On the definition of a solution to a rough differential equation

I. BAILLEUL[1][2]

Abstract. There are several ways of defining what it means for a path to solve a rough differential equation. The most commonly used notion is due to Davie; it involves a Taylor expansion property that only makes sense a priori in a given coordinate system. Bailleul's definition [2] is coordinate independent. Cass and Weidner [10] recently proved that the two definitions are actually equivalent, using deep algebraic insights on rough paths. We provide in this note an algebraic-free elementary short proof of this fact.

Résumé. Différentes notions de solution d'une équations différentielle rugueuses sont disponibles dans la littérature. La notion la plus utilisée est due à Davie ; elle met en jeu un développement de Taylor, dont le sens est lié à un choix de coordonnées. La définission donnée par Bailleul dans [2] ne dépend pas d'un choix de coordonnées. Cass et Weidner [10] ont récemment démontré que les deux définitions sont en fait équivalentes, en s'appuyant sur des résultats algébriques élaborés. On donne dans cette note une démonstration courte et élémentaire de ce fait.

1 – Main result

a) Setting – Rough paths theory was introduced by T. Lyons in [21] as a theory of controlled differential equations

$$\dot{x}_t = \mathrm{F}(x_t)\dot{h}_t \tag{1.1}$$

with $\mathrm{F} \in L(\mathbb{R}^\ell, \mathbb{R}^d)$, and a non-differentiable control h with values in $\mathbb{R}^\ell$. The path x takes here its values in $\mathbb{R}^d$. Variants of the Cauchy-Lipschitz theorem gives the well-posed character of this equation for an absolutely continuous control, and L.C. Young extended this analysis to controls that are α-Hölder, for $\alpha > \frac{1}{2}$; the solution path is then a continuous function of the control $h \in C^\alpha([0,1], \mathbb{R}^\ell)$, under proper regularity and boundedness conditions on F. Nothing better than that can be done in a deterministic setting in the Hölder class, on account of the fact that since the solution path x_t is expected to be no better than α-Hölder, so is $\mathrm{F}(x_t)$, so the equation involves making sense of the product $\mathrm{F}(x_t)\dot{h}_t$ of an α-Hölder function with an $(\alpha-1)$-Hölder distribution, either directly or in its integral form. It is known that no such product can be defined as a continuous function of its two arguments when the sum of the Hölder regularity exponents add up to a non-positive real number – see e.g. [4]. Lyons' deep insight was to realize that what really controls the dynamics in (1.3) is not the path h, but rather the data of the path together with a number of its iterated integrals

$$\int h_s^i dh_s^j, \quad \iint_{r \leqslant s} h_r^i dh_r^j dh_s^k, \ldots$$

[2]Univ. Rennes, CNRS, IRMAR - UMR 6625, F-35000 Rennes, France.

for all possible indices $1 \leqslant i, j, k, \cdots \leqslant \ell$. The rougher the control h is, the more iterated integrals you need. This can be understood from the fact that these iterated integrals are the coefficients that appear in a Taylor expansion of the solution to equation (1.3) when the control is smooth, and this leads to the notion of solution adopted in [16, 2] by Davie and Bailleul. There is no miracle though, and one cannot still make sense of these iterated integrals when the control is not sufficiently regular. Lyons' considerable feat was to extract from these 'non-existing' quantities the analytic and algebraic properties that they should satisfy and to work directly with objects enjoying these properties as new controls. These are the weak geometric Hölder p-rough paths that we encounter below. We refer the reader to the gentle introductions [22, 17, 3] for different points of views on rough differential equations; one can learn everything from scratch in the first 19 pages of [2].

b) Weak geometric Hölder p-rough paths and rough differential equations – We refer the reader to [21, 5, 2] for basics on rough paths and recall all we need in this section. Pick an integer $p \geqslant 1$. Denote by $(e_i)_{i=1..\ell}$ the canonical basis of $\mathbb{R}^\ell$, and by $T_\ell^{[p]} = \bigoplus_{i=1}^\ell (\mathbb{R}^\ell)^{\otimes i}$ the truncated tensor algebra over $\mathbb{R}^\ell$; it is equipped with the Lie bracket operation $[a, b] := ab - ba$, and endowed with a norm

$$\|a\| := \sum_{i=1}^{[p]} |a^i|^{1/i},$$

for $a =: \bigoplus_{i=1}^{[p]} a^i$; denote by $(a^I)_{I \in [\![1,\ell]\!]^k, 0 \leqslant k \leqslant [p]}$ the coordinates of a generic element a of $T_\ell^{[p]}$ in the canonical basis of $T_\ell^{[p]}$. We sometimes write $a = (a^i)_{i=1..[p]}$ instead of $a = \bigoplus_{i=1}^{[p]} a^i$. Consider $\mathbb{R}^\ell$ as a subset of $T_\ell^{[p]}$. The Lie algebra $\mathfrak{g}_\ell^{[p]}$ generated by $\mathbb{R}^\ell$ within $T_\ell^{[p]}$ is called the p-step free nilpotent Lie algebra, and its exponential $G_\ell^{[p]}$ is called the p-step free nilpotent Lie group. Recall that a weak geometric Hölder p-rough path $\mathbf{X}$ over $\mathbb{R}^\ell$ is a $\frac{1}{p}$-Hölder path with values in $G_\ell^{[p]}$. Set $\mathbf{X}_{ts} := \mathbf{X}_s^{-1} \mathbf{X}_t$, so we have $\|\mathbf{X}_{ts}\| \leqslant O(|t - s|^{1/p})$; this is equivalent to having

