma_i$ coincides with $\log$ on a sufficiently small neighbourhood of 1_H (so that σ_i play the same role in H as ξ_i in G).

Consider $f : W \mapsto H$ a continuous function such that $f(1_G) = 1_H$. Denote $Y_{nj} := f(X_{nj})$ and the associated random walk $\mathbf{Y}^n : [0, 1] \mapsto H$. The following result follows immediately by the characterisation of convergence in Theorem 5.1.1.

Proposition 5.1.11. *Suppose that for all $i, j \in \{1, \dots, k\}$*

$$D_i := \lim_{n \rightarrow \infty} n\mathbb{E} [\sigma_i(Y_{n1})]$$

and

$$C_{i,j} := \lim_{n \rightarrow \infty} n\mathbb{E} [\sigma_i(Y_{n1})\sigma_j(Y_{n1})] - \int_G \sigma_i(f(x))\sigma_j(f(x))\Pi(dx)$$

exist.

Then $\mathbf{Y}^n$ converges to the Lévy process $\mathbf{Y}$ in H with triplet (C, D, Ξ) , where Ξ is the push-forward of Π by f .

Remark 5.1.12. Observe in particular that C and D exist whenever X_{nj} is scaled by θ and $\sigma_i \circ f$ admits a Taylor expansion of order θ for all $i \in \{1, \dots, k\}$.

5.1.7 Approximating walk

In this section, given a Lévy process $\mathbf{X}$ in G , we construct an iid array X_{nj} for which the associated random walk $\mathbf{X}^n$ converges in law to $\mathbf{X}$. The array X_{nj} has the advantage that it takes values in either the Lévy measure Π of $\mathbf{X}$, or in a set which shrinks to the identity as $n \rightarrow \infty$. In certain situations, this makes the walk $\mathbf{X}^n$ significantly easier to analyse than the increments of $\mathbf{X}$, and will later help us deduce several important properties about $\mathbf{X}$ itself. In particular, the walk $\mathbf{X}^n$ will be an important tool in showing finiteness of p -variation of $\mathbf{X}$ and in the calculation of the Lévy-Khintchine formula in Theorems 5.4.1 and 5.4.5 respectively.

Throughout this section, let $\mathbf{X}$ be a Lévy process in G with triplet (A, B, Π) . For $i \in \{1, \dots, m\}$ define

$$\Gamma_i := \left\{ 0 \leq \gamma < 2 \mid \int_G |\xi_i(x)|^\gamma \Pi(dx) = \infty \right\}.$$

Define the sets of indexes

$$J = \{j \in \{1, \dots, m\} \mid A_{j,j} > 0\},$$

$$\tilde{K} = \{k \in \{1, \dots, m\} \mid 1 \notin \Gamma_k\}.$$

For $k \in \tilde{K}$ define

$$\tilde{B}_k = B_k - \int_G \xi_k(x) \Pi(dx),$$

and let $K = \{k \in \tilde{K} \mid \tilde{B}^k \neq 0\}$.

For n sufficiently large so that $\Pi(U^c) < n/2$, let

$$h_n = \inf \{h \geq 0 \mid \Pi(\{|\xi(x)| > h\} \cup U^c) \leq n/2\}.$$

Define $U_n = \{x \in U \mid |\xi(x)| \leq h_n\}$ and note that $w_n := \Pi\{U_n^c\} \leq n/2$. Remark that $\lim_{n \rightarrow \infty} h_n = 0$ which implies that U_n shrinks to 1_G as $n \rightarrow \infty$.

Define on G the probability measure $\mu_n(dx) := w_n^{-1} \mathbf{1}\{x \in U_n^c\} \Pi(dx)$. Observe that by Hölder's inequality, for all $q \geq 1$

$$\int_{U_n^c} |\xi_i(x)| \Pi(dx) \leq (n/2)^{1-1/q} \left(\int_G |\xi_i(x)|^q \Pi(dx) \right)^{1/q}. \quad (5.1.8)$$

For every $n \geq 1$, let $Y_n = Y_n^1 u_1 + \dots Y_n^m u_m$ be a $\mathfrak{g}$ -valued random variable such that for all $k \in \tilde{K}$

$$b_n^k := \mathbb{E}[Y_n^k] = (1 - w_n/n)^{-1} n^{-1} \tilde{B}^k,$$

and for all $k \notin \tilde{K}$

$$b_n^k := \mathbb{E}[Y_n^k] = (1 - w_n/n)^{-1} n^{-1} \left(B_k - \int_{U_n^c} \xi_k(x) \Pi(dx) \right),$$

and with covariances for all $i, j \in \{1, \dots, m\}$

$$\mathbb{E}[(Y_n^i - b_n^i)(Y_n^j - b_n^j)] = (1 - w_n/n)^{-1} n^{-1} A_{i,j}.$$

Remark that setting $q = 2$ in (5.1.8) implies

$$\lim_{n \rightarrow \infty} n^{-1} \int_{U_n^c} |\xi_i(x)| \Pi(dx) = 0,$$

from which it follows that $\sup_{n \geq 1} n|b_n^i| < \infty$. Moreover, it holds that

$$\lim_{n \rightarrow \infty} \mathbb{E}[(Y_n^i - b_n^i)(Y_n^j - b_n^j)] = 0.$$

It follows that we can choose Y_n such that $\exp(Y_n)$ has support in a neighbourhood V_n of 1_G , such that V_n shrinks to 1_G as $n \rightarrow \infty$. Denote by ν_n the probability measure of the G -valued random variable $\exp(Y_n)$.

Finally, let X_{n1} be the G -valued random variable associated to the probability measure $(w_n/n)\mu_n + (1 - w_n/n)\nu_n$, and let $X_{n2}, \dots, X_{nn}$ be independent copies of X_{n1} .

Lemma 5.1.13. *Let $\mathbf{X}^n$ be the random walk associated with X_{nj} . Then $\mathbf{X}^n \xrightarrow{\mathcal{D}} \mathbf{X}$ as $D_o([0, 1], G)$ -valued random variables.*

Proof. One simply needs to verify the three conditions of Theorem 5.1.1.

(1) Let $f \in C_b(G)$ be identically zero on a neighbourhood of 1_G . Then f is identically zero on U_n and V_n for all n sufficiently large, and thus

$$\lim_{n \rightarrow \infty} n\mathbb{E}[f(X_{n1})] = \lim_{n \rightarrow \infty} \int_{U_n^c} f(x) \Pi(dx) = \int_G f(x) \Pi(dx).$$

(2) Directly from the definition of μ_n and ν_n , one has for all $k \in \tilde{K}$

$$\begin{aligned} \lim_{n \rightarrow \infty} n\mathbb{E}[\xi_k(X_{n1})] &= \lim_{n \rightarrow \infty} n(w_n/n)\mu_n(\xi_k) + n(1 - w_n/n)\mathbb{E}[Y_n^k] \\ &= \lim_{n \rightarrow \infty} \int_{U_n^c} \xi_k(x) \Pi(dx) + B_k - \int_G \xi_k(x) \Pi(dx) \\ &= B_k. \end{aligned}$$

For $k \notin \tilde{K}$, one replaces $\int_G \xi_k(x) \Pi(dx)$ by $\int_{U_n^c} \xi_k(x) \Pi(dx)$ in the second line and obtains the same conclusion.

(3) Likewise, for all $i, j \in \{1, \dots, m\}$

$$\begin{aligned} \lim_{n \rightarrow \infty} n \mathbb{E} [\xi_i(X_{n1}) \xi_j(X_{n1})] &= \lim_{n \rightarrow \infty} w_n \mu_n(\xi_i \xi_j) + n(1 - w_n/n) \mathbb{E} [Y_n^i Y_n^j] \\ &= \lim_{n \rightarrow \infty} \int_{U_n^c} \xi_i(x) \xi_j(x) \Pi(dx) + A_{i,j} + n(1 - w_n/n) b_n^i b_n^j \\ &= A_{i,j} + \int_G \xi_i(x) \xi_j(x) \Pi(dx), \end{aligned}$$

where the last equality follows since $\sup_{n \geq 1} n |b_n^i| < \infty$ for all $i \in \{1, \dots, m\}$. $\square$

Lemma 5.1.14. *Let $0 < q_1, \dots, q_m \leq 2$ be real numbers such that $q_i \notin \Gamma_i$ for all $i \in \{1, \dots, m\}$, $q_i = 2$ for all $i \in J$, and $q_i \geq 1$ for all $i \in K$. Let θ be a scaling function such that $\theta(x) = \sum_{i=1}^m |\xi_i(x)|^{q_i}$ for x in a neighbourhood of 1_G . Then θ scales the array $X_{n1}, \dots, X_{nn}$.*

Proof. By the convergence $\mathbf{X}^n \xrightarrow{\mathcal{D}} \mathbf{X}$ from Lemma 5.1.13, we have

$$\sup_n n \mathbb{E} \left[\sum_{j \in J} |\xi_j(X_{n1})|^2 \right] \leq \sup_n n \mathbb{E} \left[\sum_{i=1}^m |\xi_i(X_{n1})|^2 \right] < \infty.$$

It remains to consider the terms $|\xi_i(y)|^{q_i}$ for $i \notin J$. For every $i \notin J$, observe that Y_n^i is a.s. constant with value b_n^i . If $i \notin J$ and $i \in K$, or if $i \notin J$ and $i \notin \tilde{K}$, then $1 \leq q_i \leq 2$. It follows that (5.1.8) with $q := q_i$ implies that $\sup_{n \geq 1} n |b_n^i|^{q_i} < \infty$, and thus

$$\begin{aligned} \sup_{n \geq 1} n \mathbb{E} [|\xi_i(X_{n1})|^{q_i}] &= \sup_{n \geq 1} \int_{U_n^c} |\xi_i(x)|^{q_i} \Pi(dx) + n(1 - w_n/n) \mathbb{E} [|Y_n^i|^{q_i}] \\ &\leq \int_G |\xi_i(x)|^{q_i} \Pi(dx) + \sup_n n(1 - w_n/n) |b_n^i|^{q_i} \\ &< \infty. \end{aligned}$$

Finally, if $i \notin J$ and $i \in \tilde{K} \setminus K$, then Y_n^i is a.s. zero. Thus

$$\sup_{n \geq 1} n \mathbb{E} [|\xi_i(X_{n1})|^{q_i}] = \sup_{n \geq 1} \int_{U_n^c} |\xi_i(x)|^{q_i} \Pi(dx) = \int_G |\xi_i(x)|^{q_i} \Pi(dx) < \infty,$$

which completes the proof. $\square$

The following lemma will be particularly useful in the proof of the Lévy-Khintchine formula in Theorem 5.4.5.

Lemma 5.1.15. *Let θ be a scaling function on G which scales X_{nj} . Let V be a neighbourhood of 1_N , denote $W = \text{supp}(\Pi) \cup V$, and let $f : W \mapsto \mathbb{R}$ be a bounded measurable function such that f is continuous on $\text{supp}(\Pi)$ and admits a Taylor expansion (5.1.2) on W of order θ .*

Then for all n sufficiently large, $X_{n1} \in W$ a.s. and

$$\lim_{n \rightarrow \infty} n \mathbb{E} [f(X_{n1}) - f(1_G)] = L^f,$$

where L^f does not depend on the Taylor expansion of f .

Proof. The claim that $X_{n1} \in W$ a.s. for all n sufficiently large follows immediately from the fact that the support of $\exp(Y_n)$ is contained in V_n which shrinks to 1_G as $n \rightarrow \infty$.

Recall that $h : W \mapsto \mathbb{R}$ denotes the error in the Taylor expansion of f . Since h/θ is continuous on $\text{supp}(\Pi) \setminus \{1_G\}$ and $\lim_{x \rightarrow 1_G} h(x)/\theta(x) = 0$, by the Tietze extension theorem there exists $\tilde{h} \in C_b(G)$ such that $\tilde{h} \equiv h$ on $\text{supp}(\Pi)$ and $\tilde{h} = o(\theta)$. Define $\tilde{f} \in C_b(G)$ by

$$\tilde{f} = f(1_G) + \sum_{i=1}^m b_i \xi_i + \sum_{i,j=1}^m a_{i,j} \xi_i \xi_j + \tilde{h}.$$

Then by Proposition 5.1.8,

$$\lim_{n \rightarrow \infty} n \mathbb{E} [\tilde{f}(X_{n1}) - \tilde{f}(1_G)] = L^{\tilde{f}}.$$

Note that

$$|\tilde{f}(x) - f(x)| = \mathbf{1}\{x \notin \text{supp}(\Pi)\} |\tilde{h}(x) - h(x)|,$$

and that

$$c_n := \sup_{x \in V_n, x \neq 1_G} \frac{|h(x)| + |\tilde{h}(x)|}{\theta(x)} \xrightarrow{n \rightarrow \infty} 0.$$

Since $X_{n1} \in \text{supp}(\Pi) \cup V_n$ for all $n \geq 1$, it follows that

$$\begin{aligned} \lim_{n \rightarrow \infty} n \left| \mathbb{E} [\tilde{f}(X_{n1}) - f(X_{n1})] \right| &\leq \lim_{n \rightarrow \infty} n \mathbb{E} [\mathbf{1}\{X_{n1} \notin \text{supp}(\Pi)\} |\tilde{h}(x) - h(x)|] \\ &\leq \lim_{n \rightarrow \infty} n \mathbb{E} [\theta(X_{n1})] c_n \\ &= 0. \end{aligned}$$

Finally, observe that $L^{\tilde{f}} = L^f$, from which the conclusion follows. □

5.1.8 Infinite p -variation

The purpose of this section is to establish conditions under which sample paths of a Lévy process have *infinite* p -variation, which will be used to complement our results concerning finiteness of p -variation for Lévy processes in the Lie group $G^N(\mathbb{R}^d)$ (see Theorem 5.4.1). The methods we use in this section are classical, however we provide complete arguments for all the results.

Throughout this section, let $\mathbf{X}$ be a Lévy process in G with triplet (A, B, Π) . We study first the scaling limit of the process $\log(\mathbf{X}_t)$ restricted to a neighbourhood of zero in $\mathfrak{g}$. Recall that U is an open neighbourhood of 1_G with compact closure on which $\log$ is a diffeomorphism. Denote by $\mathbf{X}^x$ the Lévy process $\mathbf{X}$ started at $x \in G$ (which is equal in law to $x\mathbf{X}^{1_G}$). Let $\tau = \inf\{t \geq 0 \mid \mathbf{X}_t \notin U\}$ be the first exit time of $\mathbf{X}$ from U .

Fix a distinguished point $y_0 \in \mathfrak{g} \setminus \log(U)$ and consider the Markov process $\mathbf{Y}$ in $\mathfrak{g}$ for which every point outside $\log(U)$ is absorbing, and such that for all $y \in \log(U)$,

$$\mathbf{Y}_t^y = \begin{cases} \log(\mathbf{X}_t^{\exp(y)}) & \text{if } t \in [0, \tau) \\ \mathbf{Y}_t^y = y_0 & \text{if } t \geq \tau. \end{cases}$$

For $\varepsilon > 0$, define the Markov process $\mathbf{Y}_t^\varepsilon := \varepsilon^{-1/2}(\mathbf{Y}_{\varepsilon t})$ for $t \in [0, 1]$. Let $\mathbf{B}$ be a centred Brownian motion on $\mathfrak{g}$ with covariance matrix $(A_{i,j})_{i,j=1}^m$ with respect to the basis $u_1, \dots, u_m$ of $\mathfrak{g}$.

Equip $C_0(\mathfrak{g})$ with the uniform topology. Consider the generators L^ε and $L_{\mathbf{B}}$ of the semigroups on $C_0(\mathfrak{g})$ associated to $\mathbf{Y}^\varepsilon$ and $\mathbf{B}$ respectively.

Lemma 5.1.16. *Let $f \in C_c^\infty(\mathfrak{g})$. Then for all $\varepsilon > 0$ sufficiently small, f is in the domain of L^ε , and $L^\varepsilon f \rightarrow L_{\mathbf{B}} f$ in $C_0(\mathfrak{g})$ as $\varepsilon \rightarrow 0$.*

Proof. Denote by ∂_i the differential operator associated with vector $u_i \in \mathfrak{g}$ (under the Euclidean structure of $\mathfrak{g}$). Note that

$$L_{\mathbf{B}} = \sum_{i,j=1}^m A_{i,j} \partial_{i,j}.$$

For every $f \in C_c^\infty(\log(U))$ define $\tilde{f} \in C_c^\infty(G)$ by $\tilde{f} \equiv f \circ \log$ on U and $\tilde{f} \equiv 0$ on $G \setminus U$.

We first claim that every $f \in C_c^\infty(\log(U))$ is in the domain of $L_{\mathbf{Y}}$, the generator of $\mathbf{Y}$, and that $L_{\mathbf{Y}} f(y) = L_{\mathbf{X}} \tilde{f}(x)$ for all $y \in \log(U)$ and $x := \exp(y)$. Indeed,

since U^c is closed and $\mathbf{X}$ is càdlàg, we have $\mathbf{X}_\tau \notin U$, from which it follows that $\mathbb{E}^y[f(\mathbf{Y}_t)] = \mathbb{E}^x[\tilde{f}(\mathbf{X}_{t \wedge \tau})]$. In particular,

$$\begin{aligned} \lim_{t \rightarrow 0} t^{-1} \mathbb{E}^y[f(\mathbf{Y}_t) - f(y)] &= \lim_{t \rightarrow 0} t^{-1} \mathbb{E}^x[\tilde{f}(\mathbf{X}_{t \wedge \tau}) - \tilde{f}(x)] \\ &= \lim_{t \rightarrow 0} t^{-1} \mathbb{E}^x \left[\int_0^t \mathbf{1}\{s < \tau\} (L_{\mathbf{X}} \tilde{f})(\mathbf{X}_s) ds \right]. \end{aligned}$$

Remark that by dominated convergence $\lim_{s \rightarrow 0} \mathbb{E}^x[\mathbf{1}\{s < \tau\} (L_{\mathbf{X}} \tilde{f})(\mathbf{X}_s)] = (L_{\mathbf{X}} \tilde{f})(x)$ for all $x \in G$, so the claim follows by Fubini's theorem.

For $f : \mathfrak{g} \mapsto \mathbb{R}$ and $\varepsilon > 0$ define $f^\varepsilon(y) := f(\varepsilon^{-1/2}y)$ for all $y \in \mathfrak{g}$. Observe that for all $y \in \mathfrak{g}$ and all $f : \mathfrak{g} \mapsto \mathbb{R}$ such that f^ε is in the domain of $L_{\mathbf{Y}}$, it holds that

$$\begin{aligned} \lim_{t \rightarrow 0} t^{-1} \mathbb{E}^y[f(\mathbf{Y}_t^\varepsilon) - f(y)] &= \varepsilon \lim_{t \rightarrow 0} (\varepsilon t)^{-1} \mathbb{E}^{\varepsilon^{1/2}y}[f^\varepsilon(\mathbf{Y}_{\varepsilon t}) - f^\varepsilon(\varepsilon^{1/2}y)] \\ &= \varepsilon (L_{\mathbf{Y}} f^\varepsilon)(\varepsilon^{1/2}y), \end{aligned}$$

so that f is in the domain of L^ε and $(L^\varepsilon f)(y) = \varepsilon (L_{\mathbf{Y}} f^\varepsilon)(\varepsilon^{1/2}y)$.

Remark moreover that if $f^\varepsilon \in C_c^\infty(\log(U))$, then for all $y \notin \log(U)$, $L_{\mathbf{Y}} f^\varepsilon(y) = 0$. In particular, we see that $L^\varepsilon f \in C_c^\infty(\mathfrak{g})$.

Recall that for all $f \in C_c^\infty(G)$ and $x \in G$,

$$\begin{aligned} L_{\mathbf{X}} f(x) &= \sum_{i=1}^m B_i(u_i f)(x) + \sum_{i,j=1}^m A_{j,i}(u_j u_i f)(x) \\ &\quad + \int_G \left[f(xy) - f(x) - \sum_{i=1}^m \xi_i(y)(u_i f)(x) \right] \Pi(dy). \end{aligned}$$

Observe that there exist smooth functions $c_p^i \in C^\infty(\log(U))$, $i, p \in \{1, \dots, m\}$, such that for all $f \in C_c^\infty(\log(U))$ and $y \in \log(U)$

$$(u_i \tilde{f})(e^y) = \sum_{p=1}^m c_p^i(y) (\partial_p f)(y), \quad (5.1.9)$$

for which it holds that $c_p^i(0) = \delta_{i,p}$. We note also that

$$\begin{aligned} (u_j u_i \tilde{f})(e^y) &= \sum_{q=1}^m c_q^j(y) \left[\partial_q \left(\sum_{p=1}^m c_p^i \partial_p f \right) \right](y) \\ &= \sum_{p,q=1}^m c_q^j(y) c_p^i(y) \partial_{q,p} f(y) + c_q^j(y) \partial_q c_p^i(y) \partial_p f(y). \end{aligned} \quad (5.1.10)$$

It follows that there exist smooth functions $C, D \in C^\infty(\log(U))$ such that for all $f \in C_c^\infty(\log(U))$ and $y \in \log(U)$

$$\sum_{i=1}^m B_i(u_i \tilde{f})(e^y) + \sum_{i,j=1}^m A_{j,i}(u_j u_i \tilde{f})(e^y) = \sum_{p=1}^m C_p(y)(\partial_p f)(y) + \sum_{p,q=1}^m D_{q,p}(y)(\partial_{q,p} f)(y),$$

for which it holds that $D_{q,p}(0) = A_{q,p}$.

Let us now fix $f \in C_c^\infty(\mathfrak{g})$. Observe that for all $y \in \mathfrak{g}$,

$$\partial_p(f^\varepsilon)(y) = \varepsilon^{-1/2}(\partial_p f)(\varepsilon^{-1/2}y) = \varepsilon^{-1/2}(\partial_p f)^\varepsilon(y), \quad (5.1.11)$$

and therefore

$$\partial_{q,p}(f^\varepsilon)(y) = \varepsilon^{-1/2} \partial_q((\partial_p f)^\varepsilon)(y) = \varepsilon^{-1}(\partial_{q,p} f)^\varepsilon(y). \quad (5.1.12)$$

Consider $\varepsilon > 0$ sufficiently small so that $f^\varepsilon \in C_c^\infty(\log(U))$. It follows that f is in the domain of L^ε and

$$\begin{aligned} (L^\varepsilon f)(y) &= \varepsilon^{1/2} \sum_{p=1}^m C_p(\varepsilon^{1/2}y) \partial_p f(y) + \sum_{p,q=1}^m D_{q,p}(\varepsilon^{1/2}y) \partial_{q,p} f(y) \\ &\quad + \int_G \varepsilon \left[\tilde{f}^\varepsilon(\exp(\varepsilon^{1/2}y)x) - \tilde{f}^\varepsilon(\exp(\varepsilon^{1/2}y)) - \sum_{i=1}^m \xi_i(x)(u_i \tilde{f}^\varepsilon)(\exp(\varepsilon^{1/2}y)) \right] \Pi(dx). \end{aligned} \quad (5.1.13)$$

Since $D_{q,p}(0) = A_{q,p}$, it suffices to show that the integral on the right side converges to zero as $\varepsilon \rightarrow 0$. Indeed, observe that, due to (5.1.9) and (5.1.11), there exists $C_1 > 0$ such that for all ε sufficiently small, $|u_i \tilde{f}^\varepsilon|_\infty \leq C_1 \varepsilon^{-1/2}$, and thus there exists $C_2 > 0$ such that absolute value of the integrand on the right side is bounded above by $C_2 \varepsilon^{1/2}$ for all $x \in G$. In particular the integrand converges to zero for all $x \in G$ as $\varepsilon \rightarrow 0$. To conclude by dominated convergence, it suffices to show that there exists $C_3 > 0$ such that for all x in a neighbourhood of 1_G and all ε sufficiently small, the absolute value of the integrand is bounded above by $C_3 |\xi(x)|^2$.

To this end, let $h \in C^\infty(G)$, $z \in G$, and define $\psi(t) := h(ze^{tY})$ for some $Y = Y_1 u_1 + \dots + Y_m u_m \in \mathfrak{g}$ with $|Y| = 1$. By Taylor's theorem,

$$|\psi(t) - \psi(0) - t\psi'(0)| \leq \frac{1}{2} \|\psi''\|_{\infty;[0,t]} t^2.$$

Furthermore, note that $\psi'(0) = Yh(z)$ and $\psi''(s) = (YYh)(ze^{sY})$ ([36], Proposition 1.88). Setting $h := \tilde{f}^\varepsilon$ and $z := \exp(\varepsilon^{1/2}y)$, we obtain

$$\begin{aligned} \left| \tilde{f}^\varepsilon(\exp(\varepsilon^{1/2}y) \exp(tY)) - \tilde{f}^\varepsilon(\exp(\varepsilon^{1/2}y)) - \sum_{i=1}^m tY_i(u_i \tilde{f}^\varepsilon)(\exp(\varepsilon^{1/2}y)) \right| \\ \leq \frac{1}{2} \|YY\tilde{f}^\varepsilon\|_\infty t^2. \end{aligned}$$

Lastly, due to (5.1.10), (5.1.11), and (5.1.12), there exists $C_4 > 0$ such that for all ε sufficiently small,

$$\left\| u_j u_i \tilde{f}^\varepsilon \right\|_\infty \leq C_4 \varepsilon^{-1}.$$

Since $\xi \equiv \log$ on U , the desired bound on the integrand follows. $\square$

For the remainder of the section, we shall assume that the starting point at $t = 0$ of every process in $\mathfrak{g}$ and G is 0 and 1_G respectively.

