A flow-based approach to rough differential equations

I. Bailleul[1]

[1]IRMAR, 263 Avenue du General Leclerc, 35042 RENNES, France, ismael.bailleul@univ-rennes1.fr

Contents

CHAPTER 1

Introduction

This course is dedicated to the study of some class of dynamics in a Banach space, index by time $\mathbb{R}_+$. Although there exists many recipes to cook up such dynamics, those generated by differential equations or vector fields on some configuration space are the most important from a historical point of view. Classical mechanics reached for example its top with the description by Hamilton of the evolution of any classical system as the solution of a first order differential equation with a universal form. The outcome, in the second half of the twentieth centary, of the study of random phenomena did not really change that state of affair, with the introduction by Itô of stochastic integration and stochastic differential equations.

Classically, one understands a differential equation as the description of a point motion, the set of all these motions being gathered into a single object called a *flow*. It is a familly $\varphi = \big(\varphi_{ts}\big)_{0 \leqslant s \leqslant t \leqslant T}$ of maps from the state space to itself, such that $\varphi_{tt} = \mathrm{Id}$, for all $0 \leqslant t \leqslant T$, and $\varphi_{ts} = \varphi_{tu} \circ \varphi_{us}$, for all $0 \leqslant s \leqslant u \leqslant t \leqslant T$. The first aim of the approach to some class of dynamics that is proposed is this course is the construction of flows, as opposed to the construction of trajectories started from some given point.

I will explain in the first part of the course a simple method for constructing a flow φ from a family $\mu = \big(\mu_{ts}\big)_{0 \leqslant s \leqslant t \leqslant T}$ of maps that almost forms a flow. The two essential points of this construction are that

    i) φ_{ts} is loosely speaking the composition of infinitely many $\mu_{t_{i+1}t_i}$ along an infinite partition $s < t_1 < \cdots < t$ of the interval $[s, t]$, with infinitesimal mesh,

    ii) φ depends continuously on μ in some sense.

Our main application of this general machinery will be to study some general class of controlled ordinary differential equations, that is differential equations of the form

$$dx_t = \sum_{i=1}^{\ell} V_i(x_t) dh_t^i,$$

where the V_i are vector fields on $\mathbb{R}^d$, say, and the controls h^i are real-valued. Giving some meaning and solving such an equation in some general framework is highly non-trivial outside the framework of absolutely continuous controls, without any extra input like probability, under the form of stochastic calculus for instance. It requires Young integration theory for controls with finite p-variation, for $1 \leqslant p < 2$, and Terry Lyons' theory of rough paths for "rougher" controls! Probabilists are well-acquainted with this kind of situation as stochastic differential equations driven by some Brownian motion are nothing but an example of the kind of problem we

intend to tackle. (With no probability!) It is the aim of this course to give you all the necessary tools to understand what is going on here, in the most elementary way as possible, while aiming at some generality.

The general machinery of approximate flows is best illustrated by looking at the classical Cauchy-Lipschitz theory. Fix some Lipschitz continuous vector fields V_i and some real-valued controls h^i of class $\mathcal{C}^1$. It will appear in our setting that a good way of understanding what it means to be a solution to the ordinary differential equation on $\mathbb{R}^n$

$$(0.1) \qquad \dot{x}_t = \sum_{i=1}^{\ell} V_i(x_t)\dot{h}_t^i =: V_i(x_t)\dot{h}_t^i,$$

is to say that the path $x_\bullet$ satisfies at any time s the Taylor-type expansion formula

$$x_t = x_s + \left(h_t^i - h_s^i\right)V_i(x_s) + o(t-s),$$

and even

$$f(x_t) = f(x_s) + \left(h_t^i - h_s^i\right)(V_i f)(x_s) + O\left(|t-s|^2\right),$$

for any function f of class $\mathcal{C}_b^2$, with $V_i f$ standing for the derivative of f in the direction of V_i. Setting $\mu_{ts}(x) := x + \left(h_t^i - h_s^i\right)V_i(x)$, the preceeding identity rewrites

$$f(x_t) = f\left(\mu_{ts}(x_s)\right) + O\left(|t-s|^2\right),$$

so the elementary map μ_{ts} provides a very accurate description of the dynamics. It almost forms a flow under mild regularity assumptions on the driving vector fields V_i, and its flow associated by the above "almost-flow to flow" machinery happens to be flow classically generated by equation (0.1).

Going back to a probabilistic setting, what insight does this machinery provide on Stratonovich stochastic differential equations

$$(0.2) \qquad \circ dx_t = V_i(x_t)\circ dw_t$$

driven by some Brownian motion w? The use of this notion of differential enables to write the following kind of Taylor-type expansion of order 2 for any function f of class $\mathcal{C}^3$.

(0.3)

$$f(x_t) = f(x_s) + \int_s^t (V_i f)(x_r)\circ dw_r$$

$$= f(x_s) + \left(w_t^i - w_s^i\right)(V_i f)(x_s) + \int_s^t\int_s^r (V_j(V_i f))(x_u)\circ dw_u\circ dw_r$$

$$= f(x_s) + \left(w_t^i - w_s^i\right)(V_i f)(x_s) + \left(\int_s^t\int_s^r \circ dw_u\circ dw_r\right)(V_j(V_i f))(x_s) + \int_s^t\int_s^r\int_s^u(\cdots)$$

For any choice of $2 < p < 3$, the Brownian increments $w_{ts}^i := w_t^i - w_s^i$ have almost-surely a size of order $(t-s)^{\frac{1}{p}}$, the iterated integrals $\int_s^t\int_s^r \circ dw_u\circ dw_r$ have size $(t-s)^{\frac{2}{p}}$, and the triple integral size $(t-s)^{\frac{3}{p}}$, with $\frac{3}{p} > 1$. What will come later out of this

formula is that a solution to equation (0.2) is precisely a path $x_\bullet$ for which one can write for any function f of class $\mathcal{C}^3$ a Taylor-type expansion of order 2 of the form

$$f(x_t) = f(x_s) + (w_t^i - w_s^i)(V_i f)(x_s) + \left(\int_s^t \int_s^r \circ dw_u \circ dw_r \right)(V_j(V_i f))(x_s) + O(|t-s|^a)$$

at any time s, for some exponent $a > 1$ independent of s. This conclusion puts forward the fact that what the dynamics really see of the Brownian control w is not only its increments w_{ts} but also its iterated integrals $\int_s^t \int_s^r \circ dw_u \circ dw_r$. The notion of p-rough path $\mathbf{X} = (X_{ts}, \mathbb{X}_{ts})_{0 \leqslant s \leqslant t \leqslant T}$ is an abstraction of this family of pairs of quantities, for $2 < p < 3$ here. This *multi-level object* satisfies some constraints of analytic type (size of its increments) and algebraic type, coming from the higher level parts of the object. As they play the role of some iterated integrals, they need to satisfy some identities consequences of the Chasles relation for elementary integrals: $\int_s^t = \int_s^u + \int_u^t$. These constraints are all what these rough paths $\mathbf{X} = (X, \mathbb{X})$ need to satisfy to give a sense to the equation

$$(0.4) \qquad\qquad dx_t = \mathrm{F}^{\otimes}(x_t)\mathbf{X}(dt)$$

for a collection $\mathrm{F} = (V_1, \ldots, V_\ell)$ of vector fields on $\mathbb{R}^n$, by defining a solution as a path $x_\bullet$ for which one can write some uniform Taylor-type expansion of order 2

$$(0.5) \qquad f(x_t) = f(x_s) + X_{ts}^i (V_i f)(x_s) + \mathbb{X}_{ts}^{jk}(V_j(V_k f))(x_s) + O(|t-s|^a),$$

for any function f of class $\mathcal{C}_b^3$. The notation $\mathrm{F}^{\otimes}$ is used here to insist on the fact that it is not only the collection F of vector fields that is used in this definition, but also the differential operators $V_j V_k$ constructed from F. The introduction and the study of p-rough paths and their collection is done in the second part of the course.

Guided by the results on flows of the first part, we shall reinterpret equation (0.4) to construct directly a flow φ solution to the equation

$$(0.6) \qquad\qquad d\varphi = \mathrm{F}^{\otimes}\mathbf{X}(dt),$$

in a sense to be made precise in the third part of the course. The recipe of construction of φ will consist in associating to F and $\mathbf{X}$ a $\mathcal{C}^1$-approximate flow $\mu = (\mu_{ts})_{0 \leqslant s \leqslant t \leqslant T}$ having everywhere a behaviour similar to that described by equation (0.5), and then to apply the theory described in the first part of the course. The maps μ_{ts} will be constructed as the time 1 maps associated with some ordinary differential equation constructed from F and $\mathbf{X}_{ts}$ in a simple way. As they will depend continuously on $\mathbf{X}$, the continuous dependence of φ on $\mathbf{X}$ will come as a consequence of point **ii)** above.

All that will be done in a deterministic setting. We shall see in the fourth part of the course how this approach to dynamics is useful in giving a fresh viewpoint on stochastic differential equations and their associated dynamics. The key point will be the fundamental fact that Brownian motion has a natural lift to a Brownian p-rough path, for any $2 < p < 3$. Once this object will be constructed by probabilistic means, the deterministic machinery for solving rough differential equations, described in the third part of the course, will enable us to associate to any realization of the Brownian rough path a solution to the rough differential equation (0.4). This solution coincides

almost-surely with the solution to the Stratonovich differential equation (0.2)! One shows in that way that this solution is a continuous function of the Brownian rough path, in striking contrast with the fact that it is only a measurable function of the Brownian path itself, with no hope for a more regular dependence in a generic setting. This fact will provide a natural and easy road to the deep results of Wong-Zakai, Stroock & Varadhan or Freidlin & Wentzell.

Several other approaches to rough differential equations are available, each with their own pros and cons. We refer the reader to the books [1] and [2] for an account of Lyons' original approach; she/he is refered to the book [3] for a thourough account of the Friz-Victoir approach, and to the lecture note [4] by Baudoin for an easier account of their main ideas and results, and to the forthcoming excellent lecture notes [5] by Friz and Hairer on Gubinelli's point of view. The present approach building on [6] does not overlap with the above ones.[1]

---

[1]Comments on these lecture notes are most welcome. Please email them at the address ismael.bailleul@univ-rennes1.fr

CHAPTER 2

Flows and approximate flows

This first part of the course will present the backbone of our approach to rough dynamics under the form of a simple recipe for constructing flows of maps on some Banach space. Although naive, it happens to be robust enough to provide a unified treatment of ordinary, rough and stochastic differential equations. We fix throughout a Banach space V.

The main technical difficulty is to deal with the non-commutative character of the space of maps from V to itself, endowed with the composition operation. To understand the part of the problem that does not come from non-commutativity, let us consider the following model problem. Suppose we are given a family $\mu = \left(\mu_{ts}\right)_{0 \leqslant s \leqslant t \leqslant 1}$ of elements of some Banach space depending continuously on s and t, and such that $\left|\mu_{ts}\right| = o_{t-s}(1)$. Is it possible to construct from μ a family $\varphi = \left(\varphi_{ts}\right)_{0 \leqslant s \leqslant t \leqslant 1}$ of elements of that Banach space, depending continuously on s and t, and such that we have

$$(0.7) \qquad\qquad \varphi_{tu} + \varphi_{us} = \varphi_{ts}$$

for all $0 \leqslant s \leqslant u \leqslant t \leqslant 1$? This additivity property plays the role of the flow property. Would the time interval $[0,1]$ be a finite discrete set $t_1 < \cdots < t_n$, the additivity property (0.7) would mean that φ_{ts} is the sum of the $\varphi_{t_{i+1}t_i}$, whose definition should be $\mu_{t_{i+1}t_i}$, as these are the only quantities we are given if no arbitrary choice is to be done. Of course, this will not turn φ into an additive map, in the sense that property (0.7) holds true, in this discrete setting, but it suggest the following attempt in the continuous setting of the time interval $[0,1]$.

Given a partition $\pi = \left\{0 < t_1 < \cdots < 1\right\}$ of $[0,1]$ and $0 \leqslant s \leqslant t \leqslant 1$, set

$$\varphi_{ts}^{\pi} = \sum_{s \leqslant t_i < t_{i+1} \leqslant t} \mu_{t_{i+1}t_i}.$$

This map almost satisfies relation (0.7) as we have

$$\varphi_{tu}^{\pi} + \varphi_{us}^{\pi} = \varphi_{ts}^{\pi} - \mu_{u^+u^-} = \varphi_{ts}^{\pi} + o_{|\pi|}(1),$$

for all $0 \leqslant s \leqslant u \leqslant t \leqslant 1$, where u^-, u^+ are the elements of π such that $u^- \leqslant u < u^+$, and $|\pi| = \max\left\{t_{i+1} - t_i\right\}$ stands for the mesh of the partition. So we expect to find a solution φ to our problem under the form φ^{π}, for a partition of $[0,1]$ of infinitesimal mesh, that is as a limit of φ^{π}'s, say along a sequence of refined partitions π_n where π_{n+1} has only one more point than π_n, say u_n. However, the sequence φ^{π_n} has no reason to converge without assuming further conditions on μ. To fix further the

setting, let us consider partitions π_n of $[0, 1]$ by dyadic times, where we exhaust first all the dyadic times multiples of 2^{-k}, in any order, before taking in the partition points multiples of $2^{-(k+1)}$. Two dyadic times s and t being given, both multiples of 2^{-k_0}, take n big enough for them to be points of π_n. Then, denoting by u_n^-, u_n^+ the two points of π_n such that $u_n^- < u_n < u_n^+$, the quantity $\varphi_{ts}^{\pi_{n+1}} - \varphi_{ts}^{\pi_n}$ will either be null if $u_n \notin [s, t]$, or

(0.8) $$\varphi_{ts}^{\pi_{n+1}} - \varphi_{ts}^{\pi_n} = \left(\mu_{u_n^+ u_n} + \mu_{u_n u_n^-} \right) - \mu_{u_n^+ u_n^-},$$

otherwise. A way to control this quantity is to assume that the map μ is *approximately additive*, in the sense that we have some positive constants c_0 and $a > 1$ such that the inequality

(0.9) $$\left| \left(\mu_{tu} + \mu_{us} \right) - \mu_{ts} \right| \leqslant c_0 \left| t - s \right|^a$$

holds for all $0 \leqslant s \leqslant u \leqslant t \leqslant 1$. Under this condition, we have

$$\left| \varphi_{ts}^{\pi_{n+1}} - \varphi_{ts}^{\pi_n} \right| \leqslant c_0 2^{-am},$$

where $\left| \pi_{n+1} \right| = 2^{-m}$. There will be 2^m such terms in the formal series $\sum_{n \geqslant 0} \left(\varphi_{ts}^{\pi_{n+1}} - \varphi_{ts}^{\pi_n} \right)$, giving a total contribution for these terms of size $2^{-(a-1)m}$, summable in m. So this sum converges to some quantity φ_{ts} which satisfies (0.7) by construction (on dyadic times only, as defined as above). Note that commutativity of the addition operation was used implicitly to write down equation (0.8).

Somewhat surprisingly, the above approach also works in the non-commutative setting of maps from V to itself under a condition which essentially amounts to replacing the addition operation and the norm $|\cdot|$ in condition (0.9) by the composition operation and the $\mathcal{C}^1$ norm. This will be the essential content of theorem 2 below, taken from the work [6].

1. $\mathcal{C}^1$-approximate flows and their associated flows

We start by defining what will play the role of an approximate flow, in the same way as μ above was understood as an approximately additive map under condition (0.9).

DEFINITION 1. *A $\mathcal{C}^1$-**approximate flow** on* V *is a family* $\mu = \left(\mu_{ts} \right)_{0 \leqslant s \leqslant t \leqslant T}$ *of* $\mathcal{C}^2$ *maps from V to itself, depending continuously on* s, t *in the topology of uniform convergence, such that*

(1.1) $$\left\| \mu_{ts} - \mathrm{Id} \right\|_{\mathcal{C}^2} = o_{t-s}(1)$$

and there exists some positive constants c_1 *and* $a > 1$, *such that the inequality*

(1.2) $$\left\| \mu_{tu} \circ \mu_{us} - \mu_{ts} \right\|_{\mathcal{C}^1} \leqslant c_1 |t - s|^a$$

holds for all $0 \leqslant s \leqslant u \leqslant t \leqslant T$.

Note that μ_{ts} is required to be $\mathcal{C}^2$ close to the identity while we ask it to be an approximate flow in a $\mathcal{C}^1$ sense. Given a partition $\pi_{ts} = \{ s = s_0 < s_1 < \cdots < s_{n-1} < s_n = t \}$ of an interval $[s, t] \subset [0, T]$, set

$$\mu_{\pi_{ts}} = \mu_{t_n t_{n-1}} \circ \cdots \circ \mu_{t_1 t_0}.$$

THEOREM 2 (Constructing flows on a Banach space). *A $\mathcal{C}^1$-approximate flow defines a unique flow $\varphi = (\varphi_{ts})_{0 \leqslant s \leqslant t \leqslant T}$ on V such that the inequality*

$$(1.3) \qquad \left\| \varphi_{ts} - \mu_{ts} \right\|_\infty \leqslant c|t-s|^a$$

holds for some positive constant c, for all $0 \leqslant s \leqslant t \leqslant T$ sufficiently close, say $t - s \leqslant \delta$. This flow satisfies the inequality

$$(1.4) \qquad \left\| \varphi_{ts} - \mu_{\pi_{ts}} \right\|_\infty \leqslant \frac{2}{1 - 2^{1-a}} \, c_1^2 T \left| \pi_{ts} \right|^{a-1}$$

for any partition π_{ts} of any interval (s,t) of mesh $\left| \pi_{ts} \right| \leqslant \delta$.

Note that the conclusion of theorem 2 holds in $\mathcal{C}^0$-norm. This loss of regularity with respect to the controls on μ given by equations (1.1) and (1.2) roughly comes from the use of uniform $\mathcal{C}^1$-estimates on some functions f_{ts} to control some increments of the form $f_{ts} \circ g_{ts} - f_{ts} \circ g'_{ts}$, for some $\mathcal{C}^0$-close maps g_{ts}, g'_{ts}. Note that if μ depends continuously on some parameter, then φ also depends continuously on that parameter, as a uniform limit of continuous functions, equation (1.8).

The remainder of this section will be dedicated to the proof of theorem 2. We shall proceed in two steps, by proving first that one can construct φ as the *uniform limit* of the μ_π's provided one can control uniformly their Lipschitz norm. This control will be proved in a second step.

1.1. First step. Let us introduce the following inductive definition to prepare the first step.

DEFINITION 3. *Let $\epsilon \in (0,1)$ be given. A partition $\pi = \{s = s_0 < s_1 < \cdots < s_{n-1} < s_n = t\}$ of (s,t) is said to be ϵ-**special** if it is either trivial or*
- *one can find an $s_i \in \pi$ sucht that $\epsilon \leqslant \frac{s_i - s}{t - s} \leqslant 1 - \epsilon$,*
- *and for any choice u of such an s_i, the partitions of $[s,u]$ and $[u,t]$ induced by π are both ϵ-special.*

A partition of any interval into sub-intervals of equal length has special type $\frac{1}{2}$. Given a partition $\pi = \{s = s_0 < s_1 < \cdots < s_{n-1} < s_n = t\}$ of (s,t) of special type ϵ and $u \in \{s_1, \ldots, s_{n-1}\}$ with $\epsilon \leqslant \frac{u-s}{t-s} \leqslant 1 - \epsilon$, the induced partitions of the intervals $[s,u]$ and $[u,t]$ are also ϵ-special. Set $m_\epsilon = \sup_{\epsilon \leqslant \beta \leqslant 1-\epsilon} \beta^a + (1-\beta)^a < 1$, and pick a constant

$$L > \frac{2c_1}{1 - m_\epsilon},$$

where c_1 is the constant that appears in the dfinition of a $\mathcal{C}^1$-approximate flow, in equation (1.2).

LEMMA 4. *Let $\mu = (\mu_{ts})_{0 \leqslant s \leqslant t \leqslant T}$ be a $\mathcal{C}^1$-approximate flow on V. Given $\epsilon > 0$, there exists a positive constant δ such that for any $0 \leqslant s \leqslant t \leqslant T$ with $t - s \leqslant \delta$, and any special partition of type ϵ of an interval $(s,t) \subset [0,T]$, we have*

$$(1.5) \qquad \left\| \mu_{\pi_{ts}} - \mu_{ts} \right\|_\infty \leqslant L|t-s|^a.$$

PROOF – We proceed by induction on the number n of sub-intervals of the partition. The case $n = 2$ is the $\mathcal{C}^0$ version of identity (1.2). Suppose the statement has been proved for $n \geqslant 2$. Fix $0 \leqslant s < t \leqslant T$ with $t - s \leqslant \delta$, and let $\pi_{ts} = \{s_0 = s < s_1 < \cdots < s_n < s_{n+1} = t\}$ be an ϵ-special partition of $[s, t]$, splitting the interval $[s, t]$ into $(n + 1)$ sub-intervals. Let u be one of the points of the partition sucht that $\epsilon \leqslant \frac{t-u}{t-s} \leqslant 1 - \epsilon$, so the two partitions π_{tu} and π_{us} are both $-\epsilon$-special, with respective cardinals no greater than n. Then

$$
\begin{aligned}
\left\| \mu_{\pi_{ts}} - \mu_{ts} \right\|_\infty &\leqslant \left\| \mu_{\pi_{tu}} \circ \mu_{\pi_{us}} - \mu_{tu} \circ \mu_{\pi_{us}} \right\|_\infty + \left\| \mu_{tu} \circ \mu_{\pi_{us}} - \mu_{ts} \right\|_\infty \\
&\leqslant \left\| \mu_{\pi_{tu}} - \mu_{tu} \right\|_\infty + \left\| \mu_{tu} \circ \mu_{\pi_{us}} - \mu_{tu} \circ \mu_{us} \right\|_\infty + \left\| \mu_{tu} \circ \mu_{us} - \mu_{ts} \right\|_\infty \\
&\leqslant L|t - u|^a + \big(1 + o_\delta(1)\big) L\, |u - s|^a + c_1 |t - s|^a,
\end{aligned}
$$

by the induction hypothesis and (1.1) and (1.2). Set $u - s = \beta(t - s)$, with $\epsilon \leqslant \beta \leqslant 1 - \epsilon$. The above inequality rewrites

$$
\left\| \mu_{\pi_{ts}} - \mu_{ts} \right\|_\infty \leqslant \left\{ \big(1 + o_\delta(1)\big)\big((1 - \beta)^a + \beta^a\big) L + c_1 \right\} |t - s|^a.
$$

In order to close the induction, we need to choose δ small enough for the condition

(1.6) $$ c_1 + \big(1 + o_\delta(1)\big) m_\epsilon L \leqslant L $$

to hold; this can be done since $m_\epsilon < 1$. $\triangleright$

As a shorthand, we shall write μ_{ts}^n for $\bigcirc_{i=0}^{n-1} \mu_{t_{i+1}t_i}$, where $s_i = s + \frac{i}{n}(t - s)$. The next proposition is to be understood as the core of our approach.

PROPOSITION 5 (Step 1). *Let* $\mu = \big(\mu_{ts}\big)_{0 \leqslant s \leqslant t \leqslant T}$ *be a* $\mathcal{C}^1$*-approximate flow on* V. *Assume the existence of a positive constant* δ *such that the maps* μ_{ts}^n, *for* $n \geqslant 2$ *and* $t - s \leqslant \delta$, *are all Lipschitz continuous, with a Lipschitz constant uniformly bounded above by some constant* c_2, *then there exists a unique flow* $\varphi = \big(\varphi_{ts}\big)_{0 \leqslant s \leqslant t \leqslant T}$ *on* V *such that the inequality*

(1.7) $$ \left\| \varphi_{ts} - \mu_{ts} \right\|_\infty \leqslant c|t - s|^a $$

holds for some positive constant c, *for all* $0 \leqslant s \leqslant t \leqslant T$ *with* $t - s \leqslant \delta$. *This flow satisfies the inequality*

(1.8) $$ \left\| \varphi_{ts} - \mu_{\pi_{ts}} \right\|_\infty \leqslant c_1 c_2 T \left| \pi_{ts} \right|^{a-1} $$

for any partition π_{ts} *of* (s, t), *of mesh* $\left| \pi_{ts} \right| \leqslant \delta$.