$$\|\mathbf{X}_{ts}^i\| \leqslant O(|t - s|^{i/p}).$$

The archetype of a weak geometric Hölder p-rough path is given by the canonical lift of a smooth $\mathbb{R}^\ell$-valued path h under the form

$$H_t := \left(1, h_t - h_0, \int_0^t (h_{u_1} - h_0) \, du_1, \ldots, \int_{0 \leqslant u_{[p]} \leqslant \cdots \leqslant u_1 \leqslant t} dh_{u_{[p]}} \cdots dh_{u_1} \right); \qquad (1.2)$$

it can be seen, after Chen [13], that H takes values in $G_\ell^{[p]}$ and that

$$H_{ts} := H_s^{-1} H_t = \left(1, h_t - h_s, \int_s^t (h_{u_1} - h_0) \, du_1, \ldots, \int_{s \leqslant u_{[p]} \leqslant \cdots \leqslant u_1 \leqslant t} dh_{u_{[p]}} \cdots dh_{u_1} \right).$$

Denote by $[\![1, \ell]\!]$ the set of integers between 1 and ℓ. Define inductively on the size of the tuple $J := \{i_1, I\} \in [\![1, \ell]\!]^{k+1}$, the element

$$e_{[J]} = [e_{i_1}, e_{[I]}]$$

of $\mathfrak{g}_\ell^{[p]}$. Given $0 \leqslant s \leqslant t \leqslant T$, set

$$\mathbf{\Lambda}_{ts} := \log \mathbf{X}_{ts} = \sum_{\substack{I \in [\![1,\ell]\!]^k \\ 0 \leqslant k \leqslant [p]}} \Lambda_{ts}^I e_{[I]} \in \mathfrak{g}_\ell^{[p]}.$$

Let $F = (V_1, \ldots, V_\ell)$ stand for a collection of ℓ smooth enough vector fields on $\mathbb{R}^d$, and $\mathbf{X}$ stand for a weak geometric Hölder p-rough path over $\mathbb{R}^\ell$. There are several definitions of a solution to the rough differential equation

$$dz_t = F(z_t) \, d\mathbf{X}_t. \tag{1.3}$$

Besides Lyons' original definition [21] and Gubinelli's formulation in terms of controlled paths [18], Davie gave in [16] a formulation of a solution in terms of Taylor expansion for the time increment of a solution. We identify freely a vector field to a first order differential operator via the expression

$$V_i f = (Df)(V_i)$$

for any C^1 function f on $\mathbb{R}^d$. For a tuple $I = (i_1, \ldots, i_k) \in [\![1, \ell]\!]^k$, we write $|I| := k$, and define a kth order differential operator setting

$$V_I = V_{i_1} \cdots V_{i_k},$$

that is to say $V_I f = V_{i_1}(\cdots (V_{i_k} f))$. With a slight abuse of notation, we write $V_I(x)$ for $(V_I \text{Id})(x)$. An $\mathbb{R}^d$-valued path $(z_t)_{0 \leqslant t \leqslant T}$ defined on some finite time interval $[0, T]$ is a solution to the rough differential equation (1.3) in the sense of Davie if one has

$$z_t = z_s + \sum_{I; |I| \leqslant [p]} X_{ts}^I (V_I)(z_s) + O(|t - s|^a), \tag{1.4}$$

for any $0 \leqslant s \leqslant t \leqslant T$, for some exponent $a > 1$. The function $O(\cdot)$ is allowed to depend on F. Two paths satisfying that condition with different exponents $a, a' > 1$ coincide if they start from the same point. This definition is a priori not well suited to make sense of a solution to a rough differential taking values in a manifold – an expression like $m + V(m)$, for a point m of a manifold M and a vector $V(m) \in T_m M$, has for instance no intrinsic meaning. Bailleul gave in [2] a more general notion of solution to a rough differential equation by requiring from a potential solution that it satisfies the estimate

$$f(z_t) = f(z_s) + \sum_{I; |I| \leqslant [p]} X_{ts}^I (V_I f)(z_s) + O_f(|t - s|^a),$$

for any real-valued sufficiently regular function f defined on the state space. The function $O_f(\cdot)$ is allowed to depend on f. The exponent a that appears here may be different from the exponent that appears in equation (1.4); this makes no difference as long as $a > 1$, as noted above. This definition makes perfect sense in a manifold setting, and taking f to be a coordinate function in the above equation makes it is clear that a solution in this sense to an equation with values in $\mathbb{R}^d$ is also a solution in the sense of Davie.

c) Main result – Cass and Weidner proved in Theorem 5.3 of their recent work [10] that the two notions are actually equivalent using some deep insights from algebra based on the use of Grossman-Larson and Connes-Kreimer Hopf algebras. That point can be proved by elementary means. We assume as usual that the vector fields V_i in (1.3) are C^γ, for some $\gamma > p$.

