Mini-course on Rough Paths (TU Wien, 2009)

P.K. Friz, Last update: 20 Jan 2009

Contents

CHAPTER 1

Rough Paths

1. On control ODEs

[1]Given $\gamma > 0$ write $\gamma = \lfloor \gamma \rfloor + \{\gamma\}$ with $\lfloor \gamma \rfloor \in \mathbb{N} \cup \{0\}$ and $\{\gamma\} \in (0,1]$. Call $Lip^{\gamma}(\mathbb{R}^e)$ the class of $\lfloor \gamma \rfloor$-times differentiable vector fields on $\mathbb{R}^e$ with all derivatives bounded and $\lfloor \gamma \rfloor^{th}$-derivative $\{\gamma\}$-Hölder continuous.

Given vector fields $V_1, ..., V_d \in Lip^1(\mathbb{R}^e)$ (= bounded & Lipschitz continuous vector fields) and $x \in C^1\left([0,T], \mathbb{R}^d\right)$, at least piecewise, then basic ODE results tell us that there exists a unique solution $y = \pi(0, y_0; x)$ to the (control) ODE

$$dy_t = V_i(y_t)\, dx_t^i \equiv V(y_t)\, dx_t, \;\; t \in [0,T]. \tag{1.1}$$

started at y_0.

Notation 1. *We write* $|dx_r| := |\dot{x}_r|\, dr$ *and* $y_{s,t} := y_t - y_s$ *for path-increments. Also,* $|V(.)| := \max\{|V_i(.)| : i = 1, ..., d\}$ *and* $|x|_{Lip,[a,b]} := \sup_{a \le s < t \le b} \frac{|x_{s,t}|}{|t-s|}$

Proposition 2. *There exists a constant* C_2 *so that*

$$|\pi(0, y_0; x)|_{Lip,[s,t]} \le C_2 |x|_{Lip,[s,t]}.$$

Proof. We have

$$\begin{aligned} |y_{s,t}| &= \left| \int_s^t V(y_r)\, dx_r \right| \\ &\le \int_s^t |V(y_r)|\, |dx_r| \\ &\le |V|_\infty |x|_{Lip,[s,t]} |t-s|. \end{aligned}$$

□

The *Lip*-(semi-)norm here can not be replaced by sup-(semi-)norm. That is, in general $\nexists C$ such that $|y_{s,t}| \le C \sup_{0 \le s < t \le T} |x_{s,t}|$.

Exercise 3. $V_1(y) = \left(1, 0, -\frac{1}{2}y^2\right), V_2(y) = \left(0, 1, \frac{1}{2}y^1\right)$ *with* $y = \left(y^1, y^2, y^3\right) \in \mathbb{R}^3$. *Let* $x_t = \frac{1}{n}\exp\left(2\pi\sqrt{-1}n^2 t\right) \in \mathbb{C} \cong \mathbb{R}^2$ *and* $y_0 = 0 \in \mathbb{R}^3$. *Clearly,* $x \to 0$ *uniformly on* $[0,1]$. *Show that* $y_1^3 \to$ *area of the unit circle as* $n \to \infty$.

Lemma 4. *For every Lipschitz continuous path* x *there exists a sequence of piecewise smooth* x^n *such that* $x^n \to x$ *pointwise and*

$$\sup_n |x^n|_{Lip,[0,T]} \le |x|_{Lip,[0,T]} < \infty.$$

[1]Corrections and comments to P.K.Friz@statslab.cam.ac.uk

PROOF. Fix a dissecion $D = (t_j)$ of $[0,T]$. We construct an approximation x^D by connecting the points x_{t_i} with geodesics in $\mathbb{R}^d$. In other words, x^D is the piecewise linear approximation to x. When $t_i \leq s \leq t \leq t_{i+1}$ for some i, then

$$x_{s,t}^D = \frac{t-s}{t_{i+1}-t_i} x_{t_i,t_{i+1}}$$

and hence $\left|x_{s,t}^D\right| \leq |t-s| \, |x|_{Lip,[0,T]}$. When $t_i \leq s \leq t_{i+1} \leq t_j \leq t \leq t_{j+1}$ then

$$\begin{aligned} \left|x_{s,t}^D\right| &\leq \left|x_{s,t_{i+1}}^D\right| + \left|x_{t_{i+1},t_j}\right| + \left|x_{t_j,t}^D\right| \\ &\leq \left(|t_{i+1}-s| + |t_j - t_{i+1}| + \left|t_{i_j} - t\right|\right) |x|_{Lip,[0,T]} \\ &\leq |t-s| \, |x|_{Lip,[0,T]} . \end{aligned}$$

Clearly, for any sequence D^n with mesh $|D^n| \to 0$ we have pointwise convergence.
□

LEMMA 5. *Let x^n be Lipschitz with $x^n \to x$ pointwise and* $\sup_n |x^n|_{Lip,[0,T]} = C < \infty$. *Then x is Lipschitz, $x^n \to x$ uniformly and* $|x|_{Lip,[0,T]} \leq \liminf_{n\to\infty} |x^n|_{Lip,[0,T]} \leq C$.

PROOF. Let t_D denote the closest (left) element in some dissection D for a given $t \in [0,T]$. We show that x^n is Cauchy w.r.t. uniform norm (which implies that x is the uniform limit). Indeed

$$\begin{aligned} |x_t^n - x_t^m| &\leq \left|x_{t_D}^n - x_{t_D}^m\right| + \left|x_t^n - x_{t_D}^n\right| + \left|x_t^m - x_{t_D}^m\right| \\ &\leq \epsilon/2 + 2C\,|D| \\ &\leq \epsilon \end{aligned}$$

for $n,m \geq N(\epsilon)$ and D with mesh $|D| \leq \epsilon/4C$. For the second statement, note that, along a subsequence, $|x^{n_k}|_{Lip,[0,T]} \to \liminf_{n\to\infty} |x^n|_{Lip,[0,T]} = C'$. Then

$$\begin{aligned} |x_{s,t}| &\leq \left|x_{s,t}^{n_k}\right| + \left|x_{s,t} - x_{s,t}^{n_k}\right| \\ &\leq |x^{n_k}|_{Lip,[0,T]} \, |t-s| + \epsilon \end{aligned}$$

for $k \geq k(\epsilon)$ and some fixed $\epsilon > 0$. Sending $k \to \infty$ we find that

$$|x_{s,t}| \leq C' \, |t-s| + \epsilon$$

and sending $\epsilon \to 0$ we find

$$|x_{s,t}| \leq C' \, |t-s| \implies |x|_{Lip,[0,T]} \leq C'.$$

□

LEMMA 6. *Let x be Lipschitz and x^n be (piecewise) C^1 with $x^n \to x$ pointwise and* $\sup_n |x^n|_{Lip,[0,T]} = C < \infty$. *Let (f^n) be a sequence of continuous paths, uniformly convergent on $[0,T]$ to some (continuous) path f. Then*

$$\int_0^T f^n dx^n \to \int_0^T f dx.$$

The integral on the right hand side can be viewed, equivalently, as Riemann-Stieltjes integral or as Lebesgue integral with integrand $f\,\dot{x}$, noting that Lipschitz paths are absolutely continuous.

Proof. Consider $f^n \equiv f$ as a first case. Fix $\epsilon > 0$. Since f is continuous on the compact $[0,T]$ it is uniformly continuous. Therefore, $\exists \delta > 0$ and finitely many intervals $I_j = (t_j, t_{j+1}]$ of length δ or less so that $\cup I_j = (0,T]$ and so that the oscillation of the continuous function f in each I_j is less than ϵ. Define the step function

$$h = \sum f(t_j) 1_{(t_j, t_{j+1}]}$$

and note $|f - h|_\infty \leq \epsilon$. From Lemma 5, x is Lipschitz and $|x|_{Lip,[0,T]} \leq C$. This implies $|dx_t| \leq C dt$ and

$$\left| \int_0^T f dx - \int_0^T h dx \right| \leq \int_0^T |f_t - h_t| \, |dx_t| \leq CT\epsilon.$$

Similarly, for any n,

$$\left| \int_0^T f dx^n - \int_0^T h dx^n \right| \leq CT\epsilon.$$

From Lemma 5, $x^n \to x$ uniformly and we have

$$\left| \int_0^T h dx^n - \int_0^T h dx \right| \leq \sum_j |f_{t_j}| \left| x^n_{t_j, t_{j+1}} - x_{t_j, t_{j+1}} \right| \to 0$$

as $n \to \infty$. By the triangle inequality and the two preceding estimates,

$$\lim\sup_{n\to\infty} \left| \int_0^T f dx^n - \int_0^T f dx \right| \leq 2CT\epsilon$$

and this hold for any $\epsilon > 0$. Sending $\epsilon \to 0$ proves the first case.The case of a sequence f^n is another (simple) application of the triangle inequality, simply insert & subtract the integral of $f dx^n$. □

Proposition 7. *Let x be a Lipschitz path and x^n be (piecewise) C^1 on $[0,T]$. Assume $x^n \to x$ pointwise and $\sup_n |x^n|_{Lip,[0,T]} < \infty$. Set $y^n := \pi(0, y_0; x^n)$. Then $\sup_n |y^n|_{Lip,[0,T]} < \infty$ and y^n converges pointwise to a unique limit, say $y = y_t$. Moreover y_t satisfies*

$$y_t - y_0 = \int_0^t V_i(y_s) dx^i_s = \int_0^t V_i(y_s) \dot{x}^i_s ds.$$

We also write $\pi(0, y_0; x)$ for this unique limit y.

Proof. From Proposition 2, $\sup_n |\pi(0, y_0; x^n)|_{Lip} \leq C_2 \sup_n |x^n|_{Lip,[0,T]} < \infty$ and hence (y^n) is equicontinuous. It is also bounded in uniform norm

$$\sup_{t\in[0,T]} |y^n_t| \leq |y_0| + T \sup_n |\pi(0, y_0; x^n)|_{Lip}.$$

Arzela-Ascoli's theorem tells us that, possibly after passing to a sub-sequence, $y_n = \pi(0, y_0; x^n)$ converges uniformly on $[0,T]$ to some continuous path y. By definition of y^n,

$$y^n_t = y_0 + \int_0^t V(y^n_s) dx^n_s.$$

Using Lipschitz regularity of V we see that $V(y^n) \to V(y)$ uniformly. Since by hypothesis $x^n \to x$ pointwise with $|x^n|_{Lip,[0,T]} \leq \sup_n |x^n|_{Lip,[0,T]} < \infty$ lemma 6 shows that as $n \to \infty$

$$\begin{aligned} y_t &= y_0 + \int_0^t V(y_s)\, dx_s \\ &= y_0 + \int_0^t V(y_s)\, \dot{x}_s ds. \end{aligned}$$

Now assume $\tilde{y}$ is another limit point of the sequence $\{\pi(0, y_0; x^n)\}$and set $\delta = y - \tilde{y}$. Then

$$|\delta_t| \leq \|V\|_{Lip^1(\mathbb{R}^e)} \int_0^t |\delta_s|\, |dx_s|$$

and from Gronwall's lemma, $y = \tilde{y}$. The result follows. □

Exercise 8. *Let x, x^n be a Lipschitz paths on $[0,T]$. Assume $x^n \to x$ pointwise and $\sup_n |x^n|_{Lip,[0,T]} < \infty$. Set $y^n = \pi(0, y_0; x^n)$ and $y = \pi(0, y_0; x)$, as constructed in the previous proposition. Show that $\sup_n |y^n|_{Lip,[0,T]} < \infty$ and $y^n \to y$ uniformly on $[0,T]$. Hint: use Lemma 4.*

It is convenient (and standard in differential geometry) to identify a vectorfield $V(y)$ with the first order differential operator $V^j(y) \frac{\partial}{\partial y^j} \equiv V^j(y)\, \partial_j$.

Lemma 9. *Let x be a Lipschitz path on $[0,T]$ and $y = \pi(0, y_0; x)$. Let f be a C^1-function on $\mathbb{R}^e$. Then*

$$f(y_t) = f(y_s) + \sum_{i=1}^d \int_s^t V_i f(y_r)\, dx_r^i.$$

Proof. When x is piecewise smooth then this is just the fundamental theorem of calculus. When x is only Lipschitz, approximate using the results above. The details are left as exercise. □

Remark 10. *$V_{i_1} \cdots V_{i_N}$ makes sense as N^{th} order differential operator provided the vector fields are $(N-1)$ times differentiable.*

Exercise 11. *(i) Check that $[V_1, V_2] := V_1 V_2 - V_2 V_1$ is in fact a 1^{st} order differential operator (and hence a vector field). It is called the Lie bracket of the vector fields V_1 and V_2.*
(ii) Compute $[V_1, V_2]$ for the vector fields given in exercise 3.
(iii) Give a geometric interpretation of the Lie bracket.

1.1. Euler scheme of order N. Define $T^N(\mathbb{R}^d) = \oplus_{k=0}^N (\mathbb{R}^d)^{\otimes k}$. For instance, $(\mathbb{R}^d)^{\otimes 2}$ is just a $(d \times d)$-matrix. By convention $(\mathbb{R}^d)^{\otimes 0} \equiv \mathbb{R}$. Let x be an $\mathbb{R}^d$-valued Lipschitz path and define k^{th} iterated integrals of the path segment $x|_{[s,t]}$ as

$$\mathbf{g}^{k,i_1,\cdots,i_k} := \int_s^t \int_s^{u_k} \ldots \int_s^{u_2} dx_{u_1}^{i_1} \ldots dx_{u_k}^{ik}.$$

and so that $\mathbf{g}^k = \left(\mathbf{g}^{k,i_1,\cdots,i_k}\right)_{i_1,\cdots,i_k \in \{1,\ldots,d\}} \in (\mathbb{R}^d)^{\otimes k}$. For later convenience set $\mathbf{g}^0 = 1 \in (\mathbb{R}^d)^{\otimes 0} \equiv \mathbb{R}$. We then define the (step-N) signature of the path segment $x|_{[s,t]}$ as

$$\mathbf{x}_{s,t} \equiv S_N(x)_{s,t} \equiv 1 + \sum_{k=1}^{N} \mathbf{g}^k \in T^N\left(\mathbb{R}^d\right).$$

Let I be the identity function on $\mathbb{R}^e$.

DEFINITION 12. *Let $N \in \mathbb{N}$. Given Lip^N-vector fields $V = (V_1, ..., V_d)$ on $\mathbb{R}^e$, $\mathbf{g} \in T^N\left(\mathbb{R}^d\right)$ and $y \in \mathbb{R}^e$ we call*

$$\mathcal{E}_{(V)}(y, \mathbf{g}) := \sum_{k=1}^{N} \sum_{\substack{i_1,...,i_k \\ \in\{1,...,d\}}} V_{i_1} \cdots V_{i_k} I(y)\, \mathbf{g}^{k,i_1,\cdots,i_k},$$

the (increment of) the step-N Euler scheme. In a typical application, $\mathbf{g}$ is the (step-N) signature of a path segment $x|_{[s,t]}$ and we call

$$\mathcal{E}_{(V)}\left(y_s, S_N(x)_{s,t}\right)$$

the (inrement of) the step-N Euler scheme for $dy = V(y)\,dx$ over the time-interval $[s,t]$.

LEMMA 13 (Euler ODE estimate). *Let $V = (V_i)_{1\leq i\leq d}$ be a collection of vector fields in* $\mathrm{Lip}^N(\mathbb{R}^e)$ *and $x : [s,t] \to \mathbb{R}^d$ be a Lipschitz path. Then there exists a constant $C_{13} = C_{13}(N)$ such that for all $s < t$,*

$$\left|\pi_{(V)}(s, y_s; x)_{s,t} - \mathcal{E}_{(V)}\left(y_s, S_N(x)_{s,t}\right)\right| \leq C_{13}\left(|V|_{\mathrm{Lip}^{\gamma-1}} \int_s^t |dx_r|\right)^{N+1}. \tag{1.2}$$

PROOF. We first show that

$$\begin{aligned} & y_{s,t} - \mathcal{E}_{(V)}\left(y_s, S_N(x)_{s,t}\right) \\ = & \sum_{\substack{i_1,...,i_N \\ \in\{1,...,d\}}} \int_{s<r_1<...<r_N<t} \left[V_{i_1} \cdots V_{i_N} I(y_{r_1}) - V_{i_1} \cdots V_{i_N} I(y_s)\right] dx_{r_1}^{i_1} \cdots dx_{r_N}^{i_N}. \end{aligned}$$

To this end, consider a smooth function f and note that for any $k \leq N-1$, $V_{i_1} \cdots V_{i_k} f \in C^1$. By iterated use of the fundamental theorem of calculus,

$$\begin{aligned} f(y_t) &= f(y_s) + \sum_{k=1}^{N-1} \sum_{\substack{i_1,...,i_k \\ \in\{1,...,d\}}} \int_{s<r_1<...<r_k<t} V_{i_1} \cdots V_{i_k} f(y_s)\, dx_{r_1}^{i_1} \cdots dx_{r_k}^{i_k} \\ &\quad + \sum_{\substack{i_1,...,i_k \\ \in\{1,...,d\}}} \int_{s<r_1<...<r_N<t} V_{i_1} \cdots V_{i_N} f(y_{r_1})\, dx_{r_1}^{i_1} \cdots dx_{r_N}^{i_N}. \end{aligned}$$

and the claim follows from specializing to $f = I$, the identity function. For the proof of (1.2) we momentarily allow all constants to depend on V. Clearly,

$$|y_{s,t}| = \left|\int_s^t V(y)\,dx\right| \leq c_1 \int_s^t |dx_r|.$$

From Lip^N-regularity of the vector fields, $V_{i_1}..V_{i_N}I(\cdot)$ is Lipschitz continuous. Hence, for all $r \in [s,t]$,

$$|V_{i_1}..V_{i_N}I(y_r) - V_{i_1}..V_{i_N}I(y_s)| \leq c_2 \left(\int_s^t |dx_r| \right).$$

and after integration,

$$\left| \int_{s<r_1<\ldots<r_N<t} [V_{i_1}..V_{i_N}I(y_{r_1}) - V_{i_1}..V_{i_N}I(y_s)]\, dx_{r_1}^{i_1} \cdots dx_{r_N}^{i_N} \right| \leq c_3 \left(\int_s^t |dx_r| \right)^{N+1}.$$

Summation over the indices finishes the estimate. At last, a scaling argument shows that

$$c_3 = c_4 \left(|V|_{\mathrm{Lip}^N} \right)^{N+1}$$

for some constant c_4 independent of V. Indeed it suffices to rewrite

$$dy = V(y)\,dx = \hat{V}(y)\,d\hat{x}$$

with $\hat{V} = V/|V|_{\mathrm{Lip}^N}$ and $\hat{x} = x\,|V|_{\mathrm{Lip}^N}$. In particular, $\left|\hat{V}\right|_{\mathrm{Lip}^N} \leq 1$ which is enough to rerun the above estimates without picking up any dependence on V. □

2. The algebra of iterated integrals

There are algebraic relations between higher iterated integrals.

Example 14. *Consider a Lipschitz path x started at $x_0 = 0 \in \mathbb{R}^d$. Since $x^i dx_r^j + x^j dx_r^i = d\left(x^i x^j\right)$ integration yields $\int_0^t x_r^i dx_r^j + \int_0^t x_r^j dx_r^i = x_t^i x_t^j$. Similarly, for arbitrary x_0,*

$$\mathbf{x}_{s,t}^{2;i,j} + \mathbf{x}_{s,t}^{2;j,i} = \int_s^t x_{s,r}^i dx_r^j + \int_s^t x_{s,r}^j dx_r^i = x_{s,t}^i x_{s,t}^j.$$

Life is simpler without indices! Recall that the symmetric part of a matrix A is given by $Sym(A) = \frac{1}{2}\left(A + A^T\right)$. Then

$$Sym\left(\mathbf{x}_{s,t}^2\right) = \frac{1}{2}\mathbf{x}_{s,t}^1 \otimes \mathbf{x}_{s,t}^1.$$

Note that $\otimes$ turns two vectors (1-tensors) v, w into a matrix $v \otimes w$ (2-tensor). Similarly, an i-tensor x and a $(k-i)$-tensor y give us a k-tensor $x \otimes y$. Let $|\cdot| \equiv |\cdot|_{(\mathbb{R}^d)^{\otimes k}}$ denote Euclidean norm on $\left(\mathbb{R}^d\right)^{\otimes k}$. Namely, for $z = \left(z^{i_1,\cdots,i_k}\right)_{i_1,\cdots,i_k \in \{1,\ldots,d\}}$ we have

$$|z|_{(\mathbb{R}^d)^{\otimes k}}^2 = \sum_{i_1,\cdots,i_k \in \{1,\ldots,d\}} \left|z^{i_1,\cdots,i_k}\right|^2.$$

The resulting family of norms $|\cdot|_{(\mathbb{R}^d)^{\otimes k}}; k = 1,2,3,\ldots$ is *compatible* in the sense that

$$|x \otimes y|_{(\mathbb{R}^d)^{\otimes k}} \leq |x|_{(\mathbb{R}^d)^{\otimes i}}\, |y|_{(\mathbb{R}^d)^{\otimes (k-i)}}.$$

Usually, there is no confusion in which tensor space computations take place and we write $|\cdot|$ instead of $|\cdot|_{(\mathbb{R}^d)^{\otimes k}}$. These norms easily lead to a norm on $T^N\left(\mathbb{R}^d\right)$. We set

$$|\mathbf{g}|^2 := \sum_{i=0}^N \left|\mathbf{g}^i\right|^2 \text{ where } \mathbf{g}^i \in \left(\mathbb{R}^d\right)^{\otimes i}.$$

The set $T_1^N(\mathbb{R}^d) \equiv \{\mathbf{g} \in T^N(\mathbb{R}^d) : \mathbf{g}^0 = 1\}$ is a group under *truncated tensor multiplication* which we now introduce. If $\mathbf{g} = 1 + \mathbf{g}^1 + ... + \mathbf{g}^N \equiv 1 + \tilde{\mathbf{g}}$ and similar for $\mathbf{h}$ then for $k = 0, ..., N$

$$(\mathbf{g} \otimes \mathbf{h})^k = \sum_{i=0}^{k} \mathbf{g}^i \otimes \mathbf{h}^{k-i}.$$

(Note that, a priori, $\mathbf{g}^N \otimes \mathbf{h}^1$ is a $(N+1)$-tensor whereas our definition sets $(N+1)$- and all higher tensor to zero. Hence why we call it *truncated*.)