PROOF – The existence and uniqueness proofs both rely on the elementary identity
(1.9)

$$
f_N \circ \cdots \circ f_1 - g_N \circ \cdots \circ g_1 = \sum_{i=1}^N \Big(g_N \circ \cdots \circ g_{N-i+1} \circ f_{N-i} - g_N \circ \cdots \circ g_{N-i+1} \circ g_{N-i} \Big) \circ f_{N-i-1} \circ \cdots \circ f_1,
$$

where the g_i and f_i are maps from V to itself, and where we use the obvious convention concerning the summand for the first and last term of the sum. In

particular, if all the maps $g_N \circ \cdots \circ g_k$ are Lipschitz continuous, with a common upper bound c' for their Lipschitz constants, then

$$(1.10) \qquad \left\| f_N \circ \cdots \circ f_1 - g_N \circ \cdots \circ g_1 \right\|_\infty \leqslant c' \sum_{i=1}^{N} \| f_i - g_i \|_\infty.$$

a) Existence. Set $D_\delta := \{ 0 \leqslant s \leqslant t \leqslant T \,;\, t - s \leqslant \delta \}$ and write $\mathbb{D}_\delta$ for the intersection of D_δ with the set of dyadic real numbers. Given $s = a2^{-k_0}$ and $t = b2^{-k_0}$ in $\mathbb{D}_\delta$, define for $n \geqslant k_0$

$$\mu_{ts}^{(n)} := \mu_{ts}^{2^n} = \mu_{s_{N(n)} s_{N(n)-1}} \circ \cdots \circ \mu_{s_1 s_0},$$

where $s_i = s + i2^{-n}$ and $s_{N(n)} = t$. Given $n \geqslant k_0$, write

$$\mu_{ts}^{(n+1)} = \overset{N(n)-1}{\underset{i=0}{\bigcirc}} \left(\mu_{s_{i+1} s_i + 2^{-n-1}} \circ \mu_{s_i + 2^{-n-1} s_i} \right)$$

and use (1.9) with $f_i = \mu_{s_{i+1} s_i + 2^{-n-1}} \circ \mu_{s_i + 2^{-n-1} s_i}$ and $g_i = \mu_{s_{i+1} s_i}$ and the fact that all the maps $\mu_{s_{N(n)} s_{N(n)-1}} \circ \cdots \circ \mu_{s_{N(n)-i+1} s_{N(n)-i}} = \mu_{s_{N(n)} s_{N(n)-i}}^i$ are Lipschitz continuous with a common Lipschitz constant c_2, by assumption, to get by (1.10) and (1.2)

$$\left\| \mu_{ts}^{(n+1)} - \mu_{ts}^{(n)} \right\|_\infty \leqslant c_2 \sum_{i=0}^{N(n)-1} \left\| \mu_{s_{i+1} s_i + 2^{-n-1}} \circ \mu_{s_i + 2^{-n-1} s_i} - \mu_{s_{i+1} s_i} \right\|_\infty \leqslant c_1 c_2 T \, 2^{-(a-1)n};$$

so $\mu^{(n)}$ converges uniformly on $\mathbb{D}_\delta$ to some continuous function φ. We see that φ satisfies inequality (1.3) on $\mathbb{D}_\delta$ as a consequence of (1.5). As φ is a uniformly continuous function of $(s,t) \in \mathbb{D}_\delta$, by (1.3), it has a unique continuous extension to D_δ, still denoted by φ. To see that it defines a flow on D_δ, notice that for dyadic times $s \leqslant u \leqslant t$, we have $\mu_{ts}^{(n)} = \mu_{tu}^{(n)} \circ \mu_{us}^{(n)}$, for n big enough; so, since the maps $\varphi_{tu}^{(n)}$ are uniformly Lipschitz continuous, we have $\varphi_{ts} = \varphi_{tu} \circ \varphi_{us}$ for such triples of times in $\mathbb{D}_\delta$, hence for all times since φ is continuous. The map φ is easily extended as a flow to the whole of $\{ 0 \leqslant s \leqslant t \leqslant T \}$. Note that φ inherits from the μ^n's their Lipschitz character, for a Lipschitz constant bounded above by c_2.

b) Uniqueness. Let ψ be any flow satisfying condition (1.3). With formulas (1.9) and (1.10) in mind, rewrite (1.3) under the form $\psi_{ts} = \mu_{ts} + O_c(|t-s|^a)$, with obvious notations. Then

$$\psi_{ts} = \psi_{s_{2^n} s_{2^n-1}} \circ \cdots \circ \psi_{s_1 s_0} = \left(\mu_{s_{2^n} s_{2^n-1}} + O_c(2^{-an}) \right) \circ \cdots \circ \left(\mu_{s_1 s_0} + O_c(2^{-an}) \right)$$

$$= \mu_{s_{2^n} s_{2^n-1}} \circ \cdots \circ \mu_{s_1 s_0} + \Delta_n = \mu_{ts}^{(n)} + \Delta_n,$$

where Δ_n is of the form of the right hand side of (1.9), so is bounded above by a constant multiple of $2^{-(a-1)n}$, since all the maps $\mu_{s_{2^n} s_{2^n-1}} \circ \cdots \circ \mu_{s_{2^n-\ell+1} s_{2^n-\ell}}$ are Lipschitz continuous with a common upper bound for their Lipschitz constants, by assumption. Sending n to infinity shows that $\psi_{ts} = \varphi_{ts}$.

c) Speed of convergence. Given any partition $\pi = \{s_0 = s < \cdots < s_n = t\}$ of (s, t), writing $\varphi_{ts} = \bigcirc_{i=0}^{n-1} \varphi_{s_{i+1}s_i}$, and using their uniformly Lipschitz character, we see as a consequence of (1.10) that we have for $|\pi_{ts}| \leqslant \delta$

$$\left\|\varphi_{ts} - \mu_{\pi_{ts}}\right\|_\infty \leqslant c_2 \sum_{i=0}^{n-1} \left\|\varphi_{s_{i+1}s_i} - \mu_{s_{i+1}s_i}\right\|_\infty \leqslant c_1 c_2 \sum_{i=0}^{n-1} |s_{i+1} - s_i|^a \leqslant c_1 c_2 T \left|\pi_{ts}\right|^{a-1}.$$

$\triangleright$

Compare what is done in the above proof with what was done in the introduction to this part of the course in a commutative setting.

1.2. Second step. The uniform Lipschitz control assumed in proposition 5 actually holds under the assumption that μ is a $\mathcal{C}^1$-approximate flow. The results of this paragraph could have been proved just after lemma 4 and do not use the result proved in the fundamental proposition 5. Recall L stands for a constant strictly greater than $\frac{2c_1}{1-m_\epsilon}$.

PROPOSITION 6 (Uniform Lipschitz controls). *Let $\mu = \left(\mu_{ts}\right)_{0 \leqslant s \leqslant t \leqslant T}$ be a $\mathcal{C}^1$-approximate flow on* V. *Then, given $\epsilon > 0$, there exists a positive constant δ such that the inequality*

$$\left\|\mu_{\pi_{ts}} - \mu_{ts}\right\|_{\mathcal{C}^1} \leqslant L|t - s|^a$$

holds for any partition π_{ts} of $[s, t]$ of special type ϵ, whenever $t - s \leqslant \delta$.

PROOF – We proceed by induction on the number n of sub-intervals of the partition as in the proof of lemma 4. The case $n = 2$ is identity (1.2). Suppose the statement has been proved for $n \geqslant 2$. Fix $0 \leqslant s < t \leqslant T$ with $t - s \leqslant \delta$, and let $\pi_{ts} = \{s_0 = s < s_1 < \cdots < s_n < s_{n+1} = t\}$ be an ϵ-special partition of $[s, t]$ of special, splitting the interval $[s, t]$ into $(n + 1)$ sub-intervals. Let u be a point of the partition with $\epsilon \leqslant \frac{u-s}{t-s} \leqslant 1 - \epsilon$, so that the two partitions π_{tu} and π_{us} are both ϵ-special, with respective cardinals no greater than n. Then, for any $x \in$ V, one can write $D_x\mu_{\pi_{ts}} - D_x\mu_{ts}$ as a telescopic sum which involve only some controlled quantities.

$$D_x\mu_{\pi_{ts}} - D_x\mu_{ts} = D_x\left(\mu_{\pi_{tu}} \circ \mu_{\pi_{us}}\right) - D_x\mu_{ts}$$

$$= \left(D_{\mu_{\pi_{us}}(x)}\mu_{\pi_{tu}} - D_{\mu_{\pi_{us}}(x)}\mu_{tu}\right)\left(D_x\mu_{\pi_{us}}\right) + \left(\left(D_{\mu_{\pi_{us}}(x)}\mu_{tu} - D_{\mu_{us}(x)}\mu_{tu}\right)\left(D_x\mu_{\pi_{us}}\right)\right)$$

$$+ \left(D_{\mu_{us}(x)}\mu_{tu}\right)\left(D_x\mu_{\pi_{us}} - D_x\mu_{us}\right) + \left(\left(D_{\mu_{us}(x)}\mu_{tu}\right)\left(D_x\mu_{us}\right) - D_x\mu_{ts}\right)$$

$$=: (1) + (2) + (3) + (4)$$

We treat each term separately using repeatedly the induction hypothesis, continuity assumption (1.1) for μ_{ts} in $\mathcal{C}^2$ topology, and lemma 4 when needed. We first have

$$\left|(1)\right| \leqslant L|t - u|^a \left(1 + o_\delta(1)\right).$$

Also,

$$\left|D_{\mu_{\pi_{us}}(x)}\mu_{tu} - D_{\mu_{us}(x)}\mu_{tu}\right| \leqslant o_{t-u}(1)\left|\mu_{\pi_{us}}(x) - \mu_{us}(x)\right| \leqslant o_{t-u}(1)\,L|u - s|^a,$$

As the term $D_x \mu_{\pi_{us}}$ has size no greater than $\big(1 + o_\delta(1)\big) + L|u - s|^a$, we have

$$\big|(2)\big| \leqslant o_\delta(1)\,|u - s|^a.$$

Last, we have the upper bound

$$\big|(3)\big| \leqslant \big(1 + o_\delta(1)\big)L|u - s|^a,$$

while $\big|(4)\big| \leqslant \big\|\mu_{tu} \circ \mu_{us} - \mu_{ts}\big\|_{\mathcal{C}^1} \leqslant c_1|t - s|^a$ by (1.2). All together, and writing $t - u = \beta(t - s)$, for some $\beta \in [\epsilon, 1 - \epsilon]$, this gives

$$\big|D_x\mu_{\pi_{ts}} - D_x\mu_{ts}\big| \leqslant \Big(\big(1 + o_\delta(1)\big)\big(\beta^a + (1 - \beta)^a\big)L + c_1 + o_\delta(1)\Big)|t - s|^a$$
$$\leqslant L\,|t - s|^a$$

for δ small enough, as $m_\epsilon < 1$. $\triangleright$

Propositions 5 and 6 together prove theorem 2. Note that an explicit choice of δ is possible as soon as one has a quantitative version of the estimate $\big\|\mu_{ts} - \mathrm{Id}\big\|_{\mathcal{C}^2} = o_{t-s}(1)$. Note also that proposition 6 provides an explicit control on the Lipschitz norm of the φ_{ts}, in terms of the Lipschitz norm of μ_{ts} and L.

2. Exercices on flows

To get a hand on the machinery of $\mathcal{C}^1$-approximate flows, we shall first see how theorem 2 gives back the classical Cauchy-Lipschitz theory of ordinary differential equations for *bounded* Lipschitz vector fields on $\mathbb{R}^d$. Working with *unbounded* Lipschitz vector fields requires a slightly different notion of *local* $\mathcal{C}^1$-approximate flow – see [6].

Theorem 2 can be understood as a non-commutative analogue of Feyel-de la Pradelle's sewing lemma [7], first introduced by Gubinelli [9] as an abstraction of a fundamental mechanism invented by Lyons [20]. Exercices 2-4 are variations on this commutative version of theorem 2, as already sketched in the introduction to this part.

1. Ordinary differential equations. Let $V_1, \ldots, V_\ell$ be $\mathcal{C}_b^2$ vector fields on $\mathbb{R}^d$ (or a Banach space), and $h_1, \ldots, h_\ell$ be real-valued $\mathcal{C}^1$ controls. Let φ stand for the flow associated with the ordinary differential equation

$$dx_t = V_i(x_t)dh_t^i.$$

a) Show that one defines a $\mathcal{C}^1$-approximate flow setting for all $x \in \mathbb{R}^d$

$$\mu_{ts}(x) = x + \big(h_t - h_s\big)^i V_i(x).$$

b) Prove that φ is equal to the flow associated to μ by theorem (2). In that sense, a path x is a solution to the above ordinary differential equation if and only if it satisfies at any time s the Taylor-type expansion formula

$$f\big(x_t\big) = f\big(x_s\big) + \big(h_t^i - h_s^i\big)\big(V_i f\big)(x_s) + o(t - s),$$

for any function f of class $\mathcal{C}_b^2$. Show that the above reasoning holds true if we only assume that the $\mathbb{R}^\ell$-valued control h is globally Lipschitz continuous. (It is actually sufficient to suppose h is α-Hölder, for some $\alpha > \frac{1}{2}$.)

c) Does anything go wrong with the above reasoning if the Lipschitz continuous vector fields V_i are not bounded?

d) Show that φ depends continuously on h in the uniform topology for φ and the Lipschitz topology for h, defined by the distance

$$d(h, h') = |h_0 - h'_0| + \mathrm{Lip}(h - h'),$$

where $\mathrm{Lip}(h - h')$ stands for the Lipschitz norm of $h - h'$. (A similar result holds if h is α-Hölder, for some $\alpha > \frac{1}{2}$, with the Lipschitz norm replaced by the α-Hölder norm.)

2. Feyel-de la Pradelle' commutative sewing lemma. Let V be a Banach space and $\mu = (\mu_{ts})_{0 \leqslant s \leqslant t \leqslant 1}$ be a V-valued continuous function. The following commutative version of theorem 2 was first proved under this form by Feyel and de la Pradelle in [**7**]; see also [**8**]. Suppose there exists some positive constants c_0 and $a > 1$ such that we have

(2.1) $$\big|(\mu_{tu} + \mu_{us}) - \mu_{ts}\big| \leqslant c_0 |t - s|^a$$

for all $0 \leqslant s \leqslant u \leqslant t \leqslant 1$. We say the μ is an *almost-additive functional* (or map).

Simplify the proof of theorem 2 to show that there exists a unique map $\varphi = (\varphi_t)_{0 \leqslant t \leqslant 1}$, whose increments $\varphi_{ts} := \varphi_t - \varphi_s$, satisfy

$$\big|\varphi_{ts} - \mu_{ts}\big| \leqslant c|t - s|^a$$

for some positive constant c and all $0 \leqslant s \leqslant t \leqslant 1$.

3. Integral products. Let $\alpha > \frac{1}{2}$ be given, and $(A_t)_{0 \leqslant t \leqslant 1}$ be an α-Hölder path with values in the space $\mathrm{L}_c(\mathrm{V})$ of continuous linear maps from V to itself. Set $A_{ts} = A_t - A_s$, for $0 \leqslant s \leqslant t \leqslant 1$, and use the notation $|\cdot|$ for the operator norm on $\mathrm{L}_c(\mathrm{V})$.

a) Use theorem (2) to show that setting $\mu_{ts} = \mathrm{Id} + A_{ts}$ determines a unique flow φ on V. A good notation for φ_{ts} is $\prod_{s \leqslant r \leqslant t} (\mathrm{Id} + dA_r) = \prod_{s \leqslant r \leqslant t} e^{dA_r}$.

b) Let $(B_t)_{0 \leqslant t \leqslant 1}$ be another $\mathrm{L}_c(\mathrm{V})$-valued α-Hölder path. Show that we define a $\mathcal{C}^1$-approximate flow μ setting $\mu_{ts} = (\mathrm{Id} + A_{ts})(\mathrm{Id} + B_{ts})$. Prove as a consequence the well-known formula

$$e^{A_0 + B_0} = \lim_n \left(e^{\frac{A_0}{2^n}} e^{\frac{B_0}{2^n}} \right)^{2^n}.$$

4. Controls and finite-variation paths. A *control* is a non-negative map $\omega = (\omega_{ts})_{0 \leqslant s \leqslant t \leqslant 1}$, null on the diagonal, and such that we have

$$\omega_{tu} + \omega_{us} \leqslant \omega_{ts}$$

for all $0 \leqslant s \leqslant u \leqslant t \leqslant 1$.

a) Show that Feyel and de la Pradelle' sewing lemma holds true if we replace $t - s$ by ω_{ts}, if we suppose that ω^a is a control.

b) Recall that a V-valued path $x = (x_t)_{0 \leqslant t \leqslant 1}$ is said to have finite p-variation, for some $p \geqslant 1$, if the following quantity is finite for all $0 \leqslant s \leqslant t \leqslant 1$:

$$|x|^p_{p-\mathrm{var};[s,t]} := \sup \sum \left| x_{t_{i+1}} - x_{t_i} \right|^p,$$

with a sum over the partition points t_i of a given partition of the interval $[s, t]$, and a supremum over the set of all partitions of $[s, t]$. Such a definition is invariant by any reparametrization of the time interval $[s, t]$. Given such a path, show that setting $\omega_{ts} = |x|^p_{p-\mathrm{var};[s,t]}$, defines a control ω.

c) Show that a path with finite p-variation can be reparametrized into a $\frac{1}{p}$-Hölder path, with 1-Hölder paths being understood as Lipschitz continuous paths. Given an $\mathbb{R}^\ell$-valued path h with finite 1-variation, set

$$\zeta_s = \inf\{t \geqslant 0 \,;\, |h|_{1-\mathrm{var};[0,t]} \geqslant s\}.$$

We define a solution $x_\bullet$ to the ordinary differential equation

$$dx_t = V_i(x_t)dh_t^i$$

driven by h as a path $x_\bullet$ such that the reparametrized path $y := x \circ \zeta$ is a solution to the ordinary differential equation

$$dy_s = V_i(y_s)d(h \circ \zeta)_t^i,$$

driven by the globally Lipschitz path $h \circ \zeta$.

d) Prove that the flow φ constructed in this case from theorem 2 depends continuously on h in the uniform norm for φ and the 1-variation topology associated with the norm $|\cdot|_{1-\mathrm{var}}$ for h. (Following the remarks of exercice 1, one can actually prove the results of questions c) and d) for paths with finite p-variation, for $1 \leqslant p < 2$.)

5. Young integral. Given another Banach space E, denote by $\mathrm{L}_c(\mathrm{V,E})$ the space of continuous linear maps from V to E equipped with the operator norm. Let α and β be positive real numbers such that $\alpha + \beta > 1$. Given any $0 < \alpha < 1$, we denote by $\mathrm{Lip}_\alpha(\mathrm{E})$ the set of α-Hölder maps. This unusual notation will be justified in the third path of the course.

a) Given an $\mathrm{L}_c(\mathrm{V,E})$-valued α-Lipschitz map $x = (x_s)_{0 \leqslant s \leqslant 1}$ and a V-valued β-Lipschitz map $y = (y_s)_{0 \leqslant s \leqslant 1}$, show that setting

$$\mu_{ts} = x_s(y_t - y_s)$$

for all $0 \leqslant s \leqslant t \leqslant 1$, defines an E-valued function μ that satisfies equation (2.1), with a constant c_0 to be made explicit.

b) The associated function φ is denoted by $\varphi_t = \int_0^t x_s dy_s$, for all $0 \leqslant t \leqslant 1$. Show that it is a continuous function of $x \in \mathrm{Lip}_\alpha\big(\mathrm{L}_c(\mathrm{V,E})\big)$ and $y \in \mathrm{Lip}_\beta(\mathrm{V})$.

6. Lipschitz dependence of φ on μ. As emphasized in the remark following theorem 2, inequality (1.8) implies that φ, understood as a function of μ, is continuous in the $\mathcal{C}^0$-norm on the sets of μ's of the form $\{\mu \,;\, (1.2)\text{ holds uniformly}\}$, equipped with the $\mathcal{C}^0$-norm. One can actually prove that it depends Lipschitz continuously on μ in the following sense.

Let $\mu = (\mu_{ts})_{0 \leqslant s \leqslant t \leqslant 1}$ and $\mu' = (\mu'_{ts})_{0 \leqslant s \leqslant t \leqslant 1}$ be $\mathcal{C}^1$-approximate flows on V, with associated flows φ and φ'. Suppose that we have

$$\big\|\mu_{ts} - \mu'_{ts}\big\|_{\mathcal{C}^1} \leqslant \epsilon \, |t - s|^{\frac{1}{p}}$$

and

$$\big\|\big(\mu_{tu} \circ \mu_{us} - \mu_{ts}\big) - \big(\mu'_{tu} \circ \mu'_{us} - \mu'_{ts}\big)\big\|_{\mathcal{C}^1} \leqslant \epsilon \, |t - s|^a$$

for a positive constant ϵ, with $a > 1$ as in the definition of the $\mathcal{C}^1$-approximate flows μ, μ', for all $0 \leqslant s \leqslant u \leqslant t \leqslant 1$. Prove that one has

$$\big\|\big(\varphi_{ts} - \mu_{ts}\big) - \big(\varphi'_{ts} - \mu'_{ts}\big)\big\|_\infty \leqslant c \epsilon \, |t - s|^a,$$

for all for $0 \leqslant s \leqslant t \leqslant 1$, for some explicit positive constant c.

Rough paths

Guide for this chapter

Hölder p-rough paths, which control the rough differential equations

$$dx_t = \mathrm{F}(x_t)\mathbf{X}(dt), \quad d\varphi_t = \mathrm{F}^{\otimes}\mathbf{X}(dt),$$

and play the role of the control h in the model classical ordinary differential equation

$$dx_t = V_i(x_t)\, dh_t^i = \mathrm{F}(x_t)\, dh_t$$

are defined in section 1.2. As $\mathbb{R}^{\ell}$-valued paths, they are not regular enough for the formula

$$\mu_{ts}(x) = x + X_{ts}^i V_i(x)$$

to define an approximate flow, as in the classical Euler scheme studied in exercice 1. The missing bit of information needed to stabilize the situation is a substitute of the non-existing iterated integrals $\int_s^t X_r^j dX_r^k$, and higher order iterated integrals, which provide a partial description of what happens to X during any time interval (s, t). A (Hölder) p-rough path is a *multi-level object* whose higher order parts provide precisely that information. We saw in the introduction that iterated integrals appear naturally in Taylor-Euler expansions of solutions to ordinary differential equations; they provide higher order numerical schemes like Milstein' second order scheme. It is an important fact that p-rough paths take values in a very special kind of algebraic structure, whose basic features are explained in section 1.1. A Hölder p-rough path will then appear as a kind of $\frac{1}{p}$-Hölder path in that space. We shall then study in section 2 the space of p-rough path for itself.

1. Definition of a Hölder p-rough path

Iterated integrals, as they appear for instance under the form $\int_s^t \int_s^y dh_r^j\, dh_u^k$ or $\int_s^t \int_s^y \int_s^r (\cdots)$, are multi-indexed quantities. A useful formalism to work with such objects is provided by the notion of tensor product. We first start our investigations by recalling some elementary facts about that notion. Eventually, all what will be used for practical computations on rough differential equations will be a product operation very similar to the product operation on polynomials. This abstract setting however greatly clarifies the meaning of these computations.

1.1. An algebraic prelude: tensor algebra over $\mathbb{R}^\ell$ and free nilpotent Lie group. Let first recall what the algebraic tensor product $U \otimes V$ of any two Banach spaces U and V is. Denote by V' the set of all continuous linear forms on V. Given $u \in U$ and $v \in V$, we define a continuous linear map on V' setting

$$(u \otimes v)(v') = (v', v)\, u,$$

for any $v' \in V'$. The algebraic tensor product $U \otimes V$ is the set of all finite linear combinations of such maps. Its elementary elements $u \otimes v$ are 1-dimensional rank maps. Note that an element of $U \otimes V$ can have several different decompositions as a sum of elementary elements; this has no consequences as they all define the same map from V' to U.

As an example, $\mathbb{R}^\ell \otimes (\mathbb{R}^\ell)'$ is the set of all linear maps from $\mathbb{R}^\ell$ to itself, that is $L(\mathbb{R}^\ell)$. We keep that interpretation for $\mathbb{R}^\ell \otimes \mathbb{R}^\ell$, as $\mathbb{R}^\ell$ and $(\mathbb{R}^\ell)'$ are canonically identified. To see which element of $L(\mathbb{R}^\ell)$ corresponds to $u \otimes v$, it suffices to look at the image of the jth vector ϵ_j of the canonical basis by the linear map $u \otimes v$; it gives the jth column of the matrix of $u \otimes v$ in the canonical basis. We have

$$(u \otimes v)(\epsilon_j) = (v, \epsilon_j)\, u.$$

The family $\big(\epsilon_{i_1} \otimes \cdots \otimes \epsilon_{i_k}\big)_{1 \leqslant i_1, \ldots, i_k \leqslant \ell}$ defines the canonical basis of $(\mathbb{R}^\ell)^{\otimes k}$.

- For $N \in \mathbb{N} \cup \{\infty\}$, write $T_\ell^{(N)}$ for the direct sum $\bigoplus_{r=0}^{N} (\mathbb{R}^\ell)^{\otimes r}$, with the convention that $(\mathbb{R}^\ell)^{\otimes 0}$ stands for $\mathbb{R}$. Denote by $\mathbf{a} = \bigoplus_{r=0}^{N} a^r$ and $\mathbf{b} = \bigoplus_{r=0}^{N} b^r$ two generic elements of $T_\ell^{(N)}$. The vector space $T_\ell^{(N)}$ is an algebra for the operations

$$\mathbf{a} + \mathbf{b} = \bigoplus_{r=0}^{N} (a^r + b^r),$$

(1.1)

$$\mathbf{ab} = \bigoplus_{r=0}^{N} c^r, \quad \text{with} \ \ c^r = \sum_{k=0}^{r} a^k \otimes b^{r-k} \in (\mathbb{R}^\ell)^{\otimes r}$$

It is called the **(truncated) tensor algebra of** $\mathbb{R}^\ell$ (if N is finite). Note the similarity between these rules and the analogue rules for addition and product of polynomials.