1. Theorem − *An $\mathbf{R}^d$-valued path z is a solution to the rough differential equation* (1.3) *in the sense of Davie if and only if it is a solution of that equation in the sense of Bailleul.*

Proof − We only need to provee that solutions to equation (1.3) in Davie' sense are solutions in Bailleul' sense. Given a globally Lipschitz vector field V on $\mathbb{R}^d$, we denote by $\exp(V)$ the time 1 map of the differential equation

$$\dot{y}_u = V(y_u),$$

that associates to x the value at time 1 of the solution to the equation started from x. Also, define inductively on the size of the tuple $\{i_1, I\} \in [\![1, \ell]\!]^{k+1}$, the vector fields

$$V_{[\{i_1, I\}]} := \big[V_{i_1}, V_{[I]}\big],$$

starting with $V_{[i]} := V_i$. Given $0 \leqslant s \leqslant t \leqslant T$, recall we write $\boldsymbol{\Lambda}_{ts}$ for $\log \mathbf{X}_{ts}$; this is an element of $\mathfrak{g}_\ell^{[p]}$. It is proved in Proposition 9 of [2] that one has

$$\left\| \exp\left(\sum_{I; |I| \leqslant [p]} \Lambda_{ts}^I V_{[I]} \right) - \left\{ \mathrm{Id} + \sum_{I; |I| \leqslant [p]} X_{ts}^I(V_I) \right\} \right\|_{L^\infty} \leqslant O\left(|t-s|^{\frac{\gamma}{p}} \right) \tag{1.5}$$

for a $O(\cdot)$ depending only on the vector fields V_i. (See also [6] for a similar statement.) A path z is thus a solution to the rough differential equation (1.3) in the sense of Davie if and only if

$$z_t = \exp\left(\sum_{I; |I| \leqslant [p]} \Lambda_{ts}^I V_{[I]} \right)(z_s) + O\big(|t-s|^{a'}\big),$$

for some exponent $1 < a' \leqslant \frac{\gamma}{p}$. Let y stand now for a solution path to the ordinary differential equation

$$\dot{y}_u = \sum_{I; |I| \leqslant [p]} \Lambda_{ts}^I V_{[I]}(y_u),$$

started from z_s at time 0. We use in the computation below the notation $O\big(|t-s|^{a'}\big)$ for a function whose value may change from line to line. For a γ-Hölder real-valued function f on $\mathbb{R}^d$ one has

$$f(z_t) = f(y_1) + O\big(|t-s|^{a'}\big)$$

$$= f(z_s) + \int_0^1 \frac{d}{du} f(y_u)\, du + O\big(|t-s|^{a'}\big)$$

$$= f(z_s) + \sum_{|I_1| \leqslant [p]} \Lambda_{ts}^{I_1} \int_0^1 V_{I_1}(y_{u_1})\, du_1 + O\big(|t-s|^{a'}\big)$$

$$= f(z_s) + \sum_{|I_1| \leqslant [p]} \Lambda_{ts}^{I_1} V_{I_1}(z_s) + \sum_{\substack{|I_1| \leqslant [p], |I_2| \leqslant [p] \\ |I_1| + |I_2| \leqslant [p]}} \int_0^1 V_{I_2} V_{I_1}(y_{u_2})\, du_2 du_1 + O\big(|t-s|^{a'}\big),$$

and, after repeating $[p]$ times the same computation giving a function as the integral of its derivative, one eventually has

$$f(z_t) = f(z_s) + \sum_{k=1}^{[p]} \frac{1}{k!} \Lambda_{ts}^{I_1} \cdots \Lambda_{ts}^{I_k} \big(V_{I_k} \cdots V_{I_1}\big)(z_s) + O\big(|t-s|^{a'}\big)$$

$$= f(z_s) + \sum_{I; |I| \leqslant [p]} X_{ts}^I(V_I f)(z_s) + O_f\big(|t-s|^{a'}\big),$$

from the fact that $\boldsymbol{\Lambda}_{ts} = \log \mathbf{X}_{ts}$. We repeat that computation in Section 2 in a more general setting. $\qquad \triangleright$

The following remarks emphasize the robust character of the above proof.

- One can weaken the regularity assumptions on the vector fields V_i by only requiring that they are $C^{[\gamma]}$ in the usual sense and the $[\gamma]$-derivative of V_i is $(\gamma - [\gamma])$-Hölder continuous, without requiring that the V_i or their derivatives be bounded. Indeed, given a solution z of (1.3) in the sense of Davie, and a time s, the classical Cauchy

Lipschitz theory ensures that $\exp\left(\sum_{I;|I|\leqslant[p]}\mathbf{\Lambda}_{ts}V_{[I]}\right)(z_s)$ is well-defined for t close enough to s, possibly depending on z_s.

- For a non-necessarily Hölder weak geometric p-rough path $\mathbf{X}$ controlled by a general control $w(s,t)$ rather than by $|t-s|^{1/p}$, one replaces $(t-s)^{1/p}$ by $w(s,t)$ in the definitions of a solution to a rough differential equation driven by $\mathbf{X}$, and requires in both definitions that the remainder is of order $w(s,t)^{ap}$, for some exponent $a>1$. The proof of Theorem 1 remains the same. The equivalence problem for the different definitions of a solution to a rough differential equation is further explored in the very recent work [7] of Brault and Lejay.