Corollary 5.1.17. *As $\varepsilon \rightarrow 0$, it holds that $\mathbf{Y}^\varepsilon \xrightarrow{\mathcal{D}} \mathbf{B}$ as $D_o([0, 1], \mathfrak{g})$ -valued random variables.*

Proof. Since $C_c^\infty(\mathfrak{g})$ is a core for $L_{\mathbf{B}}$ ([34] Proposition 17.9), the statement follows from the convergence $L^\varepsilon f \rightarrow L_{\mathbf{B}} f$ in $C_0(\mathfrak{g})$ for all $f \in C_c^\infty(\mathfrak{g})$ ([34] Theorem 17.25). $\square$

Proposition 5.1.18. *Suppose $A_{i,i} > 0$ for some $i \in \{1, \dots, m\}$. Then*

$$\mathbb{P} \left[\limsup_{t \rightarrow 0} t^{-1/2} |\xi_i(\mathbf{X}_t)| = \infty \right] = 1.$$

Proof. Let $c > 0$. Corollary 5.1.17 implies that there exist $\delta, \varepsilon_0 > 0$ such that for all $0 < \varepsilon \leq \varepsilon_0$

$$\mathbb{P} [\varepsilon^{-1/2} |\xi_i(\mathbf{X}_\varepsilon)| > c] > \delta.$$

Observe that, by the Campbell-Baker-Hausdorff formula, there exist a neighbourhood V of 1_G and a constant $C > 0$ such that for all $x, y \in V$

$$|\xi_i(x^{-1}y)| \leq |\xi_i(y) - \xi_i(x)| + C|\xi(x)|^2 + C|\xi(y)|^2. \quad (5.1.14)$$

For all integers $k \geq 1$, let $\varepsilon_k := 3^{-k} \varepsilon_0$. Corollary 5.1.17 implies that for all $\lambda > 0$

$$\lim_{k \rightarrow \infty} \mathbb{P} \left[\varepsilon_k^{-1/2} (|\xi(\mathbf{X}_{\varepsilon_k})|^2 + |\xi(\mathbf{X}_{\varepsilon_{k-1}})|^2) > \lambda \right] = 0.$$

Hence, by the first Borel–Cantelli lemma, there exists a strictly increasing sequence $(k(n))_{n \geq 1}$ such that

$$\mathbb{P} \left[\limsup_{n \rightarrow \infty} \varepsilon_{k(n)}^{-1/2} C \left(|\xi(\mathbf{X}_{\varepsilon_{k(n)}})|^2 + |\xi(\mathbf{X}_{\varepsilon_{k(n)-1}})|^2 \right) > c/2 \right] = 0. \quad (5.1.15)$$

Furthermore, since $\mathbf{X}_{\varepsilon_k, \varepsilon_{k-1}} = \mathbf{X}_{\varepsilon_k}^{-1} \mathbf{X}_{\varepsilon_{k-1}} \stackrel{\mathcal{D}}{=} \mathbf{X}_{\varepsilon_{k-1} - \varepsilon_k}$, we have for all $k \geq 1$

$$\mathbb{P} [(\varepsilon_{k-1} - \varepsilon_k)^{-1/2} |\xi_i(\mathbf{X}_{\varepsilon_k, \varepsilon_{k-1}})| > c] > \delta.$$

Note that $(\varepsilon_{k-1} - \varepsilon_k)^{-1/2} < \varepsilon_k^{-1/2}$, and that the random variables $(\mathbf{X}_{\varepsilon_k, \varepsilon_{k-1}})_{k \geq 1}$ are independent. Hence, by the second Borel–Cantelli lemma,

$$\mathbb{P} \left[\limsup_{n \rightarrow \infty} \varepsilon_{k(n)}^{-1/2} |\xi_i(\mathbf{X}_{\varepsilon_{k(n)}, \varepsilon_{k(n)-1}})| > c \right] = 1.$$

Since $\lim_{t \rightarrow 0} \mathbf{X}_t = 1_G$ a.s., it now follows from (5.1.14) and (5.1.15) that

$$\mathbb{P} \left[\limsup_{n \rightarrow \infty} \varepsilon_{k(n)}^{-1/2} |\xi_i(\mathbf{X}_{\varepsilon_{k(n)}}) - \xi_i(\mathbf{X}_{\varepsilon_{k(n)-1}})| > c/2 \right] = 1.$$

Observe that $\varepsilon_{k(n)}^{-1/2} |\xi_i(\mathbf{X}_{\varepsilon_{k(n)}}) - \xi_i(\mathbf{X}_{\varepsilon_{k(n)-1}})| > c/2$ implies that either

$$\varepsilon_{k(n)}^{-1/2} |\xi_i(\mathbf{X}_{\varepsilon_{k(n)}})| > c/4$$

or

$$\varepsilon_{k(n)}^{-1/2} |\xi_i(\mathbf{X}_{\varepsilon_{k(n)-1}})| = \sqrt{3} \varepsilon_{k(n)-1}^{-1/2} |\xi_i(\mathbf{X}_{\varepsilon_{k(n)-1}})| > c/4.$$

We thus obtain

$$\mathbb{P} \left[\limsup_{k \rightarrow \infty} \varepsilon_k^{-1/2} |\xi_i(\mathbf{X}_{\varepsilon_k})| > \frac{c}{4\sqrt{3}} \right] = 1.$$

As $c > 0$ was arbitrary, the conclusion follows. $\square$

Corollary 5.1.19. *Suppose $A_{i,i} > 0$ for some $i \in \{1, \dots, m\}$. Then*

$$\mathbb{P} \left[\sup_{\mathcal{D} \subset [0,1]} \sum_{t_k \in \mathcal{D}} |\xi_i(\mathbf{X}_{t_k, t_{k+1}})|^2 = \infty \right] = 1.$$

Proof. By Proposition 5.1.18, $\limsup_{t \rightarrow 0} t^{-1} |\xi_i(\mathbf{X}_t)|^2 = \infty$ a.s. whenever $A_{i,i} > 0$. Since $\mathbf{X}$ has stationary and independent increments, the conclusion follows from an application of the Vitali covering argument (see [6] Proposition on p.68, or [25] Theorem 13.69). $\square$

The following is a form of the classical Blumenthal–Gettoor index adapted to the setting of Lie groups. Recall the definitions of Γ_i and K from Section 5.1.7.

Proposition 5.1.20 (Blumenthal–Gettoor index). *Let $i \in \{1, \dots, m\}$ and $q > 0$. Then*

$$\mathbb{P} \left[\sup_{\mathcal{D} \subset [0,1]} \sum_{t_k \in \mathcal{D}} |\xi_i(\mathbf{X}_{t_k, t_{k+1}})|^q = \infty \right] = 1 \tag{5.1.16}$$

if either

(i) $q \in \Gamma_i$, or

(ii) $i \in K$ and $q < 1$.

Proof. Define $f \in C_c(G)$ by $f(x) = 1 - \exp(-|\xi_i(x)|^q)$. We claim that (5.1.16) holds whenever

$$\lim_{t \rightarrow 0} t^{-1} \mathbb{E}[f(\mathbf{X}_t)] = \infty. \quad (5.1.17)$$

Indeed, observe that (5.1.17) implies that $\lim_{n \rightarrow \infty} \mathbb{E}[1 - f(\mathbf{X}_{1/n})]^n = 0$. But since $\mathbf{X}$ has independent and stationary increments,

$$\mathbb{E} \left[\exp \left(- \sum_{k=1}^n |\xi_i(\mathbf{X}_{\frac{k-1}{n}, \frac{k}{n}})|^q \right) \right] = \mathbb{E} \left[\exp \left(|\xi_i(\mathbf{X}_{0, \frac{1}{n}})|^q \right) \right]^n = \mathbb{E} [1 - f(\mathbf{X}_{1/n})]^n.$$

Since for all $n \geq 1$,

$$\sup_{\mathcal{D} \subset [0,1]} \sum_{t_k \in \mathcal{D}} |\xi_i(\mathbf{X}_{t_k, t_{k+1}})|^q \geq \sum_{k=1}^n |\xi_i(\mathbf{X}_{\frac{k-1}{n}, \frac{k}{n}})|^q,$$

we obtain that

$$\mathbb{E} \left[\exp \left(- \sup_{\mathcal{D} \subset [0,1]} \sum_{t_k \in \mathcal{D}} |\xi_i(\mathbf{X}_{t_k, t_{k+1}})|^q \right) \right] = 0,$$

from which (5.1.16) follows.

It thus suffices to show (5.1.17) in either of the two cases.

(i) Let $(\psi_n)_{n \geq 1}$ be a non-decreasing sequence of non-negative functions in $C^\infty(\mathbb{R})$ such that $\lim_{n \rightarrow \infty} \psi_n(x) = |x|^q$ for all $x \in \mathbb{R}$ and such that $\psi_n \equiv 0$ on a neighbourhood V_n of zero.

For every integer $n \geq 1$, define $f_n \in C_c^\infty(G)$ by $f_n(x) = 1 - \exp(-\psi_n(\xi_i(x)))$. Note that there exists $c > 0$ such that $f_n(x) \geq c\psi_n(\xi_i(x))$ for all $x \in G$ and $n \geq 1$. Furthermore, f_n is identically zero on a neighbourhood of 1_G . It follows that for all $n \geq 1$

$$\begin{aligned} \lim_{t \rightarrow 0} t^{-1} \mathbb{E}[f_n(\mathbf{X}_t)] &= \int_G f_n(x) \Pi(dx) \\ &\geq c \int_G \psi_n(\xi_i(x)) \Pi(dx). \end{aligned}$$

Observe that $0 \leq f_n \leq f$ for all $n \geq 1$, from which it follows that

$$\liminf_{t \rightarrow 0} t^{-1} \mathbb{E}[f(\mathbf{X}_t)] \geq \sup_{n \geq 1} c \int_G \psi_n(\xi_i(x)) \Pi(dx).$$

But since $q \in \Gamma_i$, we obtain by the monotone convergence theorem that

$$\lim_{n \rightarrow \infty} \int_G \psi_n(\xi_i(x)) \Pi(dx) = \infty,$$

from which (5.1.17) follows.

(ii) Note that if $A_{i,i} > 0$ or if $q \in \Gamma_i$, then the desired result follows by Corollary 5.1.19 or by case (i) respectively. Thus suppose $A_{i,i} = 0$ and $q \notin \Gamma_i$.

Since $q < 1$, for every integer $n \geq 1$ we can find $\psi_n \in C^\infty(\mathbb{R})$ such that $|\psi_n(x)| \leq |x|^q$ for all $x \in \mathbb{R}$ and such that $\psi_n(x) = n\tilde{B}_i^{-1}x$ for all x on a neighbourhood V_n of zero.

For every integer $n \geq 1$, define $f_n \in C_c^\infty(G)$ by $f_n(x) = 1 - \exp(-\psi_n(\xi_i(x)))$. Observe that

$$f_n = n\tilde{B}_i^{-1}\xi_i - \frac{1}{2}n^2\tilde{B}_i^{-2}\xi_i^2 + O(\xi_i^3).$$

Since $A_{i,i} = 0$, it follows that for all $n \geq 1$

$$\begin{aligned} \lim_{t \rightarrow 0} t^{-1} \mathbb{E}[f_n(\mathbf{X}_t)] &= n\tilde{B}_i^{-1}B_i + \int_G \left[f_n(x) - n\tilde{B}_i^{-1}\xi_i(x) \right] \Pi(dx) \\ &= n \left(1 + n^{-1} \int_G f_n(x) \Pi(dx) \right). \end{aligned} \tag{5.1.18}$$

Note that there exists $C > 0$ such that $|f_n(x)| \leq C|\psi_n(\xi_i(x))|$ for all $x \in G$ and $n \geq 1$. In particular, since $q \notin \Gamma_i$,

$$\lim_{n \rightarrow \infty} n^{-1} \int_G |f_n(x)| \Pi(dx) \leq C \lim_{n \rightarrow \infty} n^{-1} \int_G |\xi_i(x)|^q \Pi(dx) = 0.$$

Hence, by (5.1.18),

$$\lim_{n \rightarrow \infty} \lim_{t \rightarrow 0} t^{-1} \mathbb{E}[f_n(\mathbf{X}_t)] = \infty.$$

Observe now that $f_n \leq f$ for all $n \geq 1$, from which (5.1.17) follows. $\square$

5.2 Random walks in path space

Unless otherwise stated, throughout this section we fix the time interval $[0, T]$ to be $[0, 1]$. We adopt the shorthand notation $G^N := G^N(\mathbb{R}^d)$ for $N \geq 1$, and let unspecified path spaces be defined on $[0, 1]$ and take values in G^N , for example, $C_o^{p\text{-var}}$ shall denote $C_o^{p\text{-var}}([0, 1], G^N(\mathbb{R}^d))$.

For a collection $(\mathbf{x}_j)_{j=1}^n$ of paths $\mathbf{x}_j : [0, 1] \mapsto G^N$ such that $\mathbf{x}_j(0) = 1_N$, recall their concatenation $\mathbf{x} := \mathbf{x}_1 * \dots * \mathbf{x}_n : [0, 1] \mapsto G^N$ given by

$$\mathbf{x}(t) = \begin{cases} 1_N & \text{if } t = 0 \\ \mathbf{x}(j/n)\mathbf{x}_j(n(t - j/n)) & \text{if } t \in (j/n, (j+1)/n], j \in \{0, \dots, n-1\}. \end{cases}$$

Consider an iid array $\mathbf{X}_{nj}$ of $WG\Omega_p(\mathbb{R}^d)$ -valued (or equivalently $C_o^{p\text{-var}}([0, 1], G^{[p]})$ -valued) random variables. In this section, we are interested in the question of when

the sequence of concatenations $\mathbf{X}_n := \mathbf{X}_{n1} * \dots * \mathbf{X}_{nn}$ satisfies the conditions of Corollary 3.4.8, namely, when is $(\|\mathbf{X}_n\|_{p\text{-var};[0,1]})_{n \geq 1}$ a tight collection of real random variables, and when does the following limit exist for all $M \in \mathbf{L}(\mathbb{R}^d, \mathbf{u})$

$$\lim_{n \rightarrow \infty} \mathbb{E}[M(\mathbf{X}_n)]. \quad (5.2.1)$$

In the upcoming Sections 5.2.1 and 5.2.2 we shall give general conditions, applicable to any iid array $\mathbf{X}_{nj}$, which ensure that $(\|\mathbf{X}_n\|_{p\text{-var};[0,1]})_{n \geq 1}$ is tight (see Proposition 5.2.1 and Theorem 5.2.3 below).

Note that the iid property of $\mathbf{X}_{nj}$ implies that $\mathbb{E}[M(\mathbf{X}_n)] = \mathbb{E}[M(\mathbf{X}_{n1})]^n$ for all $M \in \mathbf{L}(\mathbb{R}^d, \mathbf{u})$, so the existence and value the limit (5.2.1) depends entirely on the behaviour of the (matrix-valued) quantity $\mathbb{E}[M(\mathbf{X}_{n1})]$ as $n \rightarrow \infty$. In Section 5.2.3 we shall provide an example of a class of iid arrays $\mathbf{X}_{nj}$ for which (5.2.1) is guaranteed to exist. In Section 5.4.3 we shall moreover investigate in detail the case when the rough path $\mathbf{X}_{nj}$ depends entirely on its endpoint through a suitable path function (which will be introduced in Section 5.3). In this case, we will show that reparametrisations of $\mathbf{X}_n$ converge in law as rough paths, which a fortiori implies that the limit (5.2.1) exists. Beyond this, however, we shall not investigate general conditions under which (5.2.1) is guaranteed to exist.

5.2.1 Tightness of concatenations

For a collection of paths $(\mathbf{x}_j)_{j=1}^n$ in $C_o^{p\text{-var}}$, it holds that the p -variation of their concatenation $\mathbf{x} := \mathbf{x}_1 * \dots * \mathbf{x}_n \in C_o^{p\text{-var}}$ is bounded below by

$$\max \left(\|\mathbf{x}'\|_{p\text{-var};[0,1]}^p, \sum_{j=1}^n \|\mathbf{x}_j\|_{p\text{-var};[0,1]}^p \right) \leq \|\mathbf{x}\|_{p\text{-var};[0,1]}^p, \quad (5.2.2)$$

where $\mathbf{x}' \in D_o$ is the walk associated with the elements $x_j := \mathbf{x}_j(1)$ in G^N . The main result of this section is Proposition 5.2.1 which establishes a partial converse to (5.2.2).

Proposition 5.2.1. *Let $p \geq 1$ and let $(\mathbf{x}_j)_{j=1}^n$ be a collection of paths in $C_o^{p\text{-var}}$. Let $\mathbf{x}' \in D_o$ be the walk associated with the elements $x_j := \mathbf{x}_j(1)$, and let $\mathbf{x} := \mathbf{x}_1 * \dots * \mathbf{x}_n \in C_o^{p\text{-var}}$. Then*

$$\|\mathbf{x}\|_{p\text{-var};[0,1]}^p \leq C \left(\|\mathbf{x}'\|_{p\text{-var};[0,1]}^p + \sum_{j=1}^n \|\mathbf{x}_j\|_{p\text{-var};[0,1]}^p \right),$$

where $C = 1 + 2^p + 2 \cdot 3^{p-1}$.

The proof of Proposition 5.2.1 will follow immediately from the following lemma, which we shall use again in the sequel. We mention that the idea for the proof of the following result was taken in part from Lemma 25 of [24].

Lemma 5.2.2. *Let (E, d) be a metric space and $\mathbf{x} : [0, T] \mapsto E$ a function (not necessarily càdlàg). Let $(I_n)_{n \geq 1}$ be a countable collection of disjoint open subintervals of $[0, T]$ and set $I := \cup_{n \geq 1} I_n$. Define $\mathbf{y} : [0, T] \mapsto E$ by $\mathbf{y}_t = \mathbf{x}_t$ for $t \in [0, T] \setminus I$ and $\mathbf{y}_t = \mathbf{x}_{c_n}$ for $t \in (c_n, d_n) := I_n$. Then for all $p \geq 1$*

$$\|\mathbf{x}\|_{p\text{-var};[0,T]}^p \leq (C + 3^{p-1}) \sum_{n=1}^{\infty} \|\mathbf{x}\|_{p\text{-var};[c_n,d_n]}^p + C \|\mathbf{y}\|_{p\text{-var};[0,T]}^p,$$

where $C = 1 + 2^p + 3^{p-1}$.

Proof. Define the super-additive functions $\omega_{\mathbf{x}}(s, t) = \|\mathbf{x}\|_{p\text{-var};[s,t]}^p$ and $\omega_{\mathbf{y}}(s, t) = \|\mathbf{y}\|_{p\text{-var};[s,t]}^p$. Let $\mathcal{D} = (t_0, t_1, \dots, t_k)$ be a partition of $[0, T]$.

Denote by $J_1 = (a_1, b_1), \dots, J_m = (a_m, b_m)$ those intervals I_n which contain some partition point $t_j \in \mathcal{D}$, ordered so that $b_j < a_{j+1}$ for all $j \in \{1, \dots, m-1\}$. Call a *block* a consecutive run of partition points $t_j, t_{j+1}, \dots, t_n$ either all in J_r for some $1 \leq r \leq m$, in which case we call it red, or all outside I , in which case we call it blue. We call a consecutive pair of partition points $t_i, t_{i+1} \in \mathcal{D}$ which lie in different blocks either red-red, red-blue, or blue-red depending on their respective blocks (note there are no blue-blue pairs).

For a red block $t_j, t_{j+1}, \dots, t_n$ in J_r we have

$$\sum_{i=j+1}^n d(\mathbf{x}_{t_{i-1}}, \mathbf{x}_{t_i})^p \leq \omega_{\mathbf{x}}(a_r, b_r).$$

For a blue block $t_j, t_{j+1}, \dots, t_n$ between J_r, J_{r+1} we have

$$\begin{aligned} \sum_{i=j+1}^n d(\mathbf{x}_{t_{i-1}}, \mathbf{x}_{t_i})^p &= \sum_{i=j+1}^n d(\mathbf{y}_{t_{i-1}}, \mathbf{y}_{t_i})^p \\ &\leq \omega_{\mathbf{y}}(b_r, a_{r+1}). \end{aligned}$$

For a red-blue pair t_i, t_{i+1} with $t_i \in J_r$ and t_{i+1} between J_r, J_{r+1} , we have

$$\begin{aligned} d(\mathbf{x}_{t_i}, \mathbf{x}_{t_{i+1}})^p &\leq 2^{p-1} [d(\mathbf{x}_{t_i}, \mathbf{x}_{b_r})^p + d(\mathbf{x}_{b_r}, \mathbf{x}_{t_{i+1}})^p] \\ &\leq 2^{p-1} [\omega_{\mathbf{x}}(a_r, b_r) + \omega_{\mathbf{y}}(b_r, t_{i+1})]. \end{aligned}$$

Likewise for a blue-red pair t_i, t_{i+1} with t_i between J_r, J_{r+1} and $t_{i+1} \in J_{r+1}$, we have

$$\begin{aligned} d(\mathbf{x}_{t_i}, \mathbf{x}_{t_{i+1}})^p &\leq 2^{p-1} [d(\mathbf{x}_{t_i}, \mathbf{x}_{a_{r+1}})^p + d(\mathbf{x}_{a_{r+1}}, \mathbf{x}_{t_{i+1}})^p] \\ &\leq 2^{p-1} [\omega_{\mathbf{y}}(t_i, a_{r+1}) + \omega_{\mathbf{x}}(a_{r+1}, b_{r+1})]. \end{aligned}$$

Finally for a red-red pair t_i, t_{i+1} with $t_i \in J_r$ and $t_{i+1} \in J_{r+1}$, we have

$$\begin{aligned} d(\mathbf{x}_{t_i}, \mathbf{x}_{t_{i+1}})^p &\leq 3^{p-1} [d(\mathbf{x}_{t_i}, \mathbf{x}_{b_r})^p + d(\mathbf{x}_{b_r}, \mathbf{x}_{a_{r+1}})^p + d(\mathbf{x}_{a_{r+1}}, \mathbf{x}_{t_{i+1}})^p] \\ &\leq 3^{p-1} [\omega_{\mathbf{x}}(a_r, b_r) + \omega_{\mathbf{y}}(b_r, a_{r+1}) + \omega_{\mathbf{x}}(a_{r+1}, b_{r+1})]. \end{aligned}$$

Splitting the sum $\sum_{i=1}^k d(\mathbf{x}_{t(i-1)}, \mathbf{x}_{t(i)})^p$ into blocks and consecutive pairs in different blocks, we obtain (denoting $a_{m+1} = T$ and $b_0 = 0$)

$$\begin{aligned} \sum_{i=1}^k d(\mathbf{x}_{t(i-1)}, \mathbf{x}_{t(i)})^p &\leq \left[\sum_{r=1}^m \omega_{\mathbf{x}}(a_r, b_r) \right] + \left[\sum_{r=0}^m \omega_{\mathbf{y}}(b_r, a_{r+1}) \right] \\ &\quad + 2^{p-1} \left[\sum_{r=1}^m \omega_{\mathbf{x}}(a_r, b_r) + \sum_{r=1}^m \omega_{\mathbf{y}}(b_r, a_{r+1}) \right] \\ &\quad + 2^{p-1} \left[\sum_{r=0}^{m-1} \omega_{\mathbf{y}}(b_r, a_{r+1}) + \sum_{r=1}^m \omega_{\mathbf{x}}(a_r, b_r) \right] \\ &\quad + 3^{p-1} \left[2 \sum_{r=1}^m \omega_{\mathbf{x}}(a_r, b_r) + \sum_{r=1}^{m-1} \omega_{\mathbf{y}}(b_r, a_{r+1}) \right]. \end{aligned}$$

Since $\omega_{\mathbf{x}}$ and $\omega_{\mathbf{y}}$ are super-additive, the desired result follows. $\square$

Proof of Proposition 5.2.1. This is an immediate application of Lemma 5.2.2 to the open subintervals $I_j := ((j-1)/n, j/n) \subset [0, 1]$, $j \in \{1, \dots, n\}$, and path $\mathbf{y} := \mathbf{x}'$. $\square$

5.2.2 Tightness of the random walk

Consider now an iid array $\mathbf{X}_{nj}$ of $C_o^{p\text{-var}}$ -valued random variables with concatenations $\mathbf{X}_n := \mathbf{X}_{n1} * \dots * \mathbf{X}_{nn}$. Let $\mathbf{X}^n \in D_o$ be the random walk associated with the array $X_{nj} := \mathbf{X}_{nj}(1)$. It follows from Proposition 5.2.1 that in order to establish tightness of $(\|\mathbf{X}_n\|_{p\text{-var};[0,1]})_{n \geq 1}$, it suffices to show that $\left(\sum_{j=1}^n \|\mathbf{X}_{nj}\|_{p\text{-var};[0,1]}^p\right)_{n \geq 1}$ and $(\|\mathbf{X}^n\|_{p\text{-var};[0,1]}^p)_{n \geq 1}$ are tight.

Note that $\sum_{j=1}^n \|\mathbf{X}_{nj}\|_{p\text{-var};[0,1]}^p$ is a sum of n iid real random variables. Thus tightness of $\left(\sum_{j=1}^n \|\mathbf{X}_{nj}\|_{p\text{-var};[0,1]}^p\right)_{n \geq 1}$ depends only on the tail behaviour of the random variable $\|\mathbf{X}_{n1}\|_{p\text{-var};[0,1]}$ as $n \rightarrow \infty$.

The situation of the random walk $\mathbf{X}^n$ is more interesting. The main result of this section is Theorem 5.2.3 which provides a criterion for tightness of $(\|\mathbf{X}^n\|_{p\text{-var};[0,1]}^p)_{n \geq 1}$.

In its simplest form, Theorem 5.2.3 implies that whenever $\mathbf{X}^n$ converges in law to a Lévy process in G^N , and the iid array X_{nj} is scaled by a scaling function θ , then $(\|\mathbf{X}^n\|_{p\text{-var};[0,1]}^p)_{n \geq 1}$ is tight for all $p > \kappa \geq 1$, where κ depends only on the scaling function θ .

We mention also that Theorem 5.2.3 will also play a central role in our analysis of Lévy processes in G^N . In particular, it will be used to show that the sample paths of a Lévy process in G^N have a.s. finite p -variation.

Throughout this section, we use the shorthand notation $\mathfrak{g}^N := \mathfrak{g}^N(\mathbb{R}^d)$, the Lie algebra of G^N . Recall that $\mathfrak{g}^N$ has a natural grading of the form $\mathfrak{g}^N = \bigoplus_{k=1}^N \mathfrak{g}_k^N$, where $\mathfrak{g}_1^N = \mathbb{R}^d$ and $\mathfrak{g}_k^N = [\mathfrak{g}_{k-1}^N, \mathbb{R}^d]$ for all $k \in \{2, \dots, N\}$. Recall that $\rho^k : T^N(\mathbb{R}^d) \mapsto (\mathbb{R}^d)^{\otimes k}$ denotes the k -th level projection, so that $\rho^k : \mathfrak{g}^N \mapsto \mathfrak{g}_k^N$ is the projection associated with the grading of $\mathfrak{g}^N$. Let d_k denote the dimension of $\mathfrak{g}_k^N$ (note that $d_1 = d$) and let $m = \sum_{k=1}^N d_k$ denote the dimension of $\mathfrak{g}^N$.

Consider a basis $u_1, \dots, u_m$ of $\mathfrak{g}^N$ such that $u_1, \dots, u_d$ is a basis of $\mathbb{R}^d$, and $u_{d_1+\dots+d_{k-1}+1}, \dots, u_{d_1+\dots+d_k}$ is a basis for $\mathfrak{g}_k^N$ for all $k \in \{2, \dots, N\}$ (note how this agrees with the definition of the basis $u_1, \dots, u_d$ of $\mathbb{R}^d$ in Section 4.1.2). For every $i \in \{1, \dots, m\}$, the degree of u_i is defined to be the unique integer $k \in \{1, \dots, N\}$ such that $u_i \in \mathfrak{g}_k^N$, denoted by $\deg(u_i) = k$.

Now let $\xi_1, \dots, \xi_m \in C_c^\infty(G^N)$ and $\xi : G^N \mapsto \mathfrak{g}^N$ be smooth functions and U a neighbourhood of 1_N for which the conditions at the start of Section 5.1, with respect to the basis $u_1, \dots, u_m$, are satisfied. We moreover equip $\mathfrak{g}^N$ with the inner product for which $u_1, \dots, u_m$ is an orthonormal basis.