The exponential map $\exp : T_\ell^{(\infty)} \to T_\ell^{(\infty)}$ and the logarithm map $\log : T_\ell^{(\infty)} \to T_\ell^{(\infty)}$ are defined by the usual series

(1.2)
$$\exp(\mathbf{a}) = \sum_{n \geqslant 0} \frac{\mathbf{a}^n}{n!}, \quad \log(\mathbf{b}) = \sum_{n \geqslant 1} \frac{(-1)^n}{n} (1 - \mathbf{b})^n,$$

with the convention $\mathbf{a}^0 = 1 \in \mathbb{R} \subset T_\ell^{(\infty)}$. Denote by $\pi_N : T_\ell^{(\infty)} \to T_\ell^{(N)}$ the natural projection. We also denote by exp and log the restrictions to $T_\ell^{(N)}$ of the maps $\pi_N \circ \exp$ and $\pi_N \circ \log$ respectively. Denote by $T_\ell^{(N),1}$, resp. $T_\ell^{(N),0}$, the elements $a_0 \oplus \cdots \oplus c_N$ of $T_\ell^{(N)}$ such that $a_0 = 0$, resp. $a_0 = 1$. All the elements of $T_\ell^{(N),1}$ are

invertible, and $\exp : T_\ell^{(N),0} \to T_\ell^{(N),1}$ and $\log : T_\ell^{(N),1} \to T_\ell^{(N),0}$ are smooth reciprocal bijections. As an example, with $\mathbf{a} = 1 \oplus a^1 \oplus a^2 \in T_\ell^{2,1}$, we have

$$\log \mathbf{a} = 0 \oplus (-a^1) \oplus \left(\frac{1}{2} a^1 \otimes a^1 - a^2\right).$$

The set $T_\ell^{(N),1}$ is naturally equipped with a norm defined by the formula

$$(1.3) \qquad \|\mathbf{a}\| := \sum_{i=1}^{\ell} \|a^i\|_{\text{Eucl}}^{\frac{1}{i}},$$

where $\|a^i\|_{\text{Eucl}}$ stands for the Euclidean norm of $a^i \in (\mathbb{R}^\ell)^{\otimes i}$, identified with an element of $\mathbb{R}^{\ell^i}$ by looking at its coordinates in the canonical basis of $(\mathbb{R}^\ell)^{\otimes i}$. The choice of power $\frac{1}{i}$ comes from the fact that $T_\ell^{(N),1}$ is naturally equipped with a *dilation* operation

$$(1.4) \qquad \delta_\lambda(\mathbf{a}) = \left(1, \lambda a^1, \ldots, \lambda^N a^N\right),$$

so the norm $\|\cdot\|$ is homogeneous with respect to this dilation, in the sense that one has

$$\left\|\delta_\lambda(\mathbf{a})\right\| = |\lambda| \|\mathbf{a}\|$$

for all $\lambda \in \mathbb{R}$, and all $\mathbf{a} \in T_\ell^{(N),1}$.

The formula $[\mathbf{a}, \mathbf{b}] = \mathbf{ab} - \mathbf{ba}$, defines a Lie bracket on $T_\ell^{(N)}$. Define inductively $\mathfrak{f} := \mathfrak{f}^1 := \mathbb{R}^\ell$, considered as a subset of $T_\ell^{(\infty)}$, and $\mathfrak{f}^{n+1} = [\mathfrak{f}, \mathfrak{f}^n] \subset T_\ell^{(\infty)}$.

DEFINITION 7.   • *The Lie algebra $\mathfrak{g}_\ell^N$ generated by the $\mathfrak{f}^1, \ldots, \mathfrak{f}^N$ in $T_\ell^{(N)}$ is called the N-step free nilpotent Lie algebra*.
- *As a consequence of Baker-Campbell-Hausdorf-Dynkin formula, the subset $\exp\left(\mathfrak{g}_\ell^N\right)$ of $T_\ell^{(N),1}$ is a group for the multiplication operation. It is called the N-step nilpotent Lie group on $\mathbb{R}^\ell$ and denoted by $\mathfrak{G}_\ell^N$.*

As all finite dimensional Lie groups, the N-step nilpotent Lie group is equipped with a natural (sub-Riemannian) distance inherited from its manifold structure. Its definition rests on the fact that the element $\mathbf{a}u$ of $T_\ell^{(N)}$ is for any $\mathbf{a} \in \mathfrak{G}_\ell^{(N)}$ and $u \in \mathbb{R}^\ell \subset T_\ell^{(N)}$ a tangent vector to $\mathfrak{G}_\ell^{(N)}$ at point $\mathbf{a}$ (as u is tangent to $\mathfrak{G}_\ell^{(N)}$ at the identity and tangent vectors are transported by left translation in the group). So the ordinary differential equation

$$d\mathbf{a}_t = \mathbf{a}_t \, \dot{h}_t$$

makes sense for any $\mathbb{R}^\ell$-valued smooth control h, and defines a path in $\mathfrak{G}_\ell^{(N)}$ started from the identity. We define the size $|\mathbf{a}|$ of $\mathbf{a}$ by the formula

$$|\mathbf{a}| = \inf \int_0^1 |\dot{h}_t| \, dt,$$

where the infimum is over the set of all piecewise smooth controls h such that $\mathbf{a}_1 = \mathbf{a}$. This set is non-empty as $\mathbf{a} \in \exp\left(\mathfrak{g}^N\right)$ can be written as $\mathbf{a}_1$ for some piecewise $\mathcal{C}^1$ control, as a consequence of a theorem of sub-Riemannian geometry due to Chow; see for instance the textbook [11] for a nice account of that theorem. The distance

between any two points $\mathbf{a}$ and $\mathbf{b}$ of $\mathfrak{G}_\ell^{(N)}$ is then defined as $\left|\mathbf{a}^{-1}\mathbf{b}\right|$. It is homogeneous in the sense that if $\mathbf{a} = \exp(u)$, with $u \in \mathbb{R}^\ell \subset T_\ell^{(N)}$, then $\left|\exp(\lambda u)\right| = |\lambda||\mathbf{a}|$, for all $\lambda \in \mathbb{R}$ and all $u \in \mathbb{R}^\ell \subset T_\ell^{(N)}$.

This way of defining a distance is intrinsic to $\mathfrak{G}_\ell^{(N)}$ and classical in geometry. From an extrinsic point of view, one can also consider $\mathfrak{G}_\ell^{(N)}$ as a subset of $T_\ell^{(N)}$ and use the ambient metric to define the distance between any two points $\mathbf{a}$ and $\mathbf{b}$ of $\mathfrak{G}_\ell^{(N)}$ as $\left\|\mathbf{a}^{-1}\mathbf{b}\right\|$. It can be proved (this is elementary, see e.g. proposition 10 in Appendix A of [**12**], pp. 76-77) that the two norms $|\cdot|$ and $\|\cdot\|$ on $\mathfrak{G}_\ell^{(N)}$ are equivalent, so one can equivalently work with one or the other, depending on the context. This will be useful in defining the Brownian rough path for example.

1.2. Definition of a Hölder p-rough path. The relevance of the algebraic framework provided by the N-step nilpotent Lie group for the study of smooth paths was first noted by Chen in his seminal work [**13**]. Indeed, for any $\mathbb{R}^\ell$-valued smooth path $(x_s)_{s \geqslant 0}$, the family of iterated integrals

$$\mathfrak{X}_{ts}^N := \left(1, x_t - x_s, \int_s^t \int_s^{s_1} dx_{s_2} \otimes dx_{s_1}, \ldots, \int_{s \leqslant s_1 \leqslant \cdots \leqslant s_N \leqslant t} dx_{s_1} \otimes \cdots \otimes dx_{s_N}\right)$$

defines for all $0 \leqslant s \leqslant t$ an element of $T_\ell^{(N),1}$ with the property that if $x_\bullet$ is scaled into $\lambda x_\bullet$ then $\mathfrak{X}^N$ becomes $\delta_\lambda \mathfrak{X}^N$. We actually have $\mathfrak{X}_{ts}^N \in \mathfrak{G}_\ell^{(N)}$. To see that, notice that, as a function of t, the function $\mathfrak{X}_{ts}^N$ satisfies the differential equation

$$d\mathfrak{X}_{ts}^N = \mathfrak{X}_{ts}^N \, dx_t,$$

in $T_\ell^{(N)}$ driven by the $\mathbb{R}^\ell$-valued smooth control x, so it defines a $\mathfrak{G}_\ell^{(N)}$-valued path as an integral curve of a field of tangent vectors. The above differential equation also makes it clear the we have the following *Chen relations*

$$\mathfrak{X}_{ts}^N = \mathfrak{X}_{us}^N \, \mathfrak{X}_{tu}^N,$$

for all $0 \leqslant s \leqslant u \leqslant t$, which is nothing but the "flow" property for ordinary differential equation solutions; they imply in particular the identity

$$\mathfrak{X}_{ts}^N = \left(\mathfrak{X}_{s0}^N\right)^{-1} \mathfrak{X}_{t0}^N,$$

Rough paths and weak geometric rough paths are somehow an abstract version of this family of iterated integrals.

DEFINITION 8. *Let $2 \leqslant p$. A* **Hölder p-rough path** *on $[0, T]$ is a $T_\ell^{([p]),1}$-valued path* $\mathbf{X} : t \in [0, T] \mapsto 1 \oplus X_t^1 \oplus X_t^2 \oplus \cdots \oplus X_t^{[p]}$ *such that*

$$(1.5) \qquad\qquad \left\|X^i\right\|_{\frac{i}{p}} := \sup_{0 \leqslant s < t \leqslant T} \frac{|X_{ts}^i|}{|t - s|^{\frac{i}{p}}} < \infty,$$

for all $i = 1 \ldots [p]$, where we set $\mathbf{X}_{ts} := \mathbf{X}_s^{-1}\mathbf{X}_t$. We define the norm of $\mathbf{X}$ to be

$$(1.6) \qquad\qquad \|\mathbf{X}\| := \max_{i=1\ldots[p]} \left\|X^i\right\|_{\frac{i}{p}},$$

and a distance $d(\mathbf{X}, \mathbf{Y}) = \|\mathbf{X} - \mathbf{Y}\|$ on the set of Hölder p-rough path. A **Hölder weak geometric p-rough path on** $[0, T]$ *is a* $\mathfrak{G}_\ell^{[p]}$*-valued p-rough path.*

So a (weak geometric) Hölder p-rough path is in a way nothing but a ($\mathfrak{G}_\ell^{(N)}$ or) $T_\ell^{(N),1}$-valued $\frac{1}{p}$-Hölder continuous path, for the $\|\cdot\|$-norm introduced above and the use of $\mathbf{X}_s^{-1}\mathbf{X}_t$ in place of the usual $\mathbf{X}_t - \mathbf{X}_s$. Note that the Chen relation

$$\mathbf{X}_{ts} = \mathbf{X}_{us}\mathbf{X}_{tu}$$

is granted by the definition of $\mathbf{X}_{ts} = \mathbf{X}_s^{-1}\mathbf{X}_t$.

For $2 \leqslant p < 3$, Chen's relation is equivalent to

(i) $X_{ts}^1 = X_{tu}^1 + X_{us}^1$,
(ii) $X_{ts}^2 = X_{tu}^2 + X_{us}^1 \otimes X_{tu}^1 + X_{us}^2$.

Condition **(i)** means that $X_{ts}^1 = X_{t0}^1 - X_{s0}^1$ represents the increment of the $\mathbb{R}^\ell$-valued path $\left(X_{r0}^1\right)_{0 \leqslant r \leqslant T}$. Condition **(ii)** is nothing but the analogue of the elementary property $\int_s^t \int_s^r = \int_s^u \int_s^r + \int_u^t \int_s^u + \int_u^t \int_u^r$, satisfied by any reasonable notion of integral on $\mathbb{R}$ that satisfies the Chasles relation

$$\int_s^t = \int_s^u + \int_u^t.$$

This remark justifies thinking of the $\left(\mathbb{R}^\ell \otimes \mathbb{R}^\ell\right)$-part of a rough path as a kind of iterated integral of X^1 against itself, although this hypothetical iterated integral does not make sense in itself for lack of an integration operation for a general Hölder path in $\mathbb{R}^\ell$. In that setting, a p-rough path $\mathbf{X}$ is a weak geometric p-rough path iff the symmetric part of X_{ts}^2 is $\frac{1}{2}X_{ts}^1 \otimes X_{ts}^1$, for all $0 \leqslant s \leqslant t \leqslant T$.

Note that the space of Hölder p-rough paths is not a vector space; this prevents the use of the classical Banach space calculus.

It is clear that considering the iterated integrals of any given smooth path defines a Hölder p-rough path above it, for any $p \geqslant 2$. This *lift is not unique*, as if we are given a Hölder p-rough path $\mathbf{X} = \left(X^1, X^2\right)$, with $2 \leqslant p < 3$ say, and any $\frac{2}{p}$-Hölder continuous $(\mathbb{R}^\ell)^{\otimes 2}$-valued path $\left(M_t\right)_{0 \leqslant t \leqslant 1}$, we define a new rough path setting $M_{ts}^2 = M_t - M_s$, and

$$\mathbf{X}_{ts}' = \left(X_{ts}^1, X_{ts}^2 + M_{ts}^2\right)$$

for all $0 \leqslant s \leqslant t \leqslant 1$. Relations **(i)** and **(ii)** above are indeed easily checked.

Last, note that a Hölder p-rough path is also a Hölder q-rough path for any $p < q < [p] + 1$.

2. The metric space of Hölder p-rough paths

The distance d defined in definition 8 is actually not a distance since only the increments $\mathbf{X}_{ts} - \mathbf{Y}_{\mathbf{ts}}$ are taken into account. We define a proper metric on the set of all Hölder p-rough paths setting

$$\overline{d}(\mathbf{X}, \mathbf{Y}) = \left|X_0^1 - Y_0^1\right| + d(\mathbf{X}, \mathbf{Y}).$$

PROPOSITION 9. *The metric $\overline{d}$ turns the set of all Hölder p-rough paths into a (non-separable) complete metric space.*

PROOF – Given a Cauchy sequence of Hölder p-rough paths $^{(n)}\mathbf{X}$, there is no loss of generality in supposing that their first level starts from the same point in $\mathbb{R}^\ell$. It follows from the uniform Hölder bounds for $\left\| ^{(n)}X_{ts}^i - ^{(m)}X_{ts}^i \right\|_{\frac{i}{p}}$, and (an easily proved version of) Ascoli-Arzela theorem (for 2-parameter maps) that $^{(n)}X$ converges uniformly to some Hölder p-rough path $\mathbf{X}$. To prove the convergence of $^{(n)}X$ to $\mathbf{X}$ in d-distance, it suffices to send m to infinity in the inequality

$$\left| ^{(n)}X_{ts}^i - ^{(m)}X_{ts}^i \right| \leqslant \epsilon \, |t-s|^{\frac{i}{p}},$$

which holds for all n, m bigger than some N_ϵ, uniformly with respect to $0 \leqslant s \leqslant t \leqslant 1$.

An uncountable family of $\mathbb{R}^\ell$-valued $\frac{1}{p}$-Hölder continuous functions at pairwise $\frac{1}{p}$-Hölder distance bounded below by a positive constant is constructed in example 5.28 of [**3**]. As the set of all first levels of the set of Hölder p-rough paths is a subset of the set of $\mathbb{R}^\ell$-valued $\frac{1}{p}$-Hölder paths, this examples implies the non-separability of set of all Hölder p-rough paths. $\triangleright$

The following *interpolation result* will be useful in several places to prove rough paths convergence results at a cheap price. It roughly says that uniform bounds in a strong sense together with a convergence property in a weak sense are sufficient to prove a convergence result in a mild sense.

PROPOSITION 10. *Assume $^{(n)}\mathbf{X}$ is a sequence of Hölder p-rough paths with uniform bounds*

$$(2.1) \qquad\qquad \sup_n \left\| ^{(n)}X \right\| \leqslant C < \infty,$$

which converge pointwise, in the sense that $^{(n)}\mathbf{X}_{ts}$ converges to some $\mathbf{X}_{ts}$ for each $0 \leqslant s \leqslant t \leqslant 1$. Then the limit object $\mathbf{X}$ is a Hölder p-rough path, and $^{(n)}\mathbf{X}$ converges to $\mathbf{X}$ as a Hölder q-rough path, for any $p < q < [p] + 1$.

PROOF – (Following the solution of exercice 2.9 in [**5**]) The fact that $\mathbf{X}$ is a Hölder p-rough path is a direct consequence of the uniform bounds (2.1) and pointwise convergence:

$$\left| X_{ts}^i \right| = \lim_n \left| ^{(n)}X_{ts}^i \right| \leqslant C |t-s|^{\frac{i}{p}}.$$

Would the convergence of $^{(n)}\mathbf{X}$ to $\mathbf{X}$ be uniform, we could find a sequence ϵ_n decreasing to 0, such that, uniformly in s, t,

$$\left| X_{ts}^i - ^{(n)}X_{ts}^i \right| \leqslant \epsilon_n, \quad \left| X_{ts}^i - ^{(n)}X_{ts}^i \right| \leqslant 2C|t-s|^{\frac{i}{p}}.$$

Using the geometric interpolation $a \wedge b \leqslant a^{1-\theta} b^\theta$, with $\theta = \frac{p}{q} < 1$, we would have

$$\left| X_{ts}^i - ^{(n)}X_{ts}^i \right| \leqslant \epsilon_n^{1-\frac{p}{q}} |t-s|^{\frac{i}{q}},$$

which entails the convergence result as a Hölder q-rough path.

We proceed as follow to see that pointwise convergence suffices to get the result. Given a partition π of $[0,1]$ and any $0 \leqslant s \leqslant t \leqslant 1$, denote by $\overline{s}, \overline{t}$ the nearest points in π to s and t respectively. Writing

$$(2.2) \qquad d\Big(\mathbf{X}_{ts},\, {}^{(n)}\mathbf{X}_{ts}\Big) \leqslant d\big(\mathbf{X}_{ts}, \mathbf{X}_{\overline{t}\overline{s}}\big) + d\Big(\mathbf{X}_{\overline{t}\overline{s}},\, {}^{(n)}\mathbf{X}_{\overline{t}\overline{s}}\Big) + d\Big({}^{(n)}\mathbf{X}_{\overline{t}\overline{s}},\, {}^{(n)}\mathbf{X}_{ts}\Big)$$

and the fact that

$$\mathbf{X}_{\overline{t}\overline{s}} = \mathbf{X}_{s\overline{s}}\mathbf{X}_{ts}\mathbf{X}_{\overline{t}t}, \qquad {}^{(n)}\mathbf{X}_{\overline{t}\overline{s}} = {}^{(n)}\mathbf{X}_{s\overline{s}}{}^{(n)}\mathbf{X}_{ts}{}^{(n)}\mathbf{X}_{\overline{t}t}$$

and the uniform estimate (2.1) to see that the first and third terms in the above upper bound can be made arbitrarily small by choosing a partition with a small enough mesh, uniformly in s, t and n. The second term is dealt with the pointwise convergence assumption as it involves only finitely many points once the partition π has been chosen as above. $\qquad \triangleright$

3. Controlled paths and rough integral

It will be the set of Hölder weak geometric p-rough paths that will play the main role in the sequel, as a set of driving signals in rough differential equations. Unlike the space of Hölder p-rough paths, this set is not a linear space, nor even a(n infinite dimensional) manifold, simply a metric space for which none of the classical tools of Banach space calculus can be applied in a straightforward way. It is fortunate, however, that Gubinelli developped in [9] some intermediate spaces of rough paths which have some Banach space structure built in. Their definition requires the introduction of the notion of controlled path, which will appear as the good notion of integrand in the definition of a rough integral. The formalism of this section will be used in part IV of the course on stochastic analysis, where we shall recast the theory of rough differential equations developed in part III of the course in therms of Taylor-Euler expansion properties, using the rough integral introduced in this section. A reference Hölder p-rough path $\mathbf{X}$ is fixed throughout this section, with $2 < p < 3$.

DEFINITION 11. *An $\mathbb{R}^d$-valued path $z_\bullet$ is said to be a **path controlled by $\mathbf{X}$** if its increments $Z_{ts} = z_t - z_s$, satisfy*

$$Z_{ts} := Z'_s X_{ts} + R_{ts},$$

for all $0 \leqslant s \leqslant t \leqslant 1$, for some $\mathrm{L}\big(\mathbb{R}^\ell, \mathbb{R}^d\big)$-valued $\frac{1}{p}$-Lipschitz map $Z'_\bullet$, and some $\mathbb{R}^d$-valued $\frac{2}{p}$-Lipschitz map R.

The following example shows that a controlled path $z_\bullet$ may have several derivatives $Z'_\bullet$. Choose a Hölder p-rough path $\mathbf{X}$ with $X_{ts} = (t-s)v$, for some fixed vector $v \in \mathbb{R}^\ell$. For any path $z_\bullet$ controlled by $\mathbf{X}$, one can write

$$Z_{ts} = Z'_s(t-s)v + R^{Z'}_{ts}$$

for *any* choice of $\frac{1}{p}$-Hölder function Z' as the term $Z'_s(t-s)v$ can always be inserted in the remainder. So, strictly speaking, a controlled path is a *pair* (z, Z') with the above properties. We sometimes abuse notations and talk of the controlled path $z_\bullet$.

Using the notation $\|\cdot\|_\alpha$ to denote the α-Hölder norm of a 1 or 2-indices map, it is straightforward to see that one defines a complete metric on the set of $\mathbb{R}^d$-valued paths (z, Z') controlled by $\mathbf{X}$, together with their derivative, setting

$$\left\|(z, Z')\right\| := \|Z'\|_{\frac{1}{p}} + \|R\|_{\frac{2}{p}} + |z_0|,$$

where $\|\cdot\|_\alpha$ stands for the α-Lipschitz norm. It is elementary to see that the image of a controlled path $z_\bullet$ by an $\mathbb{R}^n$-valued $\mathcal{C}^1$ map F on $\mathbb{R}^d$ is a controlled path $F(z_\bullet)$ with derivative $D_{z_t}F \circ Z'_t$.

The definition of a controlled path involves only the first level of the rough path $\mathbf{X}$. The reference to $\mathbf{X}$ itself comes from the following crucial property of controlled paths: they admit an natural lift into a Hölder p-rough path, whose definition involves all of $\mathbf{X}$. Given two linear maps $A, B \in \mathrm{L}(\mathbb{R}^\ell, \mathbb{R}^d)$, and any $a, b \in \mathbb{R}^\ell$, we set

$$\big(A \otimes B\big)(a \otimes b) := (Aa) \otimes (Bb).$$

PROPOSITION 12. *Let (z, Z') be an $\mathbb{R}^d$-valued path controlled by $\mathbf{X}$. We define an almost-additive map setting*

$$\mu_{ts} := z_s \otimes Z_{ts} + Z'_s \otimes Z'_s \mathbb{X}_{ts},$$

for all $0 \leqslant s \leqslant t \leqslant 1$. Its associated additive map φ_{ts} is denoted by

$$\mathbb{Z}_{ts} =: \int_s^t z_{us} \otimes dz_u.$$

The pair $(z, \mathbb{Z})$ is a Hölder p-rough path.

PROOF – An elementary computation using Chen's relation $\mathbb{X}_{ts} = \mathbb{X}_{tu} + \mathbb{X}_{us} + X_{us} \otimes X_{tu}$, for any $0 \leqslant s \leqslant u \leqslant t \leqslant 1$, gives

$$\big(\mu_{tu} + \mu_{us}\big) - \mu_{ts} = Z_{us} \otimes Z_{tu} + \big(Z'_u \otimes Z'_u - Z'_s \otimes Z'_s\big)\mathbb{X}_{tu} - \big(Z'_s \otimes Z'_s\big)X_{us} \otimes X_{tu}$$

$$= Z_{us} \otimes \big(Z_{tu} - Z'_s X_{tu}\big) + O\big(|t - s|^{\frac{3}{p}}\big)$$

$$= Z_{us} \otimes \big(\big(Z'_u - Z'_s\big) \otimes X_{tu}\big) + O\big(|t - s|^{\frac{3}{p}}\big) = O\big(|t - s|^{\frac{3}{p}}\big).$$

The $\frac{2}{p}$-Hölder character of $\mathbb{Z}_{ts}$ is immediate from the identity

$$\mathbb{Z}_{ts} = \mu_{ts} + O\big(|t - s|^{\frac{3}{p}}\big),$$

while Chen's relations are straightforward to check.   ▷

Let $\big(\mathrm{F}_t\big)_{0 \leqslant s \leqslant t \leqslant 1}$ be an $\mathrm{L}(\mathbb{R}^\ell, \mathbb{R}^n)$-valued path controlled by $\mathbf{X}$. The same computation as above shows that we define an almost-additive map by the formula

$$\mathrm{F}_s X_{ts} + \mathrm{F}'_s \mathbb{X}_{ts};$$

it associated additive map is denoted by

$$\int_s^t \mathrm{F} d\mathbf{X},$$

and called the **rough integral of F with respect to $\mathbf{X}$**.