- Let E stand for a Banach space. For dynamics driven by an E-valued rough paths, following [9] and [10], one needs to define $T^{(n)}(E)$ as the completion of the truncated algebraic tensor algebra with respect to a system of cross (semi)norms. One then defines the n-step free Lie algebra over E as the closure in $T^{(n)}(E)$ of the algebraic free Lie algebra. The n-step free nilpotent Lie group is then indeed the image by the exponential map of the closed Lie algebra. An appropriate formalism where vector fields are indexed by the elements of E is required, as in [1, 6] or [10]. Identity (1.5) is proved in this setting in [1] by elementary means similar to their finite dimensional analogues. (We erronously worked in the non-complete versions of the Lie and truncated tensor algebras in [1]. The complete setting should be adopted, and nothing is changed to the story told in [1] in this extended setting. Thanks to T. Cass for pointing this out.)

- One of the nice points of Cass and Weidner's work [10] is the fact that they can also handle differential equations driven by a more general notion of rouhg path called branched rough paths. In a finite dimensional setting, one can appeal to Hairer and Kelly's result [20] stating that a solution to a rough differential equation driven by a branched rough path over $\mathbb{R}^\ell$ is also the solution to a differential equation driven by a weak geometric rough path over a larger space $\mathbb{R}^{\ell'}$, with driving vector fields built from the initial V_i. The equivalence of the two notions of solution for equations driven by weak geometric rough paths implies the equivalence of the two notions of solution for the equation driven by a branched rough path, since the Taylor expansion of $f(y_t)$ from the branched rough path and from the weak geometric rough path point of views coincide, by Theorem 5.8 in [20]. We give a self-contained proof of the equivalence of a Davie-type and Bailleul-type notion of a solution to a rough differential equation driven by a branched rough path in Section 2.

2 − Branched rough paths and their associated flows in a nutshell

We prove in this section the following analogue of Theorem 1.

2. Theorem − *An $\mathbf{R}^d$-valued path z is a solution to a rough differential equation driven by a branched rough path in the sense of Davie if and only if it is a solution of that equation in the sense of Bailleul.*

Theorem 2 gives an alternative proof of Theorem 5.3 of [10] in the setting of dynamics driven by branched rough paths. We include in this section a short self-containted introduction to finite dimensional branched rough paths before proving Theorem 2 in Section

2.2. We refer to Gubinelli's original paper [19] and Hairer and Kelly's paper [20] for alternative accounts on branched rough paths in a finite dimensional setting, and to [10] for an account of the theory in an infinite dimensional setting.

What are branched rough paths good for? Together with Equation (1.2) giving the lift of a smooth $\mathbb{R}^\ell$-valued path to the $[p]$-step free nilpotent Lie group, the definition of a weak geometric rough path **X** makes it clear that the higher order levels of **X** play the role of the non-existing integrals $\int X^i dX^j, \iint X^i dX^j dX^k$, etc. The archetypal example of a non-smooth weak geometric p-rough path is given by realizations of the Brownian rough path $\mathbf{B} = (B, \mathbb{B})$, for $2 < p < 3$, defined by

$$\mathbb{B}_{ts}^{jk} := \int_s^t (B_u - B_s) \circ dB_u.$$

Note the use of Stratonovich integration. If one defines instead the iterated integral $\mathbb{B}$ with an Itô integral, the corresponding object **B** no longer takes values in the $[p]$-step free nilpotent Lie group $G_\ell^{[p]}$. This is related to the fact that Itô integration procedure does not satisfy the usual chain rule. A different algebraic setting is needed to handle the integration theory of such controls; this is precisely what branched rough paths provide us with. The name "branched rough paths" come from the fact that these objects are indexed by trees.

2.1 − Branched rough paths

1. The starting point − Consider an ordinary controlled $\mathbb{R}^d$-valued differential equation

$$\dot{x}_r = \sum_{i=1}^\ell V_i(x_r) \dot{h}_r^i$$

driven by some smooth $\mathbb{R}^\ell$-valued path h. Given any tuple $I \in [\![1, \ell]\!]^k$, write H_{ts}^I for the iterated integral $\int_{s \leqslant s_1 \leqslant \cdots \leqslant s_k \leqslant t} dh_{s_1}^{i_1} \ldots dh_{s_k}^{i_k}$. Branched rough path appears naturally if one expands in the Taylor formula

$$f(x_t) = f(x_s) + \sum_{|I| \leqslant k_0} H_{ts}^I (V_I f)(x_s) + O\big(|t - s|^{k_0+1}\big) \tag{2.1}$$

all the iterated derivatives $V_I f$ in terms of the derivatives of its different terms. We have for instance

$$V_1\big(V_2 f\big)(x) = (D_x f)\big((V_1 V_2)(x)\big) + (D_x^2 f)\big(V_1(x), V_2(x)\big). \tag{2.2}$$