The neutral element is $e = 1 = 1 + 0 + ... + 0$ and the inverse is given by the usual power series calculus

$$(1 + \tilde{\mathbf{g}})^{-1} = 1 - \tilde{\mathbf{g}} + \tilde{\mathbf{g}}^{\otimes 2} - ...$$

(with truncation beyond level N). All group operations are smooth (after all, they are given by polynomials in the coordinates) and we see that $T_1^N(\mathbb{R}^d)$ is actually a Lie group[2].

If $\mathbf{g}$ happens to be the signature of a Lipschitz path, say $\mathbf{g} = S_N(x)_{0,1}$ we may ask for the signature of the scaled path, say $S_N(\lambda x)_{0,1}$ with $\lambda \in \mathbb{R}$. This induces the *dilatation* map

$$\delta_\lambda : (\mathbf{g}^k) \mapsto \left(\lambda^k \mathbf{g}^k\right), \qquad \lambda \in \mathbb{R}.$$

The Lie algebra[3] of $T_1^N(\mathbb{R}^d)$ can be identified with

$$T_0^N(\mathbb{R}^d) \equiv \{\tilde{\mathbf{g}} \in T^N(\mathbb{R}^d) : \tilde{\mathbf{g}}^0 = 0\}, \ \left[\tilde{\mathbf{g}}, \tilde{\mathbf{h}}\right] = \tilde{\mathbf{g}} \otimes \tilde{\mathbf{h}} - \tilde{\mathbf{h}} \otimes \tilde{\mathbf{g}}$$

and the exponential map with $\exp : T_0^N(\mathbb{R}^d) \to T_1^N(\mathbb{R}^d)$,

$$\exp(\tilde{\mathbf{g}}) = 1 + \tilde{\mathbf{g}} + \frac{1}{2!}\tilde{\mathbf{g}}^{\otimes 2} + ...$$

(again, with truncation beyond level N). We define $\log : T_1^N(\mathbb{R}^d) \to T_0^N(\mathbb{R}^d)$ by

$$\log(\mathbf{g}) = \log(1 + \tilde{\mathbf{g}}) = \tilde{\mathbf{g}} - \frac{1}{2}\tilde{\mathbf{g}}^{\otimes 2} + ...$$

Due to truncation, log is globally defined and it is easily checked that $\log = \exp^{-1}$. Thus

$$\text{Lie algebra } T_0^N(\mathbb{R}^d) \underset{\log}{\overset{\exp}{\rightleftarrows}} \text{Lie group } T_1^N(\mathbb{R}^d).$$

EXERCISE 15. *(i) Show that* $\exp(\tilde{\mathbf{g}})^{-1} = \exp(-\tilde{\mathbf{g}})$. *(ii) Given a straight path segment:* $x : [a,b] \mapsto x_0 + \frac{t-a}{b-a}v$ *with* $x_0, v \in \mathbb{R}^d$ *show that*

$$S_N(x)_{a,b} = \exp(v).$$

(iii) Conclude that $S_N(\overleftarrow{x})_{a,b} = S_N(x)_{a,b}^{-1}$ *where* $\overleftarrow{x}$ *is* x *run backwards in time.*

---

[2] A Lie group is a set which is both a group and a smooth manifold such that the group operations are smooth.

[3] A Lie algebra is a vector space L for which is given a bilinear function from $L \times L$ to L, called bracket, and denoted by $[\cdot,\cdot]$ which satisfies (1) $[a,a] = 0$ and (2) $[a,[b,c]] + [c,[a,b]] + [b,[c,a]] = 0$ for all $a,b,c \in L$.

To every Lie group G there is an associated Lie algebra. As vector space, it is usually identified with $\mathcal{T}_e G$, the tangent space of the Lie group at the unit element.

PROPOSITION 16. *Let $x:[0,T]\to\mathbb{R}^d$ be Lipschitz continuous with (step-N) signatures $\mathbf{x}_{s,t}=S_N(x)_{s,t}$. Then*

$$S_N(x)_{s,t}\otimes S_N(x)_{t,u}=S_N(x)_{s,u}. \tag{2.1}$$

PROOF. $S_N(x)^k_{s,u}$ is defined as integral over $\Delta^{(k)}_{s,u}\equiv\{s<r_1<...<r_k<u\}$. For fixed $t\in(s,u)$, ignoring sets of zero k-dimensional Lebesgues measure, $\Delta^{(k)}_{s,u}$ equals the disjoint union $A_0\cup...\cup A_k$ with

$$\begin{aligned} A_j &= \{s<r_1<...<r_j<t<r_{j+1}<....<r_k<u\} \\ &= \Delta^{(j)}_{s,t}\times\Delta^{(k-j)}_{t,u} \end{aligned}$$

Then

$$\begin{aligned} S_N(x)^{k;i_1,\dots,i_k}_{s,u} &= \sum_{j=0}^{k}\int_{A_j}dx^{i_1}...dx^{i_k} \\ &= \sum_{j=0}^{k}\int_{\Delta^{(j)}_{s,t}}dx^{i_1}...dx^{i_j}\int_{\Delta^{(k-j)}_{t,u}}dx^{i_{j+1}}...dx^{i_k} \\ &= \sum_{j=0}^{k}S_N(x)^j_{s,t}\otimes S_N(x)^{k-j}_{t,u}. \end{aligned}$$

□

REMARK 17. *Differently put, the signature of the path segment $x|_{[s,t]}$ tensor-multiplied by the signature of the path segment $x|_{[t,u]}$ equals the signature of their concatenation (= the path segment $x|_{[s,u]}$).*

It is a general fact from abstract Lie group theory that the Lie algebra describes the Lie-group in a neighbourhood of the unit element. More precisely, the exponential map is a diffeomorphism of a neighbourhood of 0 in the Lie algebra onto a neighbourhood of e in the Lie group. Even better, the group multiplication is (locally) expressed in terms of the brackets. This is the famous *Campbell-Baker-Hausdorff* formula. In our setting the CBH-formula holds globally.

PROPOSITION 18. *Let $\tilde{\mathbf{g}},\tilde{\mathbf{h}}\in T_0^N(\mathbb{R}^d)$. Then*

$$\exp\tilde{\mathbf{g}}\otimes\exp\tilde{\mathbf{h}}=\exp\left(\tilde{\mathbf{g}}+\tilde{\mathbf{h}}+\frac{1}{2}\left[\tilde{\mathbf{g}},\tilde{\mathbf{h}}\right]+...\right) \tag{2.2}$$

where ... stands for terms involving iterated brackets such as $\left[\tilde{\mathbf{g}},\left[\tilde{\mathbf{g}},\tilde{\mathbf{h}}\right]\right]$. (From definition of $[\cdot,\cdot]$, all expressions involving N or more brackets are zero.)

PROOF. The exercise below discusses $N=1,2$ directly. For the general case, let us accept the CBH-formula from Lie-group theory: then we know that (2.2) holds for $\tilde{\mathbf{g}},\tilde{\mathbf{h}}$ small enough. However, due to truncation beyond level N, both sides are actually polynomials in the coordinates $\tilde{\mathbf{g}}^{k;i_1,\dots,i_k}$ and $\tilde{\mathbf{h}}^{k;i_1,\dots,i_k}$ and hence defined everywhere. By analyticity, both sides coincide. □

EXERCISE 19 (Campbell-Baker-Hausdorff). *Check (2.2) for $N=1,2$.*

EXERCISE 20 (Zassenhaus formula). *Assume $N=2$. Check that*

$$\exp\left(\tilde{\mathbf{g}}+\tilde{\mathbf{h}}\right)=\exp\left(\tilde{\mathbf{g}}\right)\otimes\exp(\tilde{\mathbf{h}})\otimes\exp\left(-\frac{1}{2}\left[\tilde{\mathbf{g}},\tilde{\mathbf{h}}\right]\right)$$

SOLUTION 21. *It suffices to multiply both sides of (2.2 with* $\exp\left(-\frac{1}{2}\left[\tilde{\mathbf{g}},\tilde{\mathbf{h}}\right]\right)$.

EXERCISE 22. *Assume $N=2$. Given $a,b\in\mathbb{R}^d\subset\mathbb{R}^d\oplus\left[\mathbb{R}^d,\mathbb{R}^d\right]\oplus\ldots$ .Show that*

$$\exp\left(a\right)\otimes\exp\left(b\right)\otimes\exp\left(-a\right)\otimes\exp\left(-b\right)=\exp\left(\left[a,b\right]\right)$$

SOLUTION 23. *From above: $e^{\frac{1}{2}[a,b]}=e^ae^be^{-(a+b)}$ and similarly $e^{\frac{1}{2}[a,b]}=e^{a+b}e^{-a}e^{-b}$. Multiplication yields*

$$e^{[a,b]}=e^ae^be^{-a}e^{-b}.$$

We define

$$\begin{aligned} L &= L^N\left(\mathbb{R}^d\right)=\mathbb{R}^d\oplus\left[\mathbb{R}^d,\mathbb{R}^d\right]\oplus\left[\mathbb{R}^d,\left[\mathbb{R}^d,\mathbb{R}^d\right]\right]\oplus\ldots\subset T_0^N\left(\mathbb{R}^d\right), \\ G &= G^N\left(\mathbb{R}^d\right)=\exp\left(L^N\left(\mathbb{R}^d\right)\right). \end{aligned}$$

By the CBH formula, G is a subgroup of $T_1^N\left(\mathbb{R}^d\right)$ with respect to $\otimes$-multiplication. Moreover, G is closed since $G=\log^{-1}\left(L^N\left(\mathbb{R}^d\right)\right)$ and log is continuous (after all, it's a polynomial in the coordinates). It is easy to see[4] that G is a submanifold of $T_1^N\left(\mathbb{R}^d\right)$. Thus, G is a Lie subgroup of $T_1^N\left(\mathbb{R}^d\right)$. In particular, G is a Lie group in its own right and we call it the *step-N nilpotent free group over* $\mathbb{R}^d$. Its Lie algebra is given by $L^N\left(\mathbb{R}^d\right)$. By (trivial) restriction, CBH, Zassenhaus etc. hold for G.

EXERCISE 24. *(i) Show that $G^2\left(\mathbb{R}^d\right)=\left\{\mathbf{x}\in T_1^2\left(\mathbb{R}^d\right):Sym\left(\mathbf{x}^2\right)=\frac{1}{2}\mathbf{x}^1\otimes\mathbf{x}^1\right\}$.*
(ii) Show that $G^2\left(\mathbb{R}^2\right)$ has dimension 3 and is isomorphic to the 3-dimensional Heisenberg group $\mathbb{H}=\left(\mathbb{R}^3,*\right)$ where

$$\left(x,y,a\right)*\left(x',y',a'\right):=\left(x+x',y+y',a+a'+\frac{1}{2}\left(xy'-yx'\right)\right).$$

(iii) Let x be an $\mathbb{R}^2$-valued Lipschitz path. Give a geometric interpretation of (2.1).

Example 14 and part (i) of the last exercise shows that $S_2\left(x\right)_{0,1}\in G^2\left(\mathbb{R}^d\right)$ for any Lipschitz path x. This is special case of the following

THEOREM 25. *Assume x is a Lipschitz continuous $\mathbb{R}^d$-valued path. Then $t\mapsto\mathbf{x}_t=S_N\left(x\right)_{0,t}$ solves the following ODE*

$$d\mathbf{x}_t=\mathbf{x}_t\otimes\mathfrak{b}_idx^i=:U_i\left(\mathbf{x}_t\right)dx^i$$

and $\mathbf{x}_t=S_N\left(x\right)_{0,t}\in G^N\left(\mathbb{R}^d\right)$. Here, $\mathfrak{b}_1,\ldots,\mathfrak{b}_d$ is the standard basis of $\mathbb{R}^d$.

---

[4] The key is that $d\left(\log\circ\exp\right)=Id$. Then the chain-rule shows that $d\exp$ is one-to-one at every point.

In fact, one can rely on result from Lie group theory instead: A closed abstract subgroup is automatically a Lie subgroup (and hence a submanifold).

PROOF. The ODE statement is an obvious consequence of the definition of $S_N(x)_{0,t}$. Observe that vector fields $\{U_1,...,U_d\}$ are tangent to the submanifold $G \subset T_1^N(\mathbb{R}^d)$. To see this, take $\mathbf{x} \in G$ and consider $m_{\mathbf{x}} : G \to G, \mathbf{y} \mapsto \mathbf{x} \otimes \mathbf{y}$. Write $\mathcal{T}_{\mathbf{y}}G$ for the tangent space. Then the differential at $\mathbf{y} = e$ is the linear map $dm_{\mathbf{x}}|_e : \mathcal{T}_{\mathbf{e}}G \to \mathcal{T}_{\mathbf{x}}G$ given by $v \mapsto \mathbf{x} \otimes v$ and $U_i(\mathbf{x}) = dm_{\mathbf{x}}|_e \mathfrak{b}_i \in \mathcal{T}_{\mathbf{x}}G$.

Since $U_1,...,U_d$ are tangent to G, any ODE solution started in G remains in G. Let us make this last argument explicit: log provides a (global!) chart for G, that is, a diffeomorphism from G into the Euclidean space $L = L^N(\mathbb{R}^d)$, a linear subspace of $T_0^N(\mathbb{R}^d)$.Thus, the ODE for $\mathbf{x}_t$ can be written as a (coordinate-transformed) ODE on L and, trivially, the solution stays in L. Going back to the manifold, via $\log^{-1} = \exp$, the ODE solution $\mathbf{x}_t$ stays in G. □

THEOREM 26 (Chow). *For every $\mathbf{g} \in G^N(\mathbb{R}^d)$ there exists an $\mathbb{R}^d$-valued Lipschitz path x such that $S_N(x)_{0,1} = \mathbf{g}$. More precisely, G is the group generated by $\left\{\exp(v) : v \in \mathbb{R}^d\right\}$ so that every $\mathbf{g} \in G$ is the signature of a (finite number of) concatenation of straight path segments (cmp. Ex. 15).*

PROOF. $N = 1$ is trivial: $\mathbf{g}$ is a vector in $\mathbb{R}^d$ and we can take a straight line for x. When $N = 2$ an arbitrary element $\mathbf{g} \in G^N(\mathbb{R}^d)$ can be written in form

$$\mathbf{g} = \exp\left\{\sum_i a^{1;i}\mathfrak{b}_i + \sum_{i,j} a^{2;i,j}[\mathfrak{b}_i, \mathfrak{b}_j]\right\}$$

where $\{\mathfrak{b}_i : i = 1,..,d\}$ is the standard basis for $\mathbb{R}^d$ and $[\mathfrak{b}_i, \mathfrak{b}_j] = \mathfrak{b}_i \otimes \mathfrak{b}_j - \mathfrak{b}_j \otimes \mathfrak{b}_i$ forms a basis for the anti-symmetric d-dimensional matrices. Iterated use of Zassenhaus shows that $\mathbf{g}$ is the product of terms of form $\exp(c\mathfrak{b}_i)$ and $\exp(c[\mathfrak{b}_i, \mathfrak{b}_j])$ for real constants c. But every term of form $\exp(c[\mathfrak{b}_i, \mathfrak{b}_j])$, w.l.o.g. $c > 0$, expands to

$$e^{\sqrt{c}\mathfrak{b}_i}e^{\sqrt{c}\mathfrak{b}_j}e^{-\sqrt{c}\mathfrak{b}_i}e^{\sqrt{c}\mathfrak{b}_j}.$$

With a bit more effort, this argument extends to arbitrary N. □

LEMMA 27. *Assume $x^n : [0,1] \to \mathbb{R}^d$ are such that $\sup_n |x^n|_{Lip,[0,1]} = C < \infty$ and $x^n \to x$ pointwise ($\Longrightarrow$ x is Lipschitz, as seen earlier). Then for any $N \in \mathbb{N}$, $\sup_n |S_N(x^n)|_{Lip,[0,1]}, |S_N(x)|_{Lip,[0,1]} < \infty$ and*

$$S_N(x^n) \to S_N(x) \text{ pointwise on } [0,1]$$

(in fact, uniformly on $[0,1]$; see Lemma 5).

PROOF. $S_N(x)_{0,t}$ is time-t solution of a control ODE driven by x along vector fields $U_1,...,U_d$. We saw in Prop. 7/ Ex. 8 that $x^n \to x$ pointwise with uniform Lip-bounds entails $S_N(x^n)_{0,t} \to S_N(x)_{0,t}$ *provided* the vector fields are Lip^1 (which includes boundedness). The vector fields (U_i) are smooth but not bounded. But since

$$\begin{aligned} \sup_{t\in[0,1]}\left|S_N(x^n)_{0,t}\right|, \sup_{t\in[0,1]}\left|S_N(x)_{0,t}\right| &\leq \sup_n \frac{|x^n|_{Lip,[0,1]}^N}{N!} \\ &\leq C^N/N! \end{aligned}$$

we localize, that is we can safely replace (U_i) by compactly supported (smooth) vector fields $\tilde{U}_i$ so that $U_i \equiv \tilde{U}_i$ on a ball of radius $2C^N/N!$.

□

THEOREM 28 (Geodesic Existence). *For every* $\mathbf{g} \in G$,

$$\|\mathbf{g}\| := \inf \left\{ \int_0^1 |\dot{\gamma}_t| \, dt : \gamma : [0,1] \to \mathbb{R}^d \text{ Lipschitz continuous, } \gamma(0) = 0, \ S_N(\gamma)_{0,1} = \mathbf{g} \right\}$$

is finite and achieved at some minimizing Lipschitz continuous path γ^*, *i.e.*

$$\|\mathbf{g}\| = \int_0^1 |\dot{\gamma}_t^*| \, dt \text{ and } S_N(\gamma^*)_{0,1} = \mathbf{g}.$$

Moreover, for any $s < t$ *there exists a Lipschitz path* $x^{s,t} : [s,t] \to \mathbb{R}^d$ *with signature* $\mathbf{g}$ *and length* $\|\mathbf{g}\|$:

$$S_N\left(x^{s,t}\right)_{s,t} = \mathbf{g} \quad \text{and} \quad \int_s^t \left|dx^{s,t}\right| = \|\mathbf{g}\|.$$

REMARK 29. *The path* $t \in [0,1] \mapsto S_N(\gamma^*)_{0,t} \in G$ *connects the unit* e *with* $\mathbf{g} \in G$ *and is called a geodesic. Thus, strictly speaking,* γ^* *is a projected geodesic but it is (sometimes) called geodesic as well.*

PROOF. From Chow's theorem, the inf is taken over a non-empty set so that $\|\mathbf{g}\| < \infty$. By definition of inf, there is a sequence (γ^n) with signature $\mathbf{g}$ and we can assume (by reparametrization[5]) that each γ^n has a.s. constant speed $|\dot{\gamma}_t^n| \equiv c_n \leq |\gamma^n|_{Lip,[0,1]}$. On the other hand, $\left|\gamma_{s,t}^n\right| \leq \int_s^t |\dot{\gamma}_r^n| \, dr = c_n |t-s|$ and we conclude that $c_n = |\gamma^n|_{Lip,[0,1]}$. Of course, c_n is the length of the path γ_t and $c_n \downarrow \|\mathbf{g}\|$. Clearly,

$$\sup_n |\gamma^n|_{Lip,[0,1]} = \sup_n c_n < \infty$$

and from Arzela-Ascoli, after relabeling the sequence, γ^n converges uniformly to some (continuous) limit path γ^*. From the preceding lemma,

$$\mathbf{g} \equiv S_N(\gamma^n)_{0,1} \to S_N(\gamma^*)_{0,1}$$

and hence $S_N(\gamma^*)_{0,1} = \mathbf{g}$. It remains to see that

$$\|\mathbf{g}\| = \int_0^1 |\dot{\gamma}_t^*| \, dt.$$

First, $\|\mathbf{g}\| \leq \int_0^1 |\dot{\gamma}_t^*| \, dt$ is obvious from definition of $\|\mathbf{g}\|$. On the other hand, using lemma 5, and for a.e. $t \in [0,1]$,

$$\begin{aligned} |\dot{\gamma}_t^*| &\leq |\gamma^*|_{Lip} \leq \liminf_n |\gamma^n|_{Lip,[0,1]} \\ &= \liminf_n c_n \\ &= \|\mathbf{g}\|. \end{aligned}$$

Conclude that $\int_0^1 |\dot{\gamma}_t^*| \, dt \leq \|\mathbf{g}\|$. The final statement is an exercise in reparametrization. □

COROLLARY 30. *Let* $\mathbf{g}, \mathbf{h} \in G^N\left(\mathbb{R}^d\right)$. *We have*
(i) $\|\mathbf{g}\| = 0$ *iff* $\mathbf{g} = e$,
(ii) symmetry $\|\mathbf{g}\| = \left\|\mathbf{g}^{-1}\right\|$,
(iii) sub-additivity $\|\mathbf{g} \otimes \mathbf{h}\| \leq \|\mathbf{g}\| + \|\mathbf{h}\|$ *and*

[5] The signature is invariant under reparametrization.

(iv) homogenity $\|\delta_\lambda \mathbf{g}\| = |\lambda| \, \|\mathbf{g}\|$ *for all* $\lambda \in \mathbb{R}$.
(v) In particular,

$$d(\mathbf{g}, \mathbf{h}) := \left\| \mathbf{g}^{-1} \otimes \mathbf{h} \right\|$$

defines a left-invariant metric on G. Call it Carnot-Caratheodory or CC-metric.