There exists, for any $p \geqslant 3$, a notion of path controlled by a Hölder p-rough path. However, the good algebraic setting to work with these objects is not the tensor algebra introduced in this part of the course, but a Hopf algebra of labelled trees. Rough paths are replaced in that setting by *branched rough paths*. This somewhat heavier algebraic setting makes the use of branched rough paths not so convenient. Fortunately, we shall only need the results contained in this section to investigate stochastic differential equations driven by Brownian motion in part IV of the course. See Gubinelli's original work [16] on the subject, or the nice account [15] given by Hairer and Kelly to get some more insights on this question.

4. Exercices on rough paths

Exercice 7 presents a fundamental result of Lyons of primary importance in the original formulation of the theory. It essentially means that a p-rough path has a unique extension into a q-rough path, for any $q \geqslant [p] + 1$. The extension of a rough path to all higher degrees defines an object called the *signature* of the rough path, whose importance for real life data analysis is actively investigated presently. Exercices 8 and 9 emphasize the fact that rough paths naturally appear in highly oscillating systems as a class of controllers (this fact will appear clearly after reading Part III of the course). Exercises 10 an 11 deal with the question of lifting a path or a pair of rough paths into a single rough path.

7. Lyons' extension theorem [20]. Let n be a positive integer. A $T_\ell^{n,1}$-valued map $\mathbf{X} = (\mathbf{X}_{ts})_{0 \leqslant s \leqslant t \leqslant 1}$ is said to be *multiplicative* if we have

$$\mathbf{X}_{ts} = \mathbf{X}_{us}\mathbf{X}_{tu}$$

for all $0 \leqslant s \leqslant u \leqslant t \leqslant 1$. It is said to be *almost-multiplicative* if we have

$$\left| X_{ts}^k - \left(X_{us}X_{tu}\right)^k \right| \leqslant c\,|t-s|^{ka}$$

for all $0 \leqslant s \leqslant u \leqslant t \leqslant 1$ and $0 \leqslant k \leqslant n$, for some positive constants c and $a > 1$; the notation X^k stands here for the $(\mathbb{R}^\ell)^{\otimes k}$-component of an element X of $T_\ell^{n,1}$. Prove that if $\mathbf{X}$ is a $T_\ell^{(n),1}$-valued multiplicative map and Y_{ts}^{n+1} is a continuous $(\mathbb{R}^d)^{\otimes(n+1)}$-valued map such that the $T_\ell^{n+1,1}$-valued map

$$Y := \left(1, X^1, \ldots, X^n, Y^{n+1}\right)$$

is almost-multiplicative, then there exists a unique $(\mathbb{R}^d)^{\otimes(n+1)}$-valued map X_{ts}^{n+1} with

$$\left| X_{ts}^{n+1} - Y_{ts}^{n+1} \right| \leqslant c_1\,|t-s|^{(n+1)a}$$

for some positive constant c_1, such that

$$Z := \left(1, X^1, \ldots, X^n, X^{n+1}\right)$$

is a $T_\ell^{n+1,1}$-valued multiplicative map.

Starting from a Hölder p-rough path $\mathbf{X}$ and $Y^{[p]+1} = 0$, one can apply iteratively the above procedure to extend uniquely $\mathbf{X}$ into a Hölder q-rough path, for any $q \geqslant [p] + 1$, in a consistent way. This provides a $T_\ell^{\infty,1}$-valued extension of $\mathbf{X}$ called its **signature**.

8. Pure area rough path. Let x^n be the $\mathbb{R}^2$-valued path defined in complex notations by the formula

$$x_t^n = \frac{1}{n} \exp\left(2i\pi n^2 t\right),$$

for $0 \leqslant t \leqslant 1$. Let $2 < p < 3$ be given.

    a) Show that the natural lift $\mathbf{X}^n = \left(x^n, \mathbb{X}^n\right)$ of x^n to a Hölder p-rough path converges pointwise to the Hölder p-rough path $\mathbf{X} = (X, \mathbb{X})$ with $X = 0$ and

$$\mathbb{X}_{ts} = \pi\,(t - s)\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}.$$

    b) Prove the uniform bounds $\sup_n \left\| x^n \right\|_{\frac{1}{2}} < \infty$ and $\sup_n \left\| \mathbb{X}^n \right\|_1 < \infty$.

    c) Conclude by interpolation that the convergence of $\mathbf{X}^n$ to $\mathbf{X}$ takes place in the space of Hölder p-rough paths.

9. Wild oscillations. Find a widely oscillating piecewise smooth path converging to $(0, 0, t\mathrm{I})$ in the space of Hölder p-rough paths, for $3 < p < 4$. The letter I stands here for the element of $(\mathbb{R}^\ell)^{\otimes 3}$ given in the canonical basis by $\mathrm{I}_{ijk} = \delta_{ij}\delta_{jk}$.

10. Lifting α-Hölder paths to rough paths, for $\alpha > \frac{1}{2}$. Show that using the Young integral defined in exercice 5 one can lift any α-Hölder paths, with $\alpha > \frac{1}{2}$, into a Hölder p-rough path, for any $p \geqslant 2$.

11. Pairing two rough paths. The problem we address in this exercice is the following. *Given two rough paths defined on some (different) spaces, are these two rough paths "pieces" of a higher dimensional rough path?* This is a non-trivial question, when formulated in this generality, due to the fact that there is no canonical way of constructing the cross-iterated integrals between the two rough paths. However, this question has a simple answer when one of the two rough paths is actually a sufficiently regular Hölder path (this exercice), or when one can use probabilistic arguments to construct the missing iterated integrals (exercice 19).

    Let $\mathbf{X}$ be a Hölder p-rough path over $\mathbb{R}^\ell$, with $2 < p < 3$, and h be a $\frac{1}{q}$-Hölder $\mathbb{R}^d$-valued path, with $\frac{1}{p} + \frac{1}{q} > 1$; so in particular $\frac{1}{q} > \frac{1}{2}$. We describe an element of $(\mathbb{R}^\ell \times \mathbb{R}^d)^{\otimes 2}$ as a 2×2 matrix $\begin{pmatrix} A & C \\ B & D \end{pmatrix}$, with A of size $\ell \times \ell$, B of size $\ell \times d$, C of size $d \times \ell$ and D of size $d \times d$. Show that one defines a rough path $\mathbf{Z} = (Z, \mathbb{Z})$ over $\mathbb{R}^\ell \times \mathbb{R}^d$ setting

$$Z_{ts} = \left(X_{ts}, h_t - h_s\right)$$

and

$$\mathbb{Z}_{ts} = \begin{pmatrix} A_{ts} & C_{ts} \\ B_{ts} & D_{ts} \end{pmatrix},$$

with $A_{ts} = \mathbb{X}_{ts}$, $D_{ts} = \int_s^t h_{us} \otimes dh_u$ and $B_{ts} = \int_s^t h_{us} \otimes dX_u$, $C_{ts} = \int_s^t X_{us} \otimes dh_u$, where D_{ts} and C_{ts} are Young integrals and the integral B_{ts} is *defined* by the integration by parts formula

$$\int_s^t h_{us} \otimes dX_u := h_{ts} \otimes X_{ts} - \int_s^t dh_u \otimes X_{us}.$$

We say that $\mathbf{Z}$ is a **pairing of X and** h. Would could you possibly pair them if $\mathbf{X}$ were a Hölder p-rough path over $\mathbb{R}^\ell$, without any rstriction on $p > 2$, and the condition $\frac{1}{p} + \frac{1}{q} > 1$ still holds?

<div align="center">

CHAPTER 4

Flows driven by rough paths

</div>

<div style="border:1px solid black; border-radius:8px; padding:4px;">

Guide for this chapter

</div>

We have seen in part I of the course that a $\mathcal{C}^1$-approximate flow on a Banach space E defines a unique flow $\varphi = \left(\varphi_{ts}\right)_{0\leqslant s\leqslant t\leqslant 1}$ on E such that the inequality

$$(0.1) \qquad \left\| \varphi_{ts} - \mu_{ts} \right\|_\infty \leqslant c|t-s|^a$$

holds for some positive constants c and $a > 1$, for all $0 \leqslant s \leqslant t \leqslant T$ sufficiently close. The construction of φ is actually quite explicit, for if we denote by $\mu_{\pi_{ts}}$ the composition of the maps $\mu_{t_{i+1}t_i}$ along the times t_i of a partition π_{ts} of an interval $[s,t]$, the map μ_{ts} satisfies the estimate

$$(0.2) \qquad \left\| \varphi_{ts} - \mu_{\pi_{ts}} \right\|_\infty \leqslant \frac{2}{1 - 2^{1-a}} \, c_1^2 \, T \, \left| \pi_{ts} \right|^{a-1},$$

where c_1 is the constant that appears in the definition of a $\mathcal{C}^1$-approximate flow

$$(0.3) \qquad \left\| \mu_{tu} \circ \mu_{us} - \mu_{ts} \right\|_{\mathcal{C}^1} \leqslant c_1|t-s|^a.$$

It follows in particular from equation (0.1) that if μ depends continuously on some metric space-valued parameter λ, with respect to the $\mathcal{C}^0$-topology, and if identity (0.3) holds uniformly for λ moving in a bounded set say, then φ depends continuously on λ, as a uniform limit of continuous functions.

The point about the machinery of $\mathcal{C}^1$-approximate flows is that they actually pop up naturally in a number of situations, under the form of a local in time description of the dynamics under study; nothing else than a kind of Taylor expansion. This was quite clear in exercice 1 on the ordinary controlled differential equation

$$(0.4) \qquad dx_t = V_i(x_t)\, dh_t^i,$$

with $\mathcal{C}^1$ real-valued controls $h^1, \ldots, h^\ell$ and $\mathcal{C}_b^2$ vector fields $V_1, \ldots, V_\ell$ in $\mathbb{R}^d$. The 1-step Euler scheme

$$\mu_{ts}(x) = x + \left(h_t^i - h_s^i\right)V_i(x)$$

defines in that case a $\mathcal{C}^1$-approximate flow which has the awaited Taylor-type expansion, in the sense that one has

$$(0.5) \qquad f\left(\mu_{ts}(x)\right) = f(x) + \left(h_t^i - h_s^i\right)\left(V_i f\right)(x) + O\left(|t-s|^{>1}\right)$$

for any function f of class $\mathcal{C}_b^2$; but μ fails to be a flow. Its associated flow is not only a flow, it also satisfies equation (0.5) as a consequence of identity (0.1).

<div align="center">

29

</div>

We shall proceed in a very similar way to give some meaning and solve the rough differential equation on flows

$$(0.6) \qquad\qquad d\varphi = V\,dt + \mathrm{F}^{\otimes}\mathbf{X}(dt),$$

where V is a Lipschitz continuous vector field on E and $\mathrm{F} = \big(V_1, \ldots, V_\ell\big)$ is a collection of sufficiently regular vector fields on E, and $\mathbf{X}$ is a Hölder p-rough path over $\mathbb{R}^\ell$. A **solution flow** to equation (0.6) will be defined as a flow on E with a **uniform Taylor-Euler expansion** of the form

$$(0.7) \qquad f\big(\varphi_{ts}(x)\big) = f(x) + \sum_{|I| \leqslant [p]} X_{ts}^I \big(V_I f\big)(x) + O\big(|t - s|^{>1}\big),$$

where $I = \big(i_1, \ldots, i_k\big) \in [\![1, \ell]\!]^k$ is a multi-index with size $k \leqslant [p]$, and X_{ts}^I stands for the coordinates of $\mathbf{X}_{ts}$ in the canonical basis of $T_\ell^{[p],1}$. The vector field V_i is seen here as a 1st-order differential operator, and $V_I = V_{i_1} \cdots V_{i_k}$ as the kth-order differential operator obtained by applying successively the operators V_{i_n}.

For $V = 0$ and $\mathbf{X}$ the (weak geometric) p-rough path canonically associated with an $\mathbb{R}^\ell$-valued $\mathcal{C}^1$ control h, with $2 \leqslant p < 3$, equation (0.7) becomes
$$(0.8)$$
$$f\big(\varphi_{ts}(x)\big) = f(x) + \big(h_t^i - h_s^i\big)\big(V_i f\big)(x) + \left(\int_s^t \int_s^r dh_u^j\, dh_r^k\right)\big(V_j V_k f\big)(x) + O\big(|t - s|^{>1}\big),$$

which is nothing else than Taylor formula at order 2 for the solution to the ordinary differential equation (0.4) started at x at time s. Condition (0.7) is a natural analogue of (0.8) and its higher order analogues.

There is actually a simple way of constructing a map μ_{ts} which satisfies the Euler expansion (0.7). It can be defined as the time 1 map associated with an ordinary differential equation constructed form the V_i and their brackets, and where $\mathbf{X}_{ts}$ appears as a parameter under the form of its logarithm. That these maps μ_{ts} form a $\mathcal{C}^1$-approximate flow will eventually appear as a consequence of the fact that the time 1 map of a differential equation formally behaves as an exponential map, in some algebraic sense.

The notationally simpler case of flows driven by weak geometric Hölder p-rough paths, with $2 \leqslant p < 3$, is first studied in section 1 before studying the general case in section 2. The latter case does not present any additional conceptual difficulty, so a reader which who would like to get the core ideas can read section 1 only, or directly go to section 2. The two sections have been written with almost similar words on purpose.

1. Warm up: working with weak geometric Hölder p-rough paths, with $2 \leqslant p < 3$

Let V be a $\mathcal{C}_b^2$ vector field on E and $V_1, \ldots, V_\ell$ be $\mathcal{C}_b^3$ vector fields on E. Let $\mathbf{X} = (X, \mathbb{X})$ be a Hölder weak geometric p-rough path over $\mathbb{R}^\ell$, with $2 \leqslant p < 3$. Let

μ_{ts} be the well-defined time 1 map associated with the ordinary differential equation

$$(1.1) \qquad \dot{y}_u = (t-s)V(y_u) + \left(X_{ts}^i V_i + \frac{1}{2} \mathbb{X}_{ts}^{jk} [V_j, V_k] \right)(y_u), \quad 0 \leqslant u \leqslant 1;$$

it associates to any $x \in E$ the value at time 1 of the solution of the above equation started from x; it is well-defined since V and the V_i are in particular globally Lipschitz. It is a direct consequence of classical results on ordinary differential equations, and of the definition of the topology on the space of Hölder weak geometric p-rough paths, that the maps μ_{ts} depend continuously on $((s,t), \mathbf{X})$ in the uniform topology, and that

$$(1.2) \qquad \left\| \mu_{ts} - \mathrm{Id} \right\|_{\mathcal{C}^2} = o_{t-s}(1).$$

Also, considering y_u as a function of x, it is elementary to see that one has the estimate

$$(1.3) \qquad \left\| y_u - \mathrm{Id} \right\|_{\mathcal{C}^1} \leqslant c \left(1 + \|\mathbf{X}\|^3 \right) |t-s|^{1/p}, \quad 0 \leqslant u \leqslant 1,$$

for some constant depending only on V and the V_i.

1.1. From Taylor expansions to flows driven by rough paths.

The next proposition shows that μ_{ts} has precisely the kind of Taylor-Euler expansion property that we expect from a solution to a rough differential equation, as described in the introduction to that part of the course.

PROPOSITION 13. *There exists a positive constant c, depending only on V and the V_i, such that the inequality*
(1.4)
$$\left\| f \circ \mu_{ts} - \left\{ f + (t-s)Vf + X_{ts}^i (V_i f) + \mathbb{X}_{ts}^{jk} (V_j V_k f) \right\} \right\|_\infty \leqslant c \left(1 + \|\mathbf{X}\|^3 \right) \|f\|_{\mathcal{C}^3} |t-s|^{\frac{3}{p}}$$

holds for any $f \in \mathcal{C}_b^3$.

The proof of this proposition and the following one are based on the following elementary identity, obtained by applying twice the identity

$$f(y_r) = f(x) + (t-s) \int_0^r (Vf)(y_u)\, du + X_{ts}^i \int_0^r (V_i f)(y_u)\, du + \frac{1}{2} \mathbb{X}_{ts}^{jk} \int_0^r \left([V_j, V_k] f \right)(y_u)\, du,$$

first to f, then to $Vf, V_i f$ and $[V_j, V_k] f$ inside the integrals. One has

$$
f\big(\mu_{ts}(x)\big) = f(x) + (t-s)\int_0^1 (Vf)(y_u) du + X_{ts}^i \int_0^1 (V_i f)(y_{s_1}) ds_1 + \frac{1}{2}\mathbb{X}_{ts}^{jk}\int_0^1 \big([V_j, V_k] f\big)(y_u) du
$$

$$
= f(x) + (t-s)(Vf)(x) + (t-s)\int_0^1 \{(Vf)(y_u) - (Vf)(x)\} du
$$

$$
+ X_{ts}^i (V_i f)(x) + (t-s)X_{ts}^i \int_0^1\int_0^{s_1} (V V_i f)(y_{s_2})\, ds_2 ds_1
$$

$$
+ \frac{1}{2}X_{ts}^{i'} X_{ts}^i (V_{i'} V_i f)(x) + X_{ts}^i X_{ts}^{i'} \int_0^1\int_0^{s_1} \{(V_{i'} V_i f)(y_{s_2}) - (V_{i'} V_i f)(x)\}\, ds_2 ds_1
$$

$$
+ \frac{1}{2}X_{ts}^i \mathbb{X}_{ts}^{jk}\int_0^1\int_0^{s_1} \big([V_j, V_k] V_i f\big)(y_{s_2})\, ds_2 ds_1
$$

$$
+ \frac{1}{2}\mathbb{X}_{ts}^{jk}\big([V_j, V_k] f\big)(x) + \frac{1}{2}\mathbb{X}_{ts}^{jk}\int_0^1 \{\big([V_j, V_k] f\big)(y_u) - \big([V_j, V_k] f\big)(x)\}\, du.
$$

Note that since the Hölder p-rough path $\mathbf{X}$ is assumed to be weak geometric, the symmetric part of $\mathbb{X}_{ts}$ is equal to $\frac{1}{2}X_{ts} \otimes X_{ts}$, so one has

(1.5) $\quad f\big(\mu_{ts}(x)\big) = f(x) + (t-s)(Vf)(x) + X_{ts}^i (V_i f)(x) + \mathbb{X}_{ts}^{jk}(V_j V_k f)(x) + \epsilon_{ts}^f(x),$

where the remainder ϵ_{ts}^f is defined by the formula

$$
\epsilon_{ts}^f(x) := (t-s)\int_0^1 \{(Vf)(y_u) - (Vf)(x)\} du + (t-s)X_{ts}^i \int_0^1\int_0^{s_1} (V V_i f)(y_{s_2})\, ds_2 ds_1
$$

$$
+ X_{ts}^i X_{ts}^{i'} \int_0^1\int_0^{s_1} \{(V_{i'} V_i f)(y_{s_2}) - (V_{i'} V_i f)(x)\}\, ds_2 ds_1
$$

$$
+ \frac{1}{2}X_{ts}^i \mathbb{X}_{ts}^{jk}\int_0^1\int_0^{s_1} \big([V_j, V_k] V_i f\big)(y_{s_2})\, ds_2 ds_1
$$

$$
+ \frac{1}{2}\mathbb{X}_{ts}^{jk}\int_0^1 \{\big([V_j, V_k] f\big)(y_u) - \big([V_j, V_k] f\big)(x)\}\, du.
$$

PROOF OF PROPOSITION 13 – It is elementary to use estimate (1.3) and the regularity assumptions on the vector fields V, V_i to see that the remainder ϵ_{ts}^f is bounded above by a quantity of the form $c\big(1 + \|\mathbf{X}\|^3\big)\|f\|_{\mathcal{C}^3}|t - s|^{\frac{3}{p}}$, for some constant depending only on V and the V_i. $\triangleright$

A further look at formula (2.7) and estimate (1.3) also make it clear that

(1.6) $$\left\|\epsilon_{ts}^f\right\|_{\mathcal{C}^1} \leqslant c\big(1 + \|\mathbf{X}\|^3\big)|t - s|^{\frac{3}{p}},$$

for a constant c depending only on V and the V_i. This is the key remark for proving the next proposition.

PROPOSITION 14. *The family* $\big(\mu_{ts}\big)_{0 \leqslant s \leqslant t \leqslant T}$ *forms a* $\mathcal{C}^1$*-approximate flow.*

It will be convenient in the following proof to slightly abuse notations and write $V_I(x)$ for $(V_I \mathrm{Id})(x)$, for any multi-index I and point x.

PROOF – We first use formula (1.5) to write

$$\mu_{tu}\big(\mu_{us}(x)\big) = \mu_{us}(x) + (t-u)V\big(\mu_{us}(x)\big) + X_{tu}^i V_i\big(\mu_{us}(x)\big) + \mathbb{X}_{tu}^{jk}\big(V_j V_k\big)\big(\mu_{us}(x)\big) + \epsilon_{tu}^{\mathrm{Id}\,;\,[p]}\big(\mu_{us}(x)\big).$$

We deal with the term $(t-u)V\big(\mu_{us}(x)\big)$ using estimate (1.3) and the Lipschitz character of V:

$$\big|(t-u)V\big(\mu_{us}(x)\big) - (t-u)V(x)\big| \leqslant c\big(1 + \|\mathbf{X}\|^3\big)\,|u-s|^{\frac{3}{p}}.$$

The remainder $\epsilon_{tu}^{\mathrm{Id}}\big(\mu_{us}(x)\big)$ has a $\mathcal{C}^1$-norm bounded above by $c\big(1 + \|\mathbf{X}\|^3\big)^2 |t-u|^{\frac{3}{p}}$, by the remark preceeding proposition 14 and the $\mathcal{C}^1$-estimate (1.3) on μ_{us}. We develop $V_i\big(\mu_{us}(x)\big)$ to deal with the term $X_{tu}^i V_i\big(\mu_{us}(x)\big)$. As

$$V_i\big(\mu_{us}(x)\big) = V_i(x) + (u-s)\big(VV_i\big)(x) + X_{us}^{i'}\big(V_{i'}V_i\big)(x) + \mathbb{X}_{us}^{jk}\big(V_j V_k V_i\big)(x) + \epsilon_{us}^{V_i}(x)$$

we have

(1.7) $$X_{tu}^i V_i\big(\mu_{us}(x)\big) = X_{tu}^i V_i(x) + X_{us}^{i'} X_{tu}^i\big(V_{i'}V_i\big)(x) + \varepsilon_{tu,us}^{V_i}(x),$$

where the remainder $\varepsilon_{tu,us}^{V_i}$ has $\mathcal{C}^1$-norm bounded above by

(1.8) $$\left\| \varepsilon_{tu,us}^{V_i} \right\|_{\mathcal{C}^1} \leqslant c\big(1 + \|\mathbf{X}\|^3\big)\,|u-s|^{\frac{3}{p}},$$

for a constant c depending only on V and the V_n. Set

$$\varepsilon_{tu,us}(x) = \sum_{i=1}^{\ell} \varepsilon_{tu,us}^{V_i}(x).$$

The term $\mathbb{X}_{tu}^{jk}\big(V_j V_k\big)\big(\mu_{us}(x)\big)$ is simply dealt with writing

(1.9) $$\mathbb{X}_{tu}^{jk}\big(V_j V_k\big)\big(\mu_{us}(x)\big) = \mathbb{X}_{tu}^{jk}\big(V_j V_k\big)(x) + \mathbb{X}_{tu}^{jk}\Big\{ \big(V_j V_k\big)\big(\mu_{us}(x)\big) - \mathbb{X}_{tu}^{jk}\big(V_j V_k\big)(x)\Big\},$$

and using estimate (1.3) and the $\mathcal{C}_b^1$ character of $V_j V_k$ to see that the last term on the right hand side has a $\mathcal{C}^1$-norm bounded above by $c\big(1 + \|\mathbf{X}\|^3\big)\,|u-s|^{\frac{3}{p}}$. All together, this gives

$$\mu_{tu}\big(\mu_{us}(x)\big) = \mu_{us}(x) + (t-u)V(x) + X_{tu}^i V_i(x) + X_{us}^{i'} X_{tu}^i\big(V_{i'}V_i\big)(x) + \mathbb{X}_{tu}^{jk}\big(V_j V_k\big)(x) + \varepsilon_{tu,us}(x)$$

$$= x + (u-s)V(x) + X_{us}^i V_i(x) + \mathbb{X}_{us}^{jk}\big(V_j V_k\big)(x) + \epsilon_{us}^{\mathrm{Id}}(x) + (\cdots)$$

$$= x + (t-s)V(x) + X_{ts}^i V_i(x) + \mathbb{X}_{ts}^{jk}\big(V_j V_k\big)(x) + \epsilon_{us}^{\mathrm{Id}}(x) + \varepsilon_{tu,us}(x)$$

$$= \mu_{ts}(x) + \epsilon_{us}^{\mathrm{Id}}(x) + \varepsilon_{tu,us}(x),$$

so it follows from estimates (1.6) and (1.8) that μ is indeed a $\mathcal{C}^1$-approximate flow. $\triangleright$

The above proof makes it clear that one can take for constant c_1 in the $\mathcal{C}^1$-approximate flow property (1.2) for μ the constant $c\big(1 + \|\mathbf{X}\|^3\big)$, for a constant c depending only on V and the V_i.