The best way to represent the resulting sum in (2.1) is to index it by labelled rooted trees. Let $\mathcal{T}$ stand for the set of possibly empty rooted labelled trees, with labels in $\{1, \ldots, \ell\}$. Given some trees $\tau_1, \ldots, \tau_k$ in $\mathcal{T}$ and $a \in \{1, \ldots, \ell\}$, denote by $[\tau_1 \ldots \tau_k]_a$ the element of $\mathcal{T}$ obtained by attaching the trees $\tau_1, \ldots, \tau_k$ to a new root with label a. Any element of $\mathcal{T}$ can be constructed in this way starting from the empty tree **1**. We denote by $|\tau|$ the number of vertices of a tree τ, and define its symmetry factor $\sigma(\tau)$ recursively

$$\sigma\big([\tau_1^{n_1} \ldots \tau_k^{n_k}]_a\big) = n_1! \ldots n_k! \, \sigma(\tau_1)^{n_1} \ldots \sigma(\tau_k)^{n_k},$$

where $\tau_1, \ldots, \tau_k$ are distinct trees with respective multiplicities $n_1, \ldots, n_k$. This number does not depend on the label a. Denote by $\langle \mathcal{T} \rangle$ the real vector space spanned by $\mathcal{T}$ and by $\langle \mathcal{T} \rangle^*$ its algebraic dual, identified with $\langle \mathcal{T}^* \rangle$, for some copy $\mathcal{T}^*$ of $\mathcal{T}$.

Set $V(\mathbf{1}^*) := 0$. and define recursively, for $a \in [\![1, \ell]\!]$ and $\tau_1, \ldots, \tau_n$ in $\mathcal{T}$, *vector fields* on $\mathbb{R}^d$ indexed by dual trees

$$V(\bullet_a^*) := V_a, \quad V\big(([\tau_1 \ldots \tau_n]_a)^*\big) := \frac{1}{\sigma\big([\tau_1 \ldots \tau_n]_a\big)} (D^n V_a)\big(V(\tau_1^*), \ldots, V(\tau_n^*)\big); \tag{2.3}$$

vector fields on $\mathbb{R}^d$ are seen as first order differential operators. We define *differential operators* indexed by dual forests setting

$$V\big(\tau_1^* \cdots \tau_n^*\big)f = \frac{1}{\sigma\big(\tau_1 \ldots \tau_n\big)}\big(D^n f\big)\big(V(\tau_1^*), \ldots, V(\tau_n^*)\big), \tag{2.4}$$

with $\sigma\big(\tau_1 \ldots \tau_n\big) := \sigma\big([\tau_1 \ldots \tau_n]_a\big)$, for any $1 \leqslant a \leqslant \ell$. Set $|\tau_1 \cdots \tau_n| = |\tau_1| + \cdots + |\tau_n|$, and denote by $\varphi = \tau_1 \cdots \tau_n$, a generic forest, and by $\mathcal{F}$ the set of these forests. In those terms, Equation (2.1) rewrites

$$f(x_t) = f(x_s) + \sum_{\varphi \in \mathcal{F}; |\varphi| \leqslant k_0} H_{ts}^\varphi \big(V(\varphi^*)f\big)(x_s) + O\big(|t - s|^{k_0+1}\big). \tag{2.5}$$

The coefficients H_{ts}^φ are linear combinations of the initial coefficients H_{ts}^I. One has actually the recursive definition

$$H_{ts}^{\bullet a} := h_t^a - h_s^a, \qquad H_{ts}^{[\tau_1 \ldots \tau_n]_a} = \int_s^t \prod_{i=1}^n H_{us}^{\tau_i} \, dh_u^a; \tag{2.6}$$

formulas (2.5) and (2.6) go back to Cayley' seminal work [11]. In the same way as the family of iterated integrals of h lives in an algebraic structure, the free nilpotent Lie group, the family of all H^τ also lives in some algebraic structure, the dual of a Hopf algebra. We refer the reader to Sweedler's little book [25] for the basics on Hopf algebras; all that we need to know is elementary and recalled below.

2. Hopf algebraic structure − Let $(\mathcal{F}, \cdot)$ stand for the set of commuting real-valued polynomials with indeterminates the elements of $\mathcal{T}$ and polynomial multiplication operation; equip the algebraic tensor product $\mathcal{F} \otimes \mathcal{F}$ with the induced product $(a \otimes b)(c \otimes d) = (ac) \otimes (bd)$. We define a coproduct Δ on $\mathcal{F}$ as follows. Given a labelled rooted tree τ, denote by $\mathsf{Sub}(\tau)$ the set of subtrees of τ with the same root as τ. Given such a subtree s, we obtain a collection $\tau_1, \ldots, \tau_n$ of labelled rooted trees by removing s and all the adjacent edges to s from τ. Write $\tau \backslash s$ for the monomial $\tau_1 \ldots \tau_n$. One defines a linear map $\Delta : \mathcal{T} \to \mathcal{F} \otimes \mathcal{F}$ by the formula

$$\Delta \tau = \sum_{s \in \mathsf{Sub}(\tau)} (\tau \backslash s) \otimes s,$$

and extend it to $\mathcal{F}$ by linearity and by requiring that it is multiplicative

$$\Delta(\tau_1 \ldots \tau_n) := \Delta(\tau_1) \ldots \Delta(\tau_n).$$

This coproduct is coassociative, $(\Delta \otimes \mathrm{Id})\Delta = (\mathrm{Id} \otimes \Delta)\Delta$, as it should be in any Hopf algebra. Given a real-valued map $\mathbf{Y}$ on $\mathcal{T}$, we extend it into an element of the dual space $\mathcal{F}^*$ of $\mathcal{F}$ setting