Theorem 5.2.3. *Let $X_{n1}, \dots, X_{nn}$ be an iid array of G^N -valued random variables. Let $\mathbf{X}^n$ denote the associated random walk in G^N . For all $i \in \{1, \dots, m\}$, let $0 < q_i \leq 2$ be a real number, and define*

$$\kappa = \max\{1, q_1 \deg(u_1), \dots, q_m \deg(u_m)\}.$$

Then $(\mathbf{X}^n)_{n \geq 1}$ is a tight collection of $D_o([0, 1], G^N)$ -valued random variables, and for every $p > \kappa$, $(\|\mathbf{X}^n\|_{p\text{-var};[0,1]})_{n \geq 1}$ is a tight collection of real random variables, provided that the following conditions hold:

- (A) *for every $h \in [0, 1]$, $(\mathbf{X}_h^n)_{n \geq 1}$ is a tight collection of G^N -valued random variables;*
- (B) *for all $i \in \{1, \dots, m\}$, $\sup_{n \geq 1} n |\mathbb{E}[\xi_i(X_{n1})]| < \infty$;*
- (C) *the array X_{nk} is scaled by a scaling function θ , where $\theta \equiv \sum_{i=1}^m |\xi_i|^{q_i}$ on a neighbourhood of 1_N .*

Remark 5.2.4. Suppose that for a Lévy process $\mathbf{X}$ in G^N , $\mathbf{X}^n \xrightarrow{\mathcal{D}} \mathbf{X}$ as D_o -valued random variables. Then evidently $(\mathbf{X}^n)_{n \geq 1}$ is tight and condition (A) holds trivially. Furthermore, we have that condition (B) holds due to Theorem 5.1.1. Lastly, by Proposition 5.1.7, condition (C) holds upon choosing $q_i = 2$ for all $i \in \{1, \dots, m\}$.

The remainder of the section is devoted to the proof of Theorem 5.2.3, which can be split into three parts. The first part (Lemma 5.2.5 and Corollary 5.2.6) is a general tightness criterion for p -variation of strongly Markov processes. This is entirely an adaptation of the main result in [48] in which Manstavičius gave an elegant criterion to ensure that the sample paths of a strongly Markov process have a.s. finite p -variation. The second part (Lemma 5.2.7), which is the most technical part of the proof, involves establishing the required bounds to apply Corollary 5.2.6 for the case $p > N$. The third part takes up the remainder of the proof over several lemmas and involves adapting the argument to the case $p < N$. Roughly speaking, in the third part we decompose $\mathbf{X}^n$ into the lift of a walk in a lower level group, for which the previous two parts apply, and a perturbation on the higher levels, for which the p -variation can be controlled directly.

Let $\mathbf{X}$ be a strongly Markov process on $[0, T]$ taking values in a Polish space (E, d) . Define for all $h, \delta > 0$ and $s \in [0, T]$

$$\alpha(\mathbf{X}, h, \delta, s) := \sup\{\mathbb{P}^{x,s} [d(\mathbf{X}_s, \mathbf{X}_t) > \delta] \mid x \in E, t \in [0, T], s \leq t \leq s + h\}.$$

Recall that [48] Theorem 1.3 states that if there exist constants $\beta \geq 1$, and $K, \gamma, \delta_0 > 0$ such that for all $h \in [0, T]$ and $\delta \in (0, \delta_0]$,

$$\sup_{s \in [0, T]} \alpha(\mathbf{X}, h, \delta, s) \leq K \frac{h^\beta}{\delta^\gamma}, \quad (5.2.3)$$

then $\|\mathbf{X}\|_{p\text{-var}; [0, T]} < \infty$ a.s. for all $p > \gamma/\beta$.

We now state the result of [48] with two adaptations. The first is the observation that uniform bounds of the form (5.2.3) for a collection of Markov processes, along with tightness in the Skorokhod topology, leads to tightness of their p -variation. The second, which is essentially trivial, accounts for processes which are a.s. constant over fixed intervals of time.

For this purpose, for any $D([0, T], E)$ -valued random variable $\mathbf{X}$, we call a (deterministic) open interval $(s, t) \subset [0, T]$ *stationary* if $\mathbb{P}[\forall u \in (s, t), \mathbf{X}_u = \mathbf{X}_s] = 1$. We let $Z_{\mathbf{X}} \subseteq [0, T]$ denote the union of all stationary intervals, and let $R_{\mathbf{X}} = [0, T] \setminus Z_{\mathbf{X}}$ be its complement. For example, for the random walk $\mathbf{X}^n$, $R_{\mathbf{X}^n} = \{0, 1/n, \dots, (n-1)/n, 1\}$.

Lemma 5.2.5. *Let $(\mathbf{X}^n)_{n \geq 1}$ be càdlàg strongly Markov processes on $[0, T]$ taking values in a Polish space (E, d) . Suppose that*

- (a) *for all rational $h \in [0, T]$, $(\mathbf{X}_h^n)_{n \geq 1}$ is a tight collection of E -valued random variables, and*
- (b) *there exist constants $\beta \geq 1$ and $K, \gamma, \delta_0 > 0$ such that for all $n \geq 1$, $h \in [0, T]$, and $\delta \in (0, \delta_0]$,*

$$\sup_{s \in R_{\mathbf{X}^n}} \alpha(\mathbf{X}^n, h, \delta, s) \leq K \frac{h^\beta}{\delta^\gamma}.$$

Then $(\mathbf{X}^n)_{n \geq 1}$ is a tight collection of $D([0, T], E)$ -valued random variables, and for any $p > \gamma/\beta$, $(\|\mathbf{X}^n\|_{p\text{-var};[0, T]})_{n \geq 1}$ is a tight collection of real random variables.

Proof. First, note that (b) applied to small h allows us to verify the Aldous condition for the sequence $(\mathbf{X}^n)_{n \geq 1}$ (see, e.g., [38] p.188, though note one should restrict attention to sequences of stopping times τ_n taking values in $R_{\mathbf{X}^n}$ a.s., which is a trivial modification to the usual Aldous condition). Together with (a), it follows that $(\mathbf{X}^n)_{n \geq 1}$ is a tight collection of $D([0, T], E)$ -valued random variables ([38] Theorems 4.8.1 and 4.8.2).

It remains to show that $(\|\mathbf{X}^n\|_{p\text{-var};[0, T]})_{n \geq 1}$ is tight. For $\mathbf{x} \in D([0, T], E)$ and $b > 0$, define $M(\mathbf{x}) = \sup_{t \in [0, T]} d(\mathbf{x}_t, \mathbf{x}_0)$ and

$$\eta_b(\mathbf{x}) = \sup\{k \geq 0 \mid \exists (t_i)_{i=1}^{2k}, t_1 < t_2 \leq t_3 < \dots < t_{2k}, d(\mathbf{x}_{t_{2j}}, \mathbf{x}_{t_{2j-1}}) > b, 1 \leq j \leq k\}.$$

As noted above, [48] Theorem 1.3 states that (b), with $\sup_{s \in R_{\mathbf{X}^n}}$ replaced by $\sup_{s \in [0, T]}$, implies that $\|\mathbf{X}^n\|_{p\text{-var};[0, T]} < \infty$ a.s. for all $p > \gamma/\beta$ and $n \geq 1$. In fact, exactly the same proof as that of [48] Theorem 1.3 (taking into account that all stopping times considered in [48] take values in $R_{\mathbf{X}^n}$ a.s.) shows that $(\|\mathbf{X}^n\|_{p\text{-var};[0, T]})_{n \geq 1}$ is tight provided that $(M(\mathbf{X}^n))_{n \geq 1}$ and $(\eta_{\delta_0/2}(\mathbf{X}^n))_{n \geq 1}$ are tight, and provided that the uniform bounds in (b) hold.

However, M and η_b , for fixed $b > 0$, are continuous functions on the Skorokhod space, and since $(\mathbf{X}^n)_{n \geq 1}$ is tight, the desired conclusion follows. $\square$

Unless otherwise stated, we henceforth equip G^N with the homogeneous norm $\|x\| := \sum_{j=1}^N |x^j|^{1/j}$ and with a left-invariant metric d induced by some symmetric, sub-additive, homogeneous norm on G^N (recall that all homogeneous norms on G^N are equivalent, see Section 3.2).

As the increments of a random walk associated to an iid array are identically distributed and stationary, and the metric d on G^N is left-invariant, the following is now a corollary of Lemma 5.2.5.

Corollary 5.2.6. *Use the notation from Theorem 5.2.3 and suppose that (A) holds. Suppose there exists $K, \gamma > 0$ such that for all $n \geq 1$, $k \in \{1, \dots, n\}$ and $\delta \in (0, 1]$*

$$\mathbb{P} [||\mathbf{X}_{k/n}^n|| > \delta] \leq K \frac{k/n}{\delta^\gamma}.$$

Then $(\mathbf{X}^n)_{n \geq 1}$ is a tight collection of $D_o([0, 1], G^N)$ -valued random variables, and for all $p > \gamma$, $(||\mathbf{X}^n||_{p\text{-var};[0,1]})_{n \geq 1}$ is a tight collection of real random variables.

We now establish the crucial bounds needed to apply Corollary 5.2.6. For $j \in \{1, \dots, N\}$, define the set of indexes $H_j := \{i \in \{1, \dots, m\} \mid \deg(u_i) = j\}$.

Lemma 5.2.7. *Use the notation from Theorem 5.2.3 and suppose that (B) and (C) hold. Let $\gamma := N \vee \kappa$, and for $j \in \{1, \dots, N\}$, let $p_j := \max_{i \in H_j} q_i$ and denote by $\mathbf{Y}^{n,j} \in D_o([0, 1], \mathfrak{g}_j^N)$ the random walk associated with the $\mathfrak{g}_j^N$ -valued iid array $\rho^j \log(X_{nk})$.*

Then there exists $K > 0$ such that for all $n \geq 1$, $k \in \{1, \dots, n\}$ and $\delta \in (0, 1]$

$$\mathbb{P} [||\mathbf{X}_{k/n}^n|| > \delta] \leq K \frac{k/n}{\delta^\gamma}, \quad (5.2.4)$$

and, for all $j \in \{1, \dots, N\}$ such that $p_j \leq 1$,

$$\mathbb{P} [||\mathbf{Y}_{k/n}^{n,j}|| > \delta] \leq K \frac{k/n}{\delta^{p_j}}. \quad (5.2.5)$$

Remark 5.2.8. It follows from Corollary 5.2.6 and Lemma 5.2.7 that conditions (A), (B) and (C) imply that $(||\mathbf{X}^n||_{p\text{-var};[0,1]})_{n \geq 1}$ is tight for any $p > N \vee \kappa$ and that $(\mathbf{X}^n)_{n \geq 1}$ is a tight collection of $D_o([0, 1], G^N)$ -valued random variables.

Proof. We first claim that it suffices to consider the case $||X_{n1}|| \leq \varepsilon$ a.s. for all $n \geq 1$, where $\varepsilon > 0$ may be taken arbitrarily small. Indeed, let $\varepsilon > 0$ and note that there exists $c > 0$ such that $\theta(x) > c$ for all $x \in G^N$ with $||x|| > \varepsilon$. Since θ scales X_{nk} , it follows that there exists $C_1 > 0$ such that for all $n \geq 1$

$$\mathbb{P} [||X_{n1}|| \geq \varepsilon] \leq c^{-1} \mathbb{E} [\theta(X_{n1})] \leq C_1/n,$$

and hence

$$\mathbb{P} \left[\max_{1 \leq a \leq k} ||X_{na}|| \geq \varepsilon \right] = 1 - (1 - \mathbb{P} [||X_{n1}|| \geq \varepsilon])^k \leq C_1 k/n.$$

It follows that for all $n \geq 1$ and $k \in \{1, \dots, n\}$

$$\mathbb{P} [||\mathbf{X}_{k/n}^n|| > \delta] \leq \mathbb{P} [||\mathbf{X}_{k/n}^n|| > \delta, \max_{1 \leq a \leq k} ||X_{na}|| < \varepsilon] + C_1 k/n,$$

and similarly for $\mathbb{P} \left[\left| \mathbf{Y}_{k/n}^{n,j} \right| > \delta \right]$. Replacing X_{nk} by

$$X'_{nk} = \begin{cases} X_{nk} & \text{if } \|X_{nk}\| < \varepsilon \\ 1_N & \text{otherwise,} \end{cases}$$

we note that (B) and (C) imply that the same conditions hold for the iid array X'_{nk} . It thus suffices to prove the statement of the lemma for the iid array X'_{nk} instead as claimed.

We henceforth assume that $\|X_{n1}\| < \varepsilon$ a.s., where $\varepsilon > 0$ is sufficiently small so that $x \in U$ whenever $\|x\| < \varepsilon$, where we recall that U is a neighbourhood of 1_N on which $\xi|_U = \log|_U$.

We first show (5.2.5). Let $j \in \{1, \dots, N\}$ such that $p_j \leq 1$. Then, due to (C), there exists $C_2 > 0$ such that for all $n \geq 1$ and $k \in \{1, \dots, n\}$

$$\mathbb{E} \left[\sum_{i \in H_j} \sum_{a=1}^k |\xi_i(X_{na})|^{p_j} \right] \leq C_2 k/n \quad (5.2.6)$$

Since

$$\left| \mathbf{Y}_{k/n}^{n,j} \right|^{p_j} = \left| \sum_{i \in H_j} \sum_{a=1}^k \xi_i(X_{na}) \right|^{p_j} \leq \sum_{i \in H_j} \sum_{a=1}^k |\xi_i(X_{na})|^{p_j},$$

it follows from (5.2.6) and Markov's inequality that there exists $K > 0$ such that (5.2.5) holds for all $n \geq 1$, $k \in \{1, \dots, n\}$, and $\delta \in (0, 1]$.

We now show (5.2.4). For $j \in \{1, \dots, N\}$, define the set of words of length j

$$W_j := \{u_{i_1} \otimes \dots \otimes u_{i_j} \mid i_1, \dots, i_j \in \{1, \dots, d\}\}$$

and equip $T^N(\mathbb{R}^d)$ with inner product for which the set of all words $W := \cup_{j=1}^N W_j$ is an orthonormal basis. For all $w, x \in T^N(\mathbb{R}^d)$, we denote $x^w := \langle w, x \rangle$. For a word $w \in W$, let $|w|$ denote the integer $j \in \{1, \dots, N\}$ such that $w \in W_j$.

Considering $\exp : \mathfrak{g}^N \mapsto G^N$ as a polynomial map, note that (B) and (C) imply that for all $w \in T^N(\mathbb{R}^d)$

$$\sup_{n \geq 1} n \mathbb{E} [X_{n1}^w] < \infty, \quad (5.2.7)$$

and, applying the Cauchy-Schwartz inequality to terms of the form $\mathbb{E} [\xi_i(X_{n1}) \xi_h(X_{n1})]$, we have for all $w \in (\mathbb{R}^d)^{\otimes j}$

$$\sup_{n \geq 1} n \mathbb{E} [|X_{n1}^w|^{p_j}] < \infty. \quad (5.2.8)$$

Observe that for any word $w \in W_j$ and all $x_1, \dots, x_k \in T^N(\mathbb{R}^d)$, we have

$$\langle w, x_1 \dots x_k \rangle = \sum_{1 \leq a_1 \leq k} x_{a_1}^w + \sum_{w_1, w_2} \sum_{1 \leq a_1 < a_2 \leq k} x_{a_1}^{w_1} x_{a_2}^{w_2} + \dots + \sum_{w_1, \dots, w_j} \sum_{1 \leq a_1 < \dots < a_j \leq k} x_{a_1}^{w_1} \dots x_{a_j}^{w_j},$$

where $\sum_{w_1, \dots, w_r}$ indicates the summation over all combinations of words $w_1, \dots, w_r \in W$ such that $w_1 \otimes \dots \otimes w_r = w$.

Since $x \mapsto \max_{1 \leq j \leq N} |x^j|^{1/j}$ is a homogeneous norm on G^N , it suffices to show that there exists $K > 0$ such that

$$\mathbb{P} \left[\left| \sum_{1 \leq a_1 < \dots < a_r \leq k} X_{na_1}^{w_1} \dots X_{na_r}^{w_r} \right|^{1/j} > \delta \right] \leq K \frac{k/n}{\delta^\gamma} \quad (5.2.9)$$

for all $j \in \{1, \dots, N\}$, $r \in \{1, \dots, j\}$, words $w_1, \dots, w_r \in W$ such that $|w_1| + \dots + |w_r| = j$, $n \geq 1$, $k \in \{1, \dots, n\}$ and $\delta \in (0, 1]$.

To this end, let us fix $j \in \{1, \dots, N\}$, $r \in \{1, \dots, j\}$, and words $w_1, \dots, w_r \in W$ such that $|w_1| + \dots + |w_r| = j$. Consider first the case $r \in \{2, \dots, j\}$. Define

$$\gamma_j := j \vee \max\{q_i \deg(u_i) \mid i \in \{1, \dots, m\}, \deg(u_i) \leq j\}.$$

By Markov's and Jensen's inequalities (observing that $\gamma_j \leq 2j$)

$$\begin{aligned} \mathbb{P} \left[\left| \sum_{1 \leq a_1 < \dots < a_r \leq k} X_{na_1}^{w_1} \dots X_{na_r}^{w_r} \right|^{1/j} > \delta \right] &\leq \delta^{-\gamma_j} \mathbb{E} \left[\left| \sum_{1 \leq a_1 < \dots < a_r \leq k} X_{na_1}^{w_1} \dots X_{na_r}^{w_r} \right|^{\gamma_j/j} \right] \\ &\leq \delta^{-\gamma_j} \mathbb{E} \left[\left| \sum_{1 \leq a_1 < \dots < a_r \leq k} X_{na_1}^{w_1} \dots X_{na_r}^{w_r} \right|^2 \right]^{\gamma_j/2j}. \end{aligned} \quad (5.2.10)$$

To bound the last expression, consider the following polynomial in k commuting indeterminates $x_1, \dots, x_k$

$$\left(\sum_{1 \leq a_1 < \dots < a_r \leq k} x_{a_1} \dots x_{a_r} \right)^2. \quad (5.2.11)$$

The expression (5.2.11) expands into a sum of monomials of the form $x_{a_1}^{\alpha_1} \dots x_{a_r}^{\alpha_r}$, where $\alpha_j = 0, 1$ or 2 . Call the *simple degree* of such a monomial the number of $\alpha_j > 0$. The minimum simple degree is evidently r and the maximum is $2r$, and one readily sees that there exists $C_3(r) > 0$ such that for all $k \geq 1$, the number of monomials of simple degree $s \in \{r, \dots, 2r\}$ is bounded above by $C_3(r)k^s$.

Observe that for all words $w \in W$, due to (5.2.7),

$$\sup_{n \geq 1} n |\mathbb{E}[X_{n1}^w]| < C_4 < \infty.$$

Moreover, due to (5.2.8) and the assumption that $\|X_{n1}\| < \varepsilon$ a.s., we have for all $v, w \in W$

$$\sup_{n \geq 1} n \mathbb{E} [|X_{n1}^v X_{n1}^w|] \leq \sup_{n \geq 1} n \mathbb{E} [|X_{n1}^w|^2]^{1/2} \mathbb{E} [|X_{n1}^v|^2]^{1/2} < C_5 < \infty.$$

Recall that $X_{n1}, \dots, X_{nn}$ are iid. Consider a monomial $x_{a_1}^{\alpha_1} \dots x_{a_r}^{\alpha_r}$ of (5.2.11). Using each term x_a^α with $\alpha = 1$ to represent a term of the form $\mathbb{E}[X_{na}^{w_i}]$, $i \in \{1, \dots, r\}$, and each term x_a^α with $\alpha = 2$ to represent a term of the form $\mathbb{E}[X_{na}^{w_i} X_{na}^{w_h}]$, $i, h \in \{1, \dots, r\}$, it follows that for all $n \geq 1$ and $k \in \{1, \dots, n\}$

$$\mathbb{E} \left[\left(\sum_{1 \leq a_1 < \dots < a_r \leq k} X_{na_1}^{w_1} \dots X_{na_r}^{w_r} \right)^2 \right] \leq \sum_{s=r}^{2r} C_3(r) k^s (C_4/n)^{2s-2r} (C_5/n)^{2r-s}.$$

Since $r \geq 2$ by assumption and $k/n \leq 1$, there exists $C_6 > 0$ such that the right side is bounded above by $C_6(k/n)^2$ for all $n \geq 1$ and $k \in \{1, \dots, n\}$. It thus follows from (5.2.10) that for some $K > 0$

$$\mathbb{P} \left[\left| \sum_{1 \leq a_1 < \dots < a_r \leq k} X_{na_1}^{w_1} \dots X_{na_r}^{w_r} \right|^{1/j} > \delta \right] \leq K \frac{(k/n)^{2\gamma_j/2j}}{\delta^{\gamma_j}} \leq K \frac{k/n}{\delta^\gamma},$$

where the last inequality is due to $j \leq \gamma_j \leq \gamma$. This completes the case $r \in \{2, \dots, r\}$.

It remains to consider the case $r = 1$. Define now $\gamma_j := j(p_j \vee 1)$ and let $w \in W_j$. It holds that

$$\mathbb{P} \left[\left| \sum_{1 \leq a \leq k} X_{na}^w \right|^{1/j} > \delta \right] \leq \delta^{-\gamma_j} \mathbb{E} \left[\left| \sum_{1 \leq a \leq k} X_{na}^w \right|^{p_j \vee 1} \right].$$

Denote $\mu_n = \mathbb{E}[X_{n1}^w]$. Then there exist $C_7, C_8 > 0$ such that

$$\begin{aligned} \mathbb{E} \left[\left| \sum_{1 \leq a \leq k} X_{na}^w \right|^{p_j \vee 1} \right] &= \mathbb{E} \left[\left| \sum_{1 \leq a \leq k} X_{na}^w - \mu_n + \mu_n \right|^{p_j \vee 1} \right] \\ &\leq C_7 \mathbb{E} \left[\left| \sum_{1 \leq a \leq k} X_{na}^w - \mu_n \right|^{p_j \vee 1} + |k/n|^{p_j \vee 1} \right] \\ &\leq C_8 \left(\mathbb{E} \left[\sum_{1 \leq a \leq k} |X_{na}^w - \mu_n|^{p_j \vee 1} \right] + (k/n)^{p_j \vee 1} \right), \end{aligned}$$

where the second inequality is due to (5.2.7), and the final inequality is due to the Marcinkiewicz–Zygmund inequality and the fact that $p_j \leq 2$. It now follows

from (5.2.7) and (5.2.8) that

$$\begin{aligned}\mathbb{E} \left[\left| \sum_{1 \leq a \leq k} X_{na}^w \right|^{p_j \vee 1} \right] &\leq C_9 \left(k \mathbb{E} [|X_{na_1}^w|^{p_j \vee 1}] + k |\mu_n|^{p_j \vee 1} + (k/n)^{p_j \vee 1} \right) \\ &\leq C_{10} \left(k/n + kn^{-(p_j \vee 1)} + (k/n)^{p_j \vee 1} \right) \\ &\leq C_{11} (k/n).\end{aligned}$$

Since $\gamma_j \leq \gamma$, this completes the case $r = 1$ and the proof of the lemma. $\square$

As mentioned in Remark 5.2.8, we have now established the statement of Theorem 5.2.3 for $p > N$. For the case $p < N$, the main idea is to apply Lemma 5.2.7 to the iid array formed by projecting X_{nj} onto $G^{\lfloor p \rfloor}$, and then successively add back levels while controlling the p -variation.

Lemma 5.2.9. *Use the notation from Lemma 5.2.7 and suppose that (A), (B) and (C) hold. For all $j \in \{1, \dots, N\}$, it holds that $(\mathbf{Y}_h^{n,j})_{n \geq 1, h \in [0,1]}$ is a tight collection of $\mathfrak{g}_j^N$ -valued random variables.*

Proof. By Remark 5.2.8, note that $(\mathbf{X}^n)_{n \geq 1}$ is tight, from which it follows that $(\max_{1 \leq k \leq n} |\rho^j \log(X_{nk})|)_{n \geq 1}$ is tight. We may thus suppose that $|\rho^j \log(X_{n1})| \leq R$ a.s. for some large $R > 0$ and all $n \geq 1$.

Consider the decomposition $\rho^j \log(X_{nk}) = A_{nk} + B_{nk}$ where

$$A_{nk} = \rho^j \log(X_{nk}) \mathbf{1}\{\|X_{nk}\| < \varepsilon\}$$

and

$$B_{nk} = \rho^j (\log(X_{nk})) \mathbf{1}\{\varepsilon \leq \|X_{nk}\| \leq R\}.$$

We take here $\varepsilon > 0$ sufficiently small so that $\|x\| < \varepsilon$ implies $x \in U$ (recall that U is an open neighbourhood of 1_N for which $\xi|_U = \log|_U$). It suffices to prove that $(\sum_{a=1}^k B_{na})_{n \geq 1, k \in \{1, \dots, n\}}$ and $(\sum_{a=1}^k A_{na})_{n \geq 1, k \in \{1, \dots, n\}}$ are tight collections of real random variables.

Let $C_1 = C_1(\varepsilon) > 0$ be such that $C_1 \theta(x) > \rho^j (\log(x)) \mathbf{1}\{\varepsilon \leq \|x\| \leq R\}$ for all $x \in G^N$. Since θ scales X_{nk} , it holds that

$$\sup_{n \geq 1, k \in \{1, \dots, n\}} \mathbb{E} \left[\left| \sum_{a=1}^k B_{na} \right| \right] \leq \sup_{n \geq 1} C_1 n \mathbb{E} [\theta(X_{n1})] < \infty.$$

Thus, by Markov's inequality, by Markov's inequality, $(\sum_{a=1}^k B_{na})_{n \geq 1, k \in \{1, \dots, n\}}$ is tight.

Now observe that (B) and (C) imply that $\sup_{n \geq 1} n |\mathbb{E}[A_{n1}]| < \infty$. Moreover (C) implies that there exists $C_2 > 0$ such that $|\xi_i(x)|^2 \leq C_2 \theta(x)$ for all $x \in G^N$ and $i \in \{1, \dots, m\}$. Since $A_{nk} = \mathbf{1}\{|X_{na}| < \varepsilon\} \sum_{i \in H_j} \xi_i(X_{na}) u_i$, and $A_{n1}, \dots, A_{nn}$ are iid, it follows that

$$\sup_{n \geq 1, k \in \{1, \dots, n\}} \mathbb{E} \left[\left(\sum_{a=1}^k A_{na} \right)^2 \right] \leq \sup_{n \geq 1} n \sum_{a,b=1}^n |\mathbb{E}[A_{na} A_{nb}]| < \infty.$$

Thus, by Markov's inequality, $(\sum_{a=1}^k A_{na})_{n \geq 1, k \in \{1, \dots, n\}}$ is tight. $\square$

Recall that for $j \in \{1, \dots, N\}$, π^j denotes the projection $T^N(\mathbb{R}^d) \mapsto T^j(\mathbb{R}^d)$.

Corollary 5.2.10. *Use the notation from Lemma 5.2.7 and suppose that (A), (B) and (C) hold. Let $p > \kappa$ and define the $D_o([0, 1], G^{\lfloor p \rfloor \wedge N})$ -valued random variables $\mathbf{Z}^n := \pi^{\lfloor p \rfloor \wedge N} \mathbf{X}^n$. Then*

(i) $(\|\mathbf{Z}^n\|_{p\text{-var};[0,1]})_{n \geq 1}$ is tight, and

(ii) for every $j \in \{1, \dots, N\}$ such that $p_j \leq 1$, $(\|\mathbf{Y}^{n,j}\|_{p'\text{-var};[0,1]})_{n \geq 1}$ is tight for all $p' > p_j$.