Recalling proposition 13 describing the maps μ_{ts} in terms of Euler expansion, the following definition of a solution flow to a rough differential equation is to be thought of as defining a notion of solution in terms of uniform Euler expansion

$$\left\| f \circ \varphi_{ts} - \left\{ f + X_{ts}^i V_i f + \mathbb{X}_{ts}^{jk} V_j V_k f \right\} \right\|_\infty \leqslant c \, |t - s|^{>1}.$$

DEFINITION 15. *A* **flow** $\left(\varphi_{ts} \right)_{0 \leqslant s \leqslant t \leqslant T}$ *is said to* **solve the rough differential equation**

$$(1.10) \qquad\qquad d\varphi = V \, dt + \mathrm{F}^\otimes \mathbf{X}(dt)$$

if there exists a constant $a > 1$ independent of $\mathbf{X}$ and two possibly $\mathbf{X}$-dependent positive constants δ and c such that

$$(1.11) \qquad\qquad \left\| \varphi_{ts} - \mu_{ts} \right\|_\infty \leqslant c \, |t - s|^a$$

holds for all $0 \leqslant s \leqslant t \leqslant T$ with $t - s \leqslant \delta$.

If for instance $\mathbf{X}$ is the weak geometric Hölder p-rough path canonically associated with an $\mathbb{R}^\ell$-valued piecewise smooth path h, it follows from exercice 1, and the fact that the iterated integral $\int_s^t \int_s^r dh_u \otimes dh_r$ has size $|t - s|^2$, that the solution flow to the rough differential equation

$$d\varphi = V \, dt + \mathrm{F}^\otimes \mathbf{X}(dt)$$

is the flow associated with the ordinary differential equation

$$\dot{y}_t = V(y_t) dt + V_i(y_t) \, dh_t^i.$$

The following well-posedness result follows directly from theorem 2 on $\mathcal{C}^1$-approximate flows and proposition 14.

THEOREM 16. *The rough differential equation on flows*

$$d\varphi = V \, dt + \mathrm{F}^\otimes \mathbf{X}(dt)$$

has a unique solution flow; it takes values in the space of uniformly Lipschitz continuous homeomorphisms of E *with uniformly Lipschitz continuous inverses, and depends continuously on* $\mathbf{X}$.

PROOF – Applying theorem 2 on $\mathcal{C}^1$-approximate flows to μ we obtain the existence of a unique flow φ satisfying condition (2.12), for δ small enough; it further satisfies the inequality

$$(1.12) \qquad\qquad \left\| \varphi_{ts} - \mu_{\pi_{ts}} \right\|_\infty \leqslant c \left(1 + \|\mathbf{X}\|^3 \right)^2 T \, |\pi_{ts}|^{a-1},$$

for any partition π_{ts} of $[s, t] \subset [0, T]$ of mesh $|\pi_{ts}| \leqslant \delta$, as a consequence of inequality (1.8). As this bound is uniform in (s, t), and for $\mathbf{X}$ in a bounded set of the space of weak geometric Hölder p-rough paths, and since each map $\mu_{\pi_{ts}}$ is a continuous function of $\left((s, t), \mathbf{X} \right)$, the flow φ depends continuously on $\left((s, t), \mathbf{X} \right)$.

To prove that φ is a homeomorphism, note that, with the notations of part I of the course,

$$\left(\mu_{ts}^{(n)}\right)^{-1} = \mu_{s_1 s_0}^{-1} \circ \cdots \circ \mu_{s_{2^n} s_{2^n - 1}}^{-1}, \quad s_i = s + i 2^{-n}(t - s),$$

can actually be written $\left(\mu_{ts}^{(n)}\right)^{-1} = \overline{\mu}_{s_{2^n} s_{2^n - 1}} \circ \cdots \circ \overline{\mu}_{s_1 s_0}$, for the time 1 map $\overline{\mu}$ associated with the rough path $\mathbf{X}_{t-\bullet}$. As $\overline{\mu}$ enjoys the same properties as μ, the maps $\left(\mu_{ts}^{(n)}\right)^{-1}$ converge uniformly to some continuous map φ_{ts}^{-1} which satisfies by construction $\varphi_{ts} \circ \varphi_{ts}^{-1} = \mathrm{Id}$.

Recall that proposition 6 provides a uniform control of the Lipschitz norm of the maps φ_{ts}; the same holds for their inverses in view of the preceeding paragraph. We propagate this property from the set $\left\{(s, t) \in [0, T]^2 \, ; \, s \leqslant t, \, t - s \leqslant \delta\right\}$ to the whole of the $\left\{(s, t) \in [0, T]^2 \, ; \, s \leqslant t\right\}$ using the flow property of φ. $\qquad \triangleright$

REMARKS 17. (1) **Friz-Victoir approach to rough differential equations.** *The continuity of the solution flow with respect to the driving rough path $\mathbf{X}$ has the following consequence, which justifies the point of view adopted by Friz and Victoir in their works. Suppose the Hölder weak geometric p-rough path $\mathbf{X}$ is the limit in the rough path metric of the canonical Hölder weak geometric p-rough paths $\mathbf{X}^n$ associated with some piecwise smooth $\mathbb{R}^\ell$-valued paths $(x_t^n)_{0 \leqslant t \leqslant T}$. We have noticed that the solution flow φ^n to the rough differential equation*

$$d\varphi^n = V \, dt + \mathrm{F}^{\otimes} \mathbf{X}^n(dt)$$

is the flow associated with the ordinary differential equation

$$\dot{y}_u = V(y_u) du + V_i(y_u) \, d(x_u^n)^i.$$

*As $\|\varphi^n - \varphi\|_\infty = o_n(1)$, from the continuity of the solution flow with respect to the driving rough path, the flow φ appears in that case as a uniform limit of the elementary flows φ^n. A Hölder weak geometric p-rough path with the above property is called a Hölder geometric p-rough path; not all Hölder weak geometric p-rough path are Hölder geometric p-rough path [**17**], although there is little difference.*

(2) ***Time-inhomogeneous dynamics.*** *The above results have a straightforward generalization to dynamics driven by a time-dependent bounded drift $V(s; \cdot)$ which is Lipschitz continuous with respect to the time variable and $\mathcal{C}_b^2$ with respect to the space variable, uniformly with respect to time, and time-dependent vector fields $V_i(s; \cdot)$ which are Lipschitz continuous with respect to time, and $\mathcal{C}_b^3$ with respect to the space variable, uniformly with respect to time. We define in that case a $\mathcal{C}^1$-approximate flow by defining μ_{ts} as the time 1 map associated with the ordinary differential equation*

$$\dot{y}_u = (t - s)V(s; y_u) + X_{ts}^i V_i(s; y_u) + \mathbb{X}_{ts}^{jk} [V_j, V_k](y_u), \quad 0 \leqslant u \leqslant 1.$$

1.2. Classical rough differential equations. In the classical setting of rough differential equations, one is primarily interested in a notion of *solution path*, defined in terms of local Taylor-Euler expansion.

DEFINITION 18. *A **path** $\left(z_s\right)_{0 \leqslant s \leqslant T}$ is said to* **solve the rough differential equation**

$$(1.13) \qquad\qquad dz = V\,dt + \mathrm{F}\,\mathbf{X}(dt)$$

with initial condition x, if $z_0 = x$ and there exists a constant $a > 1$ independent of $\mathbf{X}$, and two possibly $\mathbf{X}$-dependent positive constants δ and c, such that

$$(1.14)$$
$$\left| f(z_t) - \left\{ f(z_s) + (t-s)(Vf)(z_s) + X_{ts}^i\big(V_i f\big)(z_s) + \mathbb{X}_{ts}^{jk}\big(V_j V_k f\big)(z_s) \right\} \right| \leqslant c\,\|f\|_{\mathcal{C}^3}\,|t-s|^a$$

holds for all $0 \leqslant s \leqslant t \leqslant T$, with $t - s \leqslant \delta$, for all $f \in \mathcal{C}_b^3$.

THEOREM 19 (Lyons' universal limit theorem). *The rough differential equation (1.13) has a unique solution path; it is a continuous function of $\mathbf{X}$ in the uniform norm topology.*

PROOF – **a) Existence.** It is clear that if $\left(\varphi_{ts}\right)_{0 \leqslant s \leqslant t \leqslant 1}$ stands for the solution flow to the equation

$$d\varphi = V\,dt + \mathrm{F}^{\otimes}\mathbf{X}(dt),$$

then the path $z_t := \varphi_{t0}(x)$ is a solution path to the rough differential equation (1.13) with initial condition x.

b) Uniqueness. Let agree to denote by $O_c(m)$ a quantity whose norm is bounded above by $c\,m$. Let α stand for the minimum of $\frac{3}{p}$ and the constant a in definition 18, and let $y_\bullet$ be any other solution path. It satisfies by proposition 13 the estimate

$$\left| y_t - \varphi_{ts}(y_s) \right| \leqslant c|t-s|^\alpha.$$

Using the fact that the maps φ_{ts} are uniformly Lipschitz continuous, with a Lipschitz constant bounded above by L say, one can write for any $\epsilon > 0$ and any integer $k \leqslant \frac{T}{\epsilon}$

$$\begin{aligned}
y_{k\epsilon} &= \varphi_{k\epsilon,(k-1)\epsilon}\big(y_{(k-1)\epsilon}\big) + O_c\big(\epsilon^\alpha\big) \\
&= \varphi_{k\epsilon,(k-1)\epsilon}\Big(\varphi_{(k-1)\epsilon,(k-2)\epsilon}\big(y_{(k-2)\epsilon}\big) + O_c\big(\epsilon^\alpha\big)\Big) + O_c\big(\epsilon^\alpha\big) \\
&= \varphi_{k\epsilon,(k-2)\epsilon}\big(y_{(k-2)\epsilon}\big) + O_{cL}\big(\epsilon^\alpha\big) + O_c\big(\epsilon^\alpha\big),
\end{aligned}$$

and see by induction that

$$\begin{aligned}
y_{k\epsilon} &= \varphi_{k\epsilon,(k-n)\epsilon}\big(y_{(k-n)\epsilon}\big) + O_{cL}\big((n-1)\epsilon^\alpha\big) + O_c\big(\epsilon^\alpha\big) \\
&= \varphi_{k\epsilon,0}(x) + O_{cL}\big(k\epsilon^\alpha\big) + o_\epsilon(1) \\
&= z_{k\epsilon} + O_{cL}\big(k\epsilon^\alpha\big) + o_\epsilon(1).
\end{aligned}$$

Taking ϵ and k so that $k\epsilon$ converges to some $t \in [0,T]$, we see that $y_t = z_t$, since $\alpha > 1$.

The continuous dependence of the solution path $z_\bullet$ with respect to $\mathbf{X}$ is transfered from φ to $z_\bullet$. $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad \triangleright$

The map that associates to the rough path $\mathbf{X}$ the solution to the rough differential equation (1.13) is called the **Ito map**.

2. The general case

We have defined in the previous section a solution to the rough differential equation

$$d\varphi = V\,dt + \mathrm{F}^\otimes \mathbf{X}(dt),$$

driven by a weak geometric Hölder p-rough path, for $2 \leqslant p < 3$, as a flow with $(s,t;x)-$uniform Taylor-Euler expansion of the form

$$f\big(\varphi_{ts}(x)\big) = f(x) + (t-s)(Vf)(x) + X_{ts}^i\big(V_if\big)(x) + \mathbb{X}_{ts}^{jk}\big(V_jV_kf\big)(x) + O\big(|t-s|^{>1}\big).$$

The definition of a solution flow in the general case will require from φ that it satisfies a similar expansion, of the form

$$(2.1)\qquad f\big(\varphi_{ts}(x)\big) = f(x) + (t-s)(Vf)(x) + \sum_{|I|\leqslant [p]} X_{ts}^I\big(V_If\big)(x) + O\big(|t-s|^{>1}\big).$$

As in the previous section, we shall obtain φ as the unique flow associated with some $\mathcal{C}^1$-approximate flow $\big(\mu_{ts}\big)_{0\leqslant s\leqslant t\leqslant 1}$, where μ_{ts} is the time 1 map associated with an ordinary differential equation constructed from the V_i and their brackets, and V and $\mathbf{X}_{ts}$. In order to avoid writing expressions with loads of indices (the $\mathbf{X}_{ts}^I$), I will first introduce in subsection 2.1 a coordinate-free way of working with rough paths and vector fields. A $\mathcal{C}^1$-approximate flow with the awaited Euler expansion will be constructed in subsection 2.2, leading to a general well-posedness result for rough differential equations on flows.

To make the crucial formula (2.7) somewhat shorter we assume in this section that $V = 0$. The reader is urged to workout by herself/himself the infinitesimal changes that have to be done in what follows in order to work with a non-null drift V. From hereon, the vector fields V_i are assumed to be of class $\mathcal{C}_b^{[p]+1}$. We denote by $\mathcal{C}_b^{[p]+1}(\mathrm{E},\mathrm{E})$ the set of $\mathcal{C}_b^{[p]+1}$ vector fields on E. We denote for by $\pi_k : T_\ell^\infty \to (\mathbb{R}^\ell)^k$ the natural projection operator and set $\pi_{\leqslant k} = \sum_{j\leqslant k}\pi_j$.

2.1. Differential operators. Let F be a continuous linear map from $\mathbb{R}^\ell$ to $\mathcal{C}_b^{[p]+1}(\mathrm{E},\mathrm{E})$ – one usually calls such a map a *vector field valued 1-form on $\mathbb{R}^\ell$*. For any $v \in \mathbb{R}^\ell$, we identify the $\mathcal{C}^{[p]+1}$ vector field $\mathrm{F}(v)$ on E with the first order differential operator

$$\mathrm{F}^\otimes(v) \,:\, g \in \mathcal{C}^1(\mathrm{E}) \mapsto (D.g)\big(\mathrm{F}(v)(\cdot)\big) \in \mathcal{C}^0(\mathrm{E});$$

in those terms, we recover the vector field $\mathrm{F}(v)$ as $\mathrm{F}^\otimes(v)\mathrm{Id}$. The map $\mathrm{F}^\otimes$ is extended to $T_\ell^{[p]+1}$ by setting

$$\mathrm{F}^\otimes(1) := \mathrm{Id} : \mathcal{C}^0(\mathrm{E}) \mapsto \mathcal{C}^0(\mathrm{E}),$$

and defining $F^\otimes(v_1 \otimes \cdots \otimes v_k)$, for all $1 \leqslant k \leqslant [p]+1$ and $v_1 \otimes \cdots \otimes v_k \in (\mathbb{R}^\ell)^{\otimes k}$, as the kth-order differential operator from $\mathcal{C}^k(\mathrm{E})$ to $\mathcal{C}^0(\mathrm{E})$, defined by the formula

$$F^\otimes(v_1 \otimes \cdots \otimes v_k) := F^\otimes(v_1) \cdots F^\otimes(v_k),$$

and by requiring linearity. So, we have the morphism property

(2.2) $$F^\otimes(\mathbf{e})\, F^\otimes(\mathbf{e}') = F^\otimes(\mathbf{e}\mathbf{e}')$$

for any $\mathbf{e}, \mathbf{e}' \in T_\ell^{[p]+1}$ with $\mathbf{e}\mathbf{e}' \in T_\ell^{[p]+1}$. This condition on $\mathbf{e}, \mathbf{e}'$ is required for if $\mathbf{e}' = v_1 \otimes \cdots \otimes v_k$ with $v_i \in \mathbb{R}^\ell$, the map $F^\otimes(\mathbf{e}')\mathrm{Id}$ from E to itself is $\mathcal{C}_b^{[p]+1-k}$, so $F^\otimes(\mathbf{e})\, F^\otimes(\mathbf{e}')$ only makes sense if $\mathbf{e}\mathbf{e}' \in T_\ell^{[p]+1}$. We also have

$$\left[F^\otimes(\mathbf{e}), F^\otimes(\mathbf{e}')\right] = F^\otimes\big([\mathbf{e}, \mathbf{e}']\big)$$

for any $\mathbf{e}, \mathbf{e}' \in T_\ell^{[p]+1}$ with $\mathbf{e}\mathbf{e}'$ and $\mathbf{e}'\mathbf{e}$ in $T_\ell^{[p]+1}$. This implies in particular that $F^\otimes(\Lambda)$ is actually a first order differential operator for any $\Lambda \in \mathfrak{g}_\ell^{[p]+1}$, that is a vector field. Note that for any $\Lambda \in \mathfrak{g}_\ell^{[p]+1}$ and $1 \leqslant k \leqslant [p]+1$, then $\Lambda^k := \pi_k(\Lambda)$ is an element of $\mathfrak{g}_\ell^{[p]}$, and the vector field $F^\otimes(\Lambda^k)\mathrm{Id}$ is $\mathcal{C}_b^{[p]+1-k}$.

We extend $F^\otimes$ to the unrestricted tensor space T_ℓ^∞ setting

(2.3) $$F^\otimes(\mathbf{e}) = F^\otimes\big(\pi_{\leqslant [p]+1}\mathbf{e}\big)$$

for any $\mathbf{e} \in T_\ell^\infty$.

Consider as a particular case the map F defined for $u \in \mathbb{R}^\ell$ by the formula

$$F(u) = u^i\, V_i(\cdot).$$

Using the formalism of this paragraph, an Euler expansion of the form

$$f\big(\varphi_{ts}(x)\big) = f(x) + \sum_{|I| \leqslant [p]} X_{ts}^I\big(V_I f\big)(x) + O\big(|t-s|^{>1}\big),$$

as in equation (2.1), becomes

$$f\big(\varphi_{ts}(x)\big) = \big(F^\otimes(\mathbf{X}_{ts})f\big)(x) + O\big(|t-s|^{>1}\big).$$

2.2. From Taylor expansions to flows driven by rough paths: bis. Let $2 \leqslant p$ be given, together with a $\mathfrak{G}_\ell^{[p]}$-valued weak-geometric Hölder p-rough path $\mathbf{X}$, defined on some time interval $[0, T]$, and some continuous linear map F from $\mathbb{R}^\ell$ to the set $\mathcal{C}_b^{[p]+1}(\mathrm{E}, \mathrm{E})$ of vector fields on E. For any $0 \leqslant s \leqslant t \leqslant T$, denote by Λ_{ts} the logarithm of $\mathbf{X}_{ts}$, and let μ_{ts} stand for the well-defined time 1 map associated with the ordinary differential equation

(2.4) $$\dot{y}_u = F^\otimes\big(\Lambda_{ts}\big)(y_u), \quad 0 \leqslant u \leqslant 1.$$

This equation is indeed an ordinary differential equation since Λ_{ts} is an element of $\mathfrak{g}_\ell^{[p]}$. For $2 \leqslant p < 3$, it reads

$$\dot{y}_u = X_{ts}^i V_i(y_u) + \frac{1}{2}\left(\mathbb{X}_{ts}^{jk} + \frac{1}{2}X_{ts}^j X_{ts}^k\right)[V_j, V_k](y_u), \quad 0 \leqslant u \leqslant 1.$$

As the tensor $X_{ts} \otimes X_{ts}$ is symmetric and the map $(j,k) \mapsto [V_j, V_k]$ is antisymmetric, this equation actually reads

$$\dot{y}_u = X_{ts}^i V_i(y_u) + \frac{1}{2} \mathbb{X}_{ts}^{jk} [V_j, V_k](y_u),$$

which is nothing else than equation (1.1), whose time 1 map defined the $\mathcal{C}^1$-approximate flow we studied in section 1.1.

It is a consequence of classical results from ordinary differential equations, and the definition of the norm on the space of weak-geometric Hölder p-rough paths, that the solution map $(r,x) \mapsto y_r$, with $y_0 = x$, depends continuously on $((s,t), \mathbf{X})$ in $\mathcal{C}^0$-norm, and satisfies the following basic estimate. The next proposition shows that μ_{ts} has precisely the kind of Taylor-Euler expansion property that we expect from a solution to a rough differential equation.

$$(2.5) \qquad \left\| y_r - \mathrm{Id} \right\|_{\mathcal{C}^1} \leqslant c \Big(1 + \|\mathbf{X}\|^{[p]} \Big) |t - s|^{\frac{1}{p}}, \quad 0 \leqslant r \leqslant 1$$

PROPOSITION 20. *There exists a positive constant c, depending only on the V_i, such that the inequality*

$$(2.6) \qquad \left\| f \circ \mu_{ts} - \mathrm{F}^{\otimes}(\mathbf{X}_{ts}) f \right\|_{\infty} \leqslant c \Big(1 + \|\mathbf{X}\|^{[p]} \Big) \|f\|_{\mathcal{C}^{[p]+1}} |t - s|^{\frac{[p]+1}{p}}$$

holds for any $f \in \mathcal{C}_b^{[p]+1}(E)$.

In the classical setting of an ordinary differential equation

$$\dot{z}_u = W(z_u), \quad z_0 = x,$$

driven by a $\mathcal{C}^1$ vector field W on E, we would get a Taylor expansion formula for $f(z_1)$ from the elementary formula

$$f(z_1) = \sum_{k=0}^{n-1} \frac{1}{k!} \left(W^{\circ k} f \right)(x) + \int_{\Delta_n} \left(W^{\circ n} f \right)(y_{s_n}) \, ds$$

obtained by induction, where $W^{\circ 0} f = f$ and $W^{\circ n+1} f := W\big(W^{\circ n} f \big)$, and

$$\Delta_n := \big\{ (s_1, \ldots, s_n) \in [0, T]^n \, ; \, s_n \leqslant \cdots \leqslant s_1 \big\},$$

with the notation ds for $ds_n \ldots ds_1$. We proceed along these lines to obtain a similar formula for the solution to the preceeding equation with

$$W = \mathrm{F}^{\otimes}\big(\boldsymbol{\Lambda}_{ts} \big) = \sum_{k_1=0}^{[p]+1} \mathrm{F}^{\otimes}\big(\Lambda_{ts}^{k_1} \big).$$

Some care however is needed to take into account the fact that the vector fields $\mathrm{F}^{\otimes}\big(\Lambda_{ts}^{k_1} \big)$ have different regularity properties.

Recall $\mathrm{F}^{\otimes}(0)$ is the null map from $\mathcal{C}^0(E)$ to itself and $\pi_0 \Lambda = 0$ for any $\Lambda \in \mathfrak{g}_\ell^{[p]}$. The proof of this proposition and the following one are based on the elementary

identity (2.7) below, obtained by applying repeatedly the identity

$$f(y_r) = f(x) + \int_0^r \left(F^{\otimes}(\boldsymbol{\Lambda}_{ts}) f \right)(y_u) \, du$$

$$= f(x) + \sum_{k_1=0}^{[p]+1} \int_0^r \left(F^{\otimes}(\boldsymbol{\Lambda}_{ts}^{k_1}) f \right)(y_u) \, du, \quad 0 \leqslant r \leqslant 1$$

together with the morphism property (2.2). As emphasized above, the above sum over k_1 is needed to take care of the different regularity properties of the maps $F^{\otimes}(\boldsymbol{\Lambda}_{ts}^{k_1}) f$.

$$f(\mu_{ts}(x)) = f(x) + \left(F^{\otimes}(\boldsymbol{\Lambda}_{ts}) f \right)(x) + \sum_{k_1+k_2 \leqslant [p]+1} \int_0^1 \int_0^{s_1} \left(F^{\otimes}(\boldsymbol{\Lambda}_{ts}^{k_2}) F^{\otimes}(\boldsymbol{\Lambda}_{ts}^{k_1}) f \right)(y_{s_2}) \, ds_2 \, ds_1$$

$$= f(x) + \left(F^{\otimes}(\boldsymbol{\Lambda}_{ts}) f \right)(x) + \int_0^1 \int_0^{s_1} \left(F^{\otimes}(\boldsymbol{\Lambda}_{ts}^{\bullet 2}) f \right)(y_{s_2}) \, ds_2 \, ds_1$$

We use here the notation $\bullet 2$ to denote the multiplication $\Lambda_{ts}^{\bullet 2} = \Lambda_{ts} \Lambda_{ts}$, not to be confused with the second level $\boldsymbol{\Lambda}_{ts}^2$ of $\boldsymbol{\Lambda}_{ts}$; the product is done here in T_ℓ^∞, and definition (2.3) used to make sense of $F^{\otimes}(\boldsymbol{\Lambda}_{ts}^{\bullet 2}) f$. Repeating $(n-1)$ times the above procedure in an iterative way, we see that

$$f(\mu_{ts}(x)) = f(x) + \sum_{k=1}^{n-1} \frac{1}{k!} \left(F^{\otimes}(\boldsymbol{\Lambda}_{ts}^{\bullet k}) f \right)(x) + \int_{\Delta_n} \left(F^{\otimes}(\boldsymbol{\Lambda}_{ts}^{\bullet n}) f \right)(y_{s_n}) \, ds$$

$$= f(x) + \sum_{k=1}^{n} \frac{1}{k!} \left(F^{\otimes}(\boldsymbol{\Lambda}_{ts}^{\bullet k}) f \right)(x) + \int_{\Delta_n} \left\{ \left(F^{\otimes}(\boldsymbol{\Lambda}_{ts}^{\bullet n}) f \right)(y_{s_n}) - \left(F^{\otimes}(\boldsymbol{\Lambda}_{ts}^{\bullet n}) f \right)(x) \right\} ds.$$

Note that $\pi_j \boldsymbol{\Lambda}_{ts}^{\bullet n} = 0$, for all $j \leqslant n-1$, and

$$\pi_{\leqslant [p]} \left(\sum_{k=1}^{[p]} \frac{1}{k!} \boldsymbol{\Lambda}_{ts}^{\bullet k} \right) = \mathbf{X}_{ts};$$

also $\pi_{\leqslant [p]} \left(\boldsymbol{\Lambda}_{ts}^{\bullet [p]} \right) = \left(X_{ts}^1 \right)^{\otimes [p]}$ is of size $|t-s|^{\frac{[p]}{p}}$. We separate the different terms in the above identity according to their size in $|t-s|$; this leads to the following expression for $f(\mu_{ts}(x))$.