$$Y^{\tau_1 \ldots \tau_n} := \langle \mathbf{Y}, \tau_1 \cdots \tau_n \rangle := \prod_{i=1}^n Y_i^\tau,$$

for a monomial $\tau_1 \cdots \tau_n$, and by linearity. One defines a convolution product on $\mathcal{F}^*$ setting

$$(\mathbf{Y} \star \mathbf{X})^\tau := (\mathbf{Y} \otimes \mathbf{X})(\Delta \tau) = \sum_{s \in \mathsf{Sub}(\tau)} Y^{\tau \backslash s} X^s, \tag{2.7}$$

for a labelled rooted tree τ. The third ingredient needed to define the Hopf algebra structure of $\mathcal{F}$ is an antipode $\mathfrak{S}$, that is a map $\mathfrak{S} : \mathcal{F} \to \mathcal{F}$ inverting Δ in the sense that $\mathrm{M}(\mathrm{Id} \otimes \mathfrak{S})\Delta = \mathrm{M}(\mathfrak{S} \otimes \mathrm{Id})\Delta = \mathrm{Id}$, where M stands for the multiplication map $\mathrm{M}(a \otimes b) = ab \in \mathcal{F}$. An explicit formula for $\mathfrak{S}$ was first obtained in [14]; see [12] for a simple and enlighting proof. From its definition, the inverse $\mathbf{a}^{-1}$ of any element $\mathbf{a}$ of $\mathcal{F}^*$ is given by the formula $\mathbf{a}^{-1} = \mathfrak{S}^* \mathbf{a}$, where $\mathfrak{S}^*$ is the dual map in $\mathcal{F}^*$ of the antipode map in $\mathcal{F}$, that is $\big(\mathbf{a}^{-1}, \tau\big) = \big(\mathbf{a}, \mathfrak{S}\tau\big)$.

Denote by $\theta = \tau_1 \dots \tau_n$, with $\tau_i \in \mathcal{T}$, a generic element of the canonical basis $\mathcal{B}$ of $\mathcal{F}$. Write $(\theta^*)_{\theta \in \mathcal{B}}$ for the dual canonical basis and use bold letters to denote generic elements of $\mathcal{F}^*$. For an element $\mathbf{a}$ of $\mathcal{F}^*$ of the special form $\mathbf{a} = \sum_{\tau' \in \mathcal{T}} a^{\tau'} (\tau')^*$, with $a^{\tau'} \in \mathbb{R}$, we have for instance

$$\mathbf{a} \star \mathbf{a} = \sum_{\theta} (\mathbf{a} \otimes \mathbf{a})(\Delta \theta)\, \theta^* = \sum_{\tau \in \mathcal{T}} (\mathbf{a} \otimes \mathbf{a})(\Delta \tau)\, \tau^*$$
$$= \sum_{\tau \in \mathcal{T}} \left(\sum a^{\tau_1} a^{\tau_2} \right) \tau^*,$$

where the inside sum is over the set of subtrees τ_1 of τ with the same root as τ, and such that $\tau \backslash \tau_1$ is a subtree τ_2 of τ. Note in particular that $\mathbf{a} \star \mathbf{a}$ is again of the form $\sum_{\tau' \in \mathcal{T}} b^{\tau'} (\tau')^*$. More generally, we have for such an $\mathbf{a} \in \mathcal{F}^*$

$$\mathbf{a}^{\star n} = \sum_{\tau \in \mathcal{T}} \left(\sum a^{\tau_1} \cdots a^{\tau_n} \right) \tau^*, \tag{2.8}$$

where the inside sum is over the set of disjoint rooted subtrees $\tau_1, \dots, \tau_n$ of τ, with respective roots $\alpha_1, \dots, \alpha_n$ such that α_{i+1} is a descendant of α_i in τ, for all $1 \leqslant i \leqslant n-1$, and any node of τ is in one of the subtrees τ_i. The convolution product (2.7) is closely related to the identities (2.3) and (2.4) defining the differential operators $V(\tau^*)$ and $V(\tau_1^* \cdots \tau_n^*)$. The following statement is proved by induction on $n + n'$.

3. Lemma – *For any dual forests $\tau_1^* \cdots \tau_n^*$ and $\sigma_1^* \cdots \sigma_{n'}^*$, one has*

$$V\big(\tau_1^* \cdots \tau_k^*\big)\, V\big(\sigma_1^* \cdots \sigma_{k'}^*\big) = V\Big(\big(\tau_1^* \cdots \tau_k^*\big) \star \big(\sigma_1^* \cdots \sigma_{k'}^*\big) \Big).$$