Proof. (i) Observe that $\mathbf{Z}^n$ is the random walk associated with the $G^{\lfloor p \rfloor \wedge N}$ -valued iid array $Z_{nk} := \pi^{\lfloor p \rfloor \wedge N} X_{nk}$. Noting that $\|\mathbf{Z}_h^n\| \leq \|\mathbf{X}_h^n\|$ for all $h \in [0, 1]$, we may apply Corollary 5.2.6 and (5.2.4) of Lemma 5.2.7 (taking $\gamma = \lfloor p \rfloor \wedge \kappa < p$) to the array Z_{nk} , from which it follows that $(\|\mathbf{Z}^n\|_{p\text{-var};[0,1]})_{n \geq 1}$ is tight.

(ii) From Lemma 5.2.7, it suffices to check that condition (a) in Lemma 5.2.5 holds for the processes $(\mathbf{Y}^{n,j})_{n \geq 1}$. However this follows from Lemma 5.2.9. $\square$

Recall that for functions $\mathbf{z} : [0, T] \mapsto G^N$ and $\mathbf{y} : [0, T] \mapsto \mathfrak{g}_N^N$, the function $\mathbf{x} = \mathbf{z} + \mathbf{y} : [0, T] \mapsto G^N$ is well-defined by $\mathbf{x}_t := \exp(\log(\mathbf{z}_t) + \mathbf{y}_t)$.

Lemma 5.2.11. *Let $\mathbf{z} : [0, T] \mapsto G^N$ and $\mathbf{y} : [0, T] \mapsto \mathfrak{g}_N^N$ be functions, and let $\mathbf{x} := \mathbf{z} + \mathbf{y}$. Then for any $p > 0$ there exists $C = C(p, N) > 0$ such that*

$$\|\mathbf{x}\|_{p\text{-var};[0,T]}^p \leq C \left(\|\mathbf{z}\|_{p\text{-var};[0,T]}^p + \|\mathbf{y}\|_{p/N\text{-var};[0,T]}^{p/N} \right).$$

Proof. Recall that we equip G^N with the homogeneous norm $\|x\| := \sum_{j=1}^N |x^j|^{1/j}$. By equivalence of homogeneous norms on G^N , it holds that there exist $C_1, C_2 > 0$ such that for all $x, y \in G^N$

$$C_1 \|x^{-1}y\| \leq d(x, y) \leq C_2 \|x^{-1}y\|.$$

Note that $(\mathbf{x}_s^{-1}\mathbf{x}_t)^j = (\mathbf{z}_s^{-1}\mathbf{z}_t)^j$ for all $j \in \{1, \dots, N-1\}$, and

$$(\mathbf{x}_s^{-1}\mathbf{x}_t)^N = (\mathbf{z}_s^{-1}\mathbf{z}_t)^N + \mathbf{y}_t^N - \mathbf{y}_s^N.$$

Thus for any partition $\mathcal{D} \subset [0, T]$, we have

$$\begin{aligned} \sum_{t_i \in \mathcal{D}} d(\mathbf{x}_{t_i}, \mathbf{x}_{t_{i+1}})^p &\leq C_2^p \sum_{t_i \in \mathcal{D}} \left(\left| (\mathbf{z}_{t_i}^{-1}\mathbf{z}_{t_{i+1}})^N + \mathbf{y}_{t_{i+1}}^N - \mathbf{y}_{t_i}^N \right|^{1/N} + \sum_{j=1}^{N-1} |(\mathbf{z}_{t_i}^{-1}\mathbf{z}_{t_{i+1}})^j|^{1/j} \right)^p \\ &\leq C_2^p \sum_{t_i \in \mathcal{D}} \left(\left| \mathbf{y}_{t_{i+1}}^N - \mathbf{y}_{t_i}^N \right|^{1/N} + \sum_{j=1}^N |(\mathbf{z}_{t_i}^{-1}\mathbf{z}_{t_{i+1}})^j|^{1/j} \right)^p \\ &\leq C_2^p 2^{0 \vee (p-1)} \sum_{t_i \in \mathcal{D}} \left[\left| \mathbf{y}_{t_{i+1}}^N - \mathbf{y}_{t_i}^N \right|^{p/N} + \left(\sum_{j=1}^N |(\mathbf{z}_{t_i}^{-1}\mathbf{z}_{t_{i+1}})^j|^{1/j} \right)^p \right] \\ &\leq C_2^p 2^{0 \vee (p-1)} \left(\|\mathbf{y}\|_{p/N\text{-var};[0,T]}^{p/N} + C_1^{-p} \|\mathbf{z}\|_{p\text{-var};[0,T]}^p \right). \end{aligned}$$

□

Lemma 5.2.12. Consider elements $x_1, \dots, x_n \in G^N$ and let $\mathbf{x} : [0, 1] \mapsto G^N$ be the associated walk. Let $1 \leq p < N$ and define $\mathbf{z} = \pi^{[p]}\mathbf{x} \in D_o([0, 1], G^{[p]})$.

For $j \in \{[p] + 1, \dots, N\}$, let $\mathbf{y}^j \in D_o([0, 1], \mathfrak{g}_j^N)$ be the walk associated with the elements $\rho^j \log(x_1), \dots, \rho^j \log(x_n) \in \mathfrak{g}_j^N$.

Then there exists $C = C(p, N) > 0$ such that

$$\|\mathbf{x}\|_{p\text{-var};[0,1]}^p \leq C \left(\|\mathbf{z}\|_{p\text{-var};[0,1]}^p + \sum_{j=[p]+1}^N \|\mathbf{y}^j\|_{p/j\text{-var};[0,1]}^{p/j} \right).$$

Proof. Consider first the level N . Decompose each $x_k = \exp(l_k + y_k^N)$ where $y_k^N := \rho^N \log(x_k)$. Note that $\mathbf{y}^N$ is the walk associated with elements $y_1^N, \dots, y_n^N \in \mathfrak{g}_N^N$.

Let $w_k = \exp(l_k) \in G^N$ and let $\mathbf{w} \in D_o([0, 1], G^N)$ be the associated walk. Since y_k^N is in the centre of $\mathfrak{g}^N$ for all $k \in \{1, \dots, n\}$, note that $\mathbf{x} = \mathbf{w} + \mathbf{y}^N$.

Let $\tilde{z}_k := \pi^{N-1}(x_k) \in G^{N-1}$ and let $\tilde{\mathbf{z}} \in D_o([0, 1], G^{N-1})$ be the associated walk. Note that $\tilde{\mathbf{z}} = \pi^{N-1}(\mathbf{x})$.

Consider $\mathbf{m} \in C_o([0, 1], G^{N-1})$ defined by

$$\mathbf{m}_t = \begin{cases} 1_N & \text{if } t = 0 \\ \mathbf{m}_{k/n} \exp(n(t - k/n) \log(\tilde{z}_k)) & \text{if } t \in (k/n, (k+1)/n], k \in \{0, \dots, n-1\}. \end{cases}$$

Observe that $\mathbf{m}_{k/n} = \tilde{z}_1 \dots \tilde{z}_k = \tilde{\mathbf{z}}_{k/n}$. Moreover, it follows readily from Proposition 5.2.1 that $\|\mathbf{m}_{k/n}\|_{p\text{-var};[0,1]} \leq C_1(p, N) \|\tilde{\mathbf{z}}\|_{p\text{-var};[0,1]}$.

Furthermore, by construction of $\mathbf{w}$, observe that the level- N lift of $\mathbf{m}$ satisfies $S_N(\mathbf{m})_{k/n} = \mathbf{w}_{k/n}$, and so $\|\mathbf{w}\|_{p\text{-var};[0,1]} \leq C_2(p, N) \|S_N(\mathbf{m})\|_{p\text{-var};[0,1]}$, from which it follows that

$$\|\mathbf{w}\|_{p\text{-var};[0,1]} \leq C_3(p, N) \|\tilde{\mathbf{z}}\|_{p\text{-var};[0,1]}$$

(remark also that $\mathbf{w}$ is the “minimal jump extension” of $\tilde{\mathbf{z}}$ in the sense of Friz and Shekhar [24], and this bound can equivalently be derived from the proof of [24] Theorem 23).

Thus, by Lemma 5.2.11, it follows that

$$\|\mathbf{x}\|_{p\text{-var};[0,1]}^p \leq C_4(p, N) \left(\|\tilde{\mathbf{z}}\|_{p\text{-var};[0,1]}^p + \|\mathbf{y}^N\|_{p/N\text{-var};[0,1]}^{p/N} \right).$$

If $N - 1 \leq p < N$, then we are done. Otherwise, we proceed inductively and apply the same procedure with $\tilde{\mathbf{z}}$ taking the role of $\mathbf{x}$. $\square$

We now have all the ingredients for the proof of Theorem 5.2.3.

Proof of Theorem 5.2.3. Observe that Corollary 5.2.6 implies that $(\mathbf{X}^n)_{n \geq 1}$ is a tight collection of $D_o([0, 1], G^N)$ -valued random variables.

Let $p > \kappa$, and define $\mathbf{Z}^n$ as in Corollary 5.2.10, and $\mathbf{Y}^{n,j}$ and p_j , for $j \in \{1, \dots, N\}$, as in Lemma 5.2.7. Observe that $p_j < p/j < 1$ for all $j \in \{[p] + 1, \dots, N\}$. It follows by Corollary 5.2.10 that $(\|\mathbf{Z}^n\|_{p\text{-var};[0,1]})_{n \geq 1}$ and $(\|\mathbf{Y}^{n,j}\|_{p/j\text{-var};[0,1]})_{n \geq 1}$ are tight for all $j \in \{[p] + 1, \dots, N\}$.

However Lemma 5.2.12 implies that

$$\|\mathbf{X}^n\|_{p\text{-var};[0,1]}^p \leq C(p, N) \left(\|\mathbf{Z}^n\|_{p\text{-var};[0,1]}^p + \sum_{j=[p]+1}^N \|\mathbf{Y}^{n,j}\|_{p/j\text{-var};[0,1]}^{p/j} \right).$$

It follows that $(\|\mathbf{X}^n\|_{p\text{-var};[0,1]})_{n \geq 1}$ is tight as desired. $\square$

5.2.3 Example - random walk connected by a rough path

Recall the shorthand notation $G^N := G^N(\mathbb{R}^d)$, and let unspecified path spaces be defined on $[0, 1]$ and take values in G^N . In this section we apply the results of the previous two sections to a class of iid arrays of $C_o^{p\text{-var}}$ -valued random variables.

More specifically, we consider a random walk $\mathbf{Y}^n$ in $\mathbb{R}^d$ associated to an iid array Y_{nj} , and we wish to connect the points of $\mathbf{Y}^n$ with independent and appropriately scaled copies of a $C_o^{p\text{-var}}$ -valued random variable $\mathbf{Z}$. We achieve this using the following construction.

Recall that every element $h \in \mathbf{L}(\mathbb{R}^d)$ canonically extends to an algebra endomorphism (which we denote by the same letter) $h : T^N(\mathbb{R}^d) \mapsto T^N(\mathbb{R}^d)$ which sends every tensor $x_1 \otimes \dots \otimes x_k$ to $h(x_1) \otimes \dots \otimes h(x_k)$. Observe in particular that h maps G^N to G^N , and for the Carnot-Carathéodory norm $\|\cdot\|$ on G^N , $\|h(x)\| \leq |h| \|x\|$ for all $x \in G^N$ (where $|h|$ is understood as the operator norm of h as an element of $\mathbf{L}(\mathbb{R}^d)$). In particular, for all $\mathbf{x} \in C^{p\text{-var}}$,

$$\|h(\mathbf{x})\|_{p\text{-var};[0,1]} \leq |h| \|\mathbf{x}\|_{p\text{-var};[0,1]}. \quad (5.2.12)$$

Let Y_{nk} be an array of $\mathbb{R}^d$ -valued random variables. Let $p \geq 1$ and fix a $C_o^{p\text{-var}}$ -valued random variable $\mathbf{Z}$, independent of the array Y_{nj} , such that $\mathbf{Z}_{0,1}^1 = z$ a.s. for a constant non-zero element $z \in \mathbb{R}^d$. For all $n \geq 1$, let $\mathbf{Z}_{n1}, \dots, \mathbf{Z}_{nn}$ be independent copies of $\mathbf{Z}$. Let $Q : \mathbb{R}^d \mapsto \mathbf{L}(\mathbb{R}^d)$ be a measurable map such that $Q^y(z) = y$ for all $y \in \mathbb{R}^d$, where we use the notation $Q^y := Q(y)$.

For every $n \geq 1$, consider the $C_o^{p\text{-var}}$ -valued random variables $\mathbf{X}_{nk} := Q^{Y_{nk}}(\mathbf{Z}_{nk})$. Define the corresponding concatenation $\mathbf{X}_n := \mathbf{X}_{n1} * \dots * \mathbf{X}_{nn}$ which is likewise a $C_o^{p\text{-var}}$ -valued random variable. Observe that the level-one projection $\rho^1 \mathbf{X}_n(k/n)$, for $k \in \{1, \dots, n\}$, coincides with the random walk $\mathbf{Y}_{k/n}^n = \sum_{i=1}^k Y_{ni}$.

Example 5.2.13 (Linear interpolation). Fix $z \in \mathbb{R}^d$ and let $Q : \mathbb{R}^d \mapsto \mathbf{L}(\mathbb{R}^d)$ be any map such that $Q^y(z) = y$ for all $y \in \mathbb{R}^d$. Define $\mathbf{z} \in C_o^{1\text{-var}}([0, 1], G^N)$ by $\mathbf{z}_t := \exp(tz)$, i.e., $\mathbf{z}$ is the level- N lift of the linear path $t \mapsto tz$.

Then for all $y \in \mathbb{R}^d$, we have $Q^y(\mathbf{z})_t^1 = ty$, i.e., $Q^y(\mathbf{z})$ is the level- N lift of the linear path $t \mapsto ty$. Hence, for an array Y_{nk} of $\mathbb{R}^d$ -valued random variables and $\mathbf{Z} = \mathbf{z}$ a.s., the $C_o^{1\text{-var}}$ -valued random variable $\mathbf{X}_n$ constructed above is precisely the level- N lift of the piecewise linear interpolation of the random walk in $\mathbb{R}^d$ associated to Y_{nj} .

Example 5.2.14 (Dilation and rotation in $\mathbb{R}^2$). Let $\mathbb{R}^d = \mathbb{R}^2$. For all $y = (y^1, y^2) \in \mathbb{R}^2$, define $Q^y \in \mathbf{L}(\mathbb{R}^2)$ on the standard basis vectors by $Q^y(1, 0) = (y^1, y^2)$ and $Q^y(0, 1) = (-y^2, y^1)$. Observe that Q^y is the (unique) dilation and rotation about the origin which sends $(1, 0)$ to y (note moreover that $Q : \mathbb{R}^2 \mapsto \mathbf{L}(\mathbb{R}^2)$ is a linear map).

Taking Y_{nk} an array of $\mathbb{R}^2$ -valued random variables, $p \geq 1$, and $\mathbf{Z}$ any $C_o^{p\text{-var}}$ -valued random variable such that $\mathbf{Z}_{0,1}^1 = (1, 0)$ a.s., it follows that the $C_o^{p\text{-var}}$ -valued random variable $\mathbf{X}_n$ corresponds to the random walk associated with Y_{nk} connected by independent, scaled and rotated copies of $\mathbf{Z}$.

Theorem 5.2.15 is the main result of this section and allows us to determine when the collection of random variables $(\mathbf{X}_n)_{n \geq 1}$ satisfies the conditions of Corollary 3.4.8.

Note that its statement does not require that $z \in \mathbb{R}^d$ is non-zero and that $Q^y(z) = y$ for all $y \in \mathbb{R}^d$ as in the above discussion.

Theorem 5.2.15. *Let Y_{nk} an iid array of $\mathbb{R}^d$ -valued random variables and let $\mathbf{Y}^n \in D_o([0, 1], \mathbb{R}^d)$ be the associated random walk.*

Let $p > 2$, $N \geq \lfloor p \rfloor$, and let $\mathbf{Z}$ be a $C_o^{p\text{-var}}$ -valued random variable independent of the array Y_{nk} .

*Let $Q : \mathbb{R}^d \mapsto \mathbf{L}(\mathbb{R}^d)$ be a measurable map. For all $n \geq 1$, define the $C_o^{p\text{-var}}$ -valued random variable $\mathbf{X}_{n1} := Q^{Y_{n1}}(\mathbf{Z})$. Let $\mathbf{X}_{n2}, \dots, \mathbf{X}_{nn}$ be iid copies of $\mathbf{X}_{n1}$, and denote their concatenation by $\mathbf{X}_n = \mathbf{X}_{n1} * \dots * \mathbf{X}_{nn}$.*

Then $(\|\mathbf{X}_n\|_{p\text{-var};[0,1]})_{n \geq 1}$ is tight and $\lim_{n \rightarrow \infty} \mathbb{E}[M(\mathbf{X}_n)]$ exists for all $M \in \mathbf{L}(\mathbb{R}^d, \mathbf{u})$ provided that the following conditions hold:

- 1) $(\mathbf{Y}^n)_{n \geq 1}$ converge in law as $D_o([0, 1], \mathbb{R}^d)$ -valued random variables to a Lévy process in $\mathbb{R}^d$;
- 2) $Q^0 = 0$, and Q is continuous and admits a Taylor expansion of order $|\cdot|^2$;
- 3) The endpoint of $\mathbf{Z}$ has an a.s. constant level-one projection, i.e., $\mathbf{Z}_{0,1}^1 = z$ a.s. for some constant $z \in \mathbb{R}^d$;
- 4) The expected signature of $\mathbf{Z}$ has a strictly positive radius of convergence, i.e.,

$$R(\text{ESig}[S(\mathbf{Z})_{0,1}]) > 0;$$

$$5) \mathbb{E} \left[\|\mathbf{Z}\|_{p\text{-var};[0,1]}^p \right] < \infty.$$

Before the proof of Theorem 5.2.15, we prove an intermediate result of independent interest. Namely, define the iid array of G^N -valued random variables $X_{nk} := \mathbf{X}_{nk}(1)$. As usual, denote by $\mathbf{X}^n$ the random walk associated to X_{nk} . Remark that $\mathbf{X}^n(k/n) = \mathbf{X}_n(k/n)$ for all $n \geq 1$ and $k \in \{1, \dots, n\}$.

Proposition 5.2.17 below, under weaker conditions than those of Theorem 5.2.15, determines the limiting distribution of the random walk $\mathbf{X}^n$. Consider the following statement

$$4') \text{ For all } i \in \{1, \dots, N\}, \mathbb{E} [|\rho^i \log(\mathbf{Z}_{0,1})|^{2/i}] < \infty,$$

Remark 5.2.16. In fact, consider the statement

$$4'') \mathbb{E} [|\text{S}(\mathbf{Z})_{0,1}^k|] < \infty \text{ for all } k \geq 1.$$

Remark that 4) implies 4''). Observe moreover that, by the coproduct identity (or equivalently the shuffle product identity) on the group-like elements $G(\mathbb{R}^d)$, for all integers $k, q \geq 1$ and $u \in (\mathbb{R}^d)^{\otimes k}$, there exists $n \geq 1$ and a linear map $\psi : T^n(\mathbb{R}^d) \mapsto \mathbb{R}$, such that $\psi(\pi^n x) = \langle u, \log(\pi^k x) \rangle^q$ for all $x \in G(\mathbb{R}^d)$. Hence 4'') implies that for all $k \geq N$, $|\log(S_k(\mathbf{Z})_{0,1})|$ has finite moments of all orders. In particular, 4) $\Rightarrow$ 4'') $\Rightarrow$ 4').

Proposition 5.2.17. *Use the notation from Theorem 5.2.15 and suppose that 1), 2), 3), and 4') hold. Let θ be a scaling function on G^N such that $\theta \equiv \sum_{i=1}^m |\xi_i|^{2/\deg(u_i)}$ in a neighbourhood of 1_N . Define the iid array $X_{nk} := \mathbf{X}_{nk}(1)$ and let $\mathbf{X}^n$ be the associated random walk.*

Then θ scales X_{nj} and $\mathbf{X}^n \rightarrow \mathbf{X}$ as $D_o([0, 1], G^N)$ -valued random variables, where $\mathbf{X}$ is a Lévy process in G^N .

Moreover, let Π be the Lévy measure of the limiting Lévy process $\mathbf{Y}$ of $\mathbf{Y}^n$. Let μ be the probability measure on G^N associated to $Z := \mathbf{Z}_{0,1}$, and let Ξ be the pushforward of $\Pi \otimes \mu$ by the map $\mathbb{R}^d \times G^N \mapsto G^N$, $(y, z) \mapsto Q^y(z)$.

Then for all $z \in G^N$ and $i, j \in \{1, \dots, m\}$, the following limits exist

$$\begin{aligned} L_i(z) &:= \lim_{n \rightarrow \infty} n \mathbb{E} [\xi_i(Q^{Y_{n1}} z)], \\ L_{i,j}(z) &:= \lim_{n \rightarrow \infty} n \mathbb{E} [\xi_i(Q^{Y_{n1}} z) \xi_j(Q^{Y_{n1}} z)], \end{aligned}$$

and it holds that $L_i(Z)$ and $L_{i,j}(Z)$ are integrable random variables, Ξ is a Lévy measure on G^N , and the triplet of $\mathbf{X}$ is (C, D, Ξ) , where

$$\begin{aligned} D_i &= \mathbb{E} [L_i(Z)], \\ C_{i,j} &= \mathbb{E} [L_{i,j}(Z)] - \int_{G^N} \xi_i(x) \xi_j(x) \Xi(dx). \end{aligned}$$

Before the proof of Proposition 5.2.17, we derive several estimates in the following lemmas, which will ultimately be used to justify dominated convergence arguments.

To fix notation, for all $x \in G^N$, define $x_i := \langle u_i, \log(x) \rangle$, so that

$$x = \exp(x_1 u_1 + \dots + x_m u_m).$$

Remark that 4') is equivalent to $\mathbb{E} [|Z_i|^{2/\deg(u_i)}] < \infty$ for all $i \in \{1, \dots, m\}$.

Let $Q_1, \dots, Q_{d^2}$ be a basis of $\mathbf{L}(\mathbb{R}^d)$, and let $r_i : \mathbb{R}^d \mapsto \mathbb{R}$ be functions such that

$$Q^y = \sum_{i=1}^{d^2} r_i(y) Q_i.$$

Note that 2) is equivalent to the statement that, for all $i \in \{1, \dots, d^2\}$, $r_i(0) = 0$, and r_i is continuous and admits a Taylor expansion of order $|\cdot|^2$. Observe that there

exist functions $q_{i,j} : \mathbb{R}^d \mapsto \mathbb{R}$, $i, j \in \{1, \dots, m\}$, which are polynomials in $r_1, \dots, r_{d^2}$, such that

$$Q^y(x)_j = \sum_{i=1}^m q_{i,j}(y)x_i.$$

Observe that if $\deg(u_i) \neq \deg(u_j)$, then $q_{i,j} = 0$, and if $\deg(u_i) = \deg(u_j)$, then $q_{i,j}$ is a homogeneous polynomial of degree $\deg(u_i)$ in $r_1, \dots, r_{d^2}$.

Lemma 5.2.18. *Use the notation from Proposition 5.2.17 and suppose 1), 2), and 4') hold. Then*

$$\mathbb{E} \left[\sup_{n \geq 1} n \mathbb{E} [\theta(Q^{Y_{n1}} Z) \mid Z] \right] < \infty.$$

Proof. We may suppose that $\theta(x) \leq \sum_{j=1}^m 1 \wedge |x_j|^{2/\deg(u_j)}$ for all $x \in G^N$. Observe that for all $y \in \mathbb{R}^d$, $z \in G^N$, and $j \in \{1, \dots, m\}$

$$\begin{aligned} 1 \wedge |(Q^y z)_j|^{2/\deg(u_j)} &= 1 \wedge \left| \sum_{i=1}^m q_{i,j}(y)z_i \right|^{2/\deg(u_j)} \\ &\leq C \sum_{i=1}^m 1 \wedge |z_i q_{i,j}(y)|^{2/\deg(u_j)}. \end{aligned}$$

Furthermore, since $q_{i,j} = 0$ if $\deg(u_i) \neq \deg(u_j)$,

$$1 \wedge |z_i q_{i,j}(y)|^{2/\deg(u_j)} \leq (1 \vee |z_i|^{2/\deg(u_i)}) (1 \wedge |q_{i,j}(y)|^{2/\deg(u_j)}).$$

Since $q_{i,j}$ is a homogeneous polynomial in $r_1, \dots, r_{d^2}$ of degree $\deg(u_j)$, and, by 2), each $r_i(y) = O(|y|)$, it follows that $|q_{i,j}(y)|^{2/\deg(u_j)} = O(|y|^2)$. Hence, there exists $C > 0$ such that $1 \wedge |q_{i,j}(y)|^{2/\deg(u_j)} \leq C(1 \wedge |y|^2)$ for all $y \in \mathbb{R}^d$.

Then, by independence of Z and Y_{n1} , we have

$$\begin{aligned} \sup_{n \geq 1} n \mathbb{E} [(1 \vee |Z_i|^{2/\deg(u_i)}) (1 \wedge |q_{i,j}(Y_{n1})|^{2/\deg(u_j)}) \mid Z] \\ \leq (1 \vee |Z_i|^{2/\deg(u_i)}) \sup_{n \geq 1} n \mathbb{E} [1 \wedge |Y_{n1}|^2]. \end{aligned}$$

Observe that, by 1), $\sup_{n \geq 1} n \mathbb{E} [1 \wedge |Y_{n1}|^2] < \infty$, and that, by 4'), $\mathbb{E} [|Z_i|^{2/\deg(u_i)}] < \infty$, from which the conclusion follows. $\square$

Lemma 5.2.19. *Use the notation from Proposition 5.2.17 and suppose 1), 2), 3), and 4') hold. Then for all $j \in \{1, \dots, m\}$*

$$\mathbb{E} \left[\sup_{n \geq 1} n |\mathbb{E} [\xi_j(Q^{Y_{n1}} Z) \mid Z]| \right] < \infty.$$

Proof. Note that the case $\deg(u_j) \geq 2$ follows from Lemma 5.2.18. Consider now $\deg(u_j) = 1$, i.e., $j \in \{1, \dots, d\}$.