Proof. Notation: for $\mathbf{g} \in G$ let $\gamma^*_{\mathbf{g}} = \gamma^*$ denote the minimizer from the geodesic existence theorem.
(i) "$\Rightarrow$" $\|\mathbf{g}\| = 0 \implies \exists$ Lipschitz $\gamma^*_{\mathbf{g}}$ with a.s. zero derivative $\implies \gamma^*_{\mathbf{g}} \equiv (const) \implies S_N\left(\gamma^*_{\mathbf{g}}\right)_{0,1} = e$. But, by construction, $S_N\left(\gamma^*_{\mathbf{g}}\right)_{0,1} = \mathbf{g}$. The "$\Leftarrow$" direction is trivial.

(ii) From the exercise below $S_N\left(\overleftarrow{\gamma^*_{\mathbf{g}}}\right)_{0,1} = \mathbf{g}^{-1}$. This implies

$$\left\|\mathbf{g}^{-1}\right\| \leq \text{length}\left(\overleftarrow{\gamma^*_{\mathbf{g}}}\right) = \text{length}\left(\gamma^*_{\mathbf{g}}\right) = \|\mathbf{g}\|.$$

The opposite inequality follows from replacing $\mathbf{g}$ by $\mathbf{g}^{-1}$.

(iii) If $\gamma^*_{\mathbf{g}}, \gamma^*_{\mathbf{h}}$ denote the resp. geodesics then, from Prop. 16,

$$\mathbf{g} \otimes \mathbf{h} = S_N\left(\gamma^*_{\mathbf{g},\mathbf{h}}\right)_{0,1}$$

where $\gamma^*_{\mathbf{g},\mathbf{h}}$ is the (Lipschitz continuous) concatenation of $\gamma^*_{\mathbf{g}}$ and $\gamma^*_{\mathbf{h}}$ with obvious length $\|\mathbf{g}\| + \|\mathbf{h}\|$. By definition, $\|\mathbf{g} \otimes \mathbf{h}\|$ must be less or equal to the length of $\gamma^*_{g,h}$.

(iv) W.l.o.g. $\lambda \neq 0$. The path $\lambda \gamma^*_{\mathbf{g}}$ has signature $\delta_\lambda \mathbf{g}$. Hence $\|\delta_\lambda \mathbf{g}\| \leq \text{length}\big(\lambda \gamma^*_{\mathbf{g}}\big) = |\lambda| \times \text{length}\big(\gamma^*_{\mathbf{g}}\big) = |\lambda| \, \|\mathbf{g}\|$. The opposite inequality follows from replacing λ by $1/\lambda$ and $\mathbf{g}$ by $\delta_\lambda \mathbf{g}$. □

Exercise 31. *Let* $\overleftarrow{x}(t) = x(1-t)$. *Show* $S_N\left(\overleftarrow{x}\right)_{0,1} = \left(S_N(x)_{0,1}\right)^{-1}$.

Solution 32. *For straight path segments this was seen in exercise 15. For x piecewise linear (=concatenation of straight path segments) this follows from Prop. 16. From Lemma 4, any Lipschitz x can be approximated by piecewise linear ones. Then use lemma 27.*

G is a subset of $T^N\left(\mathbb{R}^d\right)$ and inherits a metric and hence topology (which coincides with the manifold topology of G). In particular,

$$\mathbf{g}_n \to \mathbf{g} \text{ in } G \text{ iff } \left|\mathbf{g}^k_n - \mathbf{g}^k\right| \to 0 \forall k = 1, ..., N \tag{2.3}$$

and since in metric spaces the topology is fully described by convergence of sequences that's all we have to know. We want to check that $\mathbf{g}_n \to \mathbf{g}$ iff $d(\mathbf{g}_n, \mathbf{g})$ where d is the CC-metric constructed earlier. A similar statement is well-known in Riemannian geometry where one checks that the Riemannian distance on a manifold coincides with the natural manifold topology.

Proposition 33. *Let* $|||.|||_i$ $(i = 1, 2)$ *be continuous homogenous norms on G, that is, norms that satisfies properties (i) and (iv) and such that* $\mathbf{g} \mapsto |||\mathbf{g}|||_i$ *is continuous w.r.t.* τ. *Then there exists a constant* $c \in [1, \infty)$ *such that* $|||.|||_1 \sim |||.|||_2$ *by which we mean*

$$\frac{1}{c} |||.|||_2 \leq |||.|||_1 \leq c \, |||.|||_2 \, .$$

PROOF. As $\sim$ is transitive it suffices to show this for the particular choice

$$|||\mathbf{g}|||_1 = \max_{k=1,\ldots,N} \left|\mathbf{g}^k\right|^{1/k}.$$

The set $S = \{\mathbf{g} : |||\mathbf{g}|||_1 = 1\}$ is obviously compact w.r.t. τ : it is closed by continuity of $|||.|||_1$ and bounded in $T^N\left(\mathbb{R}^d\right)$. The map $\mathbf{g} \mapsto |||\mathbf{g}|||_2$ is also continuous and its restriction to S achieves its min and max. By property (i), the min must be strictly positive. In other words, $\exists c > 0$ such that

$$\frac{1}{c} \leq |||\mathbf{g}|||_2 \leq c \text{ when } |||\mathbf{g}|||_1 = 1.$$

Homogeneity of $|||\cdot|||_i$ $(i = 1, 2)$ easily finishes the proof. □

EXAMPLE 34. *The following are examples of continuous, homogenenous norms on G:*

$$\max\left\{\left|\mathbf{g}^k\right|^{1/k} : k = 1, ..., N\right\}, \quad \sum_{k=1}^{N} \left|\mathbf{g}^k\right|^{1/k}.$$

Also, when $\mathbf{g} = \exp(\mathbf{a})$ for $\mathbf{a} \in L$

$$\max\left\{\left|\mathbf{a}^k\right|^{1/k} : k = 1, ..., N\right\}$$

etc etc.

We now return to the "CC norm" defined in the geodesic existence theorem.

LEMMA 35. *The map $\mathbf{g} \mapsto \|\mathbf{g}\|$ is continuous (and thus constitutes a continuous homogenous norms on G).*

PROOF. We first show continuity at the unit element and assume $\mathbf{h}_n \to e$, that is, $\mathbf{h}_n^k \to 0 \forall k = 1, ..., N$. In Chow's theorem we constructed a (piecewise) linear path, say γ_n, with step-N signature equal to $\mathbf{h}_n$ and it is easy to see from that construction that the length of $\gamma_n \to 0$. But the CC norm of $\mathbf{h}_n$ is dominated by the lenght of any path with signature γ_n and so

$$\|\mathbf{h}_n\| \to 0.$$

Consider now a sequence $\mathbf{g}_n$ such that $\mathbf{g}_n \to \mathbf{g}$ in the sense of (2.3). By continuity of the group operations $\otimes$ and $(\cdot)^{-1}$, all of which are polynomial in the coordinates, $\mathbf{h}_n := \mathbf{g}_n^{-1} \otimes \mathbf{g} \to \mathbf{e}$ is an obvious consequence. From sub-additivity,

$$|\|\mathbf{g}_n\| - \|\mathbf{g}\|| \leq \left\|\mathbf{g}_n^{-1} \otimes \mathbf{g}\right\| \to 0$$

and the proof is finished. □

COROLLARY 36. *The topology induced by the CC-metric coincides with the natural topology of G as subset of $T^N\left(\mathbb{R}^d\right)$, cf. (2.3).*

PROOF. By continuity of the group operations $\otimes$ and $(\cdot)^{-1}$, $\mathbf{g}_n \to \mathbf{g}$ in the sense of (2.3) is equivalent to

$$\left|\left(\mathbf{g}^{-1} \otimes \mathbf{g}_n\right)^k\right| \to 0 \forall k = 1, ..., N.$$

This is equivalent to $\left|\left|\left|\mathbf{g}^{-1} \otimes \mathbf{g}_n\right|\right|\right| \to 0$ with respect to the continous homogenous norm given by

$$|||\mathbf{h}||| := \max\left\{\left|\mathbf{h}^k\right|^{1/k} : k = 1, ..., N\right\},$$

and by equivalence of continuous, homogenous norms this is equivalent to

$$\left|\left|\mathbf{g}^{-1}\otimes\mathbf{g}_n\right|\right| = d\left(\mathbf{g}_n,\mathbf{g}\right)\to 0.$$

□

EXERCISE 37. *Show that (G,d) is a Polish space (=complete, separable, metric space).*

EXERCISE 38. *(i) Let $x:[0,T]\to\mathbb{R}^d$ be Lipschitz so that its length is given by*

$$\int_0^T |dx| = \sup_{0\le t_1\le\ldots\le t_n\le T}\left|x_{t_i,t_{i+1}}\right|.$$

Show that x has the same length as its step-N lift $\mathbf{x}:=S_N(x)$ with respect to CC distance, i.e.

$$\int_0^T |dx| = \sup_{0\le t_1\le\ldots\le t_n\le T} d\left(\mathbf{x}_{t_i},\mathbf{x}_{t_{i+1}}\right).$$

(ii) Recall that the Cameron-Martin norm of $x:[0,T]\to\mathbb{R}^d$ can be defined as

$$|x|_{CM}^2 = \int_0^T |\dot{x}_t|^2\,dt = \sup_{0\le t_1\le\ldots\le t_n\le T}\frac{\left|x_{t_i,t_{i+1}}\right|^2}{t_{i+1}-t_i}.$$

Define a CM "norm" for $\mathbf{x}:=S_N(x)$ and show that it equals $|x|_{CM}$.

SOLUTION 39. *(i) The "$\ge$" follows readily from*

$$d\left(S_N(x)_{0,s},S_N(x)_{0,t}\right) = \left\|S_N(x)_{s,t}\right\| \ge |x_{s,t}|,$$

valid for all s,t in $[0,T]$. For the reverse inequality, first observe that by definition of the Carnot-Caratheodory homogeneous norm,

$$d\left(S_N(x)_{0,s},S_N(x)_{0,t}\right) = \left\|S_N(x)_{s,t}\right\| \le \int_s^t |dx|.$$

Apply this to $s=t_i, t=t_{i+1}$, then sum over i and take the sup over all dissections $(t_i)\subset[0,T]$.

3. Rough Path Spaces

3.1. Recalls on Hölder spaces. Set $\omega(s,t)=t-s$ and call it (Hölder) control. Let $p\in[1,\infty)$. A path y from $[0,T]$ to $\mathbb{R}^d$ is $1/p$-Hölder continuous if

$$|y_{s,t}|\le C\omega(s,t)^{1/p}$$

for some constant C. Call this class by $C^{p,\omega}\left([0,T],\mathbb{R}^d\right)$. (The class of Lipschitz paths is exactly $C^{1,\omega}\left([0,T],\mathbb{R}^d\right)$). The $1/p$-Hölder (semi-)norm on $C^{p,\omega}\left([0,T],\mathbb{R}^d\right)$ is defined by

$$|y|_{p,\omega} = \sup_{0\le s<t\le T}\frac{|y_{s,t}|}{\omega(s,t)^{1/p}}$$

and there is a $1/p$-Hölder metric given by

$$|y-\tilde{y}|_{p,\omega} = \sup_{0\le s<t\le T}\frac{|y_{s,t}-\tilde{y}_{s,t}|}{\omega(s,t)^{1/p}}.$$

To avoid trivialites (semi-norms vs norms) we only consider paths with a fixed starting point.

THEOREM 40. *(i)* $C^{p,\omega}\left([0,T],\mathbb{R}^d\right)$ *is a complete metric space under* $1/p$*-Hölder metric (in fact, it is a Banach space).*
(ii) Every $1/p$*-Hölder continuous path* $y \in C^{p,\omega}\left([0,T],\mathbb{R}^d\right)$ *can be approximated by Lipschitz paths* $y_n \in C^{1,\omega}\left([0,T],\mathbb{R}^d\right)$ *so that*

$$y_n \to y \quad \text{uniformly on } [0,T]$$

and $\sup_n |y_n|_{p,\omega} < \infty$*. In fact, we can arrange for* $\sup_n |y_n|_{p,\omega} \leq 3\,|y|_{p,\omega}$*.*
(iii) Let $p > 1$*. Define* $C^{0,p,\omega}\left([0,T],\mathbb{R}^d\right)$ *as the closure of Lipschitz paths under* $1/p$*-Hölder metric. For* $y \in C^{p,\omega}\left([0,T],\mathbb{R}^d\right)$ *we have*

$$y \in C^{0,p,\omega}\left([0,T],\mathbb{R}^d\right) \text{ iff } r(\delta;y) \equiv \sup_{\substack{0\leq s<t\leq T\\ t-s\leq\delta}} \frac{|y_{s,t}|}{\omega(s,t)^{1/p}} \to 0 \text{ as } \delta \to 0.$$

PROOF. All these statements belong to the folklore of Hölder spaces. Moreover, they are a special case ($N=1$) of Theorem 47 below, no need for a proof here. □

EXERCISE 41. *Let* $p > 1$ *and* $y(t) = t^{1/p}$*. Show that* $y \in C^{p,\omega}\left([0,T],\mathbb{R}^d\right)$ *but* $\notin C^{0,p,\omega}\left([0,T],\mathbb{R}^d\right)$*.*

REMARK 42. $C^{0,p,\omega}\left([0,T],\mathbb{R}^d\right)$ *is a Polish space but* $C^{p,\omega}\left([0,T],\mathbb{R}^d\right)$ *is not (it lacks separability).*

DEFINITION 43. *If* $y_n \to y$ *pointwise on* $[0,T]$ *and* $\sup_n |y_n|_{p,\omega} < \infty$ *we say that* $y_n \to y$ *with uniform* (p,ω)*-bounds.*

EXERCISE 44. *Generalize Lemma 5: Assume* $y_n \in C^{p,\omega}\left([0,T],\mathbb{R}^d\right)$ *and* $y_n \to y$ *with uniform* (p,ω)*-bounds. Show that* $y \in C^{p,\omega}\left([0,T],\mathbb{R}^d\right)$, $y_n \to y$ *uniformly on* $[0,T]$ *and*

$$|y|_{p,\omega} \leq \liminf_n |y_n|_{p,\omega} < \infty.$$

EXERCISE 45. *(Interpolation) Assume* $y_n \to y$ *with uniform* (p,ω)*-bounds. Show that* $y_n \to y$ *w.r.t.* $1/(p+\epsilon)$*-Hölder metric.*

3.2. Spaces of geometric (Hölder) rough paths. Again, set $\omega(s,t) = t-s$ and let $p \in [1,\infty)$. A path $\mathbf{x}$ from $[0,T]$ to $G^N\left(\mathbb{R}^d\right)$ is $1/p$-Hölder continuous if

$$\|\mathbf{x}_{s,t}\| \leq C\omega(s,t)^{1/p}$$

for some constant C. Call this class by $C^{p,\omega}\left([0,T],G^N\left(\mathbb{R}^d\right)\right)$. As before, restrict attention to paths with pinned starting point. The $1/p$-Hölder "norm" (there is no linear space here!) on $C^{p,\omega}\left([0,T],G^N\left(\mathbb{R}^d\right)\right)$ is defined by

$$\|\mathbf{x}\|_{p,\omega} = \sup_{0\leq s<t\leq T} \frac{\|\mathbf{x}_{s,t}\|}{\omega(s,t)^{1/p}}$$

and there is a $1/p$-Hölder metric based on the CC-metric,

$$d_{p,\omega}(\mathbf{x},\tilde{\mathbf{x}}) = \sup_{0\leq s<t\leq T} \frac{d(\mathbf{x}_{s,t},\tilde{\mathbf{x}}_{s,t})}{\omega(s,t)^{1/p}}.$$

We also set

$$d_\infty(\mathbf{x},\tilde{\mathbf{x}}) = \sup_{0\leq s<t\leq T} d(\mathbf{x}_{s,t},\tilde{\mathbf{x}}_{s,t}).$$

EXERCISE 46. *Show that d_∞-convergence is equivalent to convergence w.r.t.* $\tilde{d}_\infty(\mathbf{x}, \tilde{\mathbf{x}}) := \sup\{d(\mathbf{x}_t, \tilde{\mathbf{x}}_t) : 0 \le t \le T\}$.

THEOREM 47. *(i) $C^{p,\omega}([0,T], G^N(\mathbb{R}^d))$ is a complete metric space under the metric $d_{p,\omega}$.*
(ii) Every $1/p$-Hölder continuous path $\mathbf{x} \in C^{p,\omega}([0,T], G^N(\mathbb{R}^d))$ can be approximated by Lipschitz paths $x_n \in C^{1,\omega}([0,T], \mathbb{R}^d)$ in the sense that

$$S_N(x_n) \to \mathbf{x} \ \text{ uniformly on } [0,T]$$

and $\sup_n \|S_N(x_n)\| < \infty$. In fact, we can arrange for $\sup_n \|S_N(x_n)\|_{p,\omega} \le 3\|\mathbf{x}\|_{p,\omega}$
.
(iii) Assume $p > 1$. Define $C^{0,p,\omega}([0,T], G^N(\mathbb{R}^d))$ as the closure of lifted Lipschitz paths $S_N(x)$ under the metric $d_{p,\omega}$. For $\mathbf{x} \in C^{p,\omega}([0,T], G^N(\mathbb{R}^d))$ we have

$$\mathbf{x} \in C^{0,p,\omega}([0,T], G^N(\mathbb{R}^d)) \ \text{ iff } r(\delta; \mathbf{x}) \equiv \sup_{\substack{0 \le s < t \le T \\ t-s \le \delta}} \frac{\|\mathbf{x}_{s,t}\|}{\omega(s,t)^{1/p}} \to 0 \ \text{ as } \delta \to 0.$$

In particular, $C^{0,p,\omega}([0,T], G^N(\mathbb{R}^d)) \subsetneq C^{p,\omega}([0,T], G^N(\mathbb{R}^d))$.

PROOF. (i) We only show completness (the rest is easy). Assume $\mathbf{x}^n$ is Cauchy in $C^{p,\omega}([0,T], G^N(\mathbb{R}^d))$. Then for every $t \in [0,T]$, $(\mathbf{x}^n_t)$ is Cauchy in $G^N(\mathbb{R}^d)$ with limit, say, $\mathbf{x}_t$. Since $\mathbf{x}^n$ is Cauchy it follows that $\|\mathbf{x}^n\|_{p,\omega}$ is bounded (by some constant M, say). Recall $\mathbf{x}_{s,t} = \mathbf{x}_s^{-1} \otimes \mathbf{x}_t$ and continuity of group operations and $\|\cdot\|$. Thus

$$\|\mathbf{x}_{s,t}\| = \lim \|\mathbf{x}^n_{s,t}\| \le M|t-s|^{1/p}$$

and indeed $\mathbf{x} \in C^{p,\omega}([0,T], G^N(\mathbb{R}^d))$. Finally, by passing to a subsequence, we may assume that $d_{p,\omega}(\mathbf{x}^m, \mathbf{x}^{m+1}) \le 2^{-m}$. Then

$$d(\mathbf{x}^n_{s,t}, \mathbf{x}_{s,t}) \le \sum_{m \ge n} d(\mathbf{x}^m_{s,t}, \mathbf{x}^{m+1}_{s,t}) \le 2^{1-n}(t-s)^{1/p}$$

which shows $d_{p,\omega}(\mathbf{x}^n, \mathbf{x}) \to 0$, along the chosen subsequence. This shows that $\mathbf{x}$ is the only limit point of $\{\mathbf{x}^n\}$ and hence $d_{p,\omega}(\mathbf{x}^n, \mathbf{x}) \to 0$.
(ii) Same idea as in lemma 4! Given a dissection $D = \{t_j\}$ of $[0,T]$, define $\mathbf{x}^D$ by concatenation of (lifted) geodesics connecting $\mathbf{x}_{t_j}$ and $\mathbf{x}_{t_{j+1}}$. To be precise, assume $t \in [t_j, t_{j+1}]$ and define

$$\mathbf{x}^D_t = \mathbf{x}_{t_j} \otimes S_N(\gamma^*)_{t_j,t}$$

where $\gamma^* = \gamma^*_{[j]}$ is the (projected) geodesic from the Geodesic Existence Theorem, w.l.o.g. defined on $[t_j, t_{j+1}]$, which connects e with $\mathbf{x}_{t_j,t_{j+1}}$. In particular,

$$\begin{aligned} \mathbf{x}^D_{0,t} &= \mathbf{x}^D_{0,t_1} \otimes \ldots \otimes \mathbf{x}^D_{t_j,t} \\ &= S_N\left(\gamma^*_{[0]}\right)_{0,t_1} \otimes \ldots \otimes S_N\left(\gamma^*_{[j]}\right)_{t_j,t} \end{aligned}$$

and $t \mapsto \mathbf{x}^D_{0,t}$is the lift of a Lipschitz path (namely, the concatenation of $\gamma^*_{[0]}, \gamma^*_{[1]}, \ldots$).
Case 1: $t_j \le s \le t \le t_{j+1}$. Again, let $\gamma^* : [t_j, t_{j+1}] \to \mathbb{R}^d$ denote the geodesic which connects e with $\mathbf{x}_{t_j,t_{j+1}}$, w.l.o.g. with a.s. constant speed. Note that $\mathbf{x}^D_{s,t} =$

$S(\gamma^*)_{s,t}$.

$$\begin{aligned}\left\|\mathbf{x}^D_{s,t}\right\| &\leq \text{length}\left\{\gamma^*_r|_{[s,t]}\right\}\\ &= \int_s^t |\dot{\gamma}^*_r|\,dr\\ &= \frac{t-s}{t_{j+1}-t_j}\int_{t_j}^{t_{j+1}}|\dot{\gamma}^*_r|\,dr\\ &= \frac{t-s}{t_{j+1}-t_j}\left\|\mathbf{x}_{t_j,t_{j+1}}\right\|\\ &\leq \frac{t-s}{t_{j+1}-t_j}\|\mathbf{x}\|_{p,\omega}(t_{j+1}-t_j)^{1/p}\\ &\leq (t-s)^{1/p}\|\mathbf{x}\|_{p,\omega}\frac{(t-s)^{1-1/p}}{(t_{j+1}-t_j)^{1-1/p}}\\ &\leq (t-s)^{1/p}\|\mathbf{x}\|_{p,\omega}.\end{aligned}$$

Case 2: When $t_i \leq s \leq t_{i+1} \leq t_j \leq t \leq t_{j+1}$ then (using $(a+b+c)^p \leq 3^p(\max\{a,b,c\})^p \leq 3^p(a^p+b^p+c^p)$),

$$\begin{aligned}\left\|\mathbf{x}^D_{s,t}\right\|^p &\leq 3^p\left\{\left\|\mathbf{x}^D_{s,t_{i+1}}\right\|^p+\left\|\mathbf{x}_{t_{i+1},t_j}\right\|^p+\left\|\mathbf{x}^D_{t_j,t}\right\|^p\right\}\\ &\leq 3^p\left(|t_{i+1}-s|+|t_j-t_{i+1}|+\left|t_{i_j}-t\right|\right)\|\mathbf{x}\|^p_{p,\omega}\\ &= 3^p|t-s|\,\|\mathbf{x}\|^p_{p,\omega}\end{aligned}$$

so that

$$\left\|\mathbf{x}^D\right\|_{p,\omega} \leq 3\|\mathbf{x}\|_{p,\omega}.$$

Now it suffices to consider a sequence of dissections D_n with mesh $|D_n| \to 0$ to see that

$$\mathbf{x}^{D_n}_t \to \mathbf{x}_t$$

pointwise and with uniform estimates, namely $\sup_n\left\|\mathbf{x}^{D_n}\right\|_{p,\omega} \leq 3\|\mathbf{x}\|_{p,\omega}$. In fact, this easily implies that $\mathbf{x}^{D_n} \to \mathbf{x}$ uniformly on $[0,T]$.