3. Hölder branched p-rough paths – We define a Lie bracket on $\mathcal{F}^*$ setting

$$[\mathbf{a}, \mathbf{b}]_\star := \mathbf{a} \star \mathbf{b} - \mathbf{b} \star \mathbf{a},$$

for which the real vector space $\langle \mathcal{T}^* \rangle$ spanned by the dual trees is a Lie algebra. The exponential map $\exp_\star : \mathcal{F}^* \to \mathcal{F}^*$ and the logarithm map $\log_\star : \mathcal{F}^* \to \mathcal{F}^*$ are defined by the usual series

$$\exp_\star(\mathbf{a}) = \sum_{n \geqslant 0} \frac{\mathbf{a}^{\star n}}{n!}, \quad \log_\star(\mathbf{b}) = \sum_{n \geqslant 1} \frac{(\mathbf{1}^* - \mathbf{b})^{\star n}}{n},$$

with the convention $\mathbf{a}^{\star 0} = \mathbf{1}^*$, where $\mathbf{1}^*$ denote the dual of the empty tree. It is clear from the above formula for $\mathbf{a}^{\star n}$ that we have

$$\big(\exp_\star(\mathbf{a}), \tau \big) = \sum_{n=0}^{|\tau|} \frac{1}{n!} \sum a^{\tau_1} \cdots a^{\tau_n}, \tag{2.9}$$

with the same inside sum as in (2.8). Formula (2.8) justifies the convergence of the these sums, in the sense that $\big(\exp_\star(\mathbf{a}), \theta \big)$ and $\big(\log_\star(\mathbf{b}), \theta \big)$ are actually finite sums for any monomial θ, provided $(\mathbf{a}, 1) = 0$ – the pairing $(\cdot, \cdot)$ is a pairing between $\mathcal{F}^*$ and $\mathcal{F}$. Setting $\mathcal{F}_0^* := \big\{ \mathbf{a} \in \mathcal{F}^* \,;\, (\mathbf{a}, 1) = 0 \big\}$ and $\mathcal{F}_1^* := \big\{ \mathbf{a} \in \mathcal{F}^* \,;\, (\mathbf{a}, 1) = 1 \big\}$, one can see that $\exp_\star : \mathcal{F}_0^* \to \mathcal{F}_1^*$ and $\log_\star : \mathcal{F}_1^* \to \mathcal{F}_0^*$ are reciprocal bijections; the following result can be seen as a consequence. (See Sweedler's above mentioned book, or Reutenauer's book [24], Theorem 3.2.)

4. Theorem – *The pair $\big(\exp_\star(\langle \mathcal{T}^* \rangle), \star \big)$ is a group.*

Pick an integer $N \geqslant 0$. Write $\mathcal{F}^*_{(k)}$ for the vector space spanned by the monomials $(\tau_1 \ldots \tau_n)^*$, with $\sum_{i=1}^n |\tau_i| = k$, and denote by $\pi_{\leqslant N}$ the natural projection from $\mathcal{F}^*$ to the quotient space $\mathcal{F}^* \backslash \bigoplus_{k \geqslant N+1} \mathcal{F}^*_{(k)}$. Set

$$\pi_{\leqslant N}(\mathcal{F}^*) =: \mathcal{F}^*_{\leqslant N}.$$

The image of $\exp_\star(\mathfrak{g})$ by π_N is then diffeomorphic to the real vector space $\langle \mathcal{T}^*_{\leqslant N} \rangle$ spanned by $\pi_{\leqslant N}(\mathcal{T}^*)$; it is a Lie group when equipped with the operation $\pi_{\leqslant N} \circ \star$, which we still denote by $\star$. Write

$$\mathcal{G}_\ell^N \subset \mathcal{F}^*_{\leqslant N}$$

for that Lie group; it plays for branched rough paths the role that the $[p]$-step free nilpotent Lie group plays for weak geometric Hölder p-rough paths. In the same way as Chen's lifting formula (1.2) gives a geometric to the

Denote by $\bullet_a^*$ the element of the dual canonical basis, dual to $\bullet_a$. It is elementary to check that the $\mathcal{F}^*$-valued function $\mathfrak{H}_{ts} = (H_{ts}^\tau)_{\tau \in \mathcal{T}, |\tau| \leqslant N}$ associated with a smooth $\mathbb{R}^\ell$-valued path h, as defined in (2.6), considered as a function of the time parameter t, satisfies in $\pi_N(\mathcal{F}^*)$ the ordinary differential equation

$$d\mathfrak{H}_{ts} = \sum_{a=1}^\ell \mathfrak{H}_{ts} \star \left(\bullet_a^* \, dh_t^a \right);$$

this implies that $\mathfrak{H}_{ts}$ is an element of $\mathcal{G}_\ell^{(N)}$ for all times $t \geqslant s$, and that

$$\mathfrak{H}_{ts} = \mathbf{H}_s^{-1} \mathbf{H}_t,$$

with $\mathbf{H}_r := \mathfrak{H}_{r0}$, for all $r \geqslant 0$.

Definition − *Let $2 < p$. A **Hölder branched p-rough path on** $[0, T]$ is a $\mathcal{G}_\ell^{[p]}$-valued path* $\mathbf{X}$ *such that*

$$\sup_{0 \leqslant s < t \leqslant T} \frac{|X_{ts}^\varphi|}{|t - s|^{\frac{|\varphi|}{p}}} < \infty,$$

for all forests $\varphi = \tau_1 \ldots \tau_k$ with $|\varphi| := \sum_{i=1}^k |\tau_i| \leqslant [p]$, and $\mathbf{X}_{ts} := \mathbf{X}_s^{-1} \star \mathbf{X}_t$.