By 3), $Z_i = z_i$ a.s. for all $i \in \{1, \dots, d\}$. It follows that

$$\sup_{n \geq 1} n |\mathbb{E} [\xi_j(Q^{Y_{n1}} Z) \mid Z]| = \sup_{n \geq 1} n \left| \mathbb{E} \left[\xi_j \left(\sum_{i=1}^d q_{i,j}(Y_{n1}) z_i \right) \right] \right|.$$

By 2), $y \mapsto \xi_j \left(\sum_{i=1}^d q_{i,j}(y) z_i \right)$ admits a Taylor expansion of order $|\cdot|^2$. Hence, by 1), the following limit exists

$$\lim_{n \rightarrow \infty} n \mathbb{E} \left[\xi_j \left(\sum_{i=1}^d q_{i,j}(Y_{n1}) z_i \right) \right],$$

from which the conclusion follows. $\square$

Lemma 5.2.20. *Use the notation from Proposition 5.2.17 and suppose 2) and 4') hold. Then*

$$\begin{aligned} \int_{G^N} \theta(x) \Xi(dx) &= \mathbb{E} \left[\int_{\mathbb{R}^d} \theta(Q^y Z) \Pi(dy) \right] \\ &= \int_{\mathbb{R}^d} \mathbb{E} [\theta(Q^y Z)] \Pi(dy) \\ &< \infty. \end{aligned}$$

Proof. The equalities all follow by Fubini's theorem. It remains to verify that one of the integrals is finite. Indeed, we may suppose that $\theta(x) \leq \sum_{j=1}^m 1 \wedge |x_j|^{2/\deg(u_j)}$ for all $x \in G^N$. It follows that there exists $C > 0$ such that

$$\begin{aligned} \int_{\mathbb{R}^d} \mathbb{E} [\theta(Q^y Z)] \Pi(dy) &\leq C \int_{\mathbb{R}^d} \sum_{i,j=1}^m \mathbb{E} [1 \wedge |q_{i,j}(y) Z_i|^{2/\deg(u_j)}] \Pi(dy) \\ &\leq C \int_{\mathbb{R}^d} \sum_{i,j=1}^m 1 \wedge \mathbb{E} [|q_{i,j}(y) Z_i|^{2/\deg(u_i)}] \Pi(dy) \\ &< \infty, \end{aligned}$$

where the second inequality follows since $1 \wedge |\cdot|$ is a concave function on $[0, \infty)$ and that $q_{i,j} = 0$ if $\deg(u_i) \neq \deg(u_j)$, and the final inequality follows since, by 4'), $\mathbb{E} [|Z_i|^{2/\deg(u_i)}] < \infty$ for all $j \in \{1, \dots, m\}$ and, by 2), $|q_{i,j}(y)|^{2/\deg(u_i)} = O(|y|^2)$. $\square$

Proof of Proposition 5.2.17. The claim that θ scales X_{nj} follows immediately from Lemma 5.2.18. Moreover, for all $z \in G^N$ and $i, j \in \{1, \dots, m\}$, note that, by 2), the maps $y \mapsto \xi_i(Q^y z)$ and $y \mapsto \xi_i(Q^y z) \xi_j(Q^y z)$ from $\mathbb{R}^d$ to $\mathbb{R}$ are continuous and admit

Taylor expansions of order $|\cdot|^2$, from which it follows, by 1), that $L_i(z)$ and $L_{i,j}(z)$ exist.

Observe now that Ξ is indeed a Lévy measure on G^N by Lemma 5.2.20. To show that $\mathbf{X}^n \xrightarrow{\mathcal{D}} \mathbf{X}$, we check the conditions of Theorem 5.1.1, and at the same time we verify the integrability of $L_i(Z)$ and $L_{i,j}(Z)$.

(1) Let $f \in C_b(G^N)$ vanish in a neighbourhood of 0. We claim that

$$\begin{aligned} \lim_{n \rightarrow \infty} n\mathbb{E} [f(Q^{Y_{n1}} Z)] &= \lim_{n \rightarrow \infty} n\mathbb{E} [\mathbb{E} [f(Q^{Y_{n1}} Z) \mid Z]] \\ &= \mathbb{E} \left[\lim_{n \rightarrow \infty} n\mathbb{E} [f(Q^{Y_{n1}} Z) \mid Z] \right] \\ &= \mathbb{E} \left[\int_{\mathbb{R}^d} f(Q^y Z) \Pi(dy) \right] \\ &= \int_{G^N} f(x) \Xi(dx). \end{aligned}$$

Indeed, the first equality is obvious. The second follows from Lemma 5.2.18 and dominated convergence since $|f| \leq C\theta$ for some $C > 0$. The third follows from 1) and that for every $z \in G^N$, the map $y \mapsto f(Q^y z)$ is in $C_b(\mathbb{R}^d)$ and vanishes in a neighbourhood of zero. The final equality follows from Lemma 5.2.20 and Fubini's theorem.

(2) Let $i \in \{1, \dots, m\}$. We claim that

$$\begin{aligned} \lim_{n \rightarrow \infty} n\mathbb{E} [\xi_i(Q^{Y_{n1}} Z)] &= \lim_{n \rightarrow \infty} n\mathbb{E} [\mathbb{E} [\xi_i(Q^{Y_{n1}} Z) \mid Z]] \\ &= \mathbb{E} \left[\lim_{n \rightarrow \infty} n\mathbb{E} [\xi_i(Q^{Y_{n1}} Z) \mid Z] \right] \\ &= \mathbb{E} [L_i(Z)]. \end{aligned}$$

Indeed, the first equality is obvious, while the second (as well as the integrability of $L_i(Z)$) follows from Lemma 5.2.19 by dominated convergence.

(3) Let $i, j \in \{1, \dots, m\}$. We claim that

$$\begin{aligned} \lim_{n \rightarrow \infty} n\mathbb{E} [\xi_i(Q^{Y_{n1}} Z) \xi_j(Q^{Y_{n1}} Z)] &= \lim_{n \rightarrow \infty} n\mathbb{E} [\mathbb{E} [\xi_i(Q^{Y_{n1}} Z) \xi_j(Q^{Y_{n1}} Z) \mid Z]] \\ &= \mathbb{E} \left[\lim_{n \rightarrow \infty} n\mathbb{E} [\xi_i(Q^{Y_{n1}} Z) \xi_j(Q^{Y_{n1}} Z) \mid Z] \right] \\ &= \mathbb{E} [L_{i,j}(Z)]. \end{aligned}$$

Indeed, the first equality is obvious, while the second (as well as the integrability of $L_{i,j}(Z)$) follows from Lemma 5.2.18 by dominated convergence since $|\xi_i \xi_j| \leq C\theta$ for some $C > 0$. $\square$

The following is now a corollary of Proposition 5.2.17 and Theorem 5.2.3.

Corollary 5.2.21. *Use the notation from Proposition 5.2.17 and suppose 1), 2), 3), and 4') hold. Then $(\|\mathbf{X}_n\|_{p\text{-var};[0,1]})_{n \geq 1}$ is a tight collection of real random variables.*

We are now ready for the proof of Theorem 5.2.15.

Proof of Theorem 5.2.15. Note from Remark 5.2.16 that 4) in particular implies 4'). Following Corollary 5.2.21 and Proposition 5.2.1, to conclude that $(\|\mathbf{X}_n\|_{p\text{-var};[0,1]})_{n \geq 1}$ is tight, it suffices to show that $(\sum_{j=1}^n \|\mathbf{X}_{nj}\|_{p\text{-var};[0,1]}^p)_{n \geq 1}$ is tight.

Equip G^N with the Carnot-Carathéodory norm $\|\cdot\|$ and the associated metric d . For all $n \geq 1$, let $\mathbf{Z}_{n1}, \dots, \mathbf{Z}_{nn}$ be iid copies of $\mathbf{Z}$. From (5.2.12), we have

$$\|Q^{Y_{nj}}(\mathbf{Z}_{nj})\|_{p\text{-var};[0,1]}^p \leq |Q^{Y_{nj}}|^p \|\mathbf{Z}_{nj}\|_{p\text{-var};[0,1]}^p =: F_{nj}.$$

Since $\mathbf{X}_{nj}$ are iid and $\mathbf{X}_{nj} \stackrel{\mathcal{D}}{=} Q^{Y_{nj}}(\mathbf{Z}_{nj})$, it suffices to show that $(\sum_{j=1}^n F_{nj})_{n \geq 1}$ is tight.

Remark that for all $K, R > 0$

$$\mathbb{P} \left[\sum_{j=1}^n F_{nj} > K \right] \leq \mathbb{P} \left[\sum_{j=1}^n F_{nj} > K, \max_{1 \leq j \leq n} |Y_{nj}| < R \right] + \mathbb{P} \left[\max_{1 \leq j \leq n} |Y_{nj}| \geq R \right].$$

It holds, by 1), that

$$\lim_{R \rightarrow \infty} \sup_{n \geq 1} \mathbb{P} \left[\max_{1 \leq j \leq n} |Y_{nj}| \geq R \right] = 0.$$

Moreover, by Markov's inequality,

$$\sup_{n \geq 1} \mathbb{P} \left[\sum_{j=1}^n F_{nj} > K, \max_{1 \leq j \leq n} |Y_{nj}| < R \right] \leq K^{-1} \sup_{n \geq 1} n \mathbb{E} [F_{n1} \mathbf{1}\{|Y_{n1}| < R\}].$$

By 2), we have $|Q^y|^p = O(|y|^p) = o(|y|^2)$. Thus, by 1), for every $R > 0$,

$$\sup_{n \geq 1} n \mathbb{E} [|Q^{Y_{n1}}|^p \mathbf{1}\{|Y_{n1}| < R\}] < \infty.$$

By 5), $\mathbb{E} [\|\mathbf{Z}\|_{p\text{-var};[0,1]}^p] < \infty$, so that choosing R and K sufficiently large implies $\sup_{n \geq 1} \mathbb{P} \left[\sum_{j=1}^n F_{nj} > K \right]$ can be made arbitrarily small. This proves the claim that $(\|\mathbf{X}_n\|_{p\text{-var};[0,1]})_{n \geq 1}$ is tight.

It remains to show that $\lim_{n \rightarrow \infty} \mathbb{E} [M(\mathbf{X}_n)]$ exists for all $M \in \mathbf{L}(\mathbb{R}^d, \mathbf{u})$. Fix $M \in \mathbf{L}(\mathbb{R}^d, \mathbf{u}(H))$ and define the continuous bounded function $f : \mathbb{R}^d \mapsto \mathbf{L}(H)$,

$$f(y) := \mathbb{E} [(M \circ Q^y)(\mathbf{Z})],$$

so that $\mathbb{E} [f(Y_{n1})] = \mathbb{E} [M(\mathbf{X}_{n1})]$. To show that $\lim_{n \rightarrow \infty} \mathbb{E} [M(\mathbf{X}_n)]$ exists, it suffices to show that $\lim_{n \rightarrow \infty} n \mathbb{E} [f(Y_{n1}) - Id_H]$ exists. By 1) and Proposition 5.1.8, it thus

suffices to show that f admits a Taylor expansion of order $|\cdot|^2$ about the origin in $\mathbb{R}^d$.

By 4), the expected signature of $\mathbf{Z}$ has a strictly positive radius of convergence. Combined with Proposition 2.2.4, it follows that for $\lambda > 0$ sufficiently small,

$$\sum_{k \geq 0} \lambda^k \mathbb{E} [||S(\mathbf{Z})_{0,1}^k||] < \infty.$$

Let $z^k := \mathbb{E} [S(\mathbf{Z})_{0,1}^k] \in (\mathbb{R}^d)^{\otimes k}$. It follows by dominated convergence that for y in a neighbourhood of 0,

$$f(y) = \sum_{k \geq 0} (M \circ Q^y)^{\otimes k}(z^k).$$

It follows that

$$f(y) = Id_H + M(Q^y(z^1)) + M^{\otimes 2}((Q^y)^{\otimes 2} z^2) + O(|y|^3),$$

where we have use that, by 2), each $r_i(y) = O(|y|)$. Remark that $M(Q^y(z^1))$ and $M^{\otimes 2}((Q^y)^{\otimes 2} z^2)$ are linear and quadratic polynomials in $r_1(y), \dots, r_{d^2}(y)$ respectively. Since, again by 2), each r_i admits a Taylor expansion of order $|\cdot|^2$, it follows that so does $f(y)$, which completes the proof. $\square$

5.3 Path Functions

In this section, we introduce and study the concept of a path function, which shall be used to systematically connect the jumps of a càdlàg path.

Throughout the section, let (E, d) be a metric space and equip $C([0, T], E)$ and $D([0, T], E)$ with the uniform and the Skorokhod metrics respectively.

5.3.1 Definition

Definition 5.3.1. For a subset $J \subseteq E \times E$, we call $\phi : J \mapsto C([0, 1], E)$ a *path function* defined on J if

$$\phi(x, y)_0 = x \text{ and } \phi(x, y)_1 = y, \forall (x, y) \in J.$$

A path function $\phi : J \mapsto C([0, 1], E)$ is called *endpoint continuous* if ϕ is continuous and if

$$\phi(x, x)_t = x, \forall (x, x) \in J, \forall t \in [0, 1].$$

For $p \geq 1$, we say ϕ has *finite p -variation* if $\phi(x, y) \in C^{p\text{-var}}([0, 1], E)$ for all $(x, y) \in J$. Moreover, ϕ is called *p -approximating* if for every $r > 0$ there exists $C = C(r) > 0$ such that for all $(x, y) \in J$ with $d(x, y) < r$

$$\|\phi(x, y)\|_{p\text{-var};[0,1]} \leq Cd(x, y).$$

When E is a Lie group, ϕ is called *left-invariant* if there exists a subset $B \subseteq E$ such that $J = \{(x, y) \mid x^{-1}y \in B\}$ and

$$\phi(x, y)_t = x\phi(1_E, x^{-1}y)_t, \quad \forall (x, y) \in J, \quad \forall t \in [0, 1].$$

Note that for a Lie group G and a subset $B \subseteq G$, there is a canonical bijection between functions $\phi : B \mapsto C([0, 1], G)$, for which

$$\phi(x)_0 = 1_G \quad \text{and} \quad \phi(x)_1 = x, \quad \forall x \in B, \quad (5.3.1)$$

and left-invariant path functions defined on $J := \{(x, y) \mid x^{-1}y \in B\}$. Henceforth whenever we speak of a path function $\phi : B \mapsto C([0, 1], G)$ defined on a subset $B \subseteq G$, we shall mean that ϕ satisfies (5.3.1), and we shall identify ϕ with the corresponding left-invariant path function defined on J .

We mention that the endpoint continuous property of path functions is related to the continuity of the connecting function on the space $D([0, T], E)$ defined in Section 5.3.2. The p -approximating property (as one might expect) is related to p -variation properties of connected paths, and shall be discussed in Section 5.3.4.

5.3.2 The connecting function on the Skorokhod space

For a path $\mathbf{x} \in D([0, T], E)$ and a time $t \in [0, T]$, we denote by $\Delta\mathbf{x}_t := (\mathbf{x}_{t-}, \mathbf{x}_t) \in E \times E$ the jump of $\mathbf{x}$ at t , and $\|\Delta\mathbf{x}_t\| := d(\mathbf{x}_{t-}, \mathbf{x}_t)$ the size of the jump. We call t a jump time of $\mathbf{x}$ if $\|\Delta\mathbf{x}_t\| > 0$.

For a subset $J \subseteq E \times E$ and $\varepsilon \geq 0$ define the subset of càdlàg paths

$$J^\varepsilon := \{\mathbf{x} \in D([0, T], E) \mid \forall t \in [0, T], \|\Delta\mathbf{x}_t\| > \varepsilon \Rightarrow \Delta\mathbf{x}_t \in J\},$$

that is, J^ε consists of all càdlàg paths for which every jump of size greater than ε is contained in J . In particular, $\mathbf{x} \in J^0$ implies that all the jumps of $\mathbf{x}$ are in J .

In the case that E is a Lie group and $B \subseteq E$, we set $B^\varepsilon := J^\varepsilon$ where $J := \{(x, y) \mid x^{-1}y \in B\}$.

For a path function $\phi : J \mapsto C([0, 1], E)$, we shall now define the *connecting function* $\mathbf{x} \mapsto \mathbf{x}^\phi$ as a map from J^0 to $C([0, T], E)$ which serves to connect the jumps

of $\mathbf{x}$ by ϕ . The idea behind the construction is similar to that of Friz and Shekhar [24] and Williams [59]; we add sufficiently many extra time intervals to $[0, T]$ over which we traverse the jumps of $\mathbf{x}$ using ϕ .

Fix a non-increasing sequence $(r_i)_{i=1}^\infty$ of strictly positive real numbers such that $\sum_{i=1}^\infty r_i < \infty$.

Consider $\mathbf{x} \in J^0$. Let $t_1, t_2, \dots$ be the jump times of $\mathbf{x}$ ordered so that $\|\Delta \mathbf{x}_{t_1}\| \geq \|\Delta \mathbf{x}_{t_2}\| \geq \dots$ with $t_j < t_{j+1}$ if $\|\Delta \mathbf{x}_{t_j}\| = \|\Delta \mathbf{x}_{t_{j+1}}\|$. Let $0 \leq m \leq \infty$ be the number of jumps of $\mathbf{x}$. We call the sequence $(t_j)_{j=1}^m$ the *canonically ordered jump times* of $\mathbf{x}$.

Define the sequence $(n_k)_{k=0}^m$ by $n_0 = 0$, and for $1 \leq k < m + 1$ let n_k be the smallest integer such that $n_k > n_{k-1}$ and $r_{n_k} < \|\Delta \mathbf{x}_{t_k}\|$.

Remark 5.3.2. The choice to take $r_{n_k} < \|\Delta \mathbf{x}_{t_k}\|$ will play a role only in Lemma 5.3.5 to show a desirable continuity property of the map $\mathbf{x} \mapsto \mathbf{x}^\phi$ from J^0 to $C([0, T], E)$.

Let $r := \sum_{k=1}^m r_{n_k}$. Define the strictly increasing (càdlàg) function $\tau : [0, T] \mapsto [0, T + r]$ by

$$\tau(t) = t + \sum_{k=1}^m r_{n_k} \mathbf{1}\{t_k \leq t\}.$$

Note that $\tau(t-) < \tau(t)$ if and only if $t = t_k$ for some $1 \leq k < m + 1$. Moreover, note that the interval $[\tau(t_k-), \tau(t_k))$ is of length r_{n_k} .

We now define $\tilde{\mathbf{x}} \in C([0, T + r], E)$ as follows. For every $s \in [0, T + r]$, if $s = \tau(t)$ for some $t \in [0, T]$ define $\tilde{\mathbf{x}}_s = \mathbf{x}_t$, and if $s \in [\tau(t_k-), \tau(t_k))$ for some $1 \leq k < m + 1$, define

$$\tilde{\mathbf{x}}_s = \phi(\mathbf{x}_{t_k-}, \mathbf{x}_{t_k})((s - \tau(t_k-))/r_{n_k}).$$

Denote by τ_r the linear bijection from $[0, T]$ to $[0, T + r]$. We finally define the connected path

$$\mathbf{x}^\phi := \tilde{\mathbf{x}} \circ \tau_r \in C([0, T], E)$$

and the associated time change

$$\tau_{\mathbf{x}} := \tau_r^{-1} \circ \tau : [0, T] \mapsto [0, T].$$

5.3.3 Measurability and continuity in the Skorokhod topology

We equip $C([0, 1], E)$, $C([0, T], E)$, and $D([0, T], E)$ with their respective Borel σ -algebras.

Proposition 5.3.3. *Suppose E is a Polish space and $J \subseteq E \times E$ is measurable. Then J^ε is a measurable subset of $D([0, T], E)$ for all $\varepsilon \geq 0$.*

Let $\phi : J \mapsto C([0, 1], E)$ be a path function such that ϕ is measurable. Then the connecting function $\cdot^\phi : J^0 \mapsto C([0, T], E)$ is measurable.

Proof. Recall the notation of Section 5.3.2. Since E is a Polish space, it holds that the projection map $\mathbf{x} \mapsto \mathbf{x}_t$ from $D([0, T], E)$ to E is measurable for every $t \in [0, T]$ ([2] Theorem 12.5). It readily follows that the map

$$\mathbf{x} \mapsto (t_1, r_{n_1}, t_2, r_{n_2}, \dots) \quad (5.3.2)$$

from $D([0, T], E)$ to $\mathbb{R}^{\mathbb{N}}$ is measurable.

As a consequence, the map $\mathbf{x} \mapsto (\Delta \mathbf{x}_{t_1}, \Delta \mathbf{x}_{t_2}, \dots)$ from $D([0, T], E)$ to $(E \times E)^{\mathbb{N}}$ is also measurable (in case $\mathbf{x}$ has only finitely many $m < \infty$ jumps, we may choose $(\Delta \mathbf{x}_{t_1}, \Delta \mathbf{x}_{t_2}, \dots)$ to take a constant value (x, x) for some fixed $x \in E$ after m terms). Since J is measurable, it follows that J^ε is a measurable subset of $D([0, T], E)$ for all $\varepsilon \geq 0$.

Since E is a Polish space, it holds that Borel σ -algebra of $C([0, T], E)$ coincides with the σ -algebra generated by the finite-dimensional projections ([2] Example 1.3). Hence to show measurability of $\cdot^\phi$, it suffices to show that for every $s \in [0, T]$, the map $\mathbf{x} \mapsto \mathbf{x}^\phi(s)$ is measurable. Let us fix $s \in [0, T]$.

Observe that measurability of (5.3.2) implies that the map $\mathbf{x} \mapsto \tau(t)$ is also measurable for every $t \in [0, T]$. In particular, the map $\mathbf{x} \mapsto r = \tau(T) - T$ is measurable, and thus $\mathbf{x} \mapsto \tau_r(s)$ is also measurable.

Define $t' \in [0, T]$ and n' as follows: if $\tau_r(s) = \tau(t)$ for some $t \in [0, T]$, we define $t' = t$ and $n' = 0$, and if $\tau_r(s) \in [\tau(t_k-), \tau(t_k))$ for some $1 \leq k < m + 1$, we define $t' = t_k$ and $n' = n_k$. Observe that $t' = \sup\{t \in [0, T] \mid \tau(t) \leq \tau_r(s)\}$, from which it follows that $\mathbf{x} \mapsto (t', n')$ is measurable.

It now follows by measurability of ϕ that

$$\mathbf{x} \mapsto \begin{cases} \mathbf{x}_{t'} & \text{if } n' = 0 \\ \phi(\mathbf{x}_{t'-}, \mathbf{x}_{t'})((\tau_r(s) - \tau(t'-))/r_{n'}) & \text{if } n' > 0 \end{cases}$$

is measurable, where the right side is precisely $\mathbf{x}^\phi(s)$. $\square$

We next show a continuity property of $\cdot^\phi : J^0 \mapsto C([0, T], E)$, particularly in connection with weak convergence of $D([0, T], E)$ -valued random variables.

For a subset $J \subseteq E \times E$ and $\varepsilon \geq 0$, recall the definition of J^ε from Section 5.3.2. We say that a path function $\phi : J \mapsto C([0, T], E)$ *shrinks on the diagonal* in J if for every bounded set $B \subseteq J$ and $\varepsilon > 0$, there exists $\delta > 0$ such that for all $(x, y) \in B$ with $d(x, y) < \delta$

$$\sup_{t \in [0, 1]} d(\phi(x, y)_t, y) < \varepsilon.$$

Remark 5.3.4. Observe that every left-invariant, endpoint continuous path function defined on (a subset of) a Lie group G shrinks on the diagonal.

Lemma 5.3.5. *Consider $J \subseteq E \times E$ and a path function $\phi : J \mapsto C([0, 1], E)$ which shrinks on the diagonal in J . Suppose that ϕ is endpoint continuous on $K \subseteq J$.*

Let $\mathbf{x} \in K^0$ and a sequence $\mathbf{x}(n) \in J^0$ such that $\mathbf{x}(n) \rightarrow \mathbf{x}$ in the Skorokhod topology as $n \rightarrow \infty$, and such that for every $\varepsilon > 0$, there exists $n_0 > 0$ such that $\mathbf{x}(n) \in K^\varepsilon$ for all $n \geq n_0$.

Then $\lim_{n \rightarrow \infty} d_\infty(\mathbf{x}(n)^\phi, \mathbf{x}^\phi) = 0$.

Proof. Let $\varepsilon > 0$. By uniform continuity of $\mathbf{x}^\phi$, there exists $\eta > 0$ such that

$$\sup_{|s-t|<\eta} d(\mathbf{x}^\phi(t), \mathbf{x}^\phi(s)) < \varepsilon.$$

Let $t_1, t_2, \dots$ be the canonically ordered jump times of $\mathbf{x}$, and denote $[s_i, u_i) := [\tau_{\mathbf{x}}(t_i-), \tau_{\mathbf{x}}(t_i))$. For another element $\mathbf{y} \in D([0, T], E)$, let $t'_1, t'_2, \dots$ be the jump times of $\mathbf{y}$ and $[s'_i, u'_i) := [\tau_{\mathbf{y}}(t'_i-), \tau_{\mathbf{y}}(t'_i))$. Let $(n_k)_{k=0}^m$ and $(n'_k)_{k=1}^{m'}$ be the corresponding sequences for $\mathbf{x}$ and $\mathbf{y}$ respectively (where m and m' are the number of jumps of $\mathbf{x}$ and $\mathbf{y}$).

Denote by Λ^* the set of continuous, strictly increasing bijections $\lambda : [0, T] \mapsto [0, T]$ and let $Id \in \Lambda^*$ be the identity map. Equip $D([0, T], E)$ with the metric

$$\sigma(\mathbf{x}, \mathbf{y}) := \inf_{\lambda \in \Lambda^*} \max\{d_\infty(\lambda, Id), d_\infty(\mathbf{x} \circ \lambda, \mathbf{y})\}.$$

Let $\delta > 0$ (which we shall send to zero), and suppose that there exists $\lambda \in \Lambda^*$ such that $d_\infty(\mathbf{x} \circ \lambda, \mathbf{y}) < \delta$ and $d_\infty(\lambda, Id) < \delta$.

Observe that there exists an integer $k \geq 1$ such that $\lambda(t'_i) = t_i$ for all $i \in \{1, \dots, k\}$, and, denoting by $v_1 < \dots < v_k$ (resp. $v'_1 < \dots < v'_k$) the same set of points as $t_1, \dots, t_k$ (resp. $t'_1, \dots, t'_k$) ordered monotonically, it holds that $\lambda(t') \in [v_i, v_{i+1})$ for all $t' \in [v'_i, v'_{i+1})$. In particular, it holds that $d(\mathbf{y}_{t'_i-}, \mathbf{x}_{t_i-}), d(\mathbf{y}_{t'_i}, \mathbf{x}_{t_i}) < \delta$ for all $i \in \{1, \dots, k\}$.

Moreover, by choosing δ sufficiently small, we can assume that $n_i = n'_i$ for all $i \in \{1, \dots, k\}$ and that k is sufficiently large so that, by making $\sum_{j=k+1}^\infty r_j$ sufficiently small, it holds that $|\tau_{\mathbf{y}}(t') - \tau_{\mathbf{x}}(\lambda(t'))| < \eta$ for all $t' \in [v'_i, v'_{i+1})$ (this is where we have used the condition $r_{n_j} < \|\Delta \mathbf{x}_{t_j}\|$).

In particular, it holds that for all $t' \in [v'_i, v'_{i+1})$

$$\begin{aligned} d(\mathbf{y}_{\tau_{\mathbf{y}}(t')}^\phi, \mathbf{x}_{\tau_{\mathbf{y}}(t')}^\phi) &\leq d(\mathbf{y}_{\tau_{\mathbf{y}}(t')}^\phi, \mathbf{x}_{\tau_{\mathbf{x}}(\lambda(t'))}^\phi) + d(\mathbf{x}_{\tau_{\mathbf{x}}(\lambda(t'))}^\phi, \mathbf{x}_{\tau_{\mathbf{y}}(t')}^\phi) \\ &< d(\mathbf{y}_{t'}, \mathbf{x}_{\lambda(t')}) + \varepsilon \\ &< \delta + \varepsilon. \end{aligned} \tag{5.3.3}$$

This covers all points not in an interval $[s'_i, u'_i]$ (note that no continuity assumption on ϕ was needed here except that $\phi(x, y)$ itself is a continuous path for each $(x, y) \in J$).