(iii) Assume $\mathbf{x} \in C^{p,\omega}$ and $r(\delta;\mathbf{x}) \to 0$ as $\delta \to 0$. In (ii) we constructed approximations $\mathbf{x}^{D_n} = S_N\left(x^{D_n}\right)$, x^{D_n} Lipschitz, with $d_\infty\left(\mathbf{x}^{D_n},\mathbf{x}\right) \to 0$ (= uniform convergence on $[0,T]$) and with $\sup_n\left\|\mathbf{x}^{D_n}\right\|_{p,\omega} \leq 3\|\mathbf{x}\|_{p,\omega}$. Then

$$\begin{aligned}d_{p,\omega}\left(\mathbf{x}^{D_n},\mathbf{x}\right) &\leq \sup_{\substack{0\leq s<t\leq T\\ t-s\leq\delta}}\frac{d\left(\mathbf{x}^{D_n}_{s,t},\mathbf{x}_{s,t}\right)}{|t-s|^{1/p}}+\frac{1}{\delta^{1/p}}d_\infty\left(\mathbf{x}^{D_n},\mathbf{x}\right)\\ &\leq \sup_{\substack{0\leq s<t\leq T\\ t-s\leq\delta}}\frac{\left\|\mathbf{x}^{D_n}_{s,t}\right\|+\|\mathbf{x}_{s,t}\|}{|t-s|^{1/p}}+\frac{1}{\delta^{1/p}}d_\infty\left(\mathbf{x}^{D_n},\mathbf{x}\right)\\ &\leq (3+1)\,r(\delta;\mathbf{x})+\frac{1}{\delta^{1/p}}d_\infty\left(\mathbf{x}^{D_n},\mathbf{x}\right).\end{aligned}$$

For $\epsilon > 0$, choose δ small enough to make the first term $\leq \epsilon/2$ then n large enough to make the 2nd term $\leq \epsilon/2$. We established $d_{p,\omega}\left(\mathbf{x}^{D_n},\mathbf{x}\right) \to 0$ as $n \to \infty$.

Vice versa, assume there is a sequence of Lipschitz path (x^n) so that $S_N(x^n) \equiv \mathbf{x}^n$

converges w.r.t. $d_{p,\omega}$ to some limit $\mathbf{x} \in C^{p,\omega}$. Note that x^n Lipschitz implies that $\mathbf{x}^n : [0,T] \to \left(G^N\left(\mathbb{R}^d\right), d\right)$ is Lipschitz (check!). Now take $s < t$, both in $[0,T]$ with $|t-s| \leq \delta$. For all n,

$$\begin{aligned}\frac{\|\mathbf{x}_{s,t}\|}{|t-s|^{1/p}} &\leq \frac{\left\|\mathbf{x}^n_{s,t}\right\|}{|t-s|^{1/p}} + \sup_{0\leq s<t\leq T} \frac{d\left(\mathbf{x}_{s,t}, \mathbf{x}^n_{s,t}\right)}{|t-s|^{1/p}} + \\ &\leq \|\mathbf{x}^n\|_{Lip,[0,T]} |t-s|^{1-1/p} + d_{p,\omega}\left(\mathbf{x}, \mathbf{x}^n\right) \\ &\leq \|\mathbf{x}^n\|_{Lip,[0,T]} \delta^{1-1/p} + d_{p,\omega}\left(\mathbf{x}, \mathbf{x}^n\right)\end{aligned}$$

We see that

$$r\left(\delta; \mathbf{x}\right) \leq \|\mathbf{x}^n\|_{Lip,[0,T]} \delta^{1-1/p} + d_{p,\omega}\left(\mathbf{x}, \mathbf{x}^n\right)$$

and

$$\limsup_{\delta\to 0} r\left(\delta; \mathbf{x}\right) \leq d_{p,\omega}\left(\mathbf{x}, \mathbf{x}^n\right).$$

Now let $n \to \infty$ to obtain $\lim_{\delta\to 0} r\left(\delta; \mathbf{x}\right) = 0$ as required. □

PROPOSITION 48. *Assume $p > 1$. $C^{0,p,\omega}\left([0,T], G^N\left(\mathbb{R}^d\right)\right)$ a Polish space.*

PROOF. By definition, $C^{0,p,\omega}\left([0,T], G\right)$ is a closed subset of a complete metric space and hence complete. Therefore, we only need to show separability. By definition, $\forall \mathbf{x} \in C^{0,p,\omega}\left([0,T], G^N\left(\mathbb{R}^d\right)\right)$ is the $d_{p,\omega}$-limit of $S\left(x^n\right)$ for Lipschitz paths x^n. Every Lipschitz paths x^n is the pointwise limit of a C^1path $\left(x^n_m\right)_{m\geq 1}$ so that $\sup_m |x^n_m|_{Lip} < \infty$. Since $C^1\left([0,T], G^N\left(\mathbb{R}^d\right)\right)$ with Lip-norm is separable (check!) we may assume w.l.o.g. that all the $\{x^n_m\}$ were chosen from some countable subset of $C^1\left([0,T], G^N\left(\mathbb{R}^d\right)\right)$. From Lemma 27,

$$S_N\left(x^n_m\right) \to S_N\left(x^n\right) \text{ pointwise as } m \to \infty$$

and $\sup_m |S_N\left(x^n_m\right)|_{1,\omega} < \infty$. By interpolation,

$$d_{p,\omega}\left[S_N\left(x^n_m\right), S_N\left(x^n\right)\right] \to 0 \text{ as } m \to \infty.$$

In particular, we can pick $m\left(n\right)$ so that the above distance is less than $1/n$. We conclude using the triangle inequality,

$$\begin{aligned}d_{p,\omega}\left[\mathbf{x}, S_N\left(x^n_{m(n)}\right)\right] &\leq d_{p,\omega}\left[\mathbf{x}, S_N\left(x^n\right)\right] + d_{p,\omega}\left[S_N\left(x^n_{m(m)}\right), S_N\left(x^n\right)\right] \\ &\to 0 \text{ as } n \to \infty.\end{aligned}$$

□

REMARK 49. *$C^{p,\omega}\left([0,T], G^N\left(\mathbb{R}^d\right)\right)$ is not separable since $C^{p,\omega}\left([0,T], \mathbb{R}^d\right)$ is (known to be) not separable.*

EXERCISE 50. *(i) Discuss the two exercises of the preceding section in the context of this section.*
(ii) Show that $C^{p,\omega}\left([0,T], G^N\left(\mathbb{R}^d\right)\right) \subset\subset C^{p',\omega}\left([0,T], G^N\left(\mathbb{R}^d\right)\right)$ when $p < p'$.

DEFINITION 51. *Let $p \geq 1$. A path in $C^{p,\omega}\left([0,T], G^{[p]}\left(\mathbb{R}^d\right)\right)$ where $[p] \in \mathbb{N}$ is such that $[p] \leq p < [p]+1$ is called a weak geometric p-rough path with (Hölder-) control ω. A path in $C^{0,p,\omega}\left([0,T], G^{[p]}\left(\mathbb{R}^d\right)\right)$ is called a geometric p-rough path with (Hölder-) control ω.*

THEOREM 52 (Lyons-lifting map). *Let* $\mathbf{x} \in C^{p,\omega}\left([0,T], G^{[p]}\left(\mathbb{R}^d\right)\right)$ *be a weak geometric p-rough path. For fixed* $N > p$, *there exists a unique* $\mathbf{z} \in C^{p,\omega}\left([0,T], G^{N}\left(\mathbb{R}^d\right)\right)$ *so that*

$$\pi_{N,[p]}(\mathbf{z}) \equiv \mathbf{x}.$$

Here $\pi_{N,[p]}$ *denotes the natural projection*

$$\pi_{N,[p]} : G^N\left(\mathbb{R}^d\right) \to G^{[p]}\left(\mathbb{R}^d\right), 1 + \mathbf{z}^1 + ... + \mathbf{z}^N \mapsto 1 + \mathbf{z}^1 + ... + \mathbf{z}^{[p]}.$$

(We can write $S_N(\mathbf{x}) := \mathbf{z}$ *for this lift of* $\mathbf{x}$*. This is consistent with our earlier use.) Moreover, there exists* $C = C(N)$ *such that*

$$\|\mathbf{x}\|_{p,\omega} \leq \|S_N(\mathbf{x})\|_{p,\omega} \leq C \|\mathbf{x}\|_{p,\omega}.$$

PROOF. Not strictly necessary for our approach and therefore omitted. □

EXERCISE 53. *Recall that* $G^2\left(\mathbb{R}^2\right)$ *can be identified with* $\mathbb{H}$, *the 3-dimensional Heisenberg group with coordinates* $(x, y; a)$*. Show that for every real* λ, *the path* $t \mapsto (0, 0; \lambda t)$ *is an element of* $C^{2,\omega}\left([0,T], G^2\left(\mathbb{R}^2\right)\right)$*. Given any path* $t \mapsto (x(t), y(t); a(t))$ *in* $C^{2,\omega}\left([0,T], G^2\left(\mathbb{R}^2\right)\right)$, *check that the perturbation*

$$t \mapsto (x(t), y(t); a(t) + \lambda t)$$

still gives an element in $C^{2,\omega}\left([0,T], G^2\left(\mathbb{R}^2\right)\right)$*. But each such perturbation projects to the* $\mathbb{R}^2$*-valued path* $(x(t), y(t))$, *an element in* $C^{2,\omega}\left([0,T], \mathbb{R}^2\right)$ *(*$\equiv 1/2$*-Hölder path!) which therefore has uncountable many lifts to* $C^{2,\omega}\left([0,T], G^2\left(\mathbb{R}^2\right)\right)$*. This shows that one cannot replace the assumption in the last theorem by* $C^{p,\omega}\left([0,T], G^n\left(\mathbb{R}^d\right)\right)$ *for* $n < [p]$.

4. Rough Path Estimates for ODEs I

The main result in this section will require on a quantitative understanding of

(A) the difference of ODE solutions started at the same point, with different driving signals (but with common iterated integrals up to a give order);
(B) the difference of ODE solutions started at the different points but with identical driving signal;

This is the content of the following two lemmas.

LEMMA 54 (Lemma A). *Assume that*
(i) $V = (V_i)_{1\leq i\leq d}$ *is a collection of vector fields in* $\mathrm{Lip}^N(\mathbb{R}^e)$,
(ii) $s < u$ *are some elements in* $[0,T]$,
(iii) $y_s \in \mathbb{R}^e$ *(thought of as a "time−s" initial condition)*
(iv) x *and* $\tilde{x}$ *are Lipschitz paths whose restriction to* $[s,u]$ *have identical step-N signature i.e.*

$$S_N(x)_{s,u} = S_N(\tilde{x})_{s,u},$$

(v) $\ell \geq 0$ *is a bound on* $|V|_{\mathrm{Lip}^N}\left(\int_s^u |dx| + \int_s^u |d\tilde{x}|\right)$.
Then, we have for some constant $C = C(N)$,

$$\left|\pi_{(V)}(s,y_s;x)_{s,u} - \pi_{(V)}(s,y_s;\tilde{x})_{s,u}\right| \leq C\ell^{N+1}.$$

PROOF. Insert/subtract the Euler approximation of order N. □

LEMMA 55 (Lemma B). *Assume that*
(i) $V = (V_i)_{1\leq i\leq d}$ *is a collection of vector fields in* $\mathrm{Lip}^1(\mathbb{R}^e)$,
(ii) $t < u$ *are some element of* $[0,T]$,
(iii) $y_t, \tilde{y}_t \in \mathbb{R}^e$ *(thought of as "time−t" initial conditions)*
(iv) $x : [t,u] \to \mathbb{R}^d$ *is a Lipschitz path,*
(v) $\ell \geq 0$ *is a bound on* $|V|_{\mathrm{Lip}^1}\int_t^u |dx|$.
Then, if $\pi_{(V)}(t,\cdot;x)$ *denotes the (unique) solution to* $dy = V(y)\,dx$ *from some time-t initial condition, we have*

$$\left|\pi_{(V)}(t,y_t;x)_{t,u} - \pi_{(V)}(t,\tilde{y}_t;x)_{t,u}\right| \leq |y_t - \tilde{y}_t|\,.\ell\exp(\ell).$$

In particular, the flow associated to $dy = V(y)\,dx$ *is Lipschitz continuous.*

PROOF. Routine application of Gronwall's lemma. □

Recall that we keep working with the Hölder control $\omega(s,t) = t - s$.

LEMMA 56 (Davie's Lemma). *Assume* $N \geq [p]$ *and*
(i) $V = (V_i)_{1\leq i\leq d}$ *is a colllection of vector fields in* $\mathrm{Lip}^N(\mathbb{R}^e)$,
(ii) $x : [0,T] \to \mathbb{R}^d$ *is a Lipschitz path and* $\mathbf{x} := S_{[p]}(x)$ *is its canonical lift to a* $G^{[p]}(\mathbb{R}^d)$*-valued path,*
(iii) $y_0 \in \mathbb{R}^e$ *is an initial condition.*
Then there exists a constant C_1 *depending on* p, γ *(and* ***not*** *depending on the Lipschitz norm or length of* x*) such that for all* $s < t$ *in* $[0,T]$,

$$\left|\pi_{(V)}(0,y_0;x)_{s,t}\right| \leq C_1\phi_p\left(|V|_{\mathrm{Lip}^{[p]}}\|\mathbf{x}\|_{p,\omega}\,\omega(s,t)^{1/p}\right) \tag{4.1}$$

where $\phi_p(a) = \max(a, a^p)$. *Then, for any time−s initial condition* $y_s \in \mathbb{R}^e$, *and constant* $K = K(N)$,

(4.2)
$$\left|\pi_{(V)}(s, y_s; x)_{s,t} - \mathcal{E}_{(V)}\left(s, y_s; S_N(x)_{s,t}\right)\right| \leq C_2 \left(K |V|_{\mathrm{Lip}^N} \|\mathbf{x}\|_{p,\omega}\right)^{N+1} \omega(s,t)^{\theta}$$

where $\theta \equiv (N+1)/p > 1$ *and* C_2 *depends*[6] *on* γ *and* p.

Remark 57. *For every fixed* $s < t$ *in* $[0,T]$ *we can find a Lipschitz path* $x^{s,t}$: $[0,T] \to \mathbb{R}^d$ *such that*

$$S_N\left(x^{s,t}\right)_{s,t} = S_N(x)_{s,t} \quad \text{and} \quad \int_s^t \left|dx^{s,t}\right| \leq K \|\mathbf{x}\|_{p,\omega} \omega(s,t)^{1/p} \tag{4.3}$$

Indeed, it suffices to take $N = [p]$ *and take* $x^{s,t}$ *as geodesic associated to the element* $\mathbf{g} = S_{[p]}(x)_{s,t} \in G^{[p]}\left(\mathbb{R}^d\right)$, *parametrized on the interval* $[s,t]$. *The length of this curve is precisely equals to* $\|\mathbf{g}\|$, *the CC norm of* $\mathbf{g}$ *in* $G^{[p]}\left(\mathbb{R}^d\right)$, *and so*

$$\int_s^t \left|dx^{s,t}\right| = \left\|S_{[p]}(x)_{s,t}\right\| \leq \|\mathbf{x}\|_{p,\omega} \omega(s,t)^{1/p}.$$

(We can even take $K = 1$ *here). If the vector fields enjoy more regularity i.e.* Lip^N *with* $N > [p]$, *one can use the fact that the Lyons-lift of* $\mathbf{x}$ *to* $S_N(\mathbf{x})$ *comes with the estimate*

$$\|S_N(\mathbf{x})\|_{p,\omega} \leq C_N \|\mathbf{x}\|_{p,\omega}$$

*and so (**??**) holds with* $K = C_N$. *The only benefit of this is the improved "Euler"-estimate (4.2). In the proof below (4.1) will be obtained as a consequence of (4.2) which holds in particular for* $N = [p]$ *and* $K = 1$.

Remark 58. *The same arguments show that the* $\mathrm{Lip}^N, N \geq [p]$, *assumption can be weakened to* $\mathrm{Lip}^{\gamma-1}$ *where* $\gamma > p$.

Proof. W.l.o.g. $|V|_{\mathrm{Lip}^N} \leq 1$ (otherwise, rescale x ...). The ODE $dy = V(y)\,dx$ with time-s initial condition y_s has a unique solution, denoted as usual by $\pi(s, y_s; x)$. We define $\mathbf{x} = S_{[p]}(x)$, and

$$\nu(s,t) = \left(K \|\mathbf{x}\|_{p,\omega}\right)^p |t-s|$$

Then, for all $s < t$ in $[0,T]$ we define,

$$\Gamma_{s,t} = y_t - \pi\left(s, y_s; x^{s,t}\right)_t = y_{s,t} - \pi\left(s, y_s; x^{s,t}\right)_{s,t}.$$

Then, for fixed $s < t < u$ in $[0,T]$, we have

$$\Gamma_{s,u} - \Gamma_{s,t} - \Gamma_{t,u} = -\pi\left(s, y_s; x^{s,u}\right)_{s,u} + \pi\left(s, y_s; x^{s,t}\right)_{s,t} + \pi\left(t, y_t; x^{t,u}\right)_{t,u}.$$

Define $x^{s,t,u}$ to be the concatenation of $x^{s,t}$ and $x^{t,u}$ and set, for better readability,

$$\begin{aligned} A &\equiv \pi\left(s, y_s; x^{s,t,u}\right)_{s,u} - \pi\left(s, y_s; x^{s,u}\right)_{s,u} \\ B &\equiv \pi\left(t, y_t; x^{t,u}\right)_{t,u} - \pi\left(t, \pi\left(s, y_s; x^{s,t}\right)_t; x^{t,u}\right)_{t,u} \\ &= \pi\left(t, y_t; x^{t,u}\right)_{t,u} - \pi\left(t, y_t + \Gamma_{s,t}; x^{t,u}\right)_{t,u} \end{aligned}$$

which then allows to write

$$\Gamma_{s,u} - \Gamma_{s,t} - \Gamma_{t,u} = A + B. \tag{4.4}$$

[6] We do not track dependence on the dimensions d and e.

The term A is estimated by - nomen est omen - *Lemma A*, noting that

$$\int_s^u \left|dx^{s,t,u}\right| = \int_s^t \left|dx^{s,t}\right| + \int_t^u \left|dx^{t,u}\right| \leq 2K \left\|\mathbf{x}\right\|_{p,\omega} \omega\left(s,u\right)^{1/p}.$$

Similarly, *Lemma B* was tailor-made to estimate B and we are led to (recall $\theta = (N+1)/p$)

$$\begin{aligned} \left|\Gamma_{s,u} - \Gamma_{s,t} - \Gamma_{t,u}\right| &\leq c_1 \nu\left(s,u\right)^{\theta} + c_2 \left|\pi\left(s, y_s; x^{s,t}\right)_t - y_t\right| \nu\left(t,u\right)^{1/p} \exp\left(c_2 \nu\left(t,u\right)^{1/p}\right) \\ &= c_1 \nu\left(s,u\right)^{\theta} + c_2 \left|\Gamma_{s,t}\right| \nu\left(t,u\right)^{1/p} \exp\left(c_2 \nu\left(t,u\right)^{1/p}\right) \end{aligned}$$

The elementary inequality

$$1 + c_2 \nu\left(t,u\right)^{1/p} \exp\left(c_2 \nu\left(t,u\right)^{1/p}\right) \leq \exp\left(2c_2 \nu\left(s,u\right)^{1/p}\right),$$

combined with the triangle inequality, then gives

$$\left|\Gamma_{s,u}\right| \leq \left|\Gamma_{s,t}\right| \exp\left(2c_2 \nu\left(s,u\right)^{1/p}\right) + \left|\Gamma_{t,u}\right| + c_2 \nu\left(s,u\right)^{\theta}. \tag{4.5}$$

Second Step: The estimate (4.5), in conjunction with

$$\varrho\left(r\right) := \sup_{\substack{0 \leq s < t \leq T \\ \nu(s,t) \leq r}} \left|\Gamma_{s,t}\right|,$$

implies

$$\varrho\left(r\right) \leq 2\varrho\left(\frac{r}{2}\right) \exp\left(2c_1 \left(\frac{r}{2}\right)^{\frac{1}{p}}\right) + c_1 r^{\theta}. \tag{4.6}$$

An analysis of this recursion reveals the implication

$$\varrho\left(r\right) \leq c_3 r^{\theta} \exp\left(c_3 r^{1/p}\right), \quad c_3 = c_3\left(c_1\right) \tag{4.7}$$

provided that $\varrho\left(r\right) = o\left(r\right)$; but this follows simply from

$$\begin{aligned} \left|\Gamma_{s,t}\right| &= \left|y_{s,t} - \pi\left(s, y_s; x^{s,t}\right)_{s,t}\right| \\ &\lesssim \left[\left(\int_s^t \left|dx\right|\right)^{N+1} \vee \left(\int_s^t \left|dx^{s,t}\right|\right)^{N+1}\right] \quad \text{use Lemma A} \\ &\lesssim \left[\left|x\right|_{Lip}^{N+1} \left|t-s\right|^{N+1} \vee \nu\left(s,t\right)^{\theta}\right] \\ &= o\left(\nu\left(s,t\right)\right). \end{aligned}$$

The catch here is that the constant in (4.7) depends on the constant in (4.6) but not on the rate of convergence of $\varrho\left(r\right)/r \to 0$ (which plainly depends on the Lipschitz norm of x).