(2.7)

$$f(x) + \left(F^{\otimes} \left(\pi_{\leqslant [p]} \left\{ \sum_{k=1}^{n} \frac{1}{k!} \boldsymbol{\Lambda}_{ts}^{\bullet k} \right\} \right) f \right)(x) + \int_{\Delta_n} \left\{ \left(F^{\otimes} (\pi_{\leqslant [p]} \boldsymbol{\Lambda}_{ts}^{\bullet n}) f \right)(y_{s_n}) - \left(F^{\otimes} (\pi_{\leqslant [p]} \boldsymbol{\Lambda}_{ts}^{\bullet n}) f \right)(x) \right\} ds$$

$$+ \left(F^{\otimes} \left(\pi_{[p]+1} \left\{ \sum_{k=1}^{n} \frac{1}{k!} \boldsymbol{\Lambda}_{ts}^{\bullet k} \right\} \right) f \right)(x) + \int_{\Delta_n} \left\{ \left(F^{\otimes} \left(\pi_{[p]+1} \boldsymbol{\Lambda}_{ts}^{\bullet n} \right) f \right)(y_{s_n}) - \left(F^{\otimes} \left(\pi_{[p]+1} \boldsymbol{\Lambda}_{ts}^{\bullet n} \right) f \right)(x) \right\} ds$$

We denote by $\epsilon_{ts}^{f;n}(x)$ the sum of the two terms involving $\pi_{[p]+1}$ in the above line, made up of terms of size at least $|t-s|^{\frac{[p]+1}{p}}$. Note that for $n = [p]$, the integral term

in the first line involves $\pi_{\leqslant [p]}\left(\mathbf{\Lambda}_{ts}^{[p]}\right) = \left(X_{ts}^1\right)^{\otimes [p]}$ and the increment $y_{s_n} - x$, of size $|t-s|^{\frac{1}{p}}$, by estimate (2.5), so this term is of size $|t-s|^{\frac{[p]+1}{p}}$; we include it in $\epsilon_{ts}^{f\,;\,[p]}(x)$.

PROOF OF PROPOSITION 20 – Applying the above formula with $n = [p]$, we get the identity

$$f\left(\mu_{ts}(x)\right) = \left(\mathrm{F}^\otimes\left(\mathbf{X}_{ts}\right)f\right)(x) + \epsilon_{ts}^{f\,;\,[p]}(x).$$

It is clear on the formula for $\epsilon_{ts}^{f\,;\,[p]}(x)$ that its absolute value is bounded above by a constant multiple of $\left(1 + \|\mathbf{X}\|^{[p]}\right)|t-s|^{\frac{[p]+1}{p}}$, for a constant depending only on the data of the problem and f as in (2.6).   ▷

A further look at formula (2.7) makes it clear that if $2 \leqslant n \leqslant [p]$, and f is $\mathcal{C}_b^{n+1}$, the estimate

$$(2.8) \qquad \left\|\epsilon_{ts}^{f\,;\,n}\right\|_{\mathcal{C}^1} \leqslant c\left(1 + \|\mathbf{X}\|^{[p]}\right)\|f\|_{\mathcal{C}^{n+1}}|t-s|^{\frac{[p]+1}{p}},$$

holds as a consequence of formula (2.5), for a constant c depending only on the V_i.

PROPOSITION 21. *The family of maps* $\left(\mu_{ts}\right)_{0\leqslant s\leqslant t\leqslant T}$ *is a* $\mathcal{C}^1$-*approximate flow.*

PROOF – As the vector fields V_i are of class $\mathcal{C}_b^{[p]+1}$, with $[p]+1 \geqslant 3$, the identity

$$\left\|\mu_{ts} - \mathrm{Id}\right\|_{\mathcal{C}^2} = o_{t-s}(1)$$

holds as a consequence of classical results on ordinary differential equations; we turn to proving the $\mathcal{C}^1$-approximate flow property (1.2). Recall X_{ts}^m stands for $\pi_m \mathbf{X}_{ts}$. We first use for that purpose formula (2.7) to write

$$
\begin{aligned}
\mu_{tu}\left(\mu_{us}(x)\right) &= \left(\mathrm{F}^\otimes\left(\mathbf{X}_{tu}\right)\mathrm{Id}\right)\left(\mu_{us}(x)\right) + \epsilon_{tu}^{\mathrm{Id}\,;\,[p]}\left(\mu_{us}(x)\right) \\
&= \mu_{us}(x) + \sum_{m=1}^{[p]}\left(\mathrm{F}^\otimes\left(X_{tu}^m\right)\mathrm{Id}\right)\left(\mu_{us}(x)\right) + \epsilon_{tu}^{\mathrm{Id}\,;\,[p]}\left(\mu_{us}(x)\right).
\end{aligned}
$$

(2.9)

We splitted the function $\mathrm{F}^\otimes\left(\mathbf{X}_{tu}\right)\mathrm{Id}$ into a sum of functions $\mathrm{F}^\otimes\left(X_{tu}^m\right)\mathrm{Id}$ with different regularity properties, so one needs to use different Taylor expansions for each of them. One uses (2.8) and inequality (2.5) to deal with the remainder

$$\left\|\epsilon_{tu}^{\mathrm{Id}\,;\,[p]}\left(\mu_{us}(x)\right)\right\|_{\mathcal{C}^1} \leqslant c\left(1 + \|\mathbf{X}\|^{[p]}\right)^2|t-u|^{\frac{[p]+1}{p}}.$$

To deal with the term $\left(\mathrm{F}^\otimes\left(X_{tu}^m\right)\mathrm{Id}\right)\left(\mu_{us}(x)\right)$, we use formula (2.7) with $n = [p] - m$ and $f = \mathrm{F}^\otimes\left(X_{tu}^m\right)\mathrm{Id}$. Writing ds for $ds_{[p]-m}\ldots ds_1$, we have

(2.10)

$$\left(\mathrm{F}^\otimes\left(X_{tu}^m\right)\mathrm{Id}\right)\left(\mu_{us}(x)\right) = \left(\mathrm{F}^\otimes\left(X_{tu}^m\right)\mathrm{Id}\right)(x) + \left(\mathrm{F}^\otimes\left(\left\{\pi_{\leqslant [p]}\sum_{k=1}^{[p]-m}\frac{1}{k!}\mathbf{\Lambda}_{us}^{\bullet k}\right\}X_{tu}^m\right)\mathrm{Id}\right)(x) + \epsilon_{us}^{\star\,;\,p-m}(x).$$

The notation $\star$ in the above identity stands for the $\mathcal{C}_b^{[p]+2-m}$ function $F^\otimes\big(X_{tu}^m\big)\mathrm{Id}$; it has $\mathcal{C}^1$-norm controlled by (2.8). The result follows directly from (2.9) and (2.10) writing

$$\mu_{us}(x) = \Big(F^\otimes\big(\mathbf{X}_{us}\big)\mathrm{Id}\Big)(x) + \epsilon_{us}^{\mathrm{Id}\,;\,[p]}(x),$$

and using the identities $\exp\big(\mathbf{\Lambda}_{us}\big) = \mathbf{X}_{us}$ and $\mathbf{X}_{ts} = \mathbf{X}_{us}\mathbf{X}_{tu}$ in $T_\ell^{[p]}$. $\triangleright$

DEFINITION 22. *A* **flow** $(\varphi_{ts}\,;\, 0 \leqslant s \leqslant t \leqslant T)$ *is said to* **solve the rough differential equation**

$$(2.11) \qquad\qquad\qquad d\varphi = F^\otimes\,\mathbf{X}(dt)$$

if there exists a constant $a > 1$ independent of $\mathbf{X}$ and two possibly $\mathbf{X}$-dependent positive constants δ and c such that

$$(2.12) \qquad\qquad\qquad \|\varphi_{ts} - \mu_{ts}\|_\infty \leqslant c\,|t - s|^a$$

holds for all $0 \leqslant s \leqslant t \leqslant T$ with $t - s \leqslant \delta$.

This definition can be equivalently reformulated in terms of uniform Taylor-Euler expansion of the form

$$f\big(\varphi_{ts}(x)\big) = f(x) + \sum_{|I|\leqslant[p]} X_{ts}^I\big(V_I f\big)(x) + O\big(|t - s|^{>1}\big).$$

The following well-posedness result follows directly from theorem 2 and proposition 21; its proof is identical to the proof of theorem 19, without a single word to be changed, except for the power of $\|\mathbf{X}\|$ in estimate (1.12), which needs to be taken as $[p] + 1$ instead of 3.

THEOREM 23. *The rough differential equation*

$$d\varphi = F^\otimes\,\mathbf{X}(dt)$$

has a unique solution flow; it takes values in the space of uniformly Lipschitz continuous homeomorphisms of E with uniformly Lipschitz continuous inverses, and depends continuously on $\mathbf{X}$.

Remarks 17 on Friz-Victoir's approach to rough differential equations and time-inhomogeneous dynamics also hold in the general setting of this section. Section 1.2 on classical rough differential equations has a straightforward analogue in the general setting of this section. We leave the reader the pleasure to adapt it and check that Lyons' universal limit theorem holds, with exactly the same proof as given in section 1.2. Note only that we ask the vector fields V_i to be $\mathcal{C}_b^{[p]+1}$ when working with a weak geometric Hölder p-rough path; the drift vector field V is only required to be $\mathcal{C}_b^2$.

3. Exercices on flows driven by rough paths

We first see in exercice 12 how some general result on flows and $\mathcal{C}^1$-approximate flows proved in exercice 6 can be used to strengthen the result on the continuous dependence of the solution to a rough differential equation with respect to the driving rough path into a local Lipschitz dependence. The next two exercices are variations on the notion of solution flow. Roughly speaking, they are defined in terms of uniform Taylor-Euler expansion property. What happens if the driving vector fields allow for a priori higher order expansion? How robust are these expansions with respect to perturbation of the driving rough path? Exercices 13 and 14 partly answer these questions. Exercise 15 makes a crucial link between the 'differential formulation' of a solution path to a rough differential equation introduced above and the 'integral formulation' that can be set in the setting of controlled paths. The equivalence between these two formulations will be fundamental for the applications to stochastic analysis exposed in the last part of the course.

12. Local Lipschitz continuity of φ with respect to X. Use the result proved in exercice 6 to prove that the solution flow to a rough differential equation driven by **X** is a locally Lipschitz continuous function of **X**, in the uniform norm topology.

13. Taylor expansion of solution flows. Let $V_1, \ldots, V_\ell$ be $\mathcal{C}_b^{[p]+1}$ vector fields on a Banach space E, and **X** be a weak geometric Hölder p-rough path over $\mathbb{R}^\ell$, with $2 \leqslant p$. Set $F = (V_1, \ldots, V_\ell)$. The solution flow to the rough differential equation

$$d\varphi = F^{\otimes} \mathbf{X}(dt)$$

enjoys, by definition, a uniform Taylor-Euler expansion property, expressed either by writing

$$\left\| \varphi_{ts} - \mu_{ts} \right\|_\infty \leqslant c|t-s|^a$$

for the $\mathcal{C}^1$-approximate flow $(\mu_{ts})_{0 \leqslant s \leqslant t \leqslant 1}$ contructed in section 2.2, or by writing

$$\left\| f \circ \varphi_{ts} - \sum_{|I| \leqslant [p]} X_{ts}^I V_I f \right\|_\infty \leqslant c|t-s|^a.$$

What can we say if the vector fields V_i are actually more regular than $\mathcal{C}_b^{[p]+1}$?

Assume $N \geqslant [p]+2$ is given and the V_i are $\mathcal{C}_b^N$. Let **Y** be the canonical lift of **X** to a $\mathfrak{G}_\ell^N$-valued weak geometric Hölder N-rough path, given by Lyons' extension theorem proved in exercice 7. Let $\Theta_{ts} \in \mathfrak{g}_\ell^N$ stand for $\log \mathbf{Y}_{ts}$. For any $0 \leqslant s \leqslant t \leqslant 1$, let ν_{ts} be the time 1 map associated with the ordinary differential equation

$$\dot{z}_u = F^{\otimes}(\Theta_{ts})(z_u), \quad 0 \leqslant u \leqslant 1.$$

a) Prove that ν_{ts} enjoys the following Euler expansion property. For any $f \in \mathcal{C}_b^{N+1}$ we have

$$\tag{3.1} \left\| f \circ \nu_{ts} - F^{\otimes}(\mathbf{Y}_{ts}) f \right\|_\infty \leqslant c|t-s|^{\frac{N+1}{p}},$$

where the contant c depends only on the V_i and **X**.

b) Prove that $(\nu_{ts})_{0 \leqslant s \leqslant t \leqslant 1}$ is a $\mathcal{C}^1$-approximate flow.

c) Prove that φ_{ts} satisfies the high order Euler expansion formula (3.1).

14. Perturbing the signal or the dynamics? Let $2 \leqslant p$ be given and $V_1, \ldots, V_\ell$ be $\mathcal{C}_b^{[p]+1}$ vector fields on E. Let $\mathbf{X}$ be a weak geometric Hölder p-rough path over $\mathbb{R}^\ell$, and $\mathbf{a} \in \mathfrak{g}_\ell^{[p]}$ be such that $\pi_j \mathbf{a} = 0$ for all $j \leqslant [p] - 1$. Write it

$$\mathbf{a} = \sum_{|I|=[p]} a^I \mathbf{e}_{[I]},$$

where $\big(e_1, \ldots, e_\ell\big)$ stand for the canonical basis of $\mathbb{R}^\ell$, and for $I = \big(i_1, \ldots, i_k\big)$,

$$\mathbf{e}_{[I]} = \Big[e_{i_1}, \big[\ldots, \big[e_{i_{k-1}}, e_{i_k} \big] \ldots \big]$$

in $T_\ell^{[p]}$. The $\mathbf{e}_{[I]}$'s form a basis of $\mathfrak{g}_\ell^{[p]}$ with $\pi_n \mathbf{e}_{[I]} = 0$ if $n \neq |I|$. Recall the definition of $\exp : T_\ell^{[p],0} \to T_\ell^{[p],1}$ and its reciprocal log.

a) Show that one defines a weak geometric Hölder p-rough path $\overline{\mathbf{X}}$ over $\mathbb{R}^\ell$ setting

$$\overline{\mathbf{X}}_{ts} = \exp\Big(\log \mathbf{X}_{ts} + (t-s)\mathbf{a} \Big).$$

b) Show that the solution flow to the rough differential equation

$$d\psi = \mathrm{F}^\otimes \overline{\mathbf{X}}(dt)$$

coincides with the solution flow to the rough differential equation

$$d\varphi = V \, dt + \mathrm{F}^\otimes \mathbf{X}(dt),$$

where the vector field V is defined by the formula

$$V = a^I V_{[I]}.$$

15. Differential and integral formulations of a rough differential equation. Recall the setting of controlled paths investigated in exercise 11, and let $\mathrm{F} = \big(V_1, \ldots, V_\ell\big)$ be a collection of $\mathcal{C}_b^3$ vector fields on $\mathbb{R}^d$, seen as a linear map from $\mathbb{R}^\ell$ to the set of vector fields on $\mathbb{R}^d$. Let also $\mathbf{X}$ be a weak geometric Hölder p-rough path, with $2 \leqslant p < 3$. Prove that the path $(x_t)_{0 \leqslant t \leqslant 1}$ is a solution path to the rough differential equation

$$dx_t = \mathrm{F}(x_t)\,\mathbf{X}(dt),$$

in the sense of definition 18, if and only if it is a solution to the integral equation

$$x_t = x_0 + \int_0^t \mathrm{F}(x_s)\, d\mathbf{X}_s$$

(in the set of $\mathbb{R}^d$-valued paths controlled by $\mathbf{X}$), where the above integral is the rough integral defined in exercise 11. Note here that the setting of controlled paths offers the possibility to define the above integral for non-weak geometric Hölder p-rough paths, so one can also define fixed-point problems and solve rough differential equations in that setting. (See the excellent forthcoming lecture notes [?] for this point of view on rough differential equations.) Given what we have done in section 3 on controlled paths, we are bound however to working with Hölder p-rough paths, with $2 < p < 3$. Fortunatley, this will be suffcent to deal with stochastic differential equations driven by Brownian motion in the next part of the course.

CHAPTER 5

Applications to stochastic analysis

$\boxed{\text{Guide for this chapter}}$

So far, I have presented the theory of rough differential equations as a purely deterministic theory of differential equations driven by multi-scale time indexed signals. Lyons, however, constructed his theory first as a deterministic alternative to Itô's integration theory, after some hints by Föllmer in the early 80's that Itô's formula can be understood in a deterministic way, and other works (by Bichteler, Karandikhar...) on the pathwise construction of stochastic integrals. (Recall that stochastic integrals are obtained as limits *in probability* of Riemann sums, with no hope for a stronger convergence to hold as a rule.) Lyons was not only looking for a deterministic way of constructing Itô integrals, he was also looking for a way of obtaining them as *continuous* functions of their integrator! This required a notion of integrator different from the classical one... Rough paths were born as such integrators, with the rough integral of controlled integrands, defined in exercise 13, in the role of Itô integrals. What links these two notions of integrals is the following fudamental fact. *Brownian motion has a natural lift into a Hölder p-rough path*, for any $2 < p < 3$, called the **Brownian rough path**. This object is constructed in section 1 using Kolmogorov's classical regularity criterion, and used in section 2 to see that the stochastic and rough integrals coincide whenever they both make sense. This fundamental fact is used in section 3 to see that stochastic differential equations can be solved in a two step process.

   (i) **Purely probabilistic step.** Lift Brownian motion into the Brownian rough path.
  (ii) **Purely deterministic step.** Solve the rough differential equation associated with the stochastic differential equation.

This requires from the driving vector fields to be $\mathcal{C}_b^3$, for the machinery of rough differential equations to make sense, which is more demanding than the Lipschitz regularity required in the Itô setting. This constraint comes with an enormous gain yet: the solution path to the stochastic differential equation is now a *continuous* function of the driving Brownian rough path, this is Lyons' universal limit theorem, in striking contrast with the measurable character of this solution, when seen as a function of Brownian motion itself. (The twist is that the second level of the Brownian rough path is itself just a measurable function of the Brownian path.) Together with the above solution scheme for solving stochastic differential equations, this provides a simple and deep understanding of some fundamental results on diffusion processes, as section 4 on Freidlin-Wentzell theory of large deviation will demonstrate.

We follow the excellent forthcoming lecture notes [**5**] in sections 2 and 3.

1. The Brownian rough path

1.1. Definition and properties. Let $\left(B_t\right)_{0\leqslant t\leqslant 1}$ be an $\mathbb{R}^\ell$-valued Brownian motion defined on some probability space $(\Omega, \mathcal{F}, \mathbb{P})$. There is no difficulty in using Itô's theory of stochastic integrals to define the two-index continuous process

$$(1.1) \qquad \mathbb{B}^{\text{Itô}}_{ts} := \int_s^t \int_s^u dB_r \otimes dB_u = \int_s^t B_{us} \otimes dB_u.$$

This process satisfies Chen's relation

$$\mathbb{B}^{\text{Itô}}_{ts} = \mathbb{B}^{\text{Itô}}_{tu} + \mathbb{B}^{\text{Itô}}_{us} + B_{us} \otimes B_{tu}$$

for any $0 \leqslant s \leqslant u \leqslant t \leqslant 1$. As is B well-known to have almost-surely $\frac{1}{p}$-Hölder continuous sample paths, for any $p > 2$, the process

$$\mathbf{B}^{\text{Itô}} = \left(B, \mathbb{B}^{\text{Itô}}\right)$$

will appear as a Hölder p-rough path if one can show that $\mathbb{B}^{\text{Itô}}$ is almost-surely $\frac{2}{p}$-Hölder continuous. This can be done easily using *Kolmogorov's regularity criterion*, which we recall and prove for completeness. Denote for that purpose by $\mathbb{D}$ the set of dyadic rationals in $[0, 1]$ and write $\mathbb{D}_n$ for $\left\{k2^{-n} \,;\, k = 0..2^n\right\}$.

THEOREM 24 (Kolmogorov's criterion). *Let (S, d) be a metric space, and $q \geqslant 1$ and $\beta > 1/q > 0$ be given. Let also $\left(X_t\right)_{t\in\mathbb{D}}$ be an S-valued process defined on some probability space, such that one has*

$$(1.2) \qquad \left\|d(X_t, X_s)\right\|_{\mathbb{L}^q} \leqslant C\,|t - s|^\beta,$$

for some finite constant C, for all $s, t \in \mathbb{D}$. Then, for all $\alpha \in \left[0, \beta - \frac{1}{q}\right)$, there exists a random variable $\mathrm{C}_\alpha \in \mathbb{L}^q$ such that one has almost-surely

$$d\left(X_s, X_t\right) \leqslant \mathrm{C}_\alpha\,|t - s|^\alpha,$$

for all $s, t \in \mathbb{D}$; so the process X has an α-Hölder modification defined on $[0, 1]$.

PROOF – Given $s, t \in \mathbb{D}$ with $s < t$, let $m \geqslant 0$ be the only integer such that $2^{-(m+1)} \leqslant t - s < 2^{-m}$. The interval $[s, t)$ contains at most one interval $\left[r_{m+1}, r_{m+1} + 2^{-(m+1)}\right)$ with $r_{m+1} \in \mathbb{D}_{m+1}$. If so, each of the intervals $[s, r_{m+1})$ and $\left[r_{m+1} + 2^{-(m+1)}, t\right)$ contains at most one interval $\left[r_{m+2}, r_{m+2} + 2^{-(m+2)}\right)$ with $r_{m+2} \in \mathbb{D}_{m+2}$. Repeating this remark up to exhaustion of the dyadic interval $[s, t)$ by such dyadic sub-intervals, we see, using the triangle inequality, that

$$d\left(X_t, X_s\right) \leqslant 2 \sum_{n \geqslant m+1} S_n,$$

where $S_n = \sup_{u \in \mathbb{D}_n} d\left(X_u, X_{u+2^{-n}}\right)$. So we have

$$\frac{d\left(X_t, X_s\right)}{(t - s)^\alpha} \leqslant 2 \sum_{n \geqslant m+1} S_n\,2^{(m+1)\alpha} \leqslant C_\alpha$$

where $C_\alpha := 2\sum_{n \geqslant 0} 2^{n\alpha} S_n$. But as the assumption (1.2) implies

$$\mathbb{E}\left[S_n^q\right] \leqslant \mathbb{E}\left[\sum_{t \in \mathbb{D}_n} d\left(X_t, X_{t+2^{-n}}\right)^q\right] \leqslant 2^n C (2^{-n})^{q\beta},$$

we have

$$\left\|C_\alpha\right\|_{\mathbb{L}^q} \leqslant 2\sum_{n \geqslant 0} 2^{n\alpha}\|S_n\|_q \leqslant 2C \sum_{n \geqslant 0} 2^{\left(\alpha - \beta + \frac{1}{p}\right)n} < \infty,$$

so C_α is almost-surely finite. The conclusion follows in a straightforward way.   ▷

Recall the definition of the homogeneous norm on $T_\ell^{2,1}$

$$\|\mathbf{a}\| = \|1 \oplus a^1 \oplus a^2\| = |a^1| + \sqrt{|a^2|}$$

introduced in equation (1.3), with its associated distance function $d(\mathbf{a}, \mathbf{b}) = \|\mathbf{a}^{-1}\mathbf{b}\|$. To see that $\mathbf{B}^{\text{Itô}}$ is a Hölder p-rough path we need to see that it is almost-surely $\frac{1}{p}$-Hölder continuous as a $(T_\ell^{2,1}, \|\cdot\|)$-valued path. This can be obtained from Kolmologorv's criterion provided one has

$$\left\|\mathbf{B}_{ts}^{\text{Itô}}\right\|_{\mathbb{L}^q} \leqslant C \, |t-s|^{\frac{1}{2}},$$

for some constants q with $0 < \frac{1}{2} - \frac{1}{q} < \frac{1}{p}$, and C. Given the form of the norm on $T_\ell^{2,1}$, this is equivalent to requiring

$$(1.3) \qquad \left\|B_{ts}\right\|_{\mathbb{L}^q} \leqslant C \, |t-s|^{\frac{1}{2}}, \qquad \left\|\mathbb{B}_{ts}^{\text{Itô}}\right\|_{\mathbb{L}^{\frac{q}{2}}} \leqslant C \, |t-s|.$$

These two inequalities holds as a straightforward consequence of the scaling property of Brownian motion. (The random variable $\mathbb{B}_{10}^{\text{Itô}}$ is in any $\mathbb{L}^q$ as a consequence of the BDG inequality for instance.)

COROLLARY 25. *The process* $\mathbf{B}^{\text{Itô}}$ *is almost-surely a Hölder p-rough path, for any p with* $\frac{1}{3} < \frac{1}{p} < \frac{1}{2}$. *It is called the* **Itô Brownian rough path**.