Now we let $\mathbf{y} = \mathbf{x}(n)$ for some n . We may choose n sufficiently large, such that $\sigma(\mathbf{x}, \mathbf{y}) < \delta$ and such that $\Delta \mathbf{y}_{t'_i} \in K$ for all $i \in \{1, \dots, k\}$.

Due to the continuity of $\phi : K \mapsto C([0, 1], E)$ at $\Delta \mathbf{x}_{t_i} \in K$, it follows that for all $w' \in [s'_i, u'_i]$ and $i \in \{1, \dots, k\}$, there exists $w \in [s_i, u_i]$ such that $|w' - w| < \eta$ and $d(\mathbf{x}_w^\phi, \mathbf{y}_{w'}^\phi) < \varepsilon$, so that

$$d(\mathbf{y}_{w'}^\phi, \mathbf{x}_{w'}^\phi) < d(\mathbf{y}_{w'}^\phi, \mathbf{x}_w^\phi) + d(\mathbf{x}_w^\phi, \mathbf{x}_{w'}^\phi) < 2\varepsilon.$$

Finally, since ϕ shrinks on the diagonal, we may further decrease δ if necessary so that for all $w' \in [s'_j, u'_j]$ and $j > k$, it holds that $|w' - u'_j| < \eta$ and $d(\mathbf{y}_{w'}^\phi, \mathbf{y}_{u'_j}^\phi) < \varepsilon$. Now $u'_j = \tau_{\mathbf{y}}(t')$ for some $t' \in [0, T]$, and thus, by (5.3.3),

$$d(\mathbf{y}_{u'_j}^\phi, \mathbf{x}_{u'_j}^\phi) < \delta + \varepsilon,$$

from which it follows that

$$d(\mathbf{y}_{w'}^\phi, \mathbf{x}_{w'}^\phi) \leq d(\mathbf{y}_{w'}^\phi, \mathbf{y}_{u'_j}^\phi) + d(\mathbf{y}_{u'_j}^\phi, \mathbf{x}_{u'_j}^\phi) + d(\mathbf{x}_{u'_j}^\phi, \mathbf{x}_{w'}^\phi) < \delta + 3\varepsilon.$$

□

Recall from Proposition 5.3.3 that the subset $J^0 \subseteq D([0, T], E)$ and the map $\cdot^\phi : J^0 \mapsto C([0, T], E)$ are measurable provided that $J \subseteq E \times E$ and $\phi : J \mapsto C([0, 1], E)$ are measurable. Thus for any $D([0, T], E)$ -valued random variable $\mathbf{X}$ such that $\mathbf{X} \in J^0$ a.s., $\mathbf{X}^\phi$ is a $C([0, T], E)$ -valued random variable.

Proposition 5.3.6. *Suppose that $K \subseteq J \subseteq E \times E$ are measurable sets and $\phi : J \mapsto C([0, 1], E)$ is a measurable path function. Let $\mathbf{X}$ be a $D([0, T], E)$ -valued random variable such that $\mathbf{X} \in K^0$ a.s., and $(\mathbf{X}_n)_{n \geq 1}$ a collection of $D([0, T], E)$ -valued random variables such that $\mathbf{X}_n \in J^0$ a.s. and $\mathbf{X}_n \xrightarrow{\mathcal{D}} \mathbf{X}$ as $D([0, T], E)$ -valued random variables.*

Suppose that for every $\varepsilon > 0$, $\lim_{n \rightarrow \infty} \mathbb{P}[\mathbf{X}_n \notin K^\varepsilon] = 0$, and that ϕ is endpoint continuous on K and shrinks on the diagonal in J .

Then $\mathbf{X}_n^\phi \xrightarrow{\mathcal{D}} \mathbf{X}^\phi$ as $C([0, T], E)$ -valued random variables.

Proof. Use the shorthand notation $D := D([0, T], E)$ and $C := C([0, T], E)$. By Proposition 5.3.3, K^ε is a measurable subset of D for all $\varepsilon \geq 0$.

The condition $\lim_{n \rightarrow \infty} \mathbb{P}[\mathbf{X}_n \notin K^\varepsilon] = 0$ for all $\varepsilon > 0$ implies that

$$\mathbf{Y}_n := (\mathbf{X}_n, \mathbf{1}\{\mathbf{X}_n \notin K^1\}, \mathbf{1}\{\mathbf{X}_n \notin K^{1/2}\}, \dots)$$

is a sequence of $D \times \{0, 1\}^{\mathbb{N}}$ -valued random variables which converges in law to $\mathbf{Y} := (\mathbf{X}, 0, 0, \dots)$ (since the right term $(0, 0, \dots) \in \{0, 1\}^{\mathbb{N}}$ is constant, this is due to the result that $(A_n, B_n) \xrightarrow{\mathcal{D}} (A, b)$ whenever $A_n \xrightarrow{\mathcal{D}} A$ and $B \xrightarrow{\mathcal{D}} b$, where b is a constant).

Since $D \times \{0, 1\}^{\mathbb{N}}$ is a Polish space, it follows by the Skorokhod representation theorem that there exist $D \times \{0, 1\}^{\mathbb{N}}$ -valued random variables $\tilde{\mathbf{Y}}_n = (\tilde{\mathbf{X}}_n, a_{n1}, a_{n2}, \dots)$ and $\tilde{\mathbf{Y}} = (\tilde{\mathbf{X}}, a_1, a_2, \dots)$ defined on a common probability space $(\Omega, \mathcal{F}, \mathbb{Q})$ such that $\tilde{\mathbf{Y}}_n \stackrel{\mathcal{D}}{=} \mathbf{Y}_n$ and $\tilde{\mathbf{Y}} \stackrel{\mathcal{D}}{=} \mathbf{Y}$, and such that for $\mathbb{Q}$ -almost all $\omega \in \Omega$,

$$\lim_{n \rightarrow \infty} \tilde{\mathbf{Y}}_n(\omega) = \tilde{\mathbf{Y}}(\omega).$$

Note that $\tilde{\mathbf{Y}}_n \stackrel{\mathcal{D}}{=} \mathbf{Y}_n$ and $\tilde{\mathbf{Y}} \stackrel{\mathcal{D}}{=} \mathbf{Y}$ implies that $a_{nk}(\omega) = \mathbf{1}\{\tilde{\mathbf{X}}_n(\omega) \notin K^{1/k}\}$ and $a_k(\omega) = 0$ for all $n, k \geq 1$ and $\mathbb{Q}$ -almost all ω . Moreover,

$$\lim_{n \rightarrow \infty} (\mathbf{1}\{\tilde{\mathbf{X}}_n(\omega) \notin K^1\}, \mathbf{1}\{\tilde{\mathbf{X}}_n(\omega) \notin K^{1/2}\}, \dots) = (0, 0, \dots)$$

is equivalent to $\lim_{n \rightarrow \infty} \mathbf{1}\{\tilde{\mathbf{X}}_n(\omega) \notin K^\varepsilon\} = 0$ for all $\varepsilon > 0$. Thus, by Lemma 5.3.5, $\lim_{n \rightarrow \infty} d_\infty(\tilde{\mathbf{X}}_n(\omega)^\phi, \tilde{\mathbf{X}}(\omega)^\phi) = 0$ for $\mathbb{Q}$ -almost all ω .

In particular, we have that $\tilde{\mathbf{X}}_n^\phi \xrightarrow{\mathcal{D}} \tilde{\mathbf{X}}^\phi$ as C -valued random variables. However, observe that $\mathbf{X}_n^\phi \stackrel{\mathcal{D}}{=} \tilde{\mathbf{X}}_n^\phi$ and $\mathbf{X}^\phi \stackrel{\mathcal{D}}{=} \tilde{\mathbf{X}}^\phi$ due to the fact that $\mathbf{X}_n \stackrel{\mathcal{D}}{=} \tilde{\mathbf{X}}_n$ and $\mathbf{X} \stackrel{\mathcal{D}}{=} \tilde{\mathbf{X}}$, from which the desired result follows. $\square$

5.3.4 p -variation

The main result of this section is Proposition 5.3.8, which shows the important property that a p -approximating path function does not significantly increase the p -variation of a càdlàg path.

Lemma 5.3.7. *Let $p \geq 1$, $J \subseteq E \times E$, $\phi : J \mapsto C([0, T], E)$ a path function, and $\mathbf{z} \in J^0$. Let r be the size of the largest jump of $\mathbf{z}$ and suppose there exists $C > 0$ such that $\|\phi(x, y)\|_{p\text{-var};[0,1]} \leq Cd(x, y)$ for all $(x, y) \in J$ with $d(x, y) \leq r$.*

Then $\|\mathbf{z}^\phi\|_{p\text{-var};[0,T]} \leq R \|\mathbf{z}\|_{p\text{-var};[0,T]}$ where

$$R = 1 + 2^p + 3^{p-1} + C^p(1 + 2^p + 2 \cdot 3^{p-1}).$$

Proof. Let $0 \leq m \leq \infty$ be the number of jumps and $(t_j)_{j=1}^m$ be the canonically ordered jump times of $\mathbf{z}$. For $1 \leq j < m + 1$ denote

$$I_j := (c_j, d_j) := (\tau_{\mathbf{z}}(t_j-), \tau_{\mathbf{z}}(t_j))$$

and $I := \cup_{j=1}^m I_j$. Define $\mathbf{x} := \mathbf{z}^\phi$ and $\mathbf{y}_t := \mathbf{x}_t$ for all $t \in [0, T] \setminus I$ and $\mathbf{y}_t := \mathbf{x}_{c_j}$ for all $t \in (c_j, d_j)$ and $1 \leq j < m + 1$.

One readily sees that $\|\mathbf{y}\|_{p\text{-var};[0,T]} = \|\mathbf{z}\|_{p\text{-var};[0,T]}$. Moreover

$$\sum_{j=1}^m \|\mathbf{x}\|_{p\text{-var};[c_j, d_j]}^p \leq C^p \sum_{j=1}^m \|\Delta \mathbf{z}_{t_j}\|^p \leq C^p \|\mathbf{z}\|_{p\text{-var};[0,T]}^p.$$

Applying Lemma 5.2.2 to $\mathbf{y}$ and $\mathbf{x}$ concludes the proof. $\square$

The following result now follows immediately from Lemma 5.3.7 and from the definition of a p -approximating path function (Definition 5.3.1).

Proposition 5.3.8. *Let $p \geq 1$, $J \subseteq E \times E$, and $\phi : J \mapsto C([0, T], E)$ a p -approximating path function. There exists a continuous function $\psi : [0, \infty) \mapsto [0, \infty)$ such that for some $R, \varepsilon > 0$, $\psi(x) = xR$ for all $x \in [0, \varepsilon)$, and such that $\|\mathbf{x}^\phi\|_{p\text{-var};[0,T]} \leq \psi(\|\mathbf{x}\|_{p\text{-var};[0,T]})$ for all $\mathbf{x} \in J^0$.*

5.3.5 Path functions on $G^N(\mathbb{R}^d)$

Proposition 5.3.8 of the previous section in particular implies that a càdlàg path in $G^N(\mathbb{R}^d)$ of finite p -variation, for some $p < N+1$, can be transformed into an element of $C^{p\text{-var}}([0, T], G^N(\mathbb{R}^d))$ by connecting the jumps with a p -approximating path function. As a consequence, the resulting (p -rough) path can be used to drive an RDE. The purpose of this section is to establish the notation and several preliminary facts for linear RDEs driven by a path function.

We use the shorthand notation $G^N := G^N(\mathbb{R}^d)$. Recall that any function $\phi : B \mapsto C_o([0, 1], G^N)$ defined on a subset of $B \subseteq G^N$, for which $\phi(x)_1 = x$ for all $x \in B$, can be canonically identified with a left-invariant path function on $J := \{(x, y) \mid x^{-1}y \in B\}$ (see Definition 5.3.1 and the discussion thereafter).

Definition 5.3.9. Consider $1 \leq p < N+1$, linear vector fields $M \in \mathbf{L}(\mathbb{R}^d, \mathbf{L}(\mathbb{R}^e))$, a subset $B \subseteq G^N$, and $\phi : B \mapsto C_o^{p\text{-var}}([0, 1], G^N)$ a path function of finite p -variation.

We denote by M_ϕ the map

$$M_\phi : B \mapsto \mathbf{L}(\mathbb{R}^e), \quad x \mapsto M(\phi(x)).$$

Remark 5.3.10. Suppose moreover that ϕ is p -approximating and endpoint continuous. Then it holds that M_ϕ is continuous. To see this, remark that, by interpolation ([25] Lemma 8.16), $\phi : B \mapsto C_o^{p'\text{-var}}([0, 1], G^N)$ is continuous for all $p' > p$, and recall that the map $M : WG\Omega_{p'}(\mathbb{R}^d) \mapsto \mathbf{L}(\mathbb{R}^e)$ is continuous.

We now demonstrate that M_ϕ admits a Taylor expansion about the identity 1_N , provided that $\|\phi(x)\|_{p\text{-var};[0,1]}$ shrinks sufficiently fast as $x \rightarrow 1_N$. This will in particular be used in the proof of the Lévy-Khintchine formula in Theorem 5.4.5.

Recall that every $M \in \mathbf{L}(\mathbb{R}^d, \mathbf{L}(\mathbb{R}^e))$ extends canonically to a linear map $M \in \mathbf{L}(T^N(\mathbb{R}^d), \mathbf{L}(\mathbb{R}^e))$ (which we note is not an algebra homomorphism). We follow Section 5.2 for the remainder of notation.

Lemma 5.3.11. *Let $1 \leq p < \gamma < N + 1$, θ a scaling function on G^N , and $B \subseteq G^N$ a subset for which 1_N is an accumulation point.*

Let Φ be a collection of path functions defined on B such that

$$\lim_{x \rightarrow 1_N} \sup_{\phi \in \Phi} \frac{\|\phi(x)\|_{p\text{-var};[0,1]}^\gamma}{\theta(x)} = 0. \quad (5.3.4)$$

Let $M \in \mathbf{L}(\mathbb{R}^d, \mathbf{L}(\mathbb{R}^e))$. Then for all $\phi \in \Phi$, $M_\phi : B \mapsto \mathbf{L}(\mathbb{R}^e)$ admits the following Taylor expansion of order θ

$$\begin{aligned} M_\phi(x) &= Id_{\mathbb{R}^e} + \sum_{i=1}^m \xi_i(x) M(u_i) \\ &\quad + \frac{1}{2} \sum_{i,j=1}^m \xi_i(x) \xi_j(x) \mathbf{1}\{\deg(u_i) + \deg(u_j) \leq N\} M(u_i) M(u_j) + h_\phi(x), \end{aligned}$$

where $\lim_{x \rightarrow 1_N} \sup_{\phi \in \Phi} |h_\phi(x)| / \theta(x) = 0$.

Proof. Note that (5.3.4) implies in particular that $\lim_{x \rightarrow 1_N} \sup_{\phi \in \Phi} \|\phi(x)\|_{p\text{-var};[0,1]} = 0$. It follows, by Lemma 3.5.2, that for all $\phi \in \Phi$ there exists $f_\phi : B \mapsto \mathbf{L}(\mathbb{R}^e)$ such that $M_\phi(x) = M(x) + f_\phi(x)$ and

$$\limsup_{x \rightarrow 1_N} \sup_{\phi \in \Phi} \frac{|f_\phi(x)|}{\|\phi(x)\|_{p\text{-var};[0,1]}^\gamma} < \infty.$$

Since ξ_i are exponential coordinates around the identity 1_N , we have

$$\begin{aligned} M(x) &= Id_{\mathbb{R}^e} + \sum_{i=1}^m \xi_i(x) M(u_i) \\ &\quad + \frac{1}{2} \sum_{i,j=1}^m \xi_i(x) \xi_j(x) \mathbf{1}\{\deg(u_i) + \deg(u_j) \leq N\} M(u_i) M(u_j) \\ &\quad + O(|\xi(x)|^3). \end{aligned}$$

Since $|\xi|^3 = o(\theta)$, the desired result now follows from (5.3.4). $\square$

5.3.6 Globally continuous path function on $G^N(\mathbb{R}^d)$

The main result of this section is Lemma 5.3.12, which demonstrates the existence of a suitably nice 1-approximating, endpoint continuous path function ϕ defined on all of G^N .

Equip G^N with a symmetric, sub-additive, homogeneous norm $||\cdot||$ and the induced left-invariant metric $d(x, y) := ||x^{-1}y||$ (recall that all homogeneous norms on G^N are equivalent, see Section 3.2).

Recall the space $C_o^{1\text{-H\"ol}}([0, 1], G^N)$ of Lipschitz continuous paths whose starting point is the identity 1_N , equipped with the metric $d_{1\text{-H\"ol};[0,1]}$. Recall furthermore that $C_o^{0,1\text{-H\"ol}}([0, 1], G^N)$ is the closure of the set of smooth paths lifted by $S_N : C_o^{1\text{-H\"ol}}([0, 1], \mathbb{R}^d) \mapsto C_o^{1\text{-H\"ol}}([0, 1], G^N)$ (see Chapter 3, and [25] Chapters 8 and 9).

Lemma 5.3.12. *There exists a map $\phi : G^N \mapsto C_o^{0,1\text{-H\"ol}}([0, 1], G^N)$ such that for all $x \in G^N$*

- (i) $\phi(x)_1 = x$,
- (ii) the level-1 projection $\rho^1 \circ \phi(x) : [0, 1] \mapsto \mathbb{R}^d$ is piecewise linear,
- (iii) $||\phi(x)||_{1\text{-H\"ol};[0,1]} \leq C ||x||$ for a constant $C > 0$ independent of x , and
- (iv) ϕ is continuous in the 1-H\"older metric, i.e., $\lim_{n \rightarrow \infty} d_{1\text{-H\"ol};[0,1]}(\phi(x), \phi(x_n)) = 0$ whenever $\lim_{n \rightarrow \infty} d(x, x_n) = 0$.

Proof. Define the map $\psi : \mathbb{R}^m \mapsto G^N$ by $\psi(\lambda) = \exp(\lambda_m u_m) \dots \exp(\lambda_1 u_1)$. Observe that $(u_m, \dots, u_1)$ is a (strong) Malcev basis of $\mathfrak{g}^N$, and so ψ is a global diffeomorphism (see, e.g., [14] Proposition 1.2.7).

One readily sees that $x = \psi(\lambda) \mapsto \sum_{i=1}^m |\lambda_i|^{1/\deg(u_i)}$ is a homogeneous norm on G^N . Thus for some $C_1 \geq C_2 > 0$ and all $x = \psi(\lambda) \in G^N$,

$$C_1 ||x|| \leq \sum_{i=1}^m |\lambda_i|^{1/\deg(u_i)} \leq C_2 ||x||. \quad (5.3.5)$$

For every $i \in \{1, \dots, m\}$, let $\gamma_i \in C_o^{0,1\text{-H\"ol}}([0, 1], G^N)$ be a piecewise linear path such that $S_N(\gamma_i)(1) = \exp(u_i)$ and $\gamma_i(0) = 0$ (such a path always exists by Chow's theorem, [25] Theorem 7.28). For $\lambda \in \mathbb{R}^m$ and $i \in \{1, \dots, m\}$, define the path $\gamma_i^\lambda \in C_o^{0,1\text{-H\"ol}}([0, 1], G^N)$ by

$$\gamma_i^\lambda := \begin{cases} \lambda_i^{1/\deg(u_i)} \gamma_i & \text{if } \lambda_i \geq 0 \\ (-\lambda_i)^{1/\deg(u_i)} \overleftarrow{\gamma_i} & \text{if } \lambda_i < 0, \end{cases}$$

where $\overleftarrow{\gamma}_i(t) := \gamma_i(1-t) - \gamma_i(1)$ is the time reversal of γ_i .

For $x = \psi(\lambda) \in G^N$, define the piecewise linear path $\gamma(x) \in C^{1-\text{Hö}l}([0, 1], \mathbb{R}^d)$ as the concatenation of the paths $\gamma_m^\lambda, \dots, \gamma_1^\lambda$, parametrised so that γ_i^λ is traversed over the interval $[(m-i)/m, (m-i+1)/m]$. That is, $\gamma(x)$ is defined uniquely by $\gamma(x)_0 = 0$, and for all $t \in [0, 1/m]$ and $i \in \{1, \dots, m\}$

$$\gamma(x)_{(m-i)/m+t} = \gamma(x)_{(m-i)/m} + \gamma_i^\lambda(mt).$$

Finally, define $\phi(x) := S_N(\gamma(x))$ as the level- N lift of $\gamma(x)$. Observe now that (ii) holds by definition of ϕ .

Recall that for all $\alpha \geq 0$ and $\gamma \in C^{1-\text{var}}([0, 1], G^N)$, $S_N(\alpha\gamma) = \delta_\alpha S_N(\gamma)$, and that $S_N(\overleftarrow{\gamma})_1 = S_N(\overleftarrow{\gamma})_1^{-1}$. In particular $S_N(\alpha\gamma_i)_1 = \exp(\alpha^{\deg(u_i)} u_i)$ and $S_N(\alpha\overleftarrow{\gamma}_i)_1 = \exp(-\alpha^{\deg(u_i)} u_i)$. It follows from the definition of γ_i^λ that $S_N(\gamma_i^\lambda)_1 = \exp(\lambda_i u_i)$.

As a consequence of the multiplicative property of the lift S_N (Chen's identity), it holds that $\phi(x)_1 = x$, which proves (i).

It remains to show (iii) and (iv). Recall that $y \mapsto \int_0^\cdot y_t dt$ is a Banach space isomorphism from $L^\infty([0, 1], \mathbb{R}^d)$ to $C^{1-\text{Hö}l}([0, 1], \mathbb{R}^d)$ ([25] Proposition 1.37). In particular,

$$\|\gamma(x)\|_{1-\text{Hö}l;[0,1]} = \max_{1 \leq i \leq m} \|\gamma_i^\lambda\|_{1-\text{Hö}l;[0,1]}.$$

Since $\phi(x) = S_N(\gamma(x))$, it follows that ([25] Theorem 9.5) for some $C_3 = C_3(N) > 0$

$$\begin{aligned} \|\phi(x)\|_{1-\text{Hö}l;[0,1]} &\leq C_3 \|\gamma(x)\|_{1-\text{Hö}l;[0,1]} \\ &= C_3 \max_{1 \leq i \leq m} \|\gamma_i^\lambda\|_{1-\text{Hö}l;[0,1]} \\ &= C_3 \max_{1 \leq i \leq m} |\lambda_i|^{1/\deg(u_i)} \|\gamma_i\|_{1-\text{Hö}l;[0,1]} \\ &\leq C_4 \|x\|, \end{aligned}$$

where the last inequality is due to (5.3.5). This shows (iii).

Remark now that $\dot{\overleftarrow{\gamma}}_i(t) = -\dot{\gamma}_i(1-t)$. Thus for $t \in [0, 1/m]$ and $i \in \{1, \dots, m\}$,

$$\dot{\gamma}(x)_{(m-i)/m+t} = \begin{cases} m\lambda_i^{1/\deg(u_i)} \dot{\gamma}_i(mt) & \text{if } \lambda_i \geq 0 \\ -m(-\lambda_i)^{1/\deg(u_i)} \dot{\gamma}_i(1-mt) & \text{if } \lambda_i < 0, \end{cases}$$

(which is in fact a piecewise constant function in t since each γ_i is piecewise linear), from which it follows that $\dot{\gamma}(x_n) \rightarrow \dot{\gamma}(x)$ in $L^\infty([0, 1], \mathbb{R}^d)$ as $x_n \rightarrow x$. Due to the Banach space isomorphism between $L^\infty([0, 1], \mathbb{R}^d)$ and $C^{1-\text{Hö}l}([0, 1], \mathbb{R}^d)$ and the continuity of S_N in the $d_{1-\text{Hö}l}$ -metric ([25] Corollary 9.11), we obtain (iv). $\square$

A particularly nice consequence of Lemma 5.3.12 is the following corollary, which shows that the weak limit of certain $D_o([0, T], G^N)$ -valued random variables possesses sample paths of a.s. finite p -variation. In particular, we will apply this to Lévy processes in the following section.

Note that ϕ from Lemma 5.3.12 corresponds canonically to a left-invariant path function $\phi : G^N \times G^N \mapsto C([0, 1], G^N)$ (which we denote by the same symbol). Recall also the Hölder reparametrisation map $P_p : C^{p\text{-var}}([0, T], G^N) \mapsto C^{p\text{-var}}([0, T], G^N)$ from Section 3.2.2.

Corollary 5.3.13. *Let $(\mathbf{X}_n)_{n \geq 1}$ and $\mathbf{X}$ be $D_o([0, T], G^N)$ -valued random variables. Suppose that $\mathbf{X}_n \xrightarrow{\mathcal{D}} \mathbf{X}$ and that $(\|\mathbf{X}_n\|_{p\text{-var};[0, T]})_{n \geq 1}$ is tight for some $p \geq 1$.*

Then for all $p' > p$, it holds that $\|\mathbf{X}\|_{p'\text{-var};[0, T]} < \infty$ a.s., and

$$P_{p'}(\mathbf{X}_n^\phi) \xrightarrow{\mathcal{D}} P_{p'}(\mathbf{X}^\phi)$$

as $C_o^{0, p'\text{-var}}([0, T], G^N)$ -valued random variables, where ϕ is taken from Lemma 5.3.12.

Proof. Applying Proposition 5.3.6 (in which we take $K = J = G^N \times G^N$; note also that ϕ shrinks on the diagonal), it follows that $\mathbf{X}_n^\phi \xrightarrow{\mathcal{D}} \mathbf{X}^\phi$ as $C_o([0, T], G^N)$ -valued random variables.

Moreover, since ϕ is 1-approximating, and thus p -approximating, it follows by Proposition 5.3.8 that $(\|\mathbf{X}_n^\phi\|_{p\text{-var};[0, T]})_{n \geq 1}$ is tight. The desired result now follows from Proposition 3.2.4. $\square$

5.4 Lévy processes

We continue using the notation of Section 5.2. Recall in particular the shorthand notation $G^N := G^N(\mathbb{R}^d)$, the basis $u_1, \dots, u_m$ of $\mathfrak{g}^N$, and the corresponding local exponential coordinates $\xi_1, \dots, \xi_m \in C_c^\infty(G^N)$.

Throughout this section, we let $\mathbf{X}$ be a Lévy process in G^N with triplet (A, B, Π) (with respect to local coordinates $\xi_1, \dots, \xi_m$).

5.4.1 Finite p -variation of Lévy processes

The main result of this section is an (almost) complete characterisation, in terms of the triplet (A, B, Π) , of the values of $p > 1$ for which $\|\mathbf{X}\|_{p\text{-var};[0, 1]} < \infty$ a.s..

Recall the definitions of Γ_i , J , and K from Section 5.1.7. Furthermore, for $i \in \{1, \dots, m\}$ define

$$\beta_i := \inf \left\{ 0 \leq \beta \leq 2 \mid \int_G |\xi_i(x)|^\beta \Pi(dx) < \infty \right\}.$$

Note that $\beta_i = \sup\{\Gamma_i\}$ whenever Γ_i is not empty.