Third Step: (4.7) translates to, for all $s < t$ in $[0,T]$,

$$\left|\Gamma_{s,t}\right| \leq c_3 \nu\left(s,t\right)^{\theta} \exp\left(c_3 \nu\left(s,t\right)^{\frac{1}{p}}\right).$$

For *small intervals*, those $s < t$ for which $\nu\left(s,t\right) \leq 1$, trivially,

$$\left|\Gamma_{s,t}\right| \leq \left(c_3 e^{c_3}\right) \nu\left(s,t\right)^{\theta} \leq c_4 \nu\left(s,t\right)^{\frac{1}{p}}$$

and together with $\left|\pi_{(V)}\left(s, y_s; x^{s,t}\right)_{s,t}\right| \leq \nu\left(s,t\right)^{\frac{1}{p}}$ it follows that

$$\left|\pi_{(V)}\left(0, y_0; x\right)_{s,t}\right| = \left|\pi_{(V)}\left(s, y_s; x\right)_{s,t}\right| \leq c_5 \nu\left(s,t\right)^{\frac{1}{p}}.$$

Given $s<t$ with $v(s,t)>1$ we split up $[s,t]$ in $\sim v(s,t)$ intervals of form $[\tau_i,\tau_{i+1}]$ to which the previous analysis applies. In particular $\nu(\tau_i,\tau_{i+1})\leq 1$ and by the triangle inequality,

$$\left|\pi_{(V)}(0,y_0;x)_{s,t}\right|\leq c_6 v(s,t).$$

At last, this allows to improve the estimate on $\Gamma_{s,t}$ of large intervals ($\nu(s,t)\geq 1$)

$$\begin{aligned}\left|\pi_{(V)}(0,y_0;x)_{s,t}-\pi_{(V)}\left(s,y_s;x^{s,t}\right)_{s,t}\right| &\leq \left|\pi_{(V)}(0,y_0;x)_{s,t}\right|+\left|\pi_{(V)}\left(s,y_s;x^{s,t}\right)_{s,t}\right|\\ &\leq c_6 v(s,t)+v(s,t)^{1/p}\leq c_7\nu(s,t)^{\theta}\end{aligned}$$

and in fact we see that this estimate is valid for all $s<t$ in $[0,T]$.

<u>Forth Step:</u> Since $S_N\left(x^{s,t}\right)_{s,t}=S_N(x)_{s,t}$ we have

$$\begin{aligned}&\left|\pi_{(V)}(s,y_s;x)_{s,t}-\mathcal{E}_{(V)}\left(s,y_s;S_N(x)_{s,t}\right)\right|\\ \leq\;& \left|\Gamma_{s,t}\right|+\left|\pi_{(V)}\left(s,y_s;x^{s,t}\right)_{s,t}-\mathcal{E}_{(V)}\left(s,y_s;S_N\left(x^{s,t}\right)_{s,t}\right)\right|\\ \leq\;& \left|\Gamma_{s,t}\right|+c_8\left(\int_s^t\left|dx^{s,t}\right|\right)^{N+1}\\ \leq\;& c_9\nu(s,t)^{\theta}.\end{aligned}$$

The estimate from the second to the third is a straight-forward higher order Euler estimate, based on Taylor expansion of y at time s. The last estimate follows from $\left|\Gamma_{s,t}\right|\lesssim\nu(s,t)^{\theta}$, established in the third step, and the assumption on $x^{s,t}$. □

REMARK 59. *Every geometric rough path* $\mathbf{x}\in C^{p,\omega}\left([0,T],G^{[p]}\left(\mathbb{R}^d\right)\right)$ *can be approximated by lifted Lipschitz paths* $S_{[p]}(x^n)$ *(cf. Thm 47) so that* $\sup_n\left\|S_{[p]}(x^n)\right\|_{p,\omega}<\infty$*. It is then clear from (4.1) that the sequence of ODE solutions* $\left\{\pi_{(V)}(0,y_0;x^n):n\geq 1\right\}$ *is equicontinuous (and bounded). By Arzela-Ascoli, one can define any limit to be a solution to the "rough" differential equation* $dy=V(y)\,d\mathbf{x}$ *started at* y_0*; that is, we have existence of RDE solutions. With more work, and inparticular assuming one more degree of regularity* $\mathrm{Lip}^{[p]+1}(\mathbb{R}^e)$*, one can check that there is a unique RDE solution.*

We shall proceed somewhat differently below and construct RDE solution as the "closure" of the ODE solution map $x\mapsto y$ *in rough path metrics. To this end, we need further ODE estimates ...*

5. Rough Paths Estimates for ODEs II

The following theorem is Lipschitz estimate (in a certain rough path metric) of the ODE solution map

$$x\mapsto y\equiv\pi_{(V)}(0,y_0;x)$$

Recall that $\omega(s,t)=t-s$ throughout.

THEOREM 60. *Assume that*[7]
(i) $V=(V_1,\dots,V_d)$ *is a collection of* $\mathrm{Lip}^{[p]+1}$*-vector fields on* $\mathbb{R}^e$*;*
(ii) $x^1,x^2:[0,T]\to\mathbb{R}^d$ *are Lipschitz paths, canonically lifted to* $\mathbf{x}^1,\mathbf{x}^2\in C^{1,\omega}\left([0,T],G^{[p]}\left(\mathbb{R}^d\right)\right)$*, and set*

$$\nu(s,t)=|V|^p_{\mathrm{Lip}^{[p]+1}}\left(\left\|\mathbf{x}^1\right\|^p_{p,\omega}+\left\|\mathbf{x}^2\right\|^p_{p;\omega}\right)|t-s|$$

[7] The regularity assumption can be improved to $V\in Lip^{\gamma},\gamma>p$.

(iii) for all $0 \le s < t \le T$, π_k *denotes the projection from* $T^N\left(\mathbb{R}^d\right) \to \left(\mathbb{R}^d\right)^{\otimes k}$,

$$|V|_{\mathrm{Lip}^{[p]+1}}^k \left|\pi_k\left(\mathbf{x}_{s,t}^1 - \mathbf{x}_{s,t}^2\right)\right| \le \varepsilon v(s,t)^{\frac{k}{p}} \text{ for } k = 1, \ldots, [p].$$

(v) $y_0^1, y_0^2 \in \mathbb{R}^e$ *thought of time-*0 *initial conditions*

Set $y^i = \pi_{(V)}\left(0, y_0^i; x^i\right)$. *Then, for all* $0 \le s < t \le T$,

$$\left|y_{s,t}^1 - y_{s,t}^2\right| \le \left(\left|y_0^1 - y_0^2\right| + \varepsilon\right).Cv(s,t)^{1/p} \exp\left(Cv(0,T)\right).$$

where C *depends*[8] *on* p, γ.

Remark 61. *The same arguments show that the* $\mathrm{Lip}^{[p]+1}$ *assumption can be replaced by* Lip^{γ} *for any* $\gamma > p$.

Proof. (Sketch) W.l.o.g. $|V|_{\mathrm{Lip}^{[p]+1}} \le 1$ and so we may take

$$\nu(s,t) = \left(\left\|\mathbf{x}^1\right\|_{p,\omega}^p + \left\|\mathbf{x}^2\right\|_{p;\omega}^p\right)|t-s|.$$

We can find paths $x^{1,s,t}$ and $x^{2,s,t}$ such that

$$S_{[p]}\left(x^{i,s,t}\right)_{s,t} = \mathbf{x}_{s,t}^i,$$

and such that

$$\begin{aligned} \int_s^t d\left|x_r^{i,s,t}\right| &\lesssim \omega(s,t)^{1/p}, \\ \int_s^t d\left|x_r^{1,s,t} - x_r^{2,s,t}\right| &\lesssim \varepsilon\omega(s,t)^{1/p}. \end{aligned}$$

Define similarly $x^{i,s,u}$ and $x^{i,t,u}$, and then $x^{i,s,t,u}$ to be the concatenation of $x^{i,s,t}$ and $x^{i,t,u}$. Then, following the proof of Davie's lemma, we set

$$\Gamma_{s,t}^i = y_{s,t}^i - \pi\left(s, y_s^i; x^{i,s,t}\right)_{s,t}, \; i = 1, 2,$$

and also $\overline{\Gamma}_{s,t} := \Gamma_{s,t}^1 - \Gamma_{s,t}^2$. Set $\theta = ([p]+1)/p > 1$. From Davie's Lemma applied with $N = [p]$,

$$\left|\Gamma_{s,t}^i\right| \lesssim \omega(s,t)^{\theta}, \; i = 1, 2 \implies \left|\overline{\Gamma}_{s,t}\right| \lesssim c_1\omega(s,t)^{\theta}$$

We now proceed similarly as in Davie's lemma and set

$$\begin{aligned} A^i &: = \pi\left(s, y_s^i; x^{i,s,t,u}\right)_{s,u} - \pi\left(s, y_s^i; x^{i,s,u}\right)_{s,u} \\ B^i &: = \pi\left(t, y_t^i; x^{i,t,u}\right)_{t,u} - \pi\left(t, \pi\left(s, y_s^i; x^{i,s,t}\right)_t; x^{i,t,u}\right)_{t,u} \end{aligned}$$

and $\bar{A} := A^1 - A^2, \bar{B} := B^1 - B^2$. Clearly, $\overline{\Gamma}_{s,u} - \overline{\Gamma}_{s,t} - \overline{\Gamma}_{t,u} = \bar{A} + \tilde{B}$ so that

$$\left|\overline{\Gamma}_{s,u} - \overline{\Gamma}_{s,t} - \overline{\Gamma}_{t,u}\right| \le \left|\bar{A}\right| + \left|\bar{B}\right|.$$

The terms A^i arise from ODEs on $[s,u]$ with identical initial condition y_s^i, driven by different signals $x^{i,s,t,u}$ resp. $x^{i,s,u}$ which, by construction, share the same step-$[p]$ signature. The terms B^i arise from identical ODEs with different starting points. We estimate these terms by suitable technical ODE lemmas (which we would call "*Lemma* $\overline{A}$,*Lemma* $\overline{B}$"...). and since this involves the difference of A^i, B^i we need

[8] We do not track dependence on the dimensions d and e.

one more degree of differentiability (hence the assumption $V \in Lip^{[p]+1}$ ...). This leads to an estimate of form

$$\left|\bar{\Gamma}_{s,u}\right| \leq \left|\bar{\Gamma}_{s,t}\right| \exp\left(c\omega(s,u)^{1/p}\right) + \left|\bar{\Gamma}_{t,u}\right| + \left(\max_{r\in\{s,t\}}\left|y_r^1 - y_r^2\right| + \varepsilon\right) c\omega(s,u)^{\theta} \exp\left(c\omega(s,u)^{1/p}\right) \tag{5.1}$$

which can be exploited to yield the claimed estimates. □

6. Rough Differential Equations

The rough path metric $d_{p,\omega}(\mathbf{x},\tilde{\mathbf{x}})$ defines a metric on the space of Lipschitz path (modulo starting point). Indeed, given $x,\tilde{x} \in C^{1,\omega}\left([0,T],\mathbb{R}^d\right)$ where $\omega(s,t) = t-s$, their distance is defined as $d_{p,\omega}(\mathbf{x},\tilde{\mathbf{x}})$ where $\mathbf{x} := S_{[p]}(x), \tilde{\mathbf{x}} := S_{[p]}(\tilde{x})$. Theorem 60 then gives the following

Corollary 62. *Assume that $V = (V_1,...,V_d)$ is a collection of* Lip$^{[p]+1}$*-vector fields on $\mathbb{R}^e$; then the ODE solution map to $dy = V(y)\,dx$, more precisely*

$$(y_0,x) \in \mathbb{R}^e \times C^{1,\omega}\left([0,T],\mathbb{R}^d\right) \mapsto y \equiv \pi_{(V)}(0,y_0;x) \in C^{1,\omega}\left([0,T],\mathbb{R}^e\right)$$

is uniformly continuous on bounded sets with respect to the metric

$$((y_0,x),(\tilde{y}_0,\tilde{x})) \mapsto |y_0 - \tilde{y}_0| + d_{p,\omega}\left(S_{[p]}(x), S_{[p]}(\tilde{x})\right).$$

Proof. Write

$$\rho_{p,\omega}(\mathbf{x},\tilde{\mathbf{x}}) := \sup_{0\leq s\leq t\leq T}\max_{k=1,\ldots,[p]}\frac{\left|\pi_k\left(\mathbf{x}_{s,t}-\tilde{\mathbf{x}}_{s,t}\right)\right|}{|t-s|^{k/p}} \text{ and } d_{p,\omega}(\mathbf{x},\tilde{\mathbf{x}}) = \sup_{0\leq s\leq t\leq T}\frac{d\left(\mathbf{x}_{s,t},\tilde{\mathbf{x}}_{s,t}\right)}{|t-s|^{1/p}}.$$

Assume $\|\mathbf{x}\|_{p,\omega}, \|\tilde{\mathbf{x}}\|_{p,\omega} \leq R$ and consider $\mathbf{g} := \delta_{|t-s|^{-1/p}}\mathbf{x}_{s,t}$ and define similarly $\tilde{\mathbf{g}}$. Clearly, $\|\mathbf{g}\|, \|\tilde{\mathbf{g}}\| \leq R$ i.e. remain in bounded region of $G^{[p]}\left(\mathbb{R}^d\right)$. Given $\varepsilon > 0$ it suffices to show that there exists $\delta = \delta(\varepsilon,R)$ such that

$$\frac{d\left(\mathbf{x}_{s,t},\tilde{\mathbf{x}}_{s,t}\right)}{|t-s|^{1/p}} = d(\mathbf{g},\tilde{\mathbf{g}}) < \delta \implies \max_{k=1,\ldots,[p]}\left|\pi_k(\mathbf{g}-\tilde{\mathbf{g}})\right| < \varepsilon.$$

But this is easy to see. Indeed,

$$d(\mathbf{g},\tilde{\mathbf{g}}) = \left\|\mathbf{g}^{-1}\otimes\tilde{\mathbf{g}}\right\| \sim \max_{k=1,\ldots,[p]}\left|\pi_k\left(\mathbf{g}^{-1}\otimes\tilde{\mathbf{g}}\right)\right|^{1/k}$$

by equivalence of homogenous norms and so, at least for $\delta < 1$, $\left|\pi_k\left(\mathbf{g}^{-1}\otimes\tilde{\mathbf{g}}\right)\right| \lesssim \delta^k \leq \delta$ for all k. At last, estimating $\left|\pi_k(\mathbf{g}-\tilde{\mathbf{g}})\right|$ in terms of $\left|\pi_k\left(\mathbf{g}^{-1}\otimes\tilde{\mathbf{g}}\right)\right|$ is an easy consequence of continuity of the group operations. □

Knowing that the ODE solution map

$$\begin{aligned} x &\in C^{1,\omega}\left([0,T],\mathbb{R}^d\right) \\ &\mapsto y \equiv \pi_{(V)}(0,y_0;x) \in C^{1,\omega}\left([0,T],\mathbb{R}^d\right) \subset C^{p,\omega}\left([0,T],\mathbb{R}^d\right) \end{aligned}$$

is (uniformly) continuous (on bounded sets) with respect to $d_{p,\omega}$ it follows from basic analysis that $\pi_{(V)}(0,y_0;\cdot)$ has a *unique continuous* extension to the abstract

$d_{p,\omega}$-closure of Lipschitz paths. Luckily, we fully understand the resulting closure. It suffices to view, via $x \mapsto \mathbf{x} := S_{[p]}(x)$,

$$\begin{aligned} C^{1,\omega}\left([0,T],\mathbb{R}^d\right) &\subset C^{1,\omega}\left([0,T],G^{[p]}\left(\mathbb{R}^d\right)\right) \\ &\subset C^{p,\omega}\left([0,T],G^{[p]}\left(\mathbb{R}^d\right)\right) \end{aligned}$$

so that the $d_{p,\omega}$-closure of Lipschitz paths is nothing but the $C^{0,p,\omega}\left([0,T],G^{[p]}\left(\mathbb{R}^d\right)\right)$, the space of geometric p-rough path with (Hölder-) control ω. We shall write

$$\begin{aligned} \mathbf{x} &\in C^{0,p,\omega}\left([0,T],G^{[p]}\left(\mathbb{R}^d\right)\right) \\ &\mapsto y \equiv \pi_{(V)}(0,y_0;\mathbf{x}) \in C^{p,\omega}\left([0,T],\mathbb{R}^d\right) \end{aligned}$$

for this extension and call it solution to the rough differential equation (RDE), formally denoted by,

$$dy = V(y)\,d\mathbf{x}$$

started at y_0. To summarize,

Theorem 63 (T.Lyons, Universal Limit Theorem). *Let $\mathbf{x}$ be a geometric p-rough path (i.e. $\mathbf{x} \in C^{0,p,\omega}\left([0,T],G^{[p]}\left(\mathbb{R}^d\right)\right)$. Given $V \in \mathrm{Lip}^{[p]+1}$ the RDE*

$$dy = V(y)\,d\mathbf{x},$$

has a unique solution $y \equiv \pi_{(V)}(0,y_0;\mathbf{x})$ for every starting point y_0. In the case when $\mathbf{x} = S_{[p]}(x)$, x Lipschitz, this RDE solution coincides with the ODE solution to $dy = V(y)\,dx$, started at y_0. Moreover

$$(y_0,\mathbf{x}) \mapsto y$$

is (uniformly) continuous (on bounded sets).

Remark 64. *All the ODE estimates derived earlier remain valid in the limit. For instance,*

$$\left|\pi_{(V)}(0,y_0;\mathbf{x})_{s,t}\right| \leq C_1\phi_p\left(|V|_{\mathrm{Lip}^{[p]}}\|\mathbf{x}\|_{p,\omega}\,\omega(s,t)^{1/p}\right)$$

where $\phi_p(x) = \max(x,x^p)$ and for any $N \geq [p]$

$$\left|\pi_{(V)}(s,y_s;\mathbf{x})_{s,t} - \mathcal{E}_{(V)}\left(s,y_s;S_N(\mathbf{x})_{s,t}\right)\right| \leq C_2\left(K|V|_{\mathrm{Lip}^N}\|\mathbf{x}\|_{p,\omega}\right)^{N+1}\omega(s,t)^{\frac{N+1}{p}}.$$

(for $N > [p]$ the left-hand-side above involves the Lyons lift $S_N(\mathbf{x})$ of $\mathbf{x}$.)

As already remarked, the arguments presented here give, with minimal changes, the ULT under Lip^γ-regularity of the vector fields, $\gamma > p \geq 1$. Let us state the following theorem (without proof) under optimal regularity assumptions.

Theorem 65 (RDE flows). *Under the assumptions of theorem 63, in particular Lip^γ-regularity of the vector fields, $\gamma > p \geq 1$, the flow*

$$y_0 \in \mathbb{R}^e \mapsto \pi_{(V)}(0,y_0;\mathbf{x})$$

is continuously differentiable. Assuming $\mathrm{Lip}^{k-1+\gamma}(\mathbb{R}^e)$-regularity, $k \in \mathbb{N}$, implies that the flow is C^k. For every fixed T,

$$y_0 \in \mathbb{R}^e \mapsto \pi_{(V)}(0,y_0;\mathbf{x})_T$$

is then a C^k-diffeomorphism with explicit inverse given by

$$\pi_{(V)}\left(0,\cdot;\overleftarrow{\mathbf{x}}\right)_T$$

where $\overleftarrow{\mathbf{x}}(t)=\mathbf{x}(T-t)$. *Moreover,*

$$\left\|D\pi_{(V)}(0,\cdot;\mathbf{x})|_{y_0}\right\|_{p,\omega}\leq C\exp\left(C\left|V\right|^p_{\mathrm{Lip}^\gamma}\left\|\mathbf{x}\right\|^p_{p,\omega}\right)$$

where $C=C(p,\gamma)$.

THEOREM 66 (Limit theorem for RDE flows). *Assume* $\mathrm{Lip}^{k-1+\gamma}(\mathbb{R}^e)$*-regularity and that* $(\mathbf{x}^n)\subset C^{0,p,\omega}\left([0,T],G^{[p]}\left(\mathbb{R}^d\right)\right)$ *converges to some* $\mathbf{x}^\infty$. *Then for all multiindices* α *with* $|\alpha|\leq k$

$$\partial_\alpha\pi_{(V)}(0,y_0;\mathbf{x}^n)\to\partial_\alpha\pi_{(V)}(0,y_0;\mathbf{x}^\infty)$$

in $1/p$*-Hölder norm, uniformly over* y_0 *in compacts.*

CHAPTER 2

Applications to Stochastic Analysis

1. Enhanced Brownian motion as geometric rough path

1.1. Brownian Motion. Let B_t an $\mathbb{R}^d$-valued Brownian motion started at 0.

We can assume that B_t is continuous.