Note that $\mathbf{B}^{\text{Itô}}$ is not weak geometric as the symmetric part of $\mathbb{B}_{ts}^{\text{Itô}}$ is equal to $\frac{1}{2} B_{ts} \otimes B_{ts} - \frac{1}{2}(t-s)\,\text{Id}$. Note also that we may as well have used Stratonovich integral in the definition of the iterated integral

$$\mathbb{B}_{ts}^{\text{Str}} := \int_s^t \int_s^u \circ dB_r \otimes \circ dB_u = \int_s^t B_{us} \otimes \circ dB_u;$$

this does not make a big difference a priori since

$$\mathbb{B}_{ts}^{\text{Str}} = \mathbb{B}_{ts}^{\text{Itô}} + \frac{1}{2}(t-s)\text{Id}.$$

So one can define another Hölder p-rough path $\mathbf{B}_{ts}^{\text{Str}} = \left(B_{ts}, \mathbb{B}_{ts}^{\text{Str}}\right)$ above Brownian motion, called **Stratonovich Brownian rough path**. Unlike Itô Brownian rough path, it is weak geometric. (Compute the symmetric part of $\mathbb{B}_{ts}^{\text{Str}}$!) Whatever choice of Brownian rough path we do, its definition seems to involve Itô's theory of stochastic integral. It will happen to be important for applications these two rough paths can actually be constructed in a pathwise way from the Brownian path itself.

Given $n \geqslant 1$, define on the ambient probability space the σ-algebra $\mathcal{F}_n := \sigma\{B_{k2^{-n}} \, ; \, 0 \leqslant k \leqslant 2^n\}$, and let $B_\bullet^{(n)}$ stand for the continuous piecewise linear path that coincides with B at dyadic times in $\mathbb{D}_n$ and is linear in between. Denote by $B^{(n),i}$ the coordinates of $B^{(n)}$. There is no difficulty in defining

$$\mathbb{B}_{ts}^{(n)} := \int_s^t B_{us}^{(n)} \otimes dB_u^{(n)}$$

as a genuine integral as $B^{(n)}$ is piecewise linear, and one has acutally, for $j \neq k$,

$$(1.4) \qquad B_{ts}^{(n)} = \mathbb{E}\left[B_{ts} \big| \mathcal{F}_n\right], \qquad \mathbb{B}_{ts}^{(n),jk} = \mathbb{E}\left[\mathbb{B}_{ts}^{\text{Str},jk} \big| \mathcal{F}_n\right]$$

and $\mathbb{B}_{ts}^{(n),ii} = \frac{1}{2}\left(B_{ts}^{(n),i}\right)^2$.

PROPOSITION 26. *The Hölder p-rough path $\mathbf{B}^{(n)} = \big(B^{(n)}, \mathbb{B}^{(n)}\big)$ converges almost-surely to $\mathbf{B}^{\mathrm{Str}}$ in the Hölder p-rough path topology.*

PROOF – We use the interpolation result stated in proposition 10 to prove the above convergence result. The almost-sure pointwise convergence follows from the martingale convergence theorem applied to the martingales in (1.4). To get the almost-sure uniform bound

$$(1.5) \qquad\qquad\qquad \sup_n \big\| \mathbf{B}^{(n)} \big\| < \infty$$

it suffices to notice that the estimates

$$\big| B_{ts} \big| \leqslant C_p |t-s|^{\frac{1}{p}}, \qquad \big| \mathbb{B}_{ts}^{\mathrm{Str},jk} \big| \leqslant C_p^2 |t-s|^{\frac{2}{p}}$$

obtained from Kolmogorov's regularity criterion with $C_p \in \mathbb{L}^q$ for (any) $q > 2$, give

$$\big| B_{ts}^{(n)} \big| \leqslant \mathbb{E}\big[C_p \big| \mathcal{F}_n \big] |t-s|^{\frac{1}{p}}, \qquad \big| \mathbb{B}_{ts}^{(n),jk} \big| \leqslant \mathbb{E}\big[C_p^2 \big| \mathcal{F}_n \big] |t-s|^{\frac{2}{p}},$$

so the uniform estimate (1.5) follows from Doob's maximal inequality, which implies that almost-sure finite character of the maximum of the martingales $\mathbb{E}\big[C_p^{1 \text{ or } 2} \big| \mathcal{F}_n \big]$, since this maximum is integrable. ▷

1.2. How big is the Brownian rough path? The upper bound of $\big\| \mathbf{B}^{\mathrm{It\hat{o}}} \big\|_{\frac{1}{p}}$ provided by the constant $C_{\frac{1}{p}}$ of Kolmogorov's regularity result says us that $\big\| \mathbf{B}^{\mathrm{It\hat{o}}} \big\|_{\frac{1}{p}}$ is in all the $\mathbb{L}^q$ spaces. The situation is actually much better! As a first hint, notice that since $\mathbf{B}_{ts}^{\mathrm{It\hat{o}}}$ has the same distribution as $\delta_{\sqrt{t-s}} \mathbf{B}_{10}^{\mathrm{It\hat{o}}}$, and the norm of $\mathbf{B}_{10}^{\mathrm{It\hat{o}}}$ has a Gaussian tail (this is elementary), we have

$$(1.6) \qquad\qquad \mathbb{E}\left[\exp\left(\frac{\big\| \mathbf{B}_{ts}^{\mathrm{It\hat{o}}} \big\|^2}{t-s} \right) \right] = \mathbb{E}\left[\exp\left(\big\| \mathbf{B}_{10}^{\mathrm{It\hat{o}}} \big\|^2 \right) \right] < \infty$$

The following Besov embedding is useful in estimating the Hölder norm of a path from its two-point moments.

THEOREM   27 (Besov). *Given $\alpha \in \big[0, \frac{1}{2}\big)$ there exists an integer k_α and a positive constant C_α with the following properties. For any metric space (S,d) and any S-valued continuous path $(x_t)_{0 \leqslant t \leqslant 1}$ we have*

$$\| x_\bullet \|_\alpha \leqslant C_\alpha \left(\int_0^1 \int_0^1 \left(\frac{d(x_t, x_s)}{\sqrt{t-s}} \right)^{2k} ds \, dt \right)^{\frac{1}{2k}}.$$

It can be proved as a direct consequence of the famous Garsia-Rodemich-Rumsey lemma. Applied to the Brownian rough path $\mathbf{B}^{\mathrm{It\hat{o}}}$, Besov's estimate gives

$$\mathbb{E}\left[\big\| \mathbf{B}^{\mathrm{It\hat{o}}} \big\|_{\frac{1}{p}}^{2k} \right] \leqslant C_{\frac{1}{p}}^{2k} \int_0^1 \int_0^1 \mathbb{E}\left[\left(\frac{\| \mathbf{B}_{ts} \|}{\sqrt{t-s}} \right)^{2k} \right] ds \, dt = C_{\frac{1}{p}}^{2k} \, \mathbb{E}\left[\big\| \mathbf{B}_{10} \big\|^{2k} \right].$$

So it follows from (1.6) that we have for any positive constant c

$$\mathbb{E}\left[\sum_{k \geqslant k_{\frac{1}{p}}} \frac{c^k \big\| \mathbf{B}^{\mathrm{It\hat{o}}} \big\|_{\frac{1}{p}}^{2k}}{k!} \right] \leqslant \mathbb{E}\left[\exp\left(c C_{\frac{1}{p}}^2 \big\| \mathbf{B}_{10}^{\mathrm{It\hat{o}}} \big\|^2 \right) \right]$$

so $\exp \left\| \mathbf{B}^{\text{Itô}} \right\|_{\frac{1}{p}}^2$ will be integrable provided c is small enough, by (1.6).

COROLLARY 28. *The p-rough path norm of the Brownian rough path has a Gaussian tail.*

2. Rough and stochastic integral

Let $\mathbf{X}$ be any $\mathbb{R}^\ell$-valued Hölder p-rough path, with $2 < p < 3$. Recall a linear map A from $\mathbb{R}^\ell$ to $\mathbb{R}^d$ acts on $(\mathbb{R}^\ell)^{\otimes 2}$ as follows: $A(a \otimes b) = (Aa) \otimes b$. Recall also that we defined in section 3 the integral of an $\mathrm{L}(\mathbb{R}^\ell, \mathbb{R}^d)$-valued path $\left(\mathrm{F}_s\right)_{0 \leqslant s \leqslant 1}$ controlled by $\mathbf{X} = (X, \mathbb{X})$ as the well-defined limit

$$\int_0^1 \mathrm{F} \, d\mathbf{X} = \lim \sum \mathrm{F}_{t_i} X_{t_{i+1} t_i} + \mathrm{F}'_{t_i} \mathbb{X}_{t_{i+1} t_i},$$

where the sum is over the times t_i of finite partitions π of $[0, 1]$ whose mesh tends to 0. This makes sense in particular for $\mathbf{X} = \mathbf{B}^{\text{Itô}}$. At the same time, if F is adapted to the Brownian filtration, the Riemann sums $\sum \mathrm{F}_{t_i} B_{t_{i+1} t_i}$ converge in probability to the stochastic integral $\int_0^1 \mathrm{F}_s \, dB_s$, as the mesh of the partition π tends to 0. Taking subsequences if necessary, one defines simultaneously the stochastic and the rough integral on an event of probability 1. They actually coincide almost-surely if F' is adapted to the Brownian filtration! To see this, it suffices to see that $\sum \mathrm{F}'_{t_i} \mathbb{X}_{t_{i+1} t_i}$ converges in $\mathbb{L}^2$ to 0 along the subsequence of partitions used to define the stochastic integral $\int_0^1 \mathrm{F}_s dB_s$. Assume first that F' is bounded, by M say. Then, since it is adapted and F'_{t_i} is independent of $\mathbb{B}_{t_{i+1} t_i}$, an elementary conditioning gives

$$\left\| \sum \mathrm{F}'_{t_i} \mathbb{B}^{\text{Itô}}_{t_{i+1} t_i} \right\|_{\mathbb{L}^2}^2 = \sum \left\| \mathrm{F}'_{t_i} \mathbb{B}^{\text{Itô}}_{t_{i+1} t_i} \right\|_{\mathbb{L}^2}^2 \leqslant M^2 \sum \left\| \mathbb{B}_{t_{i+1} t_i} \right\|_{\mathbb{L}^2}^2 \leqslant M^2 |\pi|,$$

which proves the result in that case. If F' is not bounded, we use a localization argument and stop the process at the stopping time

$$\tau_M := \inf \left\{ u \in [0, 1] \, ; \, |\mathrm{F}'_u| > M \right\} \wedge 1.$$

The above reasoning shows in that case that we have the almost-sure equality

$$\int_0^{\tau_M} \mathrm{F} \, d\mathbf{B} = \int_0^1 \mathrm{F}_s^{\tau_M} \, dB_s,$$

from which the result follows since τ_M tends to ∞ as M increases indefinitely.

PROPOSITION 29. *Let $\left(\mathrm{F}_s\right)_{0 \leqslant s \leqslant 1}$ be an $\mathrm{L}(\mathbb{R}^\ell, \mathbb{R}^d)$-valued path controlled by $\mathbf{B}^{\text{Itô}} = (B, \mathbb{B})$, adapted to the Brownian filtration, with a derivative process F' also adapted to that filtration. Then we have almost-surely*

$$\int_0^1 \mathrm{F} \, d\mathbf{B}^{\text{Itô}} = \int_0^1 \mathrm{F}_s \, dB_s.$$

If one uses $\mathbf{B}^{\text{Str}}$ instead of $\mathbf{B}^{\text{Itô}}$ in the above rough integral, an additional well-defined term

$$(\star) := \lim_{|\pi| \searrow 0} \sum \mathrm{F}'_{t_i} \frac{1}{2} \left(t_{i+1} - t_i\right) \mathrm{Id}$$

appears in the left hand side, and we have almost-surely

$$\int_0^1 \mathrm{F} \, d\mathbf{B}^{\text{Str}} = \int_0^1 \mathrm{F} \, d\mathbf{B}^{\text{Itô}} + (\star) = \int_0^1 \mathrm{F}_s \, dB_s + (\star).$$

To identify that additional term, denote by $\mathrm{Sym}(A)$ the symmetric part of a matrix A and recall that

$$\frac{1}{2}\big(t_{i+1} - t_i\big)\mathrm{Id} = \mathrm{Sym}\big(\mathbf{B}^{\mathrm{Str}}_{t_{i+1}t_i}\big) - \mathrm{Sym}\big(\mathbf{B}^{\mathrm{Itô}}_{t_{i+1}t_i}\big) = \frac{1}{2}\mathbb{B}^{\otimes 2}_{t_{i+1}t_i} - \mathrm{Sym}\big(\mathbf{B}^{\mathrm{Itô}}_{t_{i+1}t_i}\big);$$

note also that the above reasoning showing that $\sum \mathrm{F}'_{t_i}\mathbb{B}^{\mathrm{Itô}}_{t_{i+1}t_i}$ converges to 0 in $\mathbb{L}^2$ also shows that $\sum \mathrm{F}'_{t_i}\mathrm{Sym}\big(\mathbb{B}^{\mathrm{Itô}}_{t_{i+1}t_i}\big)$ converges to 0 in $\mathbb{L}^2$. So $(\star)$ is almost-surely equal to the limit as $|\pi| \searrow 0$ of the sums

$$\frac{1}{2}\sum \mathrm{F}'_{t_i}\mathbb{B}^{\otimes 2}_{t_{i+1}t_i}.$$

Since

$$\mathrm{F}'_{t_i}\mathbb{B}_{t_{i+1}t_i} = \mathrm{F}_{t_{i+1}t_i} + \mathrm{R}_{t_{i+1}t_i}$$

for some $\frac{2}{p}$-Hölder remainder R, the above sum equals

$$\frac{1}{2}\Big(\sum \mathrm{F}_{t_{i+1}t_i}\mathbb{B}_{t_{i+1}t_i}\Big) + o_{|\pi|}(1).$$

We recognize in the right hand side sum a quantity which converges in probability to the bracket of F and B.

COROLLARY 30. *Under the assumptions of proposition 29, we have almost-surely*

$$\int_0^1 \mathrm{F}\, d\mathbf{B}^{\mathrm{Str}} = \int_0^1 \mathrm{F}_s \circ dB_s.$$

3. Rough and stochastic differential equations

Equipped with the preceeding two results, it is easy to see that the solution path to a rough differential equation driven by $\mathbf{B}^{\mathrm{Str}}$ or $\mathbf{B}^{\mathrm{Itô}}$ coincides almost-surely with the solution of the corresponding Stratonovich or Itô stochastic differential equation.

THEOREM 31. *Let* $\mathrm{F} = \big(V_1, \dots, V_\ell\big)$ *be* $\mathcal{C}^3_b$ *vector fields on* $\mathbb{R}^d$. *The solution to the rough differential equation*

$$(3.1) \qquad\qquad dx_t = \mathrm{F}(x_t)\,\mathbf{B}^{Str}(dt)$$

coincides almost-surely with the solution to the Stratonovich differential equation

$$dz_t = V_i(z_t) \circ dB_t^i.$$

A similar statement holds for the Itô Brownian rough path and solution to Itô equations.

PROOF – Recall we have seen in exercise 16 that a path is a solution to the rough differential equation (3.1) if and only if it is a solution path to the integral equation

$$x_t = x_0 + \int_0^t \mathrm{F}(x_s)\, d\mathbf{B}^{\mathrm{Str}}_s.$$

Given the result of corollary 30, the theorem will follow if we can see that $x_\bullet$ is adapted to the Brownian filtration; for if one sets $\mathrm{F}_s := \mathrm{F}(x_s)$ then its derivative $\mathrm{F}'_s = D_{x_s}\mathrm{F}'\mathrm{F}(x_s)$, with DF the differential of F with respect to x, will also be adapted. But the adaptedness of the solution $x_\bullet$ to equation (3.1) is clear from its construction in the proof of theorem 19. $\triangleright$

We obtain as a corollary of theorem 31, Lyons' universal limit theorem and the convergence result proved in proposition 26 for the rough path associated with the piecewise linear interpolation $B^{(n)}$ of B the following fundamental result, first proved by Wong and Zakai in the mid 60'.

COROLLARY 32 (Wong-Zakai theorem). *The solution path to the ordinary differential equation*

$$(3.2) \qquad dx_t^{(n)} = F\left(x_t^{(n)}\right) dB_t^{(n)}$$

converges almost-surely to the solution path to the Stratonovich differential equation

$$dx_t = F(x_t) \circ dB_t.$$

PROOF – It suffices to notice that solving the rough differential equation

$$dz_t^{(n)} = F\left(z_t^{(n)}\right) \mathbf{B}^{(n)}(dt)$$

is equivalent to solving equation (3.2). $\triangleright$

4. Freidlin-Wentzell large deviation theory

We shall close this course with a spectacular application of the continuity property of the solution map to a rough differential equation, by showing how one can recover the basics of Freidlin-Wentzell theory of large deviations for diffusion processes from a unique large deviation principle for the Stratonovich Brownian rough path. Exercise 18 also uses this continuity property to deduce Stroock-Varadhan's celebrated support theorem for diffusion laws from the corresponding statement for the Brownian rough path.

4.1. A large deviation principle for the Stratonovich Brownian rough path.
Let start this section by recalling Schilder's large deviation principle for Brownian motion.

a) Schilder's theorem. L Define for that purpose the real-valued function I on $\mathcal{C}^0\big([0,1],\mathbb{R}^d\big)$ equal to $\frac{1}{2}\|h\|_{H^1}^2 = \frac{1}{2}\int_0^1 |\dot{h}_s|^2 \, ds$ on H^1, and ∞ elsewhere. We agree to write $I(\mathcal{A})$ for $\inf\{I(h_\bullet) ; h \in \mathcal{A}\}$, for any Borel subset $\mathcal{A}$ of $\mathcal{C}^0\big([0,1],\mathbb{R}^d\big)$, endowed with the $\mathcal{C}^0$ topology.

THEOREM 33. *Let $\mathbb{P}$ stand for Wiener measure on $\mathcal{C}^0\big([0,1],\mathbb{R}^d\big)$ and B stand for the coordinate process. Given any Borel subset $\mathcal{A}$ of $\mathcal{C}^0\big([0,1],\mathbb{R}^d\big)$, we have*

$$-I\big(\overset{\circ}{\mathcal{A}}\big) \leqslant \overline{\lim} \, \varepsilon^2 \log \mathbb{P}\big(\varepsilon B_\bullet \in \mathcal{A}\big) \leqslant -I\big(\overline{\mathcal{A}}\big).$$

PROOF – The traditional proof of the lower bound is a simple application of the Cameron-Martin theorem. Indeed, if $\mathcal{A}$ is the ball of centre $h \in H^1$ with radius δ, and if we define the probability $\mathbb{Q}$ by its density

$$\frac{d\mathbb{Q}}{d\mathbb{P}} = \exp\left(-\varepsilon^{-1}\int_0^1 h_s dB_s - \frac{\varepsilon^{-2}}{2}I(h)\right)$$

with respect to $\mathbb{P}$, the process $\overline{B}_\bullet := B_\bullet - \varepsilon^{-1}h$ is a Brownian motion under $\mathbb{Q}$, and we have

$$\mathbb{P}\big(|\varepsilon B - h| \leqslant \delta\big) = \mathbb{P}\big(|\overline{B}| \leqslant \varepsilon^{-1}\delta\big) = \mathbb{E}_{\mathbb{Q}}\left[\mathbf{1}_{|\overline{B}| \leqslant \varepsilon^{-1}\delta} \exp\left(-\varepsilon^{-1}\int_0^1 h_s dB_s - \frac{\varepsilon^{-2}}{2}\mathrm{I}(h)\right)\right]$$

$$\geqslant e^{-\frac{\varepsilon^{-2}}{2}\mathrm{I}(h)} \mathbb{Q}\big(|\overline{B}| \leqslant \varepsilon^{-1}\delta\big) = e^{-\frac{\varepsilon^{-2}}{2}\mathrm{I}(h)}\big(1 - o_\varepsilon(1)\big).$$

One classically uses three facts to prove the upper bound.

(1) The piecewise linear approximation $B^{(n)}$ of B introduced above obviously satisfies the upper bound, as $B^{(n)}$ lives (as a random variable) in a finite dimensional space where it defines a Gaussian random variable.

(2) The sequence $\varepsilon B_\bullet^{(n)}$ provides an exponentially good approximation of $\varepsilon B_\bullet$, in the sense that

$$\limsup_{\varepsilon \searrow 0} \varepsilon^2 \log \mathbb{P}\big(\big|\varepsilon B^{(n)} - \varepsilon B\big|_\infty \geqslant \delta\big) \xrightarrow[m \to \infty]{} -\infty.$$

(3) The map I enjoys the following 'continuity' property. With $\mathcal{A}^\delta := \{x \,;\, /, d(x, \mathcal{A}) \leqslant \delta\}$, we have

$$\mathrm{I}(\overline{\mathcal{A}}) = \lim_{\delta \searrow 0} \mathrm{I}(\mathcal{A}^\delta).$$

The result follows from the combination of these three facts. The first and third points are easy to see. As for the second, just note that $B^{(n)} - B$ is actually made up of 2^n independent copies of a scaled Brownian bridge $2^{-\frac{n+1}{2}}\overline{B}_\bullet^k$, with each $\overline{B}^k$ defined on the dyadic interval $\big[k2^{-n}, (k+1)2^{-n}\big]$. As it suffices to look at what happens in each coordinate, classical and easy estimates on the real-valued Brownian bridge provide the result. $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\quad \triangleright$

b) Schilder's theorem for Stratonovich Brownian rough path. The extension of Schilder's theorem to the Brownian rough path requires the introduction of the function J, defined on the set of $\mathfrak{G}_\ell^2$-valued continuous paths $\mathbf{e}_\bullet = \big(e_\bullet^1, e_\bullet^2\big)$ by the formula

$$\overline{\mathrm{J}}(\mathbf{e}_\bullet) = \mathrm{I}(e_\bullet^1).$$

Recall the definition of the dilation δ_λ on T_ℓ^2, given in (1.4). Given any $0 \leqslant \frac{1}{p} < \frac{1}{2}$, one can see the distribution $\mathbf{P}_\varepsilon$ of $\delta_\varepsilon \mathbf{B}^{\mathrm{Str}}$ as a probability measure on the space of $\frac{1}{p}$-Hölder $\mathfrak{G}_\ell^2$-valued functions, with the corresponding norm.

THEOREM 34. *The family $\mathbf{P}_\varepsilon$ of probability measures on $\mathcal{C}^{\frac{1}{p}}\big([0,1], \mathfrak{G}_\ell^2\big)$ satisfies a large deviation principle with good rate function $\overline{\mathrm{J}}$.*

It should be clear to the reader that it is sufficent to prove the claim for the Brownian rough path above a 2-dimensional Brownian motion $B = \big(B^1, B^2\big)$, defined on $\mathcal{C}^0\big([0,1], \mathbb{R}^2\big)$ as the coordinate process. We shall prove this theorem as a consequence of Schilder's theorem; this would be straightforward if the second level process $\mathbb{B}$ – or rather just its anti-symmetric part – were a continuous function of the Brownian path, in uniform topology, which does not hold true of course. However, proposition 26 on the approximation of the Brownian rough path by its 'piecewise linear' counterpart makes it clear that it is almost-surely equal to a limit of continuous functional of the Brownian path. So it is tempting to try and use the following general contraction principle for large deviations. (See the book [19] by Kallenberg for an account of the basics of the theory, and a proof of this

theorem.) We state it here in our setting to avoid unnecessary generality, and define the approximated Lévy area $\mathbb{A}^m_{ts}$ as a real-valued function on the $\mathcal{C}^0\big([0,1],\mathbb{R}^2\big)$ setting

$$\mathbb{A}^m_t := \frac{1}{2} \int_0^t \left(B^1_{\frac{[ms]}{m}} \, dB^2_s - dB^1_s \, B^2_{\frac{[ms]}{m}} \right).$$

It is a continuous function of B in the uniform toplogy. The maps $\mathbb{A}^m$ converge almost-surely uniformly to the Lévy area process $\mathbb{A}_\bullet$ of B. We see the process $\mathbb{A}_\bullet$ as a map defined on the space $\mathcal{C}^0\big([0,1],\mathbb{R}^2\big)$, equal to Lévy's area process on a set of probability 1 and defined in a genuine way on H^1 using Young integrals. (Note that elements of H^1 are $\frac{1}{2}$-Hölder continuous.)

THEOREM 35. *(Extended contraction principle) If*

(1) **(Exponentially good approximation property)**

$$\limsup_\varepsilon \varepsilon^2 \log \mathbf{P}_\varepsilon\big(\big\|\mathbb{A}^m - \mathbb{A}\big\|_\infty > \delta\big) \xrightarrow[m\to\infty]{} -\infty,$$

(2) **(Uniform convergence on I-level sets)** *for each $r > 0$ we have*

$$\big\|(\mathbb{A}^m - \mathbb{A})_{\big|\{\mathrm{I}\leqslant r\}}\big\|_\infty \xrightarrow[m\to\infty]{} 0,$$

then the distribution of $\mathbb{A}_\bullet$ under $\mathbf{P}_\varepsilon$ satisfies a large deviation principle $\mathcal{C}^0\big([0,1],\mathfrak{g}^2_2\big)$ with good rate function $\inf\{\mathrm{I}(\omega)\,;\,\mathbf{a} = \mathbb{A}(\omega)\}$.