Theorem 5.4.1. *Let $\mathbf{X}$ be a Lévy process in G^N with triplet (A, B, Π) .*

(1) *Let $p > 1$ satisfy all of the following:*

- (i) $p > 2 \deg(u_j)$ for all $j \in J$;
- (ii) $p > \deg(u_k)$ for all $k \in K$;
- (iii) $p / \deg(u_i) > \beta_i$ for all $i \in \{1, \dots, m\}$.

Then $\|\mathbf{X}\|_{p\text{-var};[0,1]} < \infty$ a.s..

(2) *Let $p > 0$ satisfy any one of the following:*

- (iv) $p \leq 2 \deg(u_j)$ for some $j \in J$;
- (v) $p < \deg(u_k)$ for some $k \in K$;
- (vi) $p / \deg(u_i) \in \Gamma_i$ for some $i \in \{1, \dots, m\}$.

Then $\|\mathbf{X}\|_{p\text{-var};[0,1]} = \infty$ a.s..

Remark 5.4.2. Note the strict inequalities in (ii) and (iii) of Theorem 5.4.1. Comparing these to conditions (v) and (vi) respectively, we see that Theorem 5.4.1 does not completely determine all values of p for which $\|\mathbf{X}\|_{p\text{-var};[0,1]} < \infty$ a.s., that is, we do not cover the values $p = \deg(u_k)$ for $k \in K$, nor the values $p = \beta_i$ for any $i \in \{1, \dots, m\}$ such that $\beta_i \notin \Gamma_i$.

The strict inequalities in the conditions of (1) are essentially an artefact of our method of proof; we approximate $\mathbf{X}$ by a sequence of random walks $\mathbf{X}^n$ for which $(\|\mathbf{X}^n\|_{p\text{-var};[0,1]})_{n \geq 1}$ is tight, and thus obtain $\|\mathbf{X}\|_{p'\text{-var};[0,1]} < \infty$ a.s. for $p' > p$ by Corollary 5.3.13. Note that this method of proof, which relies crucially on interpolation and the Arzelà-Ascoli theorem ([25] Proposition 8.17), can only ever guarantee that $\|\mathbf{X}\|_{p'\text{-var};[0,1]} < \infty$ for all p' in an interval of the form (a, ∞) .

Remark 5.4.3. We can compare Theorem 5.4.1 with known results for $\mathbb{R}$; if $\mathbf{X}$ is a Lévy process in $\mathbb{R}$ with triplet $(0, B, 0)$, then evidently $\|\mathbf{X}\|_{p\text{-var};[0,1]} < \infty$ a.s. for all $p \geq 1$. Furthermore, if $\mathbf{X}$ is a Lévy process in $\mathbb{R}$ with triplet $(0, B, \Pi)$ and $1 < p < 2$, a result of Bretagnolle ([6] Theorem III) is that $\|\mathbf{X}\|_{p\text{-var};[0,1]} < \infty$ a.s. if and only if $p \notin \Gamma$, that is, $\int_{|x| < 1} |x|^p \Pi(dx) < \infty$ (note however how the result of [6] misses the case $p = 1$).

We suspect that the condition $p > 1$ in (1) can be dropped, and that $p > \deg(u_k)$ in (ii) can be replaced by $p \geq \deg(u_k)$, and that $p / \deg(u_i) > \beta_i$ in (iii) can be replaced by $p / \deg(u_i) \notin \Gamma_i$, which would complete the characterisation.

For the proof of Theorem 5.4.1, we require the following lemma.

Lemma 5.4.4. *Let $\mathbf{X}$ be a Lévy process in G^N with triplet (A, B, Π) . Assume $p > 1$ satisfies (i), (ii), and (iii) of Theorem 5.4.1.*

Let X_{nj} be the associated iid array constructed in Section 5.1.7 and $\mathbf{X}^n$ the associated random walk. Then $(\|\mathbf{X}^n\|_{p\text{-var};[0,1]})_{n \geq 1}$ is tight.

Proof. Let $1 < p' < p$ such that p' also satisfies (i), (ii), and (iii) of Theorem 5.4.1. For all $i \in \{1, \dots, m\}$, define $q_i := 2 \wedge (p' / \deg(u_i))$, and let θ be a scaling function on G^N such that $\theta \equiv \sum_{i=1}^m |\xi_i|^{q_i}$ in a neighbourhood of 1_N .

Observe that $q_i \notin \Gamma_i$ for all $i \in \{1, \dots, m\}$, $q_j = 2$ for all $j \in J$, and $q_k > 1$ for all $k \in K$. Thus, by Lemma 5.1.14, θ scales the array X_{nj} .

Moreover, by Lemma 5.1.13, $\mathbf{X}^n \xrightarrow{\mathcal{D}} \mathbf{X}$ as $D_o([0, 1], G^N)$ -valued random variables. It follows that the array X_{nj} satisfies the conditions of Theorem 5.2.3 with the above θ and $q_1, \dots, q_m$ (see also Remark 5.2.4). Since

$$p > \kappa := \max\{1, q_1 \deg(u_1), \dots, q_m \deg(u_m)\},$$

it follows that $(\|\mathbf{X}^n\|_{p\text{-var};[0,1]})_{n \geq 1}$ is tight. $\square$

Proof of Theorem 5.4.1. (1) Let X_{nj} be the iid array associated with $\mathbf{X}$ constructed in Section 5.1.7 and $\mathbf{X}^n$ the associated random walk. Let $1 < p' < p$ such that p' also satisfies (i), (ii), and (iii).

By Lemma 5.4.4, $(\|\mathbf{X}^n\|_{p'\text{-var};[0,1]})_{n \geq 1}$ is tight. Since $\mathbf{X}^n \xrightarrow{\mathcal{D}} \mathbf{X}$ as $D_o([0, 1], G^N)$ -valued random variables (Lemma 5.1.13), it follows that $\|\mathbf{X}\|_{p\text{-var};[0,1]} < \infty$ a.s. by Corollary 5.3.13.

(2) Note that for any homogeneous norm $\|\cdot\|$ on G^N , there exists $C > 0$ such that $\|x\|^{\deg(u_i)} \geq C|\xi_i(x)|$ for all $i \in \{1, \dots, m\}$ and $x \in G^N$. The desired result now follows immediately from Corollary 5.1.19 and Proposition 5.1.20. $\square$

5.4.2 The Lévy-Khintchine formula

In this section we consider the $C_o([0, 1], G^N)$ -valued random variable $\mathbf{X}^\phi$ obtained by applying a path function ϕ , and obtain a Lévy-Khintchine formula for the characteristic function $\mathbb{E}[M(\mathbf{X}^\phi)]$ in the case that for some $1 \leq p < N + 1$, $\|\mathbf{X}\|_{p\text{-var};[0,1]} < \infty$ a.s. and ϕ is p -approximating.

Recall that a Lévy process always has jumps in the support of its Lévy measure ([40] Proposition 1.4), i.e.,

$$\mathbb{P}[\forall t \in [0, 1], \mathbf{X}_{t-}^{-1} \mathbf{X}_t \in \text{supp}(\Pi) \cup \{1_N\}] = 1.$$

In particular, $\mathbf{X}$ takes values in $\text{supp}(\Pi)^0 \subseteq D([0, 1], G^N)$ a.s. (see Section 5.3.2).

It follows that for any (measurable) path function $\phi : \text{supp}(\Pi) \mapsto C_o([0, 1], G^N)$, $\mathbf{X}^\phi$ is a $C_o([0, 1], G^N)$ -valued random variable (where, as usual, we canonically identify ϕ with a left-invariant path function on $\{(x, y) \mid x^{-1}y \in \text{supp}(\Pi)\}$).

Moreover if for some $p \geq 1$, ϕ is p -approximating and $\|\mathbf{X}\|_{p\text{-var};[0,1]} < \infty$ a.s., then $\|\mathbf{X}^\phi\|_{p\text{-var};[0,1]} < \infty$ a.s. (Proposition 5.3.8).

Finally, for ϕ of finite p -variation and $M \in \mathbf{L}(\mathbb{R}^d, \mathbf{L}(\mathbb{R}^e))$, recall the definition of M_ϕ (Definition 5.3.9). Recall in particular that M_ϕ is continuous whenever ϕ is p -approximating and endpoint continuous (Remark 5.3.10).

Theorem 5.4.5 (Lévy-Khintchine formula). *Let $\mathbf{X}$ be a Lévy process in G^N with triplet (A, B, Π) . Suppose that for some $p < N + 1$, $\|\mathbf{X}\|_{p\text{-var};[0,1]} < \infty$ a.s..*

Let $\phi : \text{supp}(\Pi) \mapsto C_o^{p\text{-var}}([0, 1], G^N)$ be a p -approximating, endpoint continuous path function defined on $\text{supp}(\Pi)$.

Then for all $M \in \mathbf{L}(\mathbb{R}^d, \mathbf{u}(H))$, it holds that the function

$$M_\phi - Id_H - \sum_{i=1}^m M(u_i)\xi_i : \text{supp}(\Pi) \mapsto \mathbf{L}(H)$$

is Π -integrable, and

$$\mathbb{E} [M(\mathbf{X}^\phi)] = \exp(\Psi_{\mathbf{X}}(M)),$$

where

$$\begin{aligned} \Psi_{\mathbf{X}}(M) := & \sum_{i=1}^m B_i M(u_i) + \frac{1}{2} \sum_{i,j=1}^m A_{i,j} M(u_i) M(u_j) \\ & + \int_{G^N} \left[M_\phi(x) - Id_H - \sum_{i=1}^m \xi_i(x) M(u_i) \right] \Pi(dx). \end{aligned}$$

Remark 5.4.6. Note that by Theorem 5.4.1 part (2), the condition $\|\mathbf{X}\|_{p\text{-var};[0,1]} < \infty$ a.s. for $p < N + 1$ implies that for all i such that $\deg(u_i) > \lfloor N/2 \rfloor$, we have $A_{i,j} = 0$ for all $j \in \{1, \dots, m\}$.

Proof of Theorem 5.4.5. Since ϕ is endpoint continuous and left-invariant, remark that it shrinks on the diagonal. Let V be a neighbourhood of 1_N and $W := \text{supp}(\Pi) \cup V$. Extend ϕ to $\psi : W \mapsto C([0, 1], G^N)$ in an arbitrary way such that $\psi \equiv \phi$ on $\text{supp}(\Pi)$, and such that ψ is p -approximating and shrinks on the diagonal (for example, for all $x \in V \setminus \text{supp}(\Pi)$, define $\psi(x)$ as the path function in Lemma 5.3.12).

Let $X_{n1}, \dots, X_{nn}$ be the array constructed in Section 5.1.7 associated to $\mathbf{X}$, and let $\mathbf{X}^n$ be the associated random walk. Due to the shrinking support of the random variables Y_{nj} from Section 5.1.7, observe that

$$\text{for every } \varepsilon > 0, \mathbb{P}[\mathbf{X}^n \notin \text{supp}(\Pi)^\varepsilon] = 0 \text{ for all } n \text{ sufficiently large.} \quad (5.4.1)$$

In particular, for all n sufficiently large, $\mathbf{X}^n \in W^0$ a.s., so that $\mathbf{X}^{n,\psi}$ is well-defined.

Observe that, due to (5.4.1) and Proposition 5.3.6, $\mathbf{X}^{n,\psi} \xrightarrow{\mathcal{D}} \mathbf{X}^\psi$ as $C_o([0, 1], G^N)$ -valued random variables.

Let $p < p'' < p'$. Since $\|\mathbf{X}\|_{p\text{-var};[0,1]} < \infty$ a.s., it follows (by Theorem 5.4.1 part (2)) that p'' satisfies (i), (ii), and (iii) of Theorem 5.4.1. By Lemma 5.4.4, it follows that $(\|\mathbf{X}^n\|_{p''\text{-var};[0,1]})_{n \geq 1}$ is tight. Thus, by Proposition 5.3.8, $(\|\mathbf{X}^{n,\psi}\|_{p''\text{-var};[0,1]})_{n \geq 1}$ is also tight.

It now follows from Proposition 3.2.4 that

$$P_{p'}(\mathbf{X}^{n,\psi}) \xrightarrow{\mathcal{D}} P_{p'}(\mathbf{X}^\psi) \stackrel{\mathcal{D}}{=} P_{p'}(\mathbf{X}^\phi)$$

as $C_o^{0,p'\text{-var}}([0, 1], G^N)$ -valued random variables, where the equality in law follows from the fact that $\psi \equiv \phi$ on $\text{supp}(\Pi)$ and $\mathbf{X} \in \text{supp}(\Pi)^0$ a.s..

For all $i \in \{1, \dots, m\}$, define $q_i := 2 \wedge (p / \deg(u_i))$, and let θ be a scaling function on G^N such that $\theta \equiv \sum_{i=1}^m |\xi_i|^{q_i}$ in a neighbourhood of 1_N . By Theorem 5.4.1 part (2) and Lemma 5.1.14, it follows that θ scales the array X_{nj} .

Observe that for any $\gamma > p$, there exist $C_1, C_2 > 0$ and a neighbourhood of 1_N , $Y \subseteq W$, such that for all $x \in Y$

$$\|\psi(x)\|_{p\text{-var};[0,1]}^\gamma \leq C_1 \|x\|^\gamma \leq C_2 \sum_{i=1}^m |\xi_i(x)|^{\gamma / \deg(u_i)},$$

from which it follows that $\|\psi(\cdot)\|_{p\text{-var};[0,1]}^\gamma = o(\theta)$. In particular, taking $p < \gamma < N + 1$, it follows from Lemma 5.3.11 that for all $M \in \mathbf{L}(\mathbb{R}^d, \mathbf{L}(\mathbb{R}^e))$

$$\begin{aligned} M_\psi &= Id_{\mathbb{R}^e} + \sum_{i=1}^m \xi_i M(u_i) \\ &\quad + \frac{1}{2} \sum_{i,j=1}^m \xi_i \xi_j \mathbf{1}\{\deg(u_i) + \deg(u_j) \leq N\} M(u_i) M(u_j) + o(\theta). \end{aligned} \quad (5.4.2)$$

For $M \in \mathbf{L}(\mathbb{R}^d, \mathbf{u}(H))$, observe that M_ψ is a map from W to the unitary operators $\mathcal{U}(H)$ (and is thus bounded) and is continuous on $\text{supp}(\Pi)$. Combined with the Taylor

expansion (5.4.2), it follows from Lemma 5.1.15 that

$$\begin{aligned} \lim_{n \rightarrow \infty} n \mathbb{E} [M_\psi(X_{n1}) - Id_H] &= \sum_{i=1}^m B_i M(u_i) \\ &+ \frac{1}{2} \sum_{i,j=1}^m A_{i,j} M(u_i) M(u_j) + \int_{G^N} \left[M_\phi(x) - Id_H - \sum_{i=1}^m \xi_i(x) M(u_i) \right] \Pi(dx) \\ &= \Psi_{\mathbf{X}}(M), \end{aligned}$$

where we have used the fact that $M_\psi \equiv M_\phi$ on $\text{supp}(\Pi)$, and that $A_{i,j} = 0$ whenever $\deg(u_i) > \lfloor N/2 \rfloor$ (see Remark 5.4.6). In particular, this shows that $M_\phi - Id_H - \sum_{i=1}^m M(u_i) \xi_i$ is Π -integrable. Since the above quantities are simply finite-dimensional matrices, it follows that

$$\lim_{n \rightarrow \infty} \mathbb{E} [M_\psi(X_{n1})]^n = \exp(\Psi_{\mathbf{X}}(M)).$$

Since the array X_{nj} is iid, note that for all $n \geq 1$

$$\mathbb{E} [M(\mathbf{X}^{n,\psi})] = \mathbb{E} [M_\psi(X_{n1})]^n.$$

Furthermore, $M(P_{p'} \mathbf{x}) = M(\mathbf{x})$ for all $\mathbf{x} \in WG\Omega_{p'}(\mathbb{R}^d)$. Since $P_{p'}(\mathbf{X}^{n,\psi}) \xrightarrow{\mathcal{D}} P_{p'}(\mathbf{X}^\phi)$, and M is a continuous bounded function on $WG\Omega_{p'}(\mathbb{R}^d)$, it follows that

$$\mathbb{E} [M(\mathbf{X}^\phi)] = \exp(\Psi_{\mathbf{X}}(M)).$$

□

5.4.3 Convergence towards a Lévy process

The following theorem provides a condition under which a random walk converges in law to a Lévy process in rough path topologies when both are connected by a path function.

The result in fact follows directly from the tightness criterion of Theorem 5.2.3, and from Proposition 3.2.4, which we recall links convergence in $C_o([0, T], G^N)$, tightness of the p -variation norm, and convergence in the p -variation metric of Hölder reparametrisations.

Theorem 5.4.7. *Let X_{nj} be an iid array in G^N and $\mathbf{X}^n$ the associated random walk. Let $\mathbf{X}$ be a Lévy process in G^N such that $\mathbf{X}^n \xrightarrow{\mathcal{D}} \mathbf{X}$ as $D_o([0, 1], G^N)$ -valued random variables. Suppose that θ scales X_{nj} , where $\theta \equiv \sum_{i=1}^m |\xi_i|^{q_i}$ in a neighbourhood of 1_N for some $0 < q_i \leq 2$.*

Let $W \subseteq G^N$ be a closed subset such that $\text{supp}(\Pi) \subseteq W$ and $X_{n1} \in W$ a.s. for all $n \geq 1$. Let $p > \max\{1, q_1 \deg(u_1), \dots, q_m \deg(u_m)\}$ and $\phi : W \mapsto C_o^{p\text{-var}}([0, 1], G^N)$ a p -approximating, endpoint continuous (left-invariant) path function.

Then for every $p' > p$, $\|\mathbf{X}^\phi\|_{p'\text{-var};[0,1]} < \infty$ a.s., and $P_{p'}(\mathbf{X}^{n,\phi}) \xrightarrow{\mathcal{D}} P_{p'}(\mathbf{X}^\phi)$ as $C_o^{0,p'\text{-var}}([0, 1], G^N)$ -valued random variables.

Proof. By Theorem 5.2.3, it holds that $(\|\mathbf{X}^n\|_{p\text{-var};[0,1]})_{n \geq 1}$ is tight, and thus, by Proposition 5.3.8, $(\|\mathbf{X}^{n,\phi}\|_{p\text{-var};[0,1]})_{n \geq 1}$ is also tight.

Moreover, since ϕ is endpoint continuous on W , it follows by Proposition 5.3.6 that $\mathbf{X}^{n,\phi} \xrightarrow{\mathcal{D}} \mathbf{X}^\phi$ as $C_o([0, 1], G^N)$ -valued random variables. The conclusion now follows from Proposition 3.2.4. $\square$

Recall the flow map $U_{T \leftarrow 0}^{\mathbf{x}} \in \text{Homeo}(\mathbb{R}^e)$ for $\mathbf{x} \in WG\Omega_p(\mathbb{R}^d)$ associated with a collection of vector fields $f \in \text{Lip}_{\text{loc}}^\gamma(\mathbb{R}^e, \mathbf{L}(\mathbb{R}^d, \mathbb{R}^e))$ which satisfy the p -non-explosion condition.

As a direct consequence of the continuity of $U_{T \leftarrow 0} : WG\Omega_p(\mathbb{R}^d) \mapsto \text{Homeo}(\mathbb{R}^e)$ for all $p < \gamma$ (Theorem 3.4.1), and the invariance of $U_{T \leftarrow 0}^{\mathbf{x}}$ under reparametrisations of $\mathbf{x} \in WG\Omega_p(\mathbb{R}^d)$, we obtain the following corollary of Theorem 5.4.7.

Corollary 5.4.8. *Suppose the assumptions of Theorem 5.4.7 are verified for some $p < N + 1$. Let $\gamma > p$ and $f \in \text{Lip}_{\text{loc}}^\gamma(\mathbb{R}^e, \mathbf{L}(\mathbb{R}^d, \mathbb{R}^e))$ satisfy the p -non-explosion condition. Let $U_{1 \leftarrow 0}$ be the associated flow map.*

Then $U_{1 \leftarrow 0}^{\mathbf{X}^{n,\phi}} \xrightarrow{\mathcal{D}} U_{1 \leftarrow 0}^{\mathbf{X}^\phi}$ as $\text{Homeo}(\mathbb{R}^e)$ -valued random variables when $\text{Homeo}(\mathbb{R}^e)$ is equipped with the compact-open topology.

If moreover $f \in \text{Lip}^\gamma(\mathbb{R}^e, \mathbf{L}(\mathbb{R}^d, \mathbb{R}^e))$, then $U_{1 \leftarrow 0}^{\mathbf{X}^{n,\phi}} \xrightarrow{\mathcal{D}} U_{1 \leftarrow 0}^{\mathbf{X}^\phi}$ when $\text{Homeo}(\mathbb{R}^e)$ is equipped with the uniform topology.

Typically, the conditions of Theorem 5.4.7 are only satisfied for some $p > N + 1$ (but observe they are always satisfied for all $p > 2N$). However, using the results of Section 5.1.6, one may lift the walk to a higher level group so that Corollary 5.4.8 becomes applicable. We demonstrate this in the following examples, the first of which extends a result of Kunita [39].

Example 5.4.9 (Linear interpolation, Kunita [39]). Let $Y_{n1}, \dots, Y_{nn}$ be an iid array in $\mathbb{R}^d$ such that the associated random walk $\mathbf{Y}^n$ converges in law as a $D_o([0, 1], \mathbb{R}^d)$ -valued random variable to a Lévy process $\mathbf{Y}$ in $\mathbb{R}^d$. Observe that $1 \wedge |\cdot|^2$ scales the array Y_{nj} , so that in general, $(\|\mathbf{Y}^n\|_{p\text{-var};[0,1]})_{n \geq 1}$ is tight only for all $p > 2$.

Consider the $G^2(\mathbb{R}^d)$ -valued iid array $X_{nj} := e^{Y_{nj}}$, where, as usual, we treat $\mathbb{R}^d$ as a subspace of $\mathfrak{g}^2(\mathbb{R}^d)$, and let $\mathbf{X}^n$ be the associated random walk.

Let $N = 2$, and recall the local coordinates $\xi_1, \dots, \xi_m \in C_c^\infty(G^2(\mathbb{R}^d))$ about the identity. By definition, for all $x \in \mathbb{R}^d$ in a small neighbourhood of zero,

$$\xi_i(e^x) = \begin{cases} \langle u_i, x \rangle & \text{if } \deg(u_i) = 1 \\ 0 & \text{if } \deg(u_i) = 2. \end{cases}$$

In particular, for any scaling function θ on $G^2(\mathbb{R}^d)$ such that $\theta \equiv \sum_{i=1}^d |\xi|^2$ in a neighbourhood of the identity, it holds that θ scales X_{nj} .

Moreover, $\xi_i \circ \exp$ are smooth bounded functions on $\mathbb{R}^d$ (in fact of compact support), so the limits

$$D_i := \lim_{n \rightarrow \infty} n \mathbb{E} [\xi_i(X_{n1})]$$

and

$$C_{i,j} := \lim_{n \rightarrow \infty} n \mathbb{E} [\xi_i(X_{n1}) \xi_j(X_{n1})] - \int_{\mathbb{R}^d} \xi_i(e^x) \xi_j(e^x) \Pi(dx)$$

exist, where Π is the Lévy measure of $\mathbf{Y}$. It follows (Proposition 5.1.11, or equivalently Theorem 5.1.1), that $\mathbf{X}^n \xrightarrow{\mathcal{D}} \mathbf{X}$ as $D_o([0, 1], G^2(\mathbb{R}^d))$ -valued random variables, where $\mathbf{X}$ is the Lévy process in $G^2(\mathbb{R}^d)$ with triplet (C, D, Ξ) , where Ξ is the pushforward of Π by $\exp$.

To consider the lift of the piecewise linear interpolation of $\mathbf{Y}^n$ to $G^2(\mathbb{R}^d)$, define

$$\phi : \exp(\mathbb{R}^d) \mapsto C_o([0, 1], G^2(\mathbb{R}^d)), \quad \phi(e^x)_t = e^{tx}.$$

Clearly ϕ is a 1-approximating, endpoint continuous path function on $\exp(\mathbb{R}^d)$.

It follows that $\mathbf{X}^{n,\phi}$ is precisely (a reparametrisation of) the lift of the piecewise linear interpolation of $\mathbf{Y}^n$.

Observe now that the conditions of Theorem 5.4.7 are satisfied for all $p > 2$, from which it follows that $P_p(\mathbf{X}^{n,\phi}) \xrightarrow{\mathcal{D}} P_p(\mathbf{X}^\phi)$ as $C_o^{0,p\text{-var}}([0, 1], G^2(\mathbb{R}^d))$ -valued random variables.

In particular, taking $2 < p < 3$, it follows that ODEs driven by the piecewise linear interpolation of the random walk $\mathbf{Y}^n$ along Lip^γ vector fields, for any $\gamma > 2$, converge in law to the corresponding RDE driven by $\mathbf{X}^\phi$ (Corollary 5.4.8).

Remark 5.4.10. The previous example extends the main result of Kunita [39] (Theorem 4 and its Corollary). The main restriction of Kunita's result is the assumption that the vector fields $f_1, \dots, f_d$, along which $\mathbf{Y}^n$ drives an ODE, generate a finite dimensional Lie algebra (also, Kunita only considers smooth vector fields and establishes convergence in law under a stronger topology on $\text{Diff}(\mathbb{R}^e)$, but this discrepancy is minor, see Remark 3.4.2). This assumption essentially allows one to reduce the

problem to a random walk on a Lie group (see [39] p.340). Our approach, based on convergence in law under rough path topologies, bypasses this problem, and provides a natural interpretation of the limiting stochastic flow as the solution of an RDE.

Remark 5.4.11. We mention also a closely related result of Breuillard, Friz and Huesmann [7] who showed a result analogous to the above example in the special case that the iid array Y_{nj} is given by $Y_{n1} = n^{-1/2}Y$, where Y is a fixed random variable satisfying certain moment conditions. In this case, the limiting Lévy process $\mathbf{Y}$ is of course Brownian motion and the effect of the specific connecting path function ϕ (which was taken to be linear interpolation in the above example) becomes negligible in the limit, a point which was noted by the authors of [7].

The main difference between their result and ours is the method with which one obtains tightness in rough path topologies; the Kolmogorov–Lamperti tightness criterion is the main tool used in [7] to show tightness of $(\|\mathbf{Y}^n\|_{1/p\text{-Hö};[0,1]})_{n \geq 1}$ for sufficiently large p , which is far stronger than tightness of $(\|\mathbf{Y}^n\|_{p\text{-var};[0,1]})_{n \geq 1}$, which is what we show (and which we recall was established in large as a consequence of a result of Manstavičius [48]). Note that tightness of $(\|\mathbf{Y}^n\|_{1/p\text{-Hö};[0,1]})_{n \geq 1}$ necessarily implies that the limiting process is continuous, which automatically implies that $(\|\mathbf{Y}^n\|_{1/p\text{-Hö};[0,1]})_{n \geq 1}$ cannot be tight if the limiting Lévy process has jumps. This shows how the results of Theorems 5.2.3 and 5.4.7, even in the simple case of the above example, are not merely a technical extension of classical tightness criteria to càdlàg processes.

In the following Example 5.4.12 we demonstrate how the situation in Example 5.4.9 generalises to non-linear interpolations with essentially no extra effort.