PROPOSITION 67. *Let $\alpha \in [0, 1/2)$. Then Brownian motion is α-Hölder continuous almost surely,*

$$|B|_{\alpha\text{-}H\ddot{o}lder;[0,T]} \equiv \sup_{0 \leq s < t \leq T} \frac{|B_{s,t}|}{|t-s|^{\alpha}} < \infty \text{ a.s.}$$

PROOF. Standard. □

LEMMA 68. *There exists a constant $c > 0$ such that*

$$\forall M \geq 0 : \mathbb{P}\left[|B|_{\infty;[0,T]} \geq M\right] \leq \frac{1}{c} e^{-cM^2/T}.$$

PROOF. It suffices to consider $d = 1$. From $B \overset{\mathcal{D}}{=} (-B)$, the reflection principle[1] and Brownian scaling we see that

$$\begin{aligned} \mathbb{P}\left[|B|_{\infty;[0,T]} \geq M\right] &\leq 2\mathbb{P}[S_T \geq M] \\ &= 4\mathbb{P}(B_T \geq M) \\ &= 4\mathbb{P}\left(B_1 \geq M/T^{1/2}\right). \end{aligned}$$

The result follows from the usual tail behaviour of B_1which is a standard normal random variable. □

EXERCISE 69. *Show that the above Gauss tail estimate is equivalent to the existence of $c > 0$ small enough such that $\mathbb{E}\left[e^{c|B|^2_{\infty;[0,T]}}\right] < \infty$. (Hint: $\mathbb{E}[X] = \int_0^\infty \mathbb{P}[X \geq x]\, dx$.) How large can c be?*

1.2. Lévy's Area. For fixed $t \in [0, T]$ and $i, j \in \{1, ..., d\}$ we define *Lévy's area* as

$$A_t^{i,j} = \frac{1}{2}\left(\int_0^t B_s^i dB_s^j - B_s^j dB_s^i\right)$$

[1] The reflection principle implies $\forall a, y, t \geq 0 : \mathbb{P}[S_t \geq a, B_t \leq a - y] = \mathbb{P}[B_t \geq a + y]$. Apply this to $y = 0$ and note the trivial $\mathbb{P}[S_t \geq a, B_t \geq a] = \mathbb{P}[B_t \geq a]$.

Obviously, $A_t \in so(d)$, the anti-symmetric matrices, and it suffices to consider $i \neq j$ (or $i < j$ if you wish). From basic principles of stochastic integration,

$$A_t^{i,j} = L^2\text{-}\lim_n \frac{1}{2}\left(\sum_{t_k \in D^n : t_k < t} B_{t_k}^i B_{t_k,t_{k+1}}^j - \sum_{t_k \in D^n} B_{t_k}^j B_{t_k,t_{k+1}}^i\right). \tag{1.1}$$

Note that the "building blocks" of $A_t^{i,j}$ are of form $\int_0^t \beta_s d\tilde{\beta}_s$. Then, *conditional* on $\beta(.)$, we can view $\int_0^t \beta_s d\tilde{\beta}_s$ as if $\beta \in C[0,T] \subset L^2[0,T]$ were deterministic and

$$\int_0^t \beta_s d\tilde{\beta}_s \sim N\left(0, \int_0^t \beta_s^2 ds\right).$$

This leads to a useful integrability property of Lévy's area.

PROPOSITION 70. *There exists $c = c(d,T) > 0$ s.t. for all $t \in [0,T]$.*

$$\mathbb{E}\left[e^{c|A_t|}\right] < \infty.$$

(We could easily find the precise t dependence of c by scaling.)

PROOF. Write

$$\begin{aligned} & \mathbb{E}\left[e^{c\left|\int_0^t \beta_s d\tilde{\beta}_s\right|}\middle|\sigma\left(\beta_s : 0 \leq s \leq T\right)\right] \\ = \quad & \mathbb{E}\left[e^{c|Z|}\right] \text{ with } Z \sim N\left(0, \int_0^t \beta_s^2 ds\right) \\ \leq \quad & 2\mathbb{E}\left[e^{cZ}\right] = 2\exp\left(\frac{c^2}{2}\int_0^t \beta_s^2 ds\right) \\ \leq \quad & 2\exp\left(\frac{c^2 t}{2}|\beta|_{\infty;[0,t]}^2\right). \end{aligned}$$

Hence

$$\mathbb{E}\left[e^{c\left|\int_0^t \beta_s d\tilde{\beta}_s\right|}\right] \leq 2\mathbb{E}\exp\left(\frac{c^2 t}{2}|\beta|_{\infty;[0,t]}^2\right) < \infty$$

for $c > 0$ small enough, using the lemma above. This leads readily to exponential moments for $|A_t|$. □

1.3. Enhanced Brownian Motion. If $x : [0,T] \to \mathbb{R}^d$ is Lipschitz started at 0 then its step-2 lift satisfies

$$S_2(x)_t = \exp(x_t + a_t) \in G^2(\mathbb{R}^d)$$

with area $a_t^{i,j} = \frac{1}{2}\left(\int_0^t x_s^i dx_s^j - x_s^j dx_s^i\right)$. Recall also that

$$S_2(x)_{s,t} = S_2(x)_s^{-1} \otimes S_2(x)_t = \exp(x_{s,t} + a_{s,t})$$

where $x_{s,t} = x_t - x_s$ and $a_{s,t} \in so(d)$ is given by

$$a_{s,t}^{i,j} = \frac{1}{2}\left(\int_s^t x_{s,r}^i dx_r^j - x_{s,r}^j dx_r^i\right).$$

Similarly, we define *enhanced Brownian motion* as the $G^2(\mathbb{R}^d)$-valued process

$$\mathbf{B}_t := \exp[B_t + A_t].$$

We also write $\mathbf{B}_{s,t} = \mathbf{B}_s^{-1} \otimes \mathbf{B}_t \in G^2(\mathbb{R}^d)$ and observe that is consistent with

$$\mathbf{B}_{s,t} = \exp[B_{s,t} + A_{s,t}]$$

where $B_{s,t} = B_t - B_s$ (as usual) and $A_{s,t} \in so(d)$ a.s. equal to

$$A_{s,t}^{i,j} = \frac{1}{2}\left(\int_s^t B_{s,r}^i dB_r^j - B_{s,r}^j dB_r^i\right).$$

Why? We just recalled all this is holds for Lipschitz paths, where one can write out all iterated integrals as Riemann–Stieltjes integrals. This is still true for the Brownian case but now convergence is only in L^2-sense, see 1.1, and L^2-limits are only defined up to null-sets hence the a.s. above.

EXERCISE 71. *Show that defining* $\mathbf{B}_t \in T_1^2(\mathbb{R}^d)$ *by the enhancement* $\mathbf{B}_t^{2;ij} := \int_0^t B^i dB^j$ *(Itô-integral) for all* $i,j = 1,..,d$ *does not yield a geometric rough path. Hint: consider* $i = j$ *and compute the expectation.*

REMARK 72. *On can check that the second level of* $\mathbf{B}_t := \exp[B_t + A_t]$ *is given by the Stratonovich-integral* $\mathbf{B}_t^{2;ij} := \int_0^t B^i \circ dB^j$. *Note that the bracket* $\langle B^i, B^j \rangle = 0$ *if* $i \neq j$ *which means the Lévy-area does not care if you use Itô- or Stratonovich-integration. In short, our* $\mathbf{B}_t$ *can be viewed as Stratonovich-lift of Brownian motion.*

LEMMA 73. *(a) For all* $\lambda > 0$ *we have*

$$(\mathbf{B}_{\lambda^2 t} : t \geq 0) \stackrel{\mathcal{D}}{=} (\delta_\lambda \mathbf{B}_t : t \geq 0)$$

where δ *is the dilatation operator on* $G^2(\mathbb{R}^d)$.
(b) For all $s \geq 0$ *we have*

$$(\mathbf{B}_{s,s+t} : t \geq 0) \stackrel{\mathcal{D}}{=} (\mathbf{B}_t : t \geq 0).$$

(c) The process $(\mathbf{B}_{s,s+t} : t \geq 0)$ *is independent of* $\sigma(\mathbf{B}_r : r \leq s)$

PROOF. (a) From Brownian scaling, for any $\lambda > 0$ we have

$$(B_{\lambda^2 t})_{t \geq 0} \stackrel{\mathcal{D}}{=} (\lambda B_t)_{t \geq 0}$$

That is, speeding up time by a factor λ^2 is, in law, equivalent to spacial scaling by a factor λ. Since A is determinted as limit of a homogenuous polynomial of degree

2 in terms of Brownian increments, see (1.1), the scaling factor λ appears twice and one has

$$(B_{\lambda^2 t}, A_{\lambda^2 t})_{t\geq 0} \overset{\mathcal{D}}{=} \left(\lambda B_t, \lambda^2 A_t\right)_{t\geq 0}.$$

Now apply $\exp\colon \mathbb{R}^d \oplus so\,(d) \to G^2\left(\mathbb{R}^d\right)$.

(b) Recall that for Brownian motion, for all $s \geq 0$,

$$(B_{s,s+t} : t \geq 0) \overset{\mathcal{D}}{=} (B_t : t \geq 0)\,.$$

Then, for s fixed,

$$\begin{aligned}\left(A^{i,j}_{s,s+t}\right)_{t\geq 0} &= \frac{1}{2}\left(\int_s^{s+t} B^i_{s,r}dB^j_r - B^j_{s,r}dB^i_r\right)_{t\geq 0}\\ &= \frac{1}{2}\left(\int_s^{s+t} B^i_{s,r}dB^j_{s,r} - B^j_{s,r}dB^i_{s,r}\right)_{t\geq 0}\\ &\overset{\mathcal{D}}{=} \frac{1}{2}\left(\int_0^{t} B^i_{r}dB^j_{r} - B^j_{r}dB^i_{r}\right)_{t\geq 0}\end{aligned} \tag{1.2}$$

and the same holds for the pair

$$(B_{s,t}, A_{s,s+t})_{t\geq 0} \overset{\mathcal{D}}{=} (B_t, A_t)_{t\geq 0}\,.$$

Conclude as in part (a). For part (c) observe that, since A_r is a measurable function of $\{B_u : u \leq r\}$,

$$\sigma\left(\mathbf{B}_r : r \leq s\right) = \sigma\left(B_r, A_r : r \leq s\right) = \sigma\left(B_r : r \leq s\right).$$

On the other hand, $(B_{s,s+t}, A_{s,s+t} : t \geq 0)$ is a measurably determined by

$$\sigma\left(B_{s,r} : r \geq s\right) = \sigma\left(B_{s,s+t} : t \geq 0\right),$$

see (1.2) in particular. From defining properties of Brownian motion, $\sigma\left(B_r : r \leq s\right)$ and $\sigma\left(B_{s,s+t} : t \geq 0\right)$ are indepedent and this finishes the proof. □

REMARK 74. *We will check momentarily that we can take* $\mathbf{B}$ *to have continuous sample paths. Properties (b) & (c) are then summarized by saying the* $\mathbf{B}$ *is a left(-invariant) Brownian motion on the Liegroup* $\left(G^2\left(\mathbb{R}^d\right), \otimes\right)$.

For the next statement, note that $d\left(\mathbf{B}_s, \mathbf{B}_t\right) = \|\mathbf{B}_{s,t}\|$.

LEMMA 75. *(a) The r.v.* $\|\mathbf{B}_{0,1}\|$*has Gauss tails. (b)For every* $q \in [1, \infty)$ *there exists a constant* $c = c\,(q)$ *:*

$$\mathbb{E}\left[d\left(\mathbf{B}_s, \mathbf{B}_t\right)^q\right] = \mathbb{E}\left[\|\mathbf{B}_{s,t}\|^q\right] = c\,|t-s|^{q/2}.$$

PROOF. From Lemma 73 and homogeneity of the Carnot-Caratheodory norm,

$$\mathbf{B}_{s,t} \overset{\mathcal{D}}{=} \delta_{(t-s)^{1/2}}\mathbf{B}_{0,1} \text{ and } \|\mathbf{B}_{s,t}\| \overset{\mathcal{D}}{=} (t-s)^{1/2}\|\mathbf{B}_{0,1}\|\,.$$

Hence, it thus suffices to show that

$$\mathbb{E}\left[\|\mathbf{B}_{0,1}\|^q\right] < \infty.$$

By equivalence of homogenous norms

$$\|\mathbf{B}_{0,1}\| \sim \max\left\{|B_{0,1}|\,, |A_{0,1}|^{1/2}\right\}.$$

Clearly the Gaussian r.v. $B_{0,1}$ has Gauss tails and the same is true for $|A_{0,1}|^{1/2}$ due to Proposition 70. We conclude that $\|\mathbf{B}_{0,1}\|$ has Gauss tails which, trivially, implies finite q-moments for all $q < \infty$. □

Corollary 76. *Enhanced Brownian motion has a modification which is a.s. continuous.*

Proof. The proof of Kolmogorov's criterion adapts with no changes to stochastic processes with values in a metric space. We simply apply it to $G^2\left(\mathbb{R}^d\right)$ with Carnot-Caratheodory distance d. □

Theorem 77 (Besov-Hölder embedding). *Let f be a continuous path on $[0,1]$ with values in a metric space (E,d) and let $\alpha \in [0,1/2)$. Then there exists $C_{77} > 0$ such that*

$$\|f\|_{\alpha\text{-}H\ddot{o}lder;[0,1]} \leq C_{77} \left(\int_{[0,1]^2} \frac{\left[d\left(f_s,f_t\right)\right]^{2k}}{|t-s|^k} dsdt \right)^{\frac{1}{2k}} \tag{1.3}$$

for all $k \geq k_0(\alpha)$.

Proof. A short proof is based on the Garsia-Rodemich-Rumsey lemma. See the book *Probability: An Analytic View* by D. Stroock. □

Theorem 78. *Let $\alpha \in [0,1/2)$. Then $\|\mathbf{B}\|_{\alpha\text{-}H\ddot{o}lder;[0,T]}$ has Gauss tails. Equivalently, there exists $C_{78} > 0$ s.t.*

$$\mathbb{E}\left[\exp\left(C_{78}\|\mathbf{B}\|^2_{\alpha\text{-}H\ddot{o}lder;[0,T]}\right)\right] < \infty.$$

Proof. By Brownian scaling it suffices to consider $T = 1$. Recall that $\|\mathbf{B}_{s,t}\| = d\left(\mathbf{B}_s,\mathbf{B}_t\right)$ with CC-distance d. For $\alpha \in [0,1/2)$ fixed and $k \geq k_0(\alpha)$

$$\begin{aligned} \mathbb{E}\left[\|\mathbf{B}\|^{2k}_{\alpha\text{-Hölder}}\right] &\leq C_{77}^{2k}\mathbb{E}\left[\int_{[0,1]^2} \frac{\|\mathbf{B}_{s,t}\|^{2k}}{|t-s|^k} dsdt\right] \\ &= C_{77}^{2k}\,\mathbb{E}\|\mathbf{B}_{0,1}\|^{2k}, \end{aligned}$$

using the properties given in Lemma 73. Clearly, $\mathbb{E}\left[\exp\left(c\|\mathbf{B}\|^2_{\alpha\text{-Hölder}}\right)\right] < \infty$ iff any tail of the sum

$$\sum_k \frac{c^k}{k!}\mathbb{E}\left[\|\mathbf{B}\|^{2k}_{\alpha\text{-Hölder}}\right]$$

is finite. But

$$\begin{aligned} \sum_{k \geq k_0(\alpha)} \frac{c^k}{k!}\mathbb{E}\left[\|\mathbf{B}\|^{2k}_{\alpha\text{-Hölder}}\right] &\leq \sum_{k \geq k_0(\alpha)} \frac{\left(cC_{77}^2\right)^k}{k!}\mathbb{E}\|\mathbf{B}_{0,1}\|^{2k} \\ &\leq \mathbb{E}\left[\exp\left(cC_{77}^2\|\mathbf{B}_{0,1}\|^2\right)\right] \end{aligned}$$

which is finite for c small enough in view of Lemma 75. □

Corollary 79. *Let $\alpha \in (1/3,1/2)$. Almost surely, EBM is a geometric $(1/\alpha)$-rough path with Hölder control. Moreover, the r.v. $M_1 = \|\mathbf{B}\|_{\alpha\text{-}H\ddot{o}lder;[0,T]}$ has Gauss tails. In particular, $M_1 \in L^p$ for all $p \in [1,\infty)$.*

Remark 80. *To avoid misunderstanding (control ω vs. random sample ω) we now write*

$$\mathbf{B}_t \in C^{a\text{-}H\ddot{o}lder}\left([0,T], G^2\left(\mathbb{R}^d\right)\right) \text{ with } \alpha = 1/p \in (0,1/2)$$

and also $\|.\|_{\alpha\text{-}H\ddot{o}lder}$ instead of $\|.\|_{p,\omega}$.

2. Approximations to Enhanced Brownian Motion

We now consider a sequence D^n of *nested* dissections, that is $t \in D^n \implies t \in D^{n+1}$, with mesh $|D^n| \to 0$ as $n \to \infty$. Define $B^n(\omega) = B^{D_n}(\omega)$ as the piecewise linear approximation based on the dissection D^n. We consider the step-2 lift and write, as usual,

$$S_2(B^n) = \mathbf{B}^n =: \exp(B^n + A^n).$$

As in the last section, $\mathbf{B} = \exp(B + A)$ denotes enhanced Brownian motion.

THEOREM 81. *Almost surely and in L^2, we have pointwise convergence $\mathbf{B}_t^n \to \mathbf{B}_t$.*

PROOF. Fix $t \in [0,T]$. The statement is $d(\mathbf{B}_t^n, \mathbf{B}_t) \to 0$ (a.s. and in L^2). This is equivalent to

$$\text{(a)}\ \ |B_t^n - B_t| \to 0 \quad \text{and} \quad \text{(b)}\ \ |A_t^n - A_t| \to 0.$$

Ad (a), since $\{D^n\}$ is nested, $\mathfrak{F}_n := \sigma(B_t : t \in D^n)$ forms a filtration. We claim that a.s.

$$\mathbb{E}[B_t|\mathfrak{F}_n] = B_t^n \text{ and } \mathbb{E}[A_t|\mathfrak{F}_n] = A_t^n.$$

Using the Markov property of B,

$$\mathbb{E}[B_t|\mathfrak{F}_n] = \mathbb{E}\left[B_t|B_{t_i}, B_{t_{i+1}}\right]$$

where t_i, t_{i+1} are two neighbours in D^n with $t \in [t_i, t_{i+1}]$. It is a simple exercise of Gaussian conditioning[2] to seethat

$$\mathbb{E}\left[B_t|B_{t_i}, B_{t_{i+1}}\right] = \frac{t_{i+1} - t}{t_{i+1} - t_i} B_{t_i} + \frac{t - t_i}{t_{i+1} - t_i} B_{t_{i+1}}$$

and this is precisely equal to B_t^n. Mesh $|D^n| \to 0$ implies that B_t is $(\vee_n \mathfrak{F}_n)$-measurable and martingale convergence shows that[3]

$$B_t^n = \mathbb{E}[B_t|\mathfrak{F}_n] \to B_t \text{ a.s and in } L^2.$$

Ad (b): We first fix n and show $\mathbb{E}[A_t|\mathfrak{F}_n] = A_t^n$. It simplifies things to set $\beta = B^i, \tilde{\beta} = B^j, i \neq j$ and consider $\int_0^t \beta d\tilde{\beta}$. Let $\left\{\tilde{D}^m\right\}$ be a dissection of $[0,t]$, with t fixed, and mesh $\left|\tilde{D}^m\right| \to 0$. By L^2-continuity of $\mathbb{E}[\cdot|\mathfrak{F}_n]$ and Lemma **??**,

$$\begin{aligned}
\mathbb{E}\left[\int_0^t \beta d\tilde{\beta}|\mathfrak{F}_n\right] &= \lim_{m\to\infty} \mathbb{E}\left[\sum_{t_i \in \tilde{D}^m} \beta_{t_i}\tilde{\beta}_{t_i,t_{i+1}} \middle| \mathfrak{F}_n\right] \\
&= \lim_{m\to\infty} \sum_{t_i \in \tilde{D}^m} \mathbb{E}\left[\beta_{t_i}\tilde{\beta}_{t_i,t_{i+1}} \middle| \mathfrak{F}_n\right] \\
&= \lim_{m\to\infty} \sum_{t_i \in \tilde{D}^m} \beta_{t_i}^n \tilde{\beta}_{t_i,t_{i+1}}^n \qquad \text{(use } \beta \perp \tilde{\beta} \text{ and part (a))} \\
&= \int_0^t \beta^n d\tilde{\beta}^n \qquad \text{(by def. of Riemann-Stieltjes integral)}
\end{aligned}$$

---

[2]You may be familiar with $\mathbb{E}[B_t|B_T] = (t/T)\,B_T$.

[3]There are more elementary arguments for $B_t^n \to B$ but this one extends to the area-level.