PROOF OF THEOREM 34 – The proof amounts to proving points (1) and (2) in theorem 35. The second point is elementary if one notes that for $h \in H^1\big([0,1],\mathbb{R}^2\big)$, we have

$$\left| \int_0^t \left(h_{\frac{[ms]}{m}} - h_s \right) \otimes dh_s \right| \leqslant \|h\|_{\frac{1}{2}} \left(\frac{1}{m} \right)^{\frac{1}{2}} \|h\|_1$$

$$\leqslant \|h\|^2_{H^1} m^{-\frac{1}{2}}.$$

As for the first point, it suffices to prove that

$$\limsup_\varepsilon \varepsilon^2 \log \mathbf{P}_1 \left(\sup_{t\in[0,1]} \int_0^t \left(B^1_s - B^1_{\frac{[ms]}{m}} \right) dB^2_s \geqslant \varepsilon^{-2}\delta \right) \xrightarrow[m\to\infty]{} -\infty,$$

which we can do using elementary martingale inequalities. Indeed, denoting by M_t the martingale defined by the above stochastic integral, with bracket $\int_0^t \left| B^1_s - B^1_{\frac{[ms]}{m}} \right|^2 ds$, the classical exponential inequality gives

$$\mathbf{P}_1\left(M^*_1 \geqslant \delta\varepsilon^{-2},\ \langle M\rangle_1 \geqslant \varepsilon^{-2}\, m^{-\frac{1}{p}} \right) \leqslant \exp\left(-\frac{\delta^2\varepsilon^{-2}\, m^{\frac{1}{p}}}{2} \right),$$

while we also have

$$\mathbf{P}_1\left(\langle M\rangle_1 \geqslant \varepsilon^{-2}\, m^{-\frac{1}{p}} \right) \leqslant \mathbf{P}_1\left(\big\|B^1\big\|^2_{\frac{1}{p}}\, m^{-\frac{2}{p}} \geqslant \varepsilon^{-2}\, m^{-\frac{1}{p}} \right)$$

So the conclusion follows from the fact that the $\frac{1}{p}$-Hölder norm of B^1 has a Gaussian tail.   ▷

4.2. Freidlin-Wentzell large deviation theory for diffusion processes. All together, theorem on the rough path interpretation of Stratonovich differential equations, Lyons' universal limit theorem and the large deviation principle satisfied by the Brownian rough path prove the following basic result of Freidlin-Wentzell theory of large deviation for diffusion processes. Given some $\mathcal{C}_b^3$ vector fields $V_1, \ldots, V_\ell$ on $\mathbb{R}^d$, and $h \in H^1$, denote by y^h the solution to the well-defined controlled ordinary differential equation

$$dy_t^h = \varepsilon V_i\big(y_t^h\big) \circ dh_t^i.$$

THEOREM 36 (Freidlin-Wentzell). *Denote by $\mathbf{P}_\varepsilon$ the distribution of the solution to the Stratonovich differential equation*

$$dx_t = \varepsilon V_i(x_t) \circ dB_t^i,$$

started from some initial condition x_0. Given any $\frac{1}{p} < \frac{1}{2}$, one can consider $\mathbf{P}_\varepsilon$ as a probability measure on $\mathcal{C}^{\frac{1}{p}}\big([0,1],\mathbb{R}^d\big)$. Then the family $\mathbf{P}_\varepsilon$ satisfies a large deviation principle with good rate function

$$\mathrm{J}(z_\bullet) = \inf\big\{\mathrm{I}(h)\,;\, y_\bullet^h = z_\bullet\big\}.$$

5. Exercises on rough and stochastic analysis

Exercise 16 provides another illustration of the power of Lyons universal limit theorem and the continuity of the solution map to a rough differential equation, called the *Itô map*. It shows how to obtain a groundbreaking result of Stroock and Varadhan on the support of diffusion laws by identifying the support of the distribution of the Brownian rough path. Exercise 17 gives an interesting example of a rough path obtained as the limit of a 2-dimensional signal made up of a Brownian path and a delayed version of it. While the first level concentrates on a degenerate signal with identical coordinates and null area process as a consequence, the second level converges to a non-trivial function. Last, exercise 18 is a continuation of exercise 11 on the pairing of two rough paths.

16. Support theorem for the Brownian rough path and diffusion laws. We show in this exercise how the continuity of the Itô map leads to a deep result of Stroock and Varadhan on the support of diffusion laws. The reader unacquainted with this result may have a look at the poloshed proof given in the book by Ikeda and Watanabe [**18**] to see the benefits of the rough path approach.

a) Translating a rough path. Given a Lipschitz continuous path h and a p-rough path $\mathbf{a} = 1 \oplus a^1 \oplus a^2$, with $2 < p < 3$, check that we define another p-rough path setting

$$\tau_h(\mathbf{a})_{ts} := 1 \oplus \big(a_{ts}^1 + h_{ts}\big) \oplus \left(a_{ts}^2 + \int_s^t a_{us}^1 \otimes dh_u + \int_s^t h_{us} \otimes da_u^1 + \int_s^t h_{us} \otimes dh_u\right),$$

where the integral $\int_s^t h_{us} \otimes da_u^1$ is defined as a Young integral by the integration be parts formula

$$\int_s^t h_{us} \otimes da_u^1 := h_{ts} \otimes a_{ts}^1 - \int_s^t dh_u \otimes a_{us}^1.$$

b) Given any $\mathbb{R}^\ell$-valued coninuous path $x_\bullet$, denote as in section 1 by $x^{(n)}$ the piecewise linear coninuous interpolationof $x_\bullet$ on dyadic times of order n, and let $\mathbf{X}^{(n)}$ stand for its

associated rough path, for $2 < p < 3$. We define a map $\mathbf{X} : \mathcal{C}^0\big([0,1], \mathbb{R}^\ell\big) \to \big(\mathfrak{G}_\ell^{2,1}\big)^{[0,1]}$ setting

$$\pi_1\big(\mathbf{X}(x_\bullet)\big) := x_\bullet, \qquad \pi_2\big(\mathbf{X}(x_\bullet)\big)_t^{jk} := \limsup_n \int_0^t x_u^{(n),j} \otimes dx_u^{(n),k}.$$

So the random variable $\mathbf{X}(x_\bullet)$ is almost-surely equal to Stratonovich Browian rough path under Wiener measure $\mathbb{P}$.

(i) Show that one has $\mathbb{P}$-almost-surely

$$\mathbf{X}(x + h) = \tau_h\big(\mathbf{X}(x)\big)$$

for any Lipschitz continuous path h.

(ii) Prove that th law of the random variable $\tau_h \circ \mathbf{X}$ is equivalent to the law of $\mathbf{X}$ under $\mathbb{P}$.

Recall that the support of a probability measure on a topological space if the smallest closed setof full measure. We consider $\mathbf{X}$, under $\mathbb{P}$, as a $\mathcal{C}^{\frac{1}{p}}\big([0,1], \mathfrak{G}_\ell^2\big)$-valued random variable.

(iii) Prove that if $\mathbf{a}_\bullet$ is an element of the support of the law of $\mathbf{X}$ under $\mathbb{P}$, then $\tau_h \mathbf{a}$ as well.

c) (i) Use the same kind of arguments as in proposition 26 to show that one can ind an element $\mathbf{a}_\bullet$ in the support of the law of $\mathbf{X}$ under $\mathbb{P}$, and some Lipschitz coninuous paths $x_\bullet^{(n)}$ such that $\big\|\tau_{x^{(n)}} \mathbf{a}\big\|$ tends to 0 as $n \to \infty$.

(ii) Prove that the support of the law of $\mathbf{B}_{\bullet 0}^{\mathrm{Str}}$ in $\mathcal{C}^{\frac{1}{p}}\big([0,1], \mathfrak{G}_\ell^2\big)$ is the closure in $\frac{1}{p}$ Hölder topology of the set of of Lipschitz continuous paths.

d) **Stroock-Varadhan support theorem.** Let $\mathbf{P}$ stand for the distribution of the solution to the rough differential equation in $\mathbb{R}^d$

$$dx_t = V_i(x_t) \circ dB_t^i,$$

driven by Brownian motion and some $\mathcal{C}_b^3$ vector fields V_i. Justify that one can see $\mathbf{P}$ as a probability on $\mathcal{C}^{\frac{1}{p}}\big([0,1], \mathbb{R}^d\big)$. Let also write y^h for the solution to the ordinary differential equation

$$dy_t^h = V_i\big(y_t^h\big) dh_t^i$$

driven by a Lipschitz $\mathbb{R}^\ell$-valued path h. Prove that the support of $\mathbf{P}$ is the closure in $\mathcal{C}^{\frac{1}{p}}\big([0,1], \mathbb{R}^d\big)$ of the set of all y^h, for h ranging in the set of Lipschitz $\mathbb{R}^\ell$-valued paths.

17. Delayed Brownian motion. Let $(B_t)_{0 \leqslant t \leqslant 1}$ be a real-valued Brownian motion. Given $\epsilon > 0$, we define a 2-dimensional process setting

$$x_t = \big(B_{t-\epsilon}, B_t\big);$$

its area process

$$\mathbb{A}_{ts}^\epsilon := \frac{1}{2} \int_s^t \Big(B_{u-\epsilon, s-\epsilon} dB_u - B_{us} dB_{u-\epsilon}\Big)$$

is well-defined for $0 \leqslant t - s < \epsilon$, as $B_{\bullet - \epsilon}$ and $B_\bullet$ are independent on $[s, t]$ in that case.

1) Show that we define a rough path $\mathbf{X}^\epsilon$ setting

$$\mathbf{X}_t^\epsilon := \exp\big(x_t + \mathbb{A}_t^\epsilon\big) \in T_2^2.$$

2) Recall that d stands for the ambiant metric in T_2^2. Prove that one can find a positive constant a such that the nequality

$$\mathbb{E}\left[\exp\left(a\,\frac{d(\mathbf{X}_t^\epsilon, \mathbf{X}_s^\epsilon)^2}{t-s}\right)\right] \leqslant C < \infty$$

holds for a positive constant C independent of $0 < \epsilon \leqslant 1$ and $0 \leqslant s \leqslant t \leqslant 1$. As in section 1.2, it follows from Besov's embedding theorem that, for any $2 < p < 3$, the weak geometric Hölder p-rough path $\mathbf{X}^\epsilon$ has a Gaussian tail, with

$$\sup_{0 < \epsilon \leqslant 1} \mathbb{E}\left[\exp\left(a\|\mathbf{X}^\epsilon\|^2\right)\right] < \infty$$

for some positive constant a.

3) Define $\mathbf{1}$ as the vector of $\mathbb{R}^2$ with coordinates 1 and 1 in the canonical basis, and set

$$\mathbf{Y}_t := \exp\left(B_t\mathbf{1} - \frac{t}{2}\mathrm{Id}\right).$$

Write d_p for the distance on the set of Hölder p-rough paths given in definition 8. Prove that $\mathrm{d}_p\big(\mathbf{X}^\epsilon, \mathbf{Y}\big)$ converges to 0 in $\mathbb{L}^q$, for any $1 \leqslant q < \infty$.

18. Joint lift of a random and a deterministic rough path. Let $2 < p < 3$ and $\mathbf{X} = (X, \mathbb{X})$ be an $\mathbb{R}^d$-valued Hölder p-rough path. Denote by $\mathbf{B}$ the Itô Brownian rough path over $\mathbb{R}^\ell$. Given $j \in [\![1, d]\!]$ and $k \in [\![1, \ell]\!]$, tdefine he integral

$$\mathbb{Z}_{ts}^{jk} := \int_s^t X_{us}^k dB_u^j$$

as a genuine Itô integral, and define the integral $\int_s^t B_{us}^j dX_u^k$ by integration by part, setting

$$\mathbb{Z}_{ts}^{kj} := \int_s^t B_{us}^j dX_u^k := B_{ts}^j X_{ts}^k - \int_s^t X_{us}^k dB_u^j.$$

Prove that one defines a Hölder p-rough path $\mathbf{Z}$ over $(X, B) \in \mathbb{R}^{d+\ell}$ defining the (jk)-component of its second order level, as equal to $\mathbb{X}^{jk}$ if $1 \leqslant j, k \leqslant d$, equal to $\mathbb{B}^{jk}$ if $d + 1 \leqslant j, k \leqslant d + \ell$, and by the above formulas otherwise.

CHAPTER 6

Looking backward

1. Summary

It is now time to forget the details and summarize the main ideas.

Chapter 2 provides a toolbox for constructing flows on Banach spaces from approximate flows. The interest of working with this notion comes from the fact that approximate flows pop up naturally in a number of situations, more or less under the form of "numerical schemes", as the step-1 Euler scheme for ordinary differential equations encountered in exercise 1, or the higher order Milstein-type schemes μ_{ts} used in chapter 4 to define solutions to rough differential equations. Rough paths appear in that setting as coefficients in the numerical schemes, and as natural generalizations of multiple integrals in some Hölder scales space. The miracle that takes place here is essentially algebraic and rests on the fact that solving an ordinary differential equation is algebraically very close to an exponentiation operation. This echoes the fact that the tensor space $T_\ell^{[p]}$ in which rough paths live also has natural notions of exponential and logarithm. As a matter of fact, this "pairing" ODE-exponential-rough paths works in exactly the same way with the branched rough paths introduced in [16] by Gubinelli in order to deal with rough differential equations driven by non-weak geometric rough paths.

We have concentrated in this course on one approach to rough differential equations and rough paths. There are other approaches, with their own benefits, to start with Lyons' original formulation of his theory, as exposed in Lyons' seminal article [20] or his book [2] with Qian. Its core concept is a notion of rough integral which associates a rough path to another rough path. As you may guess, it can be shown to be a continuous functions of both its integrand and integrator. In that setting, as solution path to a rough differential equation is a fixed point to an integral equation *in the space of rough paths*; it was first solved using a Picard iteration process. Two crucial features of Lyons' original formulation were spotted by Gubinelli and Davie. The first level of a solution path $x_\bullet$ to a rough differential equation locally look like the first level of $\mathbf{X}$, and it suffices to know $x_\bullet$ and $\mathbf{X}$ to get back the entire rough path solution to the rough differential equation in Lyons' sense. This led Gubinelli to the introduction of the notion of controlled paths, which have far reaching applications to difficult problems on stochastic partial differential equations (SPDEs), as illustrated in the recent and brilliant works of Gubinelli and his co-authors. It also was one of the seeds of Hairer's groundbreaking theory of regularity structures [21], which enabled him to construct a robust solution theory for some important up to now ill-posed SPDEs coming from physics deep problems. The forthcoming lecture notes [5] by Friz and Hairer provides a very nice introduction to Gubinelli's point of view on rough paths theory.

On the other hand, Davie uncovered in [22] the fact that one can characterize the first level of a solution path to a rough differential equation in Lyons' sense in terms of

numerical schemes of Milstein type. Has was able in that setting to prove sharp well-posedness and existence results, under essentially optimal regularity conditions on the driving vector fields, for p-rough paths with $2 \leqslant p < 3$. No notion of integral is needed in this approach, and we only work with $\mathbb{R}^d$-valued paths, as opposed to Lyons' formulation.

His ideas were reworked and generalized by Friz and Victoir [23], who defined solution paths to rough differential equations driven by some rough path $\mathbf{X}$ as limits of solution paths to controlled ordinary differential equations, in which the rough path canonically associated with the (smooth or absolutely continuous) control converges in a rough paths sense to $\mathbf{X}$. The book [3] by Friz and Victoir provides a thorough account of their approach. See also the short 2009 lecture notes [24] of Friz. (My presentation in section 4 of the material on Freidlin-Wentzell large deviation theory follows his approach.) The point of view presented in these notes builds on Davie's approach and on the inspiring work [7] of Feyel and de la Pradelle; it is mainly taken from the article [6].

I hope you enjoyed the tour.[1]

---

[1]Please do not hesitate to send me any comments or suggestions to the email address ismaelbailleul@univ-rennes1.fr.

Bibliography

[1] Lyons, T.J. and Caruana, M. and Lévy, Th. Differential equations driven by rough paths. *Lecture Notes in Mathematics*, 1908, Springer 2007.

[2] Lyons, T. and Qian, Z. System control and rough paths. *Oxford Mathematical Monographs*, Oxford University Press 2002.

[3] Friz, P. and Victoir, N. Multidimensional stochastic processes as rough paths. *CUP, Cambridge Studies in Advanced Mathematics*, 120, 2010.

[4] Baudoin, F., Rough paths theory. *Lecture notes*, http://fabricebaudoin.wordpress.com/category/rough-paths-theory/, 2013.

[5] Friz, P. and Hairer, M., A short course on rough paths. *Lect. Notes Math.*, www.hairer.org/notes/RoughPaths.pdf, 2014.

[6] Bailleul, I., Flows driven by rough paths. arXiv:1203.0888, 2013.

[7] Feyel, D. and de La Pradelle, A. Curvilinear integrals along enriched paths. *Electron. J. Probab.*, 11:860–892, 2006.

[8] Feyel, D. and de La Pradelle, A. and Mokobodzki, G. A non-commutative sewing lemma. *Electron. Commun. Probab.*, 13:24–34, 2008.

[9] Gubinelli, M., Controlling rough paths. *J. Funct. Anal.*, 216:86–140, 2004.

[10] Lyons, T.. Differential equations driven by rough signals. *Rev. Mat. Iberoamericana*, 14 (2):215–310, 1998.

[11] Montgomery, R., A tour of subriemannian geometries, their geodesics and applications. *Mathematical Surveys and Monographs*, 91, 2002.

[12] Lejay, A., Yet another introduction to rough paths. *Séminaire de Probabilités*, LNM 1979:1–101, 2009.

[13] Chen, K.T. Iterated path integrals. *Bull. Amer. Math. Soc.*, 83(5):831–879, 1977.

[14] Gubinelli, M., Ramification of rough paths. *J. Diff. Equations*, 248:693–721, 2010.

[15] Hairer, M. and Kelly, D., Geometric versus non-geometric rough paths. arXiv:1210.6294, 2012.

[16] Gubinelli, M., Ramification of rough paths. *J. Diff. Eq.*, 248: 693–721, 2010.

[17] Friz, P. and Victoir, N. A note on the notion of geometric rough paths. *Probab. Theory Related Fields*, 136 (3):395–416, 2006.

[18] Ikeda, N. and Watanabe, S. Stochastic differential equations and diffusion processes. *North-Holland Publishing Company*, 1988.

[19] Kallenberg, O., Foundations of modern probabiity. *Probability and its Applications*, Srpinger, 2002.

[20] Lyons, T., Differential equations driven by rough signals. *Rev. Mat. Iberoamericana*, 14 (2):215–310, 1998.

[21] Hairer, M., A theory of regularity structures. arXiv: 1303.5113, 2013.

[22] Davie, A., Differential equations driven by rough paths: an approach via discrete approximation. *AMRX*, 2008.

[23] Friz, P. and Victoir, N., Euler estimates for rough differential equations. *J. of Diff. Eq.*,244(2): 388–412, 2008.

[24] Friz, P., Mini-course on rough paths (TU Wien, 2009). *http://page.math.tu-berlin.de/ friz/ Vienna2009/roughpaths.htm*, 2009.

2. A guided tour of the litterature

The litterature on rough paths theory is increasing rapidly. To help you find your way in this bush, I comment below on a few references whose reading may be helpful to get a better view of the domain. Reference numbers refer to the bibliography following this paragraph, while stared references refer to the above bibliography.

Lecture notes. Here are collected a few references that aim at giving a pedagogical presentation of rough paths theory, from different point of views.

  - Lejay gave in [1] a self-contained and easily accessible account of the theory of Young differential equations, which correspond to rough differential equations driven by p-rough paths, with $1 < p < 2$.

  - The lecture notes [2] by Lejay provides a well-motivated and detailed study of the algebraic setting in which rough paths theory *needs* to be formulated. It is easily readable.

  - The approach of Friz-Victoir to rough differential equations, as described in remark 17, was put forward in [23]*; it is developed thouroughly in their monograph [3]*. The lecture notes [24]* by Friz, and [4]* by Baudoin, provide an easy access to that approach.

  - We warmly recommend the reading of the forthcoming lecture notes [5]* of Friz and Hairer on the theory of rough paths and rough differential equations seen from the point of view of controlled paths. Although it does not lend itself to an easy access when the roughness index p is greater than 3, this approach is the seed of the very exciting development of a new framework for handling SPDEs which were previously untractable. Have a look for instance at the (hard) work [3, 4] of Hairer on regularity structures, or the somewhat more "down-to-earth" work [5] of Gubinelli, Imkeller and Perkowski to see how ideas from controlled paths can enable you to do some forbidden operation: multiplying two distributions!

Historical works. I have chosen to put forward here a few references that illustrate the development of the theory.

  - Lyons' amazing seminal work [6] is a must.

  - One owes to Strichartz [7] a far reaching generaiztion of the well-known *Baker-Campbell-Dynkin-Hausdorff* formula expressing the multiplication in a Lie group as an operation in the Lie algebra. This fantastic paper was the basis of basis of groudbreaking works by Castell [8], Ben Arous and others on *Taylor expansions* for stochastic differential equations, and can somehow be seen as a precursor to the approach to rough differential equations put forward in this course.

  - The other paper that inspired our flow-based approach is the *sewing lemma* proved in [7, 8] by Feyel, de la Pradelle and Mokobodzki.

  - The work [22]* by Davie showing that one could understand rough differential equations in terms of numerical schemes – or Taylor expansions – was also instrumental in the development of Friz-Victoir's approach to the subject, as well as to the present approach.

Today. The theory of rough paths is presently experiencing a fantastic development in all sorts of directions. Just a few of them: Malliavin related business, differential geometry and machine learning, to testify of the diversity of directions that are actively being investigated.
.2cm

- - There has been much industry in proving that one can use Malliavin calculus methods and rough paths theory to show that solutions of rough differential equations driven by Gaussian rough paths have a density at any fixed time under some bracket-type conditions on the driving vector fields and some non-degeneracy conditions on the Gaussian noise. For two landmark results in this direction, see the works [**9**] of Cass and Friz, and what may be a temporarily final point [**10**] by Cass, Hairer, Litterer and Tindel.

- - Rough path theory has an inherent geometric content built in, to start with the nilpotent Lie group on which rough paths live. Given that most all of physics takes place on finite or infinite dimensional manifolds (configuration spaces), it is natural to try and give an intrinsic notion of rough path on a manifold. Starting with the seminal work [**11**] of Cass, Litterer and Lyons that makes a first step in this direction, this important question is being investigated. See the work [**12**] of Cass, Driver and Litterer, for a nice reworking of the ideas of [**11**], giving an intrinsic notion of rough path on a compact finite dimensional manifold, and my own work [**13**] for a general framework for dealing with rough integrators on Banach manifolds.

- - Lyons' group in Oxford is presently exploring the potential application of the use of signature (the set of all iterated integrals of a rough path) to investigate learning questions! See for instance the works [**14**] by Levin, Lyons, Ni and [**15**], by Gyurko, Lyons, Kontkowski and Field.

Bibliography

[1] Lejay, A., Controlled differential equations as Young integrals: a simple approach. *J. Diff. Equations*, 249:1777–1798, 2010.

[2] Lejay, A., Yet another introduction to rough paths. *Sém. Probab. XLII, LNM 1979*, 1–101, 2009.

[3] Hairer, M., A theory of regularity structures. arXiv:1303.5113, 2013.

[4] Hairer, M., Introduction to Regularity Structures. arXiv:1401.3014, 2014.

[5] Gubinelli, M. and Imkeller, P. and Perkowski, N., Paracontrolled distributions and singular PDEs. arXiv:1210.2684, 2013.

[6] Lyons, T., Differential equations driven by rough signals. *Rev. Mat. Iberoamericana*, 14(2):215–310, 1998.

[7] Strichartz, R., The Campbell-Baker-Hausdorff-Dynkin formula and solutions of differential equations. *J. Funct. Analysis*, 72:320–345, 1987.

[8] Castell, F., Asymptotic expansion of stochastic flows. *Prob. Th. Rel. Fields*, 96:225–239, 1993.

[9] Cass, T. and Friz, P., Densities for rough differential equations under Hormander's condition. *Annals of Mathematics*, 171:2115–2141, 2010.

[10] Cass, T. and Hairer, M. and Litterer, C. and Tindel, S., Smoothness of the density for solutions to Gaussian Rough Differential Equations. arXiv:1209.3100, 2012.

[11] Cass, T. and Litterer, C. and Lyons, T., Rough paths on manifolds. *New trends in stochastic analysis and related topics*, 33–88, World Scientific Publishing, 2011.

[12] Cass, T. and Driver, B. and Litterer, C., Constrained rough paths. arXiv:1402.4529, 2014.

[13] Bailleul, I., Rough integrators on Banach manifolds. arXiv:1403.3285, 2014.

[14] Levin, D. and Lyons, T. and Ni, H., Learning from the past, predicting the statistics for the future, learning an evolving system. arXiv:1309.0260, 2013.

[15] Gyurko, L. and Lyons, T. and Kontkowski, M. and Field, J., EXtracting information from the signature of a financial stream. arXiv:1307.7244, 2013.

Index