Example 5.4.12 (Non-linear interpolation). As in Example 5.4.9, let Y_{nj} be an iid array in $\mathbb{R}^d$ such that $\mathbf{Y}^n \xrightarrow{\mathcal{D}} \mathbf{Y}$.

Let $\psi : \mathbb{R}^d \mapsto C_o^{q\text{-var}}([0, 1], \mathbb{R}^d)$ be a q -approximating endpoint continuous path function for some $1 \leq q < 2$. Define the map $f : \mathbb{R}^d \mapsto G^2(\mathbb{R}^d)$ by

$$f(x) = S_2(\psi(x))_1.$$

Consider the iid array $X_{nj} := f(Y_{nj})$.

Observe that there exist $C_1, C_2, C_3 > 0$ such that for all i with $\deg(u_i) = 2$ and all $x \in \mathbb{R}^d$ within a neighbourhood of 0,

$$|\xi_i(f(x))| \leq C_1 \|f(x)\|^2 \leq C_2 \|\psi(x)\|_{q\text{-var};[0,1]}^2 \leq C_3 |x|^2. \quad (5.4.3)$$

It follows that X_{nj} is scaled by any scaling function θ on $G^2(\mathbb{R}^d)$, such that $\theta \equiv \sum_{i=1}^d |\xi_i|^2$ in a neighbourhood of the identity.

We now make the following assumption on ψ and Y_{n1} : for all $i, j \in \{1, \dots, m\}$

$$D_i := \lim_{n \rightarrow \infty} n \mathbb{E} [\xi_i(f(Y_{n1}))]$$

and

$$C_{i,j} := \lim_{n \rightarrow \infty} n \mathbb{E} [\xi_i(f(Y_{n1})) \xi_j(f(Y_{n1}))] - \int_{\mathbb{R}^d} \xi_i(f(x)) \xi_j(f(x)) \Pi(dx)$$

exist. This occurs, for example, whenever every $\xi_i \circ f$ admits a Taylor expansion of order $|\cdot|^2$, but in general will depend on the array Y_{nj} and the path function ψ (note how (5.4.3) only implies that $\xi_i \circ f = O(|\cdot|^2)$, which is weaker than the existence of a Taylor expansion).

Under this assumption, it follows (Proposition 5.1.11) that the random walk $\mathbf{X}^n$ associated with the array X_{nj} converges in law to the Lévy process $\mathbf{X}$ with triplet (C, D, Ξ) , where Ξ is the pushforward of Π by f .

Define now $\phi : f(\mathbb{R}^d) \mapsto C_o^{q\text{-var}}([0, 1], G^2(\mathbb{R}^d))$ by

$$\phi(x) = S_2(\psi(x^1)).$$

It immediately follows that ϕ is a path function (by definition of f) and that ϕ is q -approximating. Moreover, considering ψ as a continuous map from $\mathbb{R}^d$ into $C_o^{q'\text{-var}}([0, 1], \mathbb{R}^d)$ for some $q < q' < 2$, it follows that ϕ is endpoint continuous (by continuity of S_2).

Exactly as in Example 5.4.9, it follows that for all $p > 2$, $P_p(\mathbf{X}^{n,\phi}) \xrightarrow{\mathcal{D}} P_p(\mathbf{X}^\phi)$ as $C_o^{0,p\text{-var}}([0, 1], G^2(\mathbb{R}^d))$ -valued random variables. Remark that $\mathbf{X}^{n,\phi}$ is (a reparametrisation of) the lift of $\mathbf{Y}^{n,\psi}$, which in turn is (a reparametrisation of) the random walk $\mathbf{Y}^n$ interpolated by the path function ψ .

It follows, taking $2 < p < 3$, that ODEs (or RDEs if $q > 1$) driven by $\mathbf{Y}^{n,\psi}$ along Lip^γ vector fields, for any $\gamma > 2$, converge in law to the corresponding RDE driven by $\mathbf{X}^\phi$ (Corollary 5.4.8).

Remark 5.4.13. McShane [49] considered non-linear interpolations of the increments of Brownian motion and showed that the corresponding ODEs converge (in L^r) to the associated Stratonovich SDE with an adjusted drift. We note that the interpolating functions considered in Example 5.4.12 contain the non-linear interpolations considered by McShane [49] as a special case. In particular, one can readily verify that $C_{i,j}$ and D_i exist when Y_{nk} are the increments of a Brownian motion and when ψ is the interpolation function considered by McShane (see [49], p.285).

In a similar way, the following Example 5.4.14 can be compared to the non-standard approximation of Brownian motion considered by Sussman [56] (see [25] Section 13.3.4 and [20] for a concise description of the results of [49] and [56] in the language of rough paths).

Of course our results are of a different nature to those of [49] and [56], as we do not consider our random walk to be the increments of an underlying process, and thus only consider convergence in law (and not in probability or L^r).

Example 5.4.14 (Perturbed walk). As in Examples 5.4.9 and 5.4.12, let Y_{nj} be an iid array in $\mathbb{R}^d$ such that $\mathbf{Y}^n \xrightarrow{\mathcal{D}} \mathbf{Y}$.

Let $N \geq 2$ and fix a path $\gamma \in C^{1\text{-var}}([0, 1], \mathbb{R}^d)$ such that $S_N(\gamma)_{0,1} =: \exp(v)$, where v is an element of $\mathfrak{g}^N(\mathbb{R}^d)$ such that $v^k = 0$ for all $k < N$ (that is, v is in the centre of $\mathfrak{g}^N(\mathbb{R}^d)$).

In this example we wish to consider the random path $\mathbf{Z}^n \in C^{1\text{-var}}([0, 1], \mathbb{R}^d)$ defined by linearly joining the points of $\mathbf{Y}^n$, and, between each linear chord, running along the path $n^{-1/N}\gamma$.

Define the closed subset

$$W := \exp(\mathbb{R}^d) \exp(\mathbb{R}^+ v) = \{\exp(y) \exp(\lambda v) \mid y \in \mathbb{R}^d, \lambda \geq 0\} \subseteq G^N(\mathbb{R}^d).$$

Note that for all $x \in W$ there exist unique $y \in \mathbb{R}^d$ and $\lambda \geq 0$ such that $x = \exp(y) \exp(\lambda v)$.

Define correspondingly $\phi : W \mapsto C^{1\text{-var}}([0, 1], G^N(\mathbb{R}^d))$ by

$$\phi(\exp(y) \exp(\lambda v))_t = \begin{cases} \exp(2ty) & \text{if } t \in [0, 1/2] \\ \exp(y) S_N(\lambda^{1/N} \gamma)_{2t-1} & \text{if } t \in (1/2, 1]. \end{cases}$$

Consider the $G^N(\mathbb{R}^d)$ -valued iid array $X_{nj} := \exp(Y_{nj}) \exp(n^{-1}v)$ and the associated random walk $\mathbf{X}^n$. Observe that $\mathbf{X}^{n,\phi}$ is (a reparametrisation of) the level- N lift of the path $\mathbf{Z}^n$ described above.

Observe that ϕ is endpoint continuous and that there exists $C > 0$ such that for all $\lambda \in \mathbb{R}$

$$\|\lambda^{1/N} \gamma\|_{1\text{-var};[0,1]} = |\lambda|^{1/N} \|\gamma\|_{1\text{-var};[0,1]} \leq C \|\exp(\lambda v)\|,$$

from which it follows that ϕ is also 1-approximating.

We now claim that $\mathbf{X}^n \xrightarrow{\mathcal{D}} \mathbf{X}$ for a Lévy process $\mathbf{X}$ in $G^N(\mathbb{R}^d)$. A straightforward way to show this is to take local coordinates $\sigma_1, \dots, \sigma_d \in C_c^\infty(\mathbb{R}^d)$ so that $\sigma := \sum_{i=1}^d \sigma_i u_i$ is the identity in a neighbourhood of zero. Write the triplet of $\mathbf{Y}$ as (A, B, Π) with respect to $\sigma_1, \dots, \sigma_d$.

Define the function

$$f_n : \mathbb{R}^d \mapsto \mathfrak{g}^N(\mathbb{R}^d), \quad f_n(y) = \xi(e^y e^{v/n}).$$

Note that, since v is in the centre of $\mathfrak{g}^N(\mathbb{R}^d)$, there exists a neighbourhood of zero $V \subset \mathbb{R}^d$ and $n_0 > 0$ such that for all $n > n_0 \geq 0$

$$f_n(y) = \sum_{i=1}^d \sigma_i(y) u_i + \xi(e^{v/n}) + h_n(y),$$

where $h_n \equiv 0$ on V . It follows that $(f_n)_{n \geq 1}$ is uniformly controlled by $1 \wedge |\cdot|^2$ (in fact it is uniformly controlled by any scaling function on $\mathbb{R}^d$, see Section 5.1.5), and thus by Proposition 5.1.8,

$$\lim_{k \rightarrow \infty} \sup_{n \geq 1} k \left| \mathbb{E} [f_n(Y_{k1}) - \xi(e^{v/n})] - Q_n \right| = 0,$$

where

$$Q_n := \sum_{i=1}^d B_i u_i + \int_{\mathbb{R}^d} h_n(y) \Pi(dy) \in \mathfrak{g}^N(\mathbb{R}^d).$$

Observe now that for all $y \in \mathbb{R}^d$, $\lim_{n \rightarrow \infty} h_n(y) = \xi(e^y) - \sigma(y)$, so that by dominated convergence,

$$\lim_{n \rightarrow \infty} Q_n = \sum_{i=1}^d B_i u_i + \int_{\mathbb{R}^d} (\xi(e^y) - \sigma(y)) \Pi(dy) =: Q \in \mathfrak{g}^N(\mathbb{R}^d),$$

from which it follows that

$$\lim_{n \rightarrow \infty} n \mathbb{E} [f_n(Y_{n1}) - f_n(0)] = Q.$$

Since $n f_n(0) = v$ for all n sufficiently large, it follows that

$$D_i := \lim_{n \rightarrow \infty} n \mathbb{E} [\xi_i(X_{n1})]$$

exists for all $i \in \{1, \dots, m\}$.

Furthermore, letting Ξ denote the pushforward of Π by $\exp$, one can show in exactly the same way that

$$\lim_{n \rightarrow \infty} n \mathbb{E} [f(X_{n1})] = \int_{G^N(\mathbb{R}^d)} f(x) \Xi(dx)$$

for every $f \in C_b(G^N(\mathbb{R}^d))$ which is identically zero on a neighbourhood of 1_N , and

$$C_{i,j} := \lim_{n \rightarrow \infty} n \mathbb{E} [\xi_i(X_{n1}) \xi_j(X_{n1})] - \int_{G^N(\mathbb{R}^d)} \xi_i(x) \xi_j(x) \Xi(dx)$$

exists for all $i, j \in \{1, \dots, m\}$. It follows (Theorem 5.1.1) that $\mathbf{X}^n \xrightarrow{\mathcal{D}} \mathbf{X}$ as claimed, where $\mathbf{X}$ is the Lévy process with triplet (C, D, Ξ) .

Likewise, one readily checks that X_{nj} is scaled by θ where $\theta \equiv \sum_{i=1}^d |\xi_i|^2 + \sum_{i \in H_N} |\xi_i|$ in a neighbourhood of 1_N , where we recall $H_N := \{i \in \{1, \dots, m\} \mid \deg(u_i) = N\}$.

It now follows by Theorem 5.4.7 that for all $p > N$, $P_p(\mathbf{X}^{n,\phi}) \xrightarrow{\mathcal{D}} P_p(\mathbf{X}^\phi)$ as $C^{0,p\text{-var}}([0, 1], G^N(\mathbb{R}^d))$ -valued random variables. In particular, taking $N < p < N+1$, it follows that ODEs driven by the random path $\mathbf{Z}^n$ described above, along Lip^γ vector fields, for any $\gamma > N$, converge in law to the corresponding RDE driven by $\mathbf{X}^\phi$ (Corollary 5.4.8).

Chapter 6

Several open problems

We conclude the document with a list of several open problems which we believe to be of interest and which remain a subject of further research for the author. We group the problems with the chapter of the current work to which they are most closely related.

Chapter 2

- **Characteristic function for infinite dimensional V .** In Section 2.3.2, we defined a characteristic function for probability measures on the group-like elements $G(V)$ when V was a finite dimensional space. The finite-dimensionality of V was used only to ensure that $G(V)$ was a Polish space (which remains true whenever V is separable and metrizable), and, more crucially, to show in Theorem 2.3.8 that the points of $G(V)$ were separated by a family of unitary representations. This leads to the following question: if V is a (separable) Banach space, are the points of $G(V)$ still separated by the unitary representations induced by linear maps $M : V \mapsto \mathfrak{u}$? An affirmative answer would allow us to extend the characteristic function to $G(V)$ -valued random variables. This would in turn lead immediately to the same solutions of the moment problem for $G(V)$ -valued random variables as studied in Sections 4.1 and 4.2.
- **Bochner's theorem and the existence moment problem.** In the one dimensional case, Bochner's theorem determines when a function ψ arises as the characteristic function of a probability measure μ on $\mathbb{R}$ (or equivalently, on $G(\mathbb{R})$). In the multidimensional case, we ask the following question: given a map ψ , which sends every linear map $M : \mathbb{R}^d \mapsto \mathfrak{u}(H)$ to an operator in $\mathbf{L}(H)$,

what are necessary and sufficient conditions for there to exist a probability measure μ on $G(\mathbb{R}^d)$ such that $\psi(M) = \mu(M)$?

A related problem is to classify the elements of $P(\mathbb{R}^d) = \prod_{k \geq 0} (\mathbb{R}^d)^{\otimes k}$ which arise as the expected signature of a probability measure on $G(\mathbb{R}^d)$. This is of course the analogue of the classical “existence moment problem” for probability measures on $\mathbb{R}$ (different from the “determinate moment problem” considered in Chapter 4).

A special case for both of these problems, which still remains unsolved, is to classify the elements of $E(\mathbb{R}^d)$ (that is, the elements of $P(\mathbb{R}^d)$ with an infinite radius of convergence) which are the expected signature of a probability measure on $G(\mathbb{R}^d)$.

Chapter 3

- **Recovering solutions to RDEs from the signature.** Recall from Section 3.4 that, due to the work of Hambly and Lyons [30] and later Boedihardjo, Geng, Lyons and Yang [4], the solution map $U_{T \leftarrow 0}^{\mathbf{x}}$ associated to an RDE driven along vector fields $f \in \text{Lip}^\gamma(\mathbb{R}^e, \mathbf{L}(\mathbb{R}^d, \mathbb{R}^e))$ depends only on the signature of $\mathbf{x}$. Equivalently, recalling that the signature determines the solution to a linear RDE in terms of an absolutely convergent series, the result of Boedihardjo et al. [4] states that it is enough to understand the effect of $\mathbf{x}$ on all linear systems to uniquely determine its effect on all *non-linear* systems. As it stands, however, this correspondence is currently only theoretical and brings about the following question: is there an effective method to determine $U_{T \leftarrow 0}^{\mathbf{x}}$ directly from the signature of $\mathbf{x}$?

This question is closely related to the problem of recovering $\mathbf{x}$ itself from its signature, which has been a long standing and difficult problem even in the case of smooth paths, see Lyons and Xu [43], [44].

Chapter 4

- **Densities of solutions to RDEs.** Cass and Friz [8] first gave general conditions under which an RDE driven by a Gaussian rough path along smooth vector fields $f \in C^\infty(\mathbb{R}^e, \mathbf{L}(\mathbb{R}^d, \mathbb{R}^e))$ satisfying Hörmander’s condition admits a density in $\mathbb{R}^e$. Their work involved an elegant combination of Malliavin calculus

and rough paths theory, which has since brought about further developments; one of main applications of the greedy sequence and the function $N_{\kappa,[0,T],p}(\mathbf{x})$ introduced by Cass, Litterer and Lyons [10] has been to show smoothness of densities for RDEs driven by Gaussian rough paths, a result recently achieved by Cass, Hairer, Litterer and Tindel [9].

It is natural to ask whether one can show existence and smoothness of densities for RDEs driven by non-Gaussian processes. A specific problem is provided by Markovian rough paths. It can be shown that for all $N \geq 1$, the level- N lift of a Markovian rough path $\mathbf{X}^a$ always admits a density with respect to the Haar measure on $G^N(\mathbb{R}^d)$. However, to the author's best knowledge, it is an open problem whether the analogous result of Cass and Friz [8] holds in this setting, that is: given smooth vector fields $f \in C^\infty(\mathbb{R}^e, \mathbf{L}(\mathbb{R}^d, \mathbb{R}^e))$ satisfying Hörmander's condition, does the RDE driven by $\mathbf{X}^a$ along f admit a density in $\mathbb{R}^e$? A solution to this problem may help to further identify the important elements needed in the study of densities for RDEs driven by more general processes.

Bibliography

- [1] Richard Arens. The space L^ω and convex topological rings. *Bull. Amer. Math. Soc.*, 52:931–935, 1946.
- [2] Patrick Billingsley. *Convergence of probability measures*. Wiley Series in Probability and Statistics: Probability and Statistics. John Wiley & Sons, Inc., New York, second edition, 1999. A Wiley-Interscience Publication.
- [3] Horatio Boedihardjo and Xi Geng. The uniqueness of signature problem in the non-markov setting. arXiv:1401.6165, July 2014. Preprint.
- [4] Horatio Boedihardjo, Xi Geng, Terry Lyons, and Danyu Yang. The signature of a rough path: Uniqueness. arXiv:1406.7871, August 2014. Preprint.
- [5] V. I. Bogachev. *Measure theory. Vol. I, II*. Springer-Verlag, Berlin, 2007.
- [6] Jean Bretagnolle. p -variation de fonctions aléatoires. II. Processus à accroissements indépendants. In *Séminaire de Probabilités, VI (Univ. Strasbourg, année universitaire 1970-1971)*, pages 64–71. Lecture Notes in Math., Vol. 258. Springer, Berlin, 1972.
- [7] Emmanuel Breuillard, Peter Friz, and Martin Huesmann. From random walks to rough paths. *Proc. Amer. Math. Soc.*, 137(10):3487–3496, 2009.
- [8] Thomas Cass and Peter Friz. Densities for rough differential equations under Hörmander’s condition. *Ann. of Math. (2)*, 171(3):2115–2141, 2010.
- [9] Thomas Cass, Martin Hairer, Christian Litterer, and Samy Tindel. Smoothness of the density for solutions to Gaussian rough differential equations. *Ann. Probab.*, 43(1):188–239, 2015.
- [10] Thomas Cass, Christian Litterer, and Terry Lyons. Integrability and tail estimates for Gaussian rough differential equations. *Ann. Probab.*, 41(4):3026–3050, 2013.

- [11] Thomas Cass and Marcel Ogrodnik. Tail estimates for markovian rough paths. arXiv:1411.5189, May 2015. Preprint.
- [12] Kuo-Tsai Chen. Integration of paths, geometric invariants and a generalized Baker-Hausdorff formula. *Ann. of Math. (2)*, 65:163–178, 1957.
- [13] Kuo-Tsai Chen. Integration of paths—a faithful representation of paths by non-commutative formal power series. *Trans. Amer. Math. Soc.*, 89:395–407, 1958.
- [14] Lawrence J. Corwin and Frederick P. Greenleaf. *Representations of nilpotent Lie groups and their applications. Part I*, volume 18 of *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, Cambridge, 1990. Basic theory and examples.
- [15] Joachim Cuntz. Excision in periodic cyclic theory for topological algebras. In *Cyclic cohomology and noncommutative geometry (Waterloo, ON, 1995)*, volume 17 of *Fields Inst. Commun.*, pages 43–53. Amer. Math. Soc., Providence, RI, 1997.
- [16] A. M. Davie. Differential equations driven by rough paths: an approach via discrete approximation. *Appl. Math. Res. Express. AMRX*, (2):Art. ID abm009, 40, 2007.
- [17] Jan J. Dijkstra and Jan van Mill. Erdős space and homeomorphism groups of manifolds. *Mem. Amer. Math. Soc.*, 208(979):vi+62, 2010.
- [18] Seán Dineen. *Complex analysis in locally convex spaces*, volume 57 of *North-Holland Mathematics Studies*. North-Holland Publishing Co., Amsterdam, 1981. Notas de Matemática [Mathematical Notes], 83.
- [19] Philip Feinsilver. Processes with independent increments on a Lie group. *Trans. Amer. Math. Soc.*, 242:73–121, 1978.
- [20] Peter Friz and Harald Oberhauser. Rough path limits of the Wong-Zakai type with a modified drift term. *J. Funct. Anal.*, 256(10):3236–3256, 2009.
- [21] Peter Friz and Nicolas Victoir. A note on the notion of geometric rough paths. *Probab. Theory Related Fields*, 136(3):395–416, 2006.
- [22] Peter Friz and Nicolas Victoir. On uniformly subelliptic operators and stochastic area. *Probab. Theory Related Fields*, 142(3-4):475–523, 2008.

- [23] Peter K. Friz and Martin Hairer. *A course on rough paths*. Universitext. Springer, Cham, 2014. With an introduction to regularity structures.
- [24] Peter K. Friz and Atul Shekhar. General rough integration, levy rough paths and a levy–kintchine type formula. arXiv:1212.5888, November 2014. Preprint.
- [25] Peter K. Friz and Nicolas B. Victoir. *Multidimensional stochastic processes as rough paths*, volume 120 of *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, Cambridge, 2010.
- [26] Masatoshi Fukushima, Yoichi Oshima, and Masayoshi Takeda. *Dirichlet forms and symmetric Markov processes*, volume 19 of *de Gruyter Studies in Mathematics*. Walter de Gruyter & Co., Berlin, extended edition, 2011.
- [27] Antonio Giambruno and Angela Valenti. On minimal $\ast$ -identities of matrices. *Linear and Multilinear Algebra*, 39(4):309–323, 1995.
- [28] Antonio Giambruno and Mikhail Zaicev. *Polynomial identities and asymptotic methods*, volume 122 of *Mathematical Surveys and Monographs*. American Mathematical Society, Providence, RI, 2005.
- [29] Alexandre Grothendieck. Produits tensoriels topologiques et espaces nucléaires. *Mem. Amer. Math. Soc.*, 1955(16):140, 1955.
- [30] Ben Hambly and Terry Lyons. Uniqueness for the signature of a path of bounded variation and the reduced path group. *Ann. of Math. (2)*, 171(1):109–167, 2010.
- [31] Keisuke Hara and Masanori Hino. Fractional order Taylor’s series and the neo-classical inequality. *Bull. Lond. Math. Soc.*, 42(3):467–477, 2010.
- [32] J. R. Holub. Tensor product mappings. *Math. Ann.*, 188:1–12, 1970.
- [33] G. A. Hunt. Semi-groups of measures on Lie groups. *Trans. Amer. Math. Soc.*, 81:264–293, 1956.
- [34] Olav Kallenberg. *Foundations of modern probability*. Probability and its Applications (New York). Springer-Verlag, New York, 1997.
- [35] Alexander S. Kechris. *Classical descriptive set theory*, volume 156 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, 1995.

- [36] Anthony W. Knap. *Lie groups beyond an introduction*, volume 140 of *Progress in Mathematics*. Birkhäuser Boston, Inc., Boston, MA, 1996.
- [37] Konrad Knopp. Ein einfaches Verfahren zur Bildung stetiger nirgends differenzierbarer Funktionen. *Math. Z.*, 2(1-2):1–26, 1918.
- [38] Vassili N. Kolokoltsov. *Markov processes, semigroups and generators*, volume 38 of *de Gruyter Studies in Mathematics*. Walter de Gruyter & Co., Berlin, 2011.
- [39] Hiroshi Kunita. Some problems concerning Lévy processes on Lie groups. In *Stochastic analysis (Ithaca, NY, 1993)*, volume 57 of *Proc. Sympos. Pure Math.*, pages 323–341. Amer. Math. Soc., Providence, RI, 1995.
- [40] Ming Liao. *Lévy processes in Lie groups*, volume 162 of *Cambridge Tracts in Mathematics*. Cambridge University Press, Cambridge, 2004.
- [41] Eugene Lukacs. *Characteristic functions*. Hafner Publishing Co., New York, 1970. Second edition, revised and enlarged.
- [42] Terry Lyons and Zhongmin Qian. *System control and rough paths*. Oxford Mathematical Monographs. Oxford University Press, Oxford, 2002.
- [43] Terry Lyons and Weijun Xu. Hyperbolic development and inversion of signature. arXiv:1507.00286, July 2015. Preprint.
- [44] Terry Lyons and Weijun Xu. Inverting the signature of a path. arXiv:1406.7833, July 2015. Preprint.
- [45] Terry J. Lyons. Differential equations driven by rough signals. *Rev. Mat. Iberoamericana*, 14(2):215–310, 1998.
- [46] Terry J. Lyons, Michael Caruana, and Thierry Lévy. *Differential equations driven by rough paths*, volume 1908 of *Lecture Notes in Mathematics*. Springer, Berlin, 2007.
- [47] Anastasios Mallios. *Topological algebras. Selected topics*, volume 124 of *North-Holland Mathematics Studies*. North-Holland Publishing Co., Amsterdam, 1986. Notas de Matemática [Mathematical Notes], 109.
- [48] Martynas Manstavičius. p -variation of strong Markov processes. *Ann. Probab.*, 32(3A):2053–2066, 2004.

- [49] E. J. McShane. Stochastic differential equations and models of random processes. In *Proceedings of the Sixth Berkeley Symposium on Mathematical Statistics and Probability (Univ. California, Berkeley, Calif., 1970/1971), Vol. III: Probability theory*, pages 263–294. Univ. California Press, Berkeley, Calif., 1972.
- [50] Christophe Reutenauer. *Free Lie algebras*, volume 7 of *London Mathematical Society Monographs. New Series*. The Clarendon Press Oxford University Press, New York, 1993.
- [51] Walter Rudin. *Functional analysis*. McGraw-Hill Book Co., New York-Düsseldorf-Johannesburg, 1973. McGraw-Hill Series in Higher Mathematics.
- [52] Helmut H. Schaefer. *Topological vector spaces*. Springer-Verlag, New York-Berlin, 1971. Third printing corrected, Graduate Texts in Mathematics, Vol. 3.
- [53] Daniel W. Stroock and S. R. S. Varadhan. Diffusion processes with continuous coefficients. I. *Comm. Pure Appl. Math.*, 22:345–400, 1969.
- [54] Daniel W. Stroock and S. R. S. Varadhan. Diffusion processes with continuous coefficients. II. *Comm. Pure Appl. Math.*, 22:479–530, 1969.
- [55] Daniel W. Stroock and S. R. S. Varadhan. Limit theorems for random walks on Lie groups. *Sankhyā Ser. A*, 35(3):277–294, 1973.
- [56] Héctor J. Sussmann. Limits of the Wong-Zakai type with a modified drift term. In *Stochastic analysis*, pages 475–493. Academic Press, Boston, MA, 1991.
- [57] François Trèves. *Topological vector spaces, distributions and kernels*. Academic Press, New York, 1967.
- [58] Christian Valqui. Universal extension and excision for topological algebras. *K-Theory*, 22(1-2):145–160, 2001.
- [59] David R. E. Williams. Path-wise solutions of stochastic differential equations driven by Lévy processes. *Rev. Mat. Iberoamericana*, 17(2):295–329, 2001.