After exchanging the roles of β and $\tilde{\beta}$ and subtraction we find $\mathbb{E}[A_t|\mathfrak{F}_n] = A_t^n$ as claimed. The final reasoning is as above: A_t is $(\vee_n \mathfrak{F}_n)$-measurable, this follows from (1.1), and by martingale convergence

$$A_t^n = \mathbb{E}[A_t|\mathfrak{F}_n] \to A_t \text{ a.s and in } L^2.$$

□

THEOREM 82. *For $\alpha \in [0, 1/2)$ there exists a random variable M_1 with Gauss tails such that*

$$\sup_n \|\mathbf{B}^n\|_{\alpha\text{-}Hölder}, \|\mathbf{B}\|_{\alpha\text{-}Hölder} \leq M_1 < \infty \text{ a.s.}$$

PROOF. We keep the notation of the last proof where we established

$$B_t^n = \mathbb{E}[B_t|\mathfrak{F}_n] \text{ and } A_t^n = \mathbb{E}[A_t|\mathfrak{F}_n].$$

Simple algebra (attention $A_{s,t} \neq A_t - A_s$!) yields

$$B_{s,t}^n = \mathbb{E}[B_{s,t}|\mathfrak{F}_n] \text{ and } A_{s,t}^n = \mathbb{E}[A_{s,t}|\mathfrak{F}_n]. \tag{2.1}$$

We focus on one component in the matrix $A_{s,t}$, say $A_{s,t}^{i,j}$ with $i \neq j$. Clearly,

$$\left|A_{s,t}^{i,j}\right| \leq |A_{s,t}| \leq \left(|B_{s,t}| \vee |A_{s,t}|^{1/2}\right)^2 \sim \|\mathbf{B}_{s,t}\|^2$$

where $\sim$ is a reminder of the Lipschitz equivalence of homogenous norms on $G^2(\mathbb{R}^d)$. From Theorem 78, $\|\mathbf{B}_{s,t}\|^2 \leq C_1 (t-s)^{2\alpha}$ for a non-negative r.v. $C_1 \in L^p$ for all $p < \infty$, so that

$$-C_1 (t-s)^{2\alpha} \leq A_{s,t}^{i,j} \leq C_1 (t-s)^{2\alpha}.$$

Conditioning w.r.t. $\mathfrak{F}_n$ yields

$$-C_2 (t-s)^{2\alpha} \leq \mathbb{E}[A_{s,t}^{i,j}|\mathfrak{F}_n] \leq C_2 (t-s)^{2\alpha}$$

where $C_2 = \sup\{\mathbb{E}[C_1|\mathfrak{F}_n] : n \geq 1\} \in L^p$ for all $p < \infty$, using Doob's maximal inequality. From (2.1) the last line rewrites

$$-C_2 (t-s)^{2\alpha} \leq A_{s,t}^{n;i,j} \leq C_2 (t-s)^{2\alpha}$$

where C_2 is independent of n and it is a small step to[4]

$$\sup_n \left|A_{s,t}^n\right| \leq C_2 (t-s)^{2\alpha}.$$

The same reasoning, slightly easier in fact, shows that

$$\sup_n \left|B_{s,t}^n\right| \leq C_2 (t-s)^{\alpha}.$$

Putting everything together

$$\left\|\mathbf{B}_{s,t}^n\right\| \sim \left|B_{s,t}^n\right| \vee \left|A_{s,t}^n\right|^{1/2} \leq C_2 (t-s)^{\alpha}$$

which is precisely the required estimate on $\|\mathbf{B}^n\|_{\alpha\text{-Hölder}}$, uniform over $n \geq 1$. □

[4]Depending on the $so(d)$-norm, C_2 may change in some insignificant way.

REMARK 83. *Let $\alpha \in (1/3, 1/2)$. From the last section, EBM $\mathbf{B}$ is a geometric $(1/\alpha)$-rough path with Hölder control*

$$\mathbf{B} \in C^{a\text{-Hölder}}\left([0,T], G^2\left(\mathbb{R}^d\right)\right) \text{ with } [1/\alpha] = 2.$$

We now found a sequence of Lipschitz paths, namely B^n, so that

$$S_2(B^n) \to \mathbf{B} \text{ a.s.}$$

pointwise with uniform α-Hölder bounds. On the other hand, for a.e. $\mathbf{B} = \mathbf{B}(\omega)$ Theorem 47 gives Lipschitz paths $\{x^n\}$ so that $S_2(x^n) \to \mathbf{B}$. However, these x^n are based on knowledge of path and Lévy's area, both contained in $\mathbf{B}$, whereas B^n was constructed only knowning the path B. What is the difference? For instance, we had to rule out another null-set in the a.s.-convergence of $S_2(B^n) \to \mathbf{B}$ (where?). More interestingly, think of a lift of B to $\tilde{\mathbf{B}}$ given by

$$\tilde{\mathbf{B}} = \exp\left(B + \tilde{A}\right) \neq \mathbf{B} = \exp(B + A).$$

For simplicity, let $d = 2$ and identify $so(d)$ with scalars. Then an example is given by $\tilde{A}_t = A_t + t$. Note $\tilde{\mathbf{B}} \in C^{a\text{-Hölder}}\left([0,T], G^2\left(\mathbb{R}^2\right)\right)$, $\alpha \in (1/3, 1/2)$, since

$$\left\|\tilde{\mathbf{B}}_{s,t}\right\| \sim |B_{s,t}| + \left|\tilde{A}_{s,t}\right|^{1/2} = |B_{s,t}| + |A_{s,t} + (t-s)|^{1/2} \leq |B_{s,t}| + |A_{s,t}|^{1/2} + |t-s|^{1/2}.$$

Theorem 47 applies equally well to $\tilde{\mathbf{B}}$, but there will systematic "spinning" in the approximations, say $\{\tilde{x}^n\}$, to make sure that the right area $\tilde{A}_t = A_t + t$ is achieved. The non-uniqueness for the lift of a (Brownian) path B is not surprising and has, in fact, nothing to with probability, see Theorem 52. In any case, what we have seen is that piecewise-linear approximations lead to an area which is consistent with the one from Itô- resp. Stratonovich- integration theory. We also note that to simulate $\mathbf{B}$ we don't have to sample to Lévy-area but can trust the convergence of the piecewise-linear approximations. Finally, $S_2(B^n) \to \mathbf{B}$ allows us to identify $\pi(0, y_0; \mathbf{B}(\omega))$ as a classical Stratonovich solution of an SDE. More on this in the following section.

3. RDEs driven by EBM

We saw that a.e. realization $\mathbf{B} = \mathbf{B}(\omega)$ is a weak geometric $1/\alpha$-rough path for $1/\alpha \in (2,3)$. Given vector fields $V_1, ..., V_d \in Lip^3(\mathbb{R}^e)$ we then have *pathwise* (that is for fixed ω) existence and uniqueness of the RDE solution, $y(\omega) = \pi(0, y_0; \mathbf{B}(\omega))$ on $[0,T]$. We quote a classical result from Stratonovich SDE theory.

PROPOSITION 84 (Wong-Zakai). *Assume Lip^3-vector fields $V = (V_1, ..., V_d)$ Then the Stratonovich SDE*

$$dY = V(Y) \circ dB$$

started at y_0 has a unique (up to indistinguishability) continuous solution Y on $[0,T]$. Moreover, consider a sequence of (not necessarily nested) dissections $\{D^n\}$ with mesh $|D^n| \to 0$. Then the following convergence holds in probability and uniformly on $[0,T]$,

$$\pi\left(0, y_0; B^{D_n}\right) \to Y.$$

(As usual, B^{D_n} stands for the piecewise linear approximations based on the dissection D^n and $\pi\left(0, y_0; B^{D_n}\right)$ is the unique solution of the control ODE $\dot{y} = V(y)\dot{B}^{D_n}$.)

REMARK 85. *We did not develop enough Stratonovich theory to give a full proof but you should consider the special case* $d = 2$ *with* V_1, V_2 *from Exercise 3. The convergence à la Wong-Zakai is, in this case, nothing else than* $\mathbf{B}_t^n \to \mathbf{B}_t$ *in probability and uniformly in* $[0,T]$*. In the last section, Theorem 81, we showed that, by using nested dissections, we can achieve a.s.-convergence and uniform convergence follows, of course, from our uniform Hölder bounds.*

COROLLARY 86. *The pathlevel RDE solution* $y(\omega) = \pi(0, y_0; \mathbf{B}(\omega))$ *is a version of the unique strong solution to the above Stratonovich SDE.*

PROOF. We use nested dissections $\{D^n\}$ with $|D^n| \to 0$. From Theorem 81 and the universal limit theorem we see that

$$\pi\left(0, y_0; B^{D_n}\right) \to \pi(0, y_0; \mathbf{B}) \text{ a.s.}$$

and uniformly on $[0,T]$. On the other hand, Wong-Zakai tells us that

$$\pi\left(0, y_0; B^{D_n}\right) \to Y \text{ in probability}$$

and uniformly on $[0,T]$. It then follows that $Y = \pi(0, y_0; \mathbf{B})$ a.s. □

REMARK 87. $\pi(0, y_0; \mathbf{B}(\omega))$ *is a "nice" version in the sense that the map* $y_0 \mapsto \pi(0, y_0; \mathbf{B}(\omega))$ *is immediatley well-defined for a.e.* ω*. (This is not the case in the usual theory of SDEs!). We shall come back to this later.*

4. Stroock-Varadhan's support theorem for SDEs

4.1. Support of Brownian motion. [5]Almost surely, d-dimensional Brownian motion $B \in C^{\alpha}\left([0,T],\mathbb{R}^d\right)$ for $\alpha \in [0,1/2)$. By interpolation, we also have that a.s. B takes values in the Polish space $C^{0,\alpha}\left([0,T],\mathbb{R}^d\right)$, in fact in the closed subspace of paths started at 0, write $C^{0,\alpha}\left([0,T],\mathbb{R}^d\right)$. Then B can then be viewed as $C^{0,\alpha}$-valued random variable and its law of B is a Borel probability measure on $C^{0,\alpha}$ (check!).

Definition 88. *Let μ be a Borel probability measure on some Polish space (E,δ). The (topological) support (w.r.t. the topology induced by δ) of μ is the smallest closed set of full measure.*

Theorem 89 (Cameron-Martin). *Let B be $\mathbb{R}^d$-Brownian motion on $[0,T]$. Define $T_h(B) = B + h$ where h is a Cameron-Martin path (=the indefinite integral $\int_0^{\cdot}$ of some function in $L^2\left([0,T];\mathbb{R}^d\right)$). Then the law of $T_h(B)$ is equivalent to the law of B.*

Proof. *Any* book on Brownian motion. □

Remark 90. *Let us record some simple properties of T_h. It is a continuous map of $C^{0,\alpha}\left([0,T],\mathbb{R}^d\right)$ into itself and bijective with inverse T_{-h}. In particular, the image of any open sets under T_h is again open.*

Corollary 91. *Let h be a Cameron-Martin path and $x \in C^{0,\alpha}\left([0,T],\mathbb{R}^d\right)$. Then $x \in$ supp(law of B) implies $T_h(x) \in$ supp(law of B).*

Proof. Write $\mathcal{N}(x)$ for all open neighbourhoods of x. To show that $T_h(x)$ is in the support, it suffices to show that

$$\forall V \in \mathcal{N}(T_h(x)) : \mathbb{P}(B \in V) > 0.$$

Fix $V \in \mathcal{N}(T_h(x))$. By continuity, there exists $U \in \mathcal{N}(x)$ so that $T_h(U) \subset V$. From the above remark, $T_h(U) \in \mathcal{N}(T_h(x))$. Thus

$$\begin{aligned} \mathbb{P}(B \in V) &\geq \mathbb{P}(B \in T_h(U)) \\ &= \mathbb{P}(T_{-h}B \in U) \end{aligned}$$

and from Cameron-Martin the last expression if positive iff $\mathbb{P}(B \in U)$ is positive. But this is true since $U \in \mathcal{N}(x)$ and x is in the support. □

Corollary 92. *Let $\alpha \in (0,1/2)$. The topological support of the law of Brownian motion on $[0,T]$ in α-Hölder topology is precisely $C_0^{0,\alpha}\left([0,T];\mathbb{R}^d\right)$.*

Proof. Almost surely, $B(\omega) \in C_0^{0,\alpha}\left([0,T];\mathbb{R}^d\right)$ which is closed in α-Hölder topology. Therefore,

$$\text{supp}\,(\text{law of } B) \subset C_0^{0,\alpha}\left([0,T];\mathbb{R}^d\right).$$

Vice-versa, the support contains (trivially!) one point, say

$$x \in C_0^{0,\alpha}\left([0,T];\mathbb{R}^d\right).$$

From the properties of the space $C^{0,\alpha}$, see Theorem 40, there are Lipschitz paths $\{x^n\}$ with $x^n(0) = 0$ so that

$$x - x^n = T_{-x^n}(x) \to 0 \text{ in } \alpha\text{-Hölder topology.}$$

[5]Corrections and comments to P.K.Friz@statslab.cam.ac.uk

Since every Lipschitz path is a Cameron Martin path, $T_{-x^n}(x) \in$ support for all n. By definition, the support is closed (in α-Hölder topology) and therefore $0 \in$ support. But then any translate $T_h(0) = h \in$ support, for all Lipschitz (in fact, Cameron Martin) path h. Since Lipschitz paths are dense in $C^{0,\alpha}$, taking the closure yields

$$C^{0,\alpha}\left([0,T];\mathbb{R}^d\right) \subset \text{supp}\,(\text{law of } B).$$

□

4.2. Intermezzo: Translation of rough paths. We just used the translation map $T_h(x) = x + h$ for $\mathbb{R}^d$-valued paths x and h. Assume both x and h are Lipschitz, started at 0, and consider the step-2 lift: $\mathbf{x} \equiv S_2(x)$, and $T_h(\mathbf{x}) \equiv S_2(T_h(x))$. From definition of S_2,

$$T_h(\mathbf{x}) = 1 + \left(\mathbf{x}^1 + h\right) + \left(\mathbf{x}^2 + \int_0^{\cdot} x \otimes dh + \int_0^{\cdot} h \otimes dx + S_2(h)\right).$$

PROPOSITION 93. *The map $\mathbf{x} \mapsto T_h(\mathbf{x})$ can be extended to a continuous map of $C^{0,\alpha}\left([0,1], G^2\left(\mathbb{R}^d\right)\right)$ into itself and T_h also denotes this extension. It is bijective with inverse T_{-h}. In particular, the image of any open sets under T_h is again open.*

PROOF. Exercise. □

EXERCISE 94. *Prove the last proposition. Hint: $d_{\alpha\text{-Hölder}}(\mathbf{x}^n, \mathbf{x}) \to 0$ is equivalent to usual α-Hölder convergence of the path-level, $\mathbf{x}^{n;1} \to \mathbf{x}^1$, and convergence of the second level in the sense*

$$\sup_{0 \le s < t \le T} \frac{\left|\mathbf{x}^{n;2}_{s,t} - \mathbf{x}^2_{s,t}\right|^{1/2}}{|t-s|^{\alpha}} \to 0.$$

EXERCISE 95. *Assume $T_{-x^n}(\mathbf{x}) \to 0$. Show that this is not equivalent to $S_2(x^n) \to \mathbf{x}$ and neither implies the other. (The convergence is always w.r.t. $d_{\alpha\text{-Hölder}}$.)*

4.3. Support of Enhanced Brownian motion. Almost surely, EBM $\mathbf{B}(\omega) \in C^{\alpha}\left([0,T], G^2\left(\mathbb{R}^d\right)\right)$ for $\alpha \in [0, 1/2)$. By interpolation, we also have that a.s. $\mathbf{B}$ takes values in the Polish space $C^{0,\alpha}\left([0,T], G^2\left(\mathbb{R}^d\right)\right)$ and in fact in the close subspace of paths starting at the unit element of G^2. (Write $C^{0,\alpha}_0\left([0,T], G^2\left(\mathbb{R}^d\right)\right)$ for this space.) Similar to the case of Brownian motion, we view $\mathbf{B}$ as $C^{0,\alpha}\left([0,T], G^2\left(\mathbb{R}^d\right)\right)$-valued random variable; the law of $\mathbf{B}$ is a Borel probability measure on $C^{0,\alpha}\left([0,T], G^2\left(\mathbb{R}^d\right)\right)$ (check!).

THEOREM 96 (Cameron-Martin for BM on G^2). *Let $\mathbf{B}$ be $G^2\left(\mathbb{R}^d\right)$-valued enhanced Brownian motion on $[0,T]$. Let h be a Lipschitz path (& hence a Cameron Martin path). Then the law of $T_h(\mathbf{B})$ is equivalent to the law of $\mathbf{B}$.*

PROOF. We can assume that the underlying probability space is Wiener space $C\left([0,T], \mathbb{R}^d\right)$ equipped with Wiener-measure $\mathbb{W}$. In particular, Brownian motion is the coordinate map $B(t,\omega) = \omega(t)$. The law of B is $\mathbb{W}$ and we write $\mathbb{W}^h$ for the law of $B + h$. From Cameron-Martin we know that the measures are equivalent, $\mathbb{W} \sim \mathbb{W}^h$. Now, $\mathbf{B}$ is a measurable map from $C\left([0,T], \mathbb{R}^d\right) \to C^{0,\alpha}\left([0,T], G^2\left(\mathbb{R}^d\right)\right)$. It is easy to see that (using Stratonovich calculus or, more elementary, the L^2-convergent Riemann-Stieltjes sum for the area) that

$$\mathbf{B}(\cdot + h) = T_h\mathbf{B} \text{ a.s.}$$

and hence the law of $T_h \mathbf{B}$ is $\mathbf{B}_* \mathbb{W}^h$ = short notation for the image measure of $\mathbb{W}^h$ under $\mathbf{B}$. On the other hand, the law of $\mathbf{B}$ is $\mathbf{B}_* \mathbb{W}$. Equivalence of measure implies equivalence of image measures, and we find $\mathbf{B}_* \mathbb{W} \sim \mathbf{B}_* \mathbb{W}^h$. The proof is now easily finished. □

REMARK 97. *$T_h(\mathbf{B})$ can be defined for any Cameron-Martin path and the above holds true for any Cameron Martin path.*

COROLLARY 98. *Let h be a Lipschitz (and hence Cameron Martin) and $\mathbf{x} \in$ supp(law of $\mathbf{B}$). Then $T_h(\mathbf{x}) \in$ supp(law of $\mathbf{B}$).*

PROOF. With the properties of $\mathbf{x} \mapsto T_h(\mathbf{x})$ we established in Proposition 93 and the Cameron-Martin theorem for BM on G^2, the proof given earlier for Brownian motion (Corollary 91) adapts with no changes. □

LEMMA 99. *Let $\alpha \in (0, 1/2)$. There exists $\mathbf{x} \in$ supp(law of $\mathbf{B}$) and $\{x^n\}$ Lipschitz with $x^n(0) = 0$ so that*

$$\|T_{-x^n} \mathbf{x}\|_{\alpha\text{-}H\ddot{o}lder} \to 0.$$

PROOF. If B^n denotes the piecewise linear approximation based on a nested sequence of dissections, we saw that

$$S(B^n) \to \mathbf{B} \text{ a.s. (pointwise)}$$

with uniform α-Hölder bounds. In fact, the essential observation was that

$$\mathbb{E}\left[\int_0^t \beta d\tilde{\beta} | \mathfrak{F}_n\right] = \int_0^t \beta^n d\tilde{\beta}^n$$

where $\mathfrak{F}_n = \sigma\left(\beta_t, \tilde{\beta}_t : t \in D^n\right)$. The arguments given in section 2 also give

$$\begin{aligned}
\mathbb{E}\left[\int_0^t \beta d\tilde{\beta} \,\middle|\, \sigma(\beta_t : t \in D^n) \vee \sigma(\tilde{\beta}_t : t \in [0,T])\right] &= \int_0^t \beta^n d\tilde{\beta} \\
\mathbb{E}\left[\int_0^t \beta d\tilde{\beta} \,\middle|\, \sigma(\beta_t : t \in [0,T]) \vee \sigma(\tilde{\beta}_t : t \in D^n)\right] &= \int_0^t \beta d\tilde{\beta}^n.
\end{aligned}$$

(both integrals on the r.h.s. make sense as Riemann-Stieltjes integral) and also

$$T_{-B^n} \mathbf{B} \to 0 \text{ a.s. (pointwise)}$$

with uniform α-Hölder bounds. The usual interpolation finishes the proof. Indeed, we could have started with $\tilde{\alpha} \in (\alpha, 1/2)$, get uniform $\tilde{\alpha}$-bounds and use interpolation to obtain $T_{-B^n} \mathbf{B} \to 0$ in α-Hölder topology. This statements holds a.s. and we can take any $\mathbf{x} = \mathbf{B}(\omega)$ for ω in a set of full measure. □

REMARK 100. *As discussed in the last section,$\|T_{-B^n} \mathbf{B}\|_{\alpha\text{-}H\ddot{o}lder} \to 0$ does not follow from $d_{\alpha\text{-}H\ddot{o}lder}(\mathbf{B}, S(B_n)) \to 0$.*

COROLLARY 101. *Let $\alpha \in (0, 1/2)$. The topological support of the law of $G^2(\mathbb{R}^d)$-valued Enhanced Brownian motion on $[0,T]$ w.r.t. $d_{\alpha\text{-}H\ddot{o}lder}$ is precisely $C_0^{0,\alpha}([0,T]; G^2(\mathbb{R}^d))$.*

PROOF. Same argument as for d-dimensional Brownian motion! □

4.4. Support for SDE solutions.

THEOREM 102 (Stroock-Varadhan). *Assume* Lip^3*-vector fields* $V = (V_1, ..., V_d)$. *Consider the unique (up to indistinguishability) Stratonovich SDE solution on* $[0,T]$ *to*

$$dY = V_i(Y) \circ dB^i$$

started at y_0. *Let* $\alpha \in (1/3, 1/2)$. *A.e. path is* α*-Hölder continuous and the topological support of the law of* Y *w.r.t.* α*-Hölder topology is precisely the* α*-Hölder closure of all control ODE solutions,*

$$\mathcal{S} \equiv \{\pi(0, y_0; h) : h \text{ Lipschitz on } [0,T]\}$$

where $\pi(0, y_0; h)$ *stands for the solution of* $\dot{y} = V_i(y)\dot{h}^i$ *started at* y_0 *as usual.*

PROOF. Since $\mathbf{B}(\omega) \in C^{0,\alpha}$ we can find Lipschitz paths $\{B^n(\omega)\}$ (for instance, piecewise linear or geodesic approximations) so that $d_{\alpha\text{-Hölder}}(S_2(B^n), \mathbf{B}) \to 0$ a.s.. From the ULT, $\pi(0, y_0; \mathbf{B}(\omega))$ is the limit of $\pi(0, y_0; B^n(\omega)) \in \mathcal{S}$ w.r.t. α-Hölder topology and it follows that

$$\text{supp (law of } \pi(0, y_0; \mathbf{B})) \subset \overline{\mathcal{S}}$$

where the closure is taken in α-Hölder topology. For the other inclusion we need that for every Lipschitz h and every $\epsilon > 0$, the event $A = \{|\pi(0, y_0; \mathbf{B}) - \pi(0, y_0; h)|_{\alpha\text{-Hölder}} < \epsilon\}$ has positive probability. Note that $\pi(0, y_0; h) \equiv \pi(0, y_0; S_2(h))$. From the ULT, the map $\mathbf{x} \mapsto \pi(0, y_0; \mathbf{x})$ is continuous in the respective Hölder topologies. In particular, continuity at the point $\mathbf{x} = S_2(h) \equiv h$ implies the existence of $\delta = \delta(h, \epsilon)$ such that

$$d_{\alpha\text{-Hölder}}(\mathbf{x}, S_2(h)) < \delta \implies |\pi(0, y_0; \mathbf{x}) - \pi(0, y_0; h)|_{\alpha\text{-Hölder}} < \epsilon.$$

and, combined with the support description of EBM, this finishes the proof,

$$\begin{aligned} \mathbb{P}(\text{event } A) &\geq \mathbb{P}(d_{\alpha\text{-Hölder}}(\mathbf{B}, S_2(h)) < \delta) \\ &> 0 \text{ (from Corollary 101).} \end{aligned}$$

□

REMARK 103. *The result is easily adatped to an additional drift term* $V_0(Y)\,dt$.

5. Freidlin-Wentzell large deviations for SDEs

5.1. Recalls on Large Deviations. [6]Let $\mathcal{X}$ be a topological space. A *rate function* is a lower semicontinuous mapping $I : \mathcal{X} \to [0, \infty]$, i.e. a mapping so that all level sets $\{x \in \mathcal{X} : I(x) \leq \Lambda\}$ are closed. A *good rate function* is a rate function for which all level sets are compacts subsetc of $\mathcal{X}$. The set $\mathcal{D}_I := \{x \in E : I(x) < \infty\}$ is called *domain of* I. Given $A \subset E$ we also set

$$I(A) = \inf_{x \in A} I(x).$$

Unless otherwise stated, we assume that probability measures on $\mathcal{X}$ are defined on the Borel sets, i.e. the smallest σ-algebra generated by the open sets in $\mathcal{X}$.

[6]E.g. [DZ]: Dembo-Zeitouni, Large Deviation Techniques and Applications, Springer

DEFINITION 104. *A family* $\{\mu_\varepsilon : \varepsilon > 0\}$ *of probability measures on* $\mathcal{X}$ *satisfies the large deviation principle (LDP) with good rate function* I *if, for every Borel set* A,

$$-I(A^\circ) \leq \liminf_{\varepsilon\to 0} \varepsilon^2 \log \mu_\varepsilon(A) \leq \limsup_{\varepsilon\to 0} \varepsilon^2 \log \mu_\varepsilon(A) \leq -I(\bar{A}).$$

REMARK 105. *By reparametrizing the family of probability measures one can (equivalently) consider* $\varepsilon \log \mu_\varepsilon(A)$.

A large deviation principle behaves nicely under a continuous map. We have

THEOREM 106 (Contraction Principle). *Let* $\mathcal{X}$ *such* $\mathcal{Y}$ *be Hausdorff topological spaces. Suppose* $f : \mathcal{X} \to \mathcal{Y}$ *is a continuous map. If* $\{\mu_\varepsilon\}$ *satisfies a LDP on* $\mathcal{X}$ *with good rate function* I*, then the image measures* $\{f_*\mu_\varepsilon\}$*, where* $f_*\mu_\varepsilon \equiv \mu_\varepsilon \circ f^{-1}$*, satisfy a LDP on* $\mathcal{Y}$ *with good rate function*

$$J(y) = \inf\{I(x) : x \in \mathcal{X} \text{ and } f(x) = y\}.$$

DEFINITION 107. *A family* $\{\mu_\varepsilon : \varepsilon > 0\}$ *of probability measures on a topological space* $\mathcal{X}$ *is exponentially tight if for every* $M < \infty$*, there exists a compact*[7] *set* K_M *such that*

$$\limsup_{\varepsilon\to 0} \varepsilon^2 \log \mu_\varepsilon(K_M^c) < -M.$$

THEOREM 108 (Inverse Contraction Principle). *Let* $\mathcal{X}$ *such* $\mathcal{Y}$ *be Hausdorff topological spaces. Suppose* $g : \mathcal{Y} \to \mathcal{X}$ *is a continuous injection and that* $\{\nu_\varepsilon\}$ *is an exponentially tight family of probability measures on* $\mathcal{Y}$*. If* $\{g_*\nu_\varepsilon\}$ *satisfies the LDP in* $\mathcal{X}$ *with the rate function* $I : \mathcal{X} \to [0, \infty]$*, then* $\{\nu_\varepsilon\}$ *satisfies the LDP in* $\mathcal{Y}$ *with the good rate function* $I' \equiv I \circ g$.

PROOF. Thm 4.2.2. in [DZ], combined with Proposition 4.1.5 in [DZ] and the remark that $\mathcal{D}_I \subset g(\mathcal{Y})$. □

EXAMPLE 109. *A good example to have in mind is the continuous embedding* $\iota : C^{\alpha\text{-}H\ddot{o}l}([0,1]) \hookrightarrow C([0,1])$*. Scaled Brownian motion* εB *satisfies a LDP on* $C([0,1])$ *with uniform topology ("Schilder's theorem"). For any* $\alpha < 1/2$*, the laws of* εB *are exponenitally tight in* $C^{\alpha\text{-}H\ddot{o}l}([0,1])$*; an easy consequence of Fernique and compact embeddings of Hölder spaces. It follows that* εB *satisfies a LDP on* $C^{\alpha\text{-}H\ddot{o}l}([0,1])$ *with respect to* α*-Hölder topology.*

THEOREM 110 (Extended Contraction Principle). *Let* $\{\mu_\varepsilon\}$ *be a family of probability measures that satisfies the LDP with good rate function* I *on a Hausdorff topological space* $\mathcal{X}$*. For* $m = 1, 2, \ldots,$ *let* $f^m : \mathcal{X} \to \mathcal{Y}$ *be continuous maps, with* $(\mathcal{Y}, d)$ *a separable metric space. Assume there exists a measurable map* $f : \mathcal{X} \to \mathcal{Y}$ *such that for every* $\Lambda < \infty$

$$\lim_{m\to\infty} \sup_{\{x : I(x) \leq \Lambda\}} d(f^m(x), f(x)) = 0.$$

Assume that $\{f_*^m \mu_\varepsilon\}$ *are exponentially good approximations of* $\{f_*\mu_\varepsilon\}$ *and in the sense that*[8]

$$\lim_{m\to\infty} \limsup_{\varepsilon\to 0} \varepsilon^2 \log \mu_\varepsilon(\{x : d(f^m(x), f(x)) > \delta\}) = -\infty.$$

Then $\{f_*\mu_\varepsilon\}$ *satisfies the LDP in* $\mathcal{Y}$ *with the good rate function* $I' \equiv \inf\{I(x) : y = f(x)\}$.

[7]Since $\overline{K_M}^c \subset K_M^c$ it is enough to require that K_M is precompact.

[8]The separability on $\mathcal{Y}$ guarantees measurability of $\{x : d(f^m(x), f(x)) > \delta\}$.

DZ, Thm 4.2.23. □

5.2. Schilder's theorem. Let B denote d-dimensional standard Brownian motion on $[0,1]$.If $P_\varepsilon \equiv \mathbb{P}_*(\varepsilon B)$ denotes the law of εB, viewed as Borel measure on $C_0([0,1],\mathbb{R}^d)$, the next theorem can be summarized in saying that $(P_\varepsilon)_{\varepsilon>0}$ satisfies a large deviation principle (short: LDP) on the space $C_0([0,1],\mathbb{R}^d)$ with rate function I. (When no confusion arises, we shall simply say that $(\varepsilon B)_{\varepsilon>0}$ satisfies a LDP.) We recall that the *Cameron-Martin space* for d-dimensional Brownian is given by

$$\mathcal{H} = \left\{ h \in C_0([0,1],\mathbb{R}^d) \text{ absolutely continuous, } \dot{h} \in L^2[0,1],\mathbb{R}^d \right\}$$

and has the Hilbert structure given by $\langle h,g\rangle_{\mathcal{H}} = \left\langle \dot{h},\dot{g}\right\rangle_{L^2}$.All subsequent large deviation statements will involve the good rate function

$$I(h) = \begin{cases} \frac{1}{2}\langle h,h\rangle_{\mathcal{H}} & \text{if } h \in \mathcal{H} \\ +\infty & \text{otherwise} \end{cases}.$$

Theorem 111 (Schilder). *For any measurable $A \subset C_0([0,1],\mathbb{R}^d)$ we have*

$$-I(A^\circ) \le \liminf_{\varepsilon\to 0} \varepsilon^2 \log \mathbb{P}[\varepsilon B \in A] \le \limsup_{\varepsilon\to 0} \varepsilon^2 \log \mathbb{P}[\varepsilon B \in A] \le -I(\bar{A}). \tag{5.1}$$

Here, $A^\circ, \bar{A}$ denote the interiori resp. closure of A with respect to uniform topology.

Proof. Any book on large deviations. □

5.3. Schilder's Theorem for EBM and LDP for Lévy's area. As in the last section, the inverse contraction principle implies that a LDP for Enhanced Brownian motion in (homogenous) α-Hölder topology is a simple of the corresponding LDP in uniform topology. Our main task is therefore to establish a LDP for EBM in uniform topology, or - in other words - a LDP for Brownian and Lévy's area. As the reader may realize, the difficulty lies in the discontinuity of Lévy's area

$$\left(\beta,\tilde{\beta}\right) \mapsto A := \frac{1}{2}\left(\int_0^\cdot \beta d\tilde{\beta} - \int_0^\cdot \tilde{\beta} d\beta\right),$$

as a function[9] of $\left(\beta,\tilde{\beta}\right) \in C([0,1],\mathbb{R}^2)$ equipped with uniform (or α-Hölder) topology, and this prevents us from obtaining large deviations for Lévy's area from Schilder's theorem via the contraction prinicple. What we will establish instead is that A can be "exponentially well" approximated by continuous maps $\left(\beta,\tilde{\beta}\right) \mapsto A^n = A^n\left(\beta,\tilde{\beta}\right)$; for each n, the contraction principle then gives a LDP for A^n and one hopes to send $n\to\infty$ in the corresponding lower and upper bounds. A general result of large deviation theory, the *extended contraction principle* is a formalization of this idea. It requires us to check two conditions (cf. thm 110) and this is what we now do.

Lemma 112. *The $so(d)$-valued approximations to Lévy's area process given by*

$$A_t^m := \frac{1}{2}\int_0^t B_{[ms]/m} \otimes dB_s - \int_0^t dB_s \otimes B_{[ms]/m}$$

[9]Equipping $C([0,1],\mathbb{R}^2)$ with Wiener-measure, stochastic integration defines A as a measurable map.

gives rise to exponentially good approximations to $\{\delta_\varepsilon \mathbf{B}\}$ *in the sense that*

$$\lim_{m\to\infty} \overline{\lim}_{\varepsilon\to 0} \varepsilon^2 \log \mathbb{P}\left[d_\infty\left(\exp\left(\varepsilon B + \varepsilon^2 A^m\right), \delta_\varepsilon \mathbf{B}\right) \geq \delta\right] = -\infty.$$

PROOF. We first note that

$$\begin{aligned} d_\infty\left(e^{\varepsilon B + \varepsilon^2 A^m}, \delta_\varepsilon \mathbf{B}\right) &= \varepsilon \sup_{t\in[0,1]} d\left(e^{B_t + A_t^m}, e^{B_t + A_t}\right) \\ &\sim \varepsilon \sup_t \left|A_t^m - A_t\right|^{1/2} \end{aligned}$$

With view on a generic building block of the area, let $\left(\beta, \tilde{\beta}\right)$ be a two-dimensional Brownian motion. Replacing δ by a constant times δ^2, if needed, it is is enough to show that

$$\lim_{m\to\infty} \overline{\lim}_{\varepsilon\to 0} \varepsilon^2 \log \mathbb{P}\left[\sup_{t\in[0,1]} \varepsilon^2 \left|\int_0^t \beta_s d\tilde{\beta}_s - \int_0^t \beta_{[ms]/m} d\tilde{\beta}_s\right| \geq \delta\right] \to -\infty.$$

We introduce the continuous martingale

$$M_t := \int_0^t \left(\beta_s - \beta_{[ms]/m}\right) d\tilde{\beta}_s, \quad \langle M\rangle_t = \int_0^t \left(\beta_s - \beta_{[ms]/m}\right)^2 ds.$$

Pick $\alpha \in (0, 1/2)$. By the exponential martingale inequality

$$\mathbb{P}\left[\sup_{t\in[0,1]} M_t \geq \delta\varepsilon^{-2} \text{ and } \langle M\rangle_1 \leq \varepsilon^{-2} m^{-\alpha}\right] \leq \exp\left(-\frac{1}{2}\frac{\left(\delta\varepsilon^{-2}\right)^2}{\varepsilon^{-2} m^{-\alpha}}\right) = \exp\left(-\delta^2\varepsilon^{-2} m^\alpha/2\right).$$

On the other hand,

$$\begin{aligned} \mathbb{P}\left[\langle M\rangle_1 > \varepsilon^{-2} m^{-\alpha}\right] &= \mathbb{P}\left[\int_0^1 \left(\beta_s - \beta_{[ms]/m}\right)^2 ds > \varepsilon^{-2} m^{-\alpha}\right] \\ &\leq \mathbb{P}\left[|\beta|^2_{\alpha\text{-Höl};[0,1]} \left(\frac{1}{m}\right)^{2\alpha} > \varepsilon^{-2} m^{-\alpha}\right] \end{aligned}$$

so that

$$\mathbb{P}\left[\sup_{t\in[0,1]} M_t \geq \delta\varepsilon^{-2}\right] \leq \exp\left(-\delta^2\varepsilon^{-2} m^\alpha/2\right) + \mathbb{P}\left[|\beta|^2_{\alpha\text{-Höl};[0,1]} > \varepsilon^{-2} m^\alpha\right]$$

We conclude with the fact that the α-Hölder norm of β has a Gauss tail. □

LEMMA 113. *Define* $A(h)_t^m = \frac{1}{2}\int_0^t h_{[ms]/m} \otimes dh_s - \int_0^t dh_s \otimes h_{[ms]/m}$ *for any* $h \in \mathcal{H}$. *Then, for all* $\Lambda > 0$,

(5.2) $$\lim_{m\to\infty} \sup_{\{h\in\mathcal{H}: I(h)\leq\Lambda\}} d_\infty\left(\exp\left(h + A(h)^m\right), S_2(h)\right) = 0.$$

PROOF. We first note that

$$d_\infty\left(e^{h + A(h)^m}, S_2(h)\right) = \sup_{t\in[0,1]} d\left(e^{h_t + A(h)_t^m}, e^{h + A(h)_t}\right) \sim \sup_t \left|A(h)_t^m - A(h)_t\right|^{1/2}$$

where

$$A(h)_t = \frac{1}{2}\int_0^t h \otimes dh - \int_0^t dh \otimes h.$$

To see convergence to zero, uniform on level sets of $h \mapsto I(h) = |h|_{\mathcal{H}}^2/2$, it suffices to note that for any $t \in [0,1]$

$$\begin{aligned}\left|\int_0^t h_{[ms]/m} \otimes dh_s - \int_0^t h_s \otimes dh_s\right| &= \left|\int_0^t \left|h_{[ms]/m} - h_s\right| \otimes dh_s\right| \\ &\leq |h|_{1/2\text{-Höl};[0,1]} \left(\frac{1}{m}\right)^{1/2} |h|_{1\text{-var};[0,1]} \\ &\leq |h|_{\mathcal{H}}^2 \left(\frac{1}{m}\right)^{1/2}.\end{aligned}$$

□

Theorem 114 (Schilder, EBM). *(i) The family $(\delta_\varepsilon \mathbf{B} : \varepsilon > 0)$ satisfies a large deviation principle in uniform topology. More precisely, viewing $P_\varepsilon := \mathbb{P}_* (\delta_\varepsilon \mathbf{B})$ as Borel measure on the Polish space $\left(C_0\left([0,1], G^2\left(\mathbb{R}^d\right)\right), d_\infty\right)$, the family $(P_\varepsilon : \varepsilon > 0)$ satisfies a LDP on this space with good rate function*

$$J(\mathbf{y}) = I(\pi_1(\mathbf{y}))$$

where π_1 denotes the projection of $G^2\left(\mathbb{R}^d\right) \to \mathbb{R}^d$.
(ii) For any $\alpha \in [0,1/2)$, the family $(\delta_\varepsilon \mathbf{B} : \varepsilon > 0)$ satisfies a large deviation in homogenous α-Hölder topology. More precisely, viewing $P_\varepsilon := \mathbb{P}_* (\delta_\varepsilon \mathbf{B})$ as Borel measure on the Polish space $(C_0^{0,\alpha\text{-Höl}}\left([0,1], G^2\left(\mathbb{R}^d\right)\right), d_{\alpha\text{-Höl}})$, the family $(P_\varepsilon : \varepsilon > 0)$ satisfies a LDP on this space with good rate function J.

Proof. (i) By the extended contraction principle, the previous two lemmas show that $\delta_\varepsilon \mathbf{B}$ satisfy a LDP with good rate funtion

$$J(h) = \inf\{I(h) : \ h \text{ such that } S_2(h) = y\} = I(\pi_1(h)).$$

(ii) By the inverse contraction principle we only need to show that that $\{\delta_\varepsilon \mathbf{B}\}$ is exponentially tight in α-Hölder topology. But this follows from the compact embedding of

$$C^{\alpha'\text{-Hölder}}\left([0,1], G^2\left(\mathbb{R}^d\right)\right) \hookrightarrow C^{0,\alpha\text{-Hölder}}\left([0,1], G^2\left(\mathbb{R}^d\right)\right)$$

and Gauss tails of $\|\mathbf{B}\|_{\alpha'\text{-Hölder}}$ where $\alpha < \alpha' < 1/2$,

$$\exists c > 0 : \mathbb{P}\left[\|\mathbf{B}\|_{\alpha'\text{-Höl}} > l\right] \leq \exp\left(-cl^2\right).$$

Indeed, defining the following precompact sets in α-Hölder topology,

$$K_M = \left\{x : |x|_{\alpha'\text{-Hölder}} \leq \sqrt{M/c}\right\}.$$

we find the required estimate,

$$\begin{aligned}\varepsilon^2 \log\left[(\delta_\varepsilon \mathbf{B})_* \mathbb{P}\right](K_M^c) &= \varepsilon^2 \log \mathbb{P}\left[\|\delta_\varepsilon \mathbf{B}\|_{\alpha'\text{-Hölder}} > \sqrt{M/c}\right] \\ &= \varepsilon^2 \log \mathbb{P}\left[\|\mathbf{B}\|_{\alpha'\text{-Hölder}} > \sqrt{\frac{M}{c\varepsilon^2}}\right] \leq -M.\end{aligned}$$

□

Remark 115. *(Markovian approach). $\mathbf{B}$ is a hypoelliptic diffusion on $G^2\left(\mathbb{R}^d\right)$ which admits a density with respect to Haar-measure. It is well-known that*

$$\log p_t(x,y) \sim -\frac{1}{2} d(x,y)^2$$

where $x, y \in G^2(\mathbb{R}^d)$. By following Varadhan's arguments of his famous 1967 paper, one finds that

$$\left(\mathbf{B}\left(\varepsilon^2 t\right)\right)$$

satisfies a LDP with rate function given by

$$\tilde{I}(\mathbf{h}) = \frac{1}{2} \lim \sup_{|D| \to 0} \sum_{t_i \in D} \frac{d\left(\mathbf{h}_{t_i}, \mathbf{h}_{t_{i-1}}\right)^2}{t_i - t_{i-1}} \in [0, \infty]$$

By uniqueness of rate functions, $\tilde{I}(\mathbf{h}) = I(\pi_1(\mathbf{h}))$ which of course can be seen be elementary means (exercise 38).

5.4. Large Deviations for SDE solutions.

THEOREM 116 (Freidlin-Wentzell). *Assume Lip^3-vector fields $V = (V_1, ..., V_d)$. Consider the unique (up to indistinguishability) Stratonovich SDE solution on $[0,1]$ to*

$$dY^\varepsilon = V_i(Y)^\varepsilon \circ \varepsilon dB^i$$

started at y_0. Let $\alpha \in (1/3, 1/2)$. Then Y^ε satisfies a LDP (in α-Hölder topology) with rate function given by

$$J(y) = \inf\left\{I(h) : \pi_{(V)}(0, y_0; h) = y\right\}.$$

PROOF. The Stratonovich solution is given by the random RDE solution $\pi_{(V)}(0, y_0; \delta_\varepsilon \mathbf{B})$. It depends continuously on $\delta_\varepsilon \mathbf{B}$ which satisfies a LDP in ε. Conclude with the contraction principle. □

6. Comments and References

These lecture notes are based on the forthcoming textbook [FV]. Limit theorems for stochastic flows appear in [IW, Ma]; with rough path theory this is a trivial consequence of theorem 66 and rough path convergence of piecewise linear approximations to Brownian motion. The Stroock-Varadhan support theorem is discussed in the books [IW, St]; for Freidlin-Wentzell theory see [DZ, DS, Va]. The rough path approach to support and LD was pioneered in the paper [LQZ].

[DS] Deuschel-Stroock: Large Deviations
[DZ] Dembo-Zeitouni; Large Deviation Techniques, Springer
[FV] Friz-Victoir; Multidimensional Stochastic Processes as Rough Paths, CUP (2008)
[IW] Ikeda-Watanabe; Stochastic Differential Equations, North Holland (1989)
[LQZ] Ledoux-Qian-Zhang; Support Theorem and Large Deviations via Rough Paths, SPA (2002)
[Ma] Malliavin; Stochastic Analysis; Springer (1997)
[St] Stroock; Markov Processes from Ito's perspective; PUP (2003)
[Va] Varadhan; Large Deviations; Siam