The norm of $\mathbf{X}$ is defined as

$$\|\mathbf{X}\| := \max_{\varphi \in \mathcal{F} \, ; \, |\varphi| \leqslant [p]} \sup_{0 \leqslant s < t \leqslant T} \frac{|X_{ts}^\varphi|}{|t - s|^{\frac{|\varphi|}{p}}}; \tag{2.10}$$

we also define a distance $d(\mathbf{X}, \mathbf{Y}) = \|\mathbf{X} - \mathbf{Y}\|$ on the nonlinear set of Hölder branched p-rough path. Given $0 \leqslant s \leqslant t \leqslant T$ and a Hölder branched p-rough path $\mathbf{X}$ defined on the time interval $[0, T]$, we define an element of $\langle \mathcal{T}^*_{\leqslant [p]} \rangle$ setting

$$\mathbf{\Lambda}_{ts} := \log_\star \mathbf{X}_{ts};$$

by definition

$$\exp_\star(\mathbf{\Lambda}_{ts}) = \mathbf{X}_{ts} \in \mathcal{G}_\ell^{[p]}.$$

4. Rough differential equations driven by Hölder branched p-rough paths − Let $\mathbf{X}$ be a Hölder branched p-rough path over $\mathbf{R}^\ell$, and $\mathrm{F} = (V_1, \ldots, V_\ell)$ be γ-Hölder vector fields on $\mathbb{R}^d$, for $\gamma > p$. A path z in $\mathbb{R}^d$ is said to be a solution path to the rough differential equation

$$dz_t = \mathrm{F}(z_t) \, d\mathbf{X}_t \tag{2.11}$$

in the sense of Davie if one has

$$x_t = x_s + \sum_{\varphi \in \mathcal{F}; |\varphi| \leqslant [p]} \big(V(\mathbf{X}_{ts})\mathrm{Id}\big)(x_s) + O\big(|t-s|^a\big).$$

for some constant $a > 1$. (This is actually Gubinelli's definition [19] rather than Davie's definition.) The path z is a solution to Equation (2.11) in the sense of Bailleul if one has

$$f(x_t) = f(x_s) + \sum_{\varphi \in \mathcal{F}; |\varphi| \leqslant [p]} \big(V(\mathbf{X}_{ts})f\big)(x_s) + O\big(|t-s|^a\big),$$

for some constant $a > 1$, for every real-valued function γ-Hölder function f on $\mathbb{R}^d$. (The $O(\cdot)$ term is allowed to depend on f.) Theorem 2 states the equivalence of these two notions of definition. Conditions for the existence of a unique solution are given in Gubinelli' seminal paper [19].

2.2 – Proof of Theorem 2

We only need to prove that a solution in Davie' sense is a solution in Bailleul' sense. We proceed as in Section 1 and prove that the result comes from a Taylor expansion property satisfied by the time 1 map of an approximate dynamics and the morphism property stated in Lemma 3.

Recall we write $\boldsymbol{\Lambda}_{ts}$ for $\log_\star \mathbf{X}_{ts}$. Given some times $0 \leqslant s \leqslant t \leqslant T$, let μ_{ts} stand for the well-defined time 1 map associated with the ordinary differential equation

$$\dot{y}_u = V\big(\boldsymbol{\Lambda}_{ts}\big)(y_u). \tag{2.12}$$

(The assumption that $\gamma > p$ and the vector fields V_i are γ-Lipschitz ensures that all the vector field $V(\boldsymbol{\Lambda}_{ts})$ is globally Lipscthiz continuous.)

5. Lemma – *There is a constant $a > 1$ with the following property. For any γ-Hölder real-valued function f on $\mathbb{R}^d$, one has*

$$f \circ \mu_{ts} = f + V\big(\mathbf{X}_{ts}\big)f + O_f\big(|t-s|^a\big),$$

for a remainder term $O_f\big(|t-s|^a\big)$ that may depend on f, and that has a C^1-norm bounded above by a constant multiple of $|t-s|^a$.

Proof – The Taylor expansion property follows from the morphism property of the V-map from Lemma 3, writing

$$f(y_1) = f(y_0) + \int_0^1 \big\{V(\boldsymbol{\Lambda}_{ts})f\big\}(y_{s_1})\, ds_1$$

$$= f(y_0) + \big\{V(\boldsymbol{\Lambda}_{ts})f\big\}(x) + \int_0^1 \int_0^{s_1} \big\{V(\boldsymbol{\Lambda}_{ts})V(\boldsymbol{\Lambda}_{ts})f\big\}(y_{s_2})\, ds_2 ds_1$$

$$= f(y_0) + \big\{V(\boldsymbol{\Lambda}_{ts})f\big\}(x) + \int_0^1 \int_0^{s_1} \big\{V(\boldsymbol{\Lambda}_{ts}^{\star 2})f\big\}(y_{s_2})\, ds_2 ds_1$$

$$= f(y_0) + \big\{V(\boldsymbol{\Lambda}_{ts})f\big\}(x) + \frac{1}{2}\big\{V(\boldsymbol{\Lambda}_{ts}^{\star 2})f\big\}(x) + \int_0^1 \int_0^{s_1} \int_0^{s_2} \big\{V(\boldsymbol{\Lambda}_{ts}^{\star 3})f\big\}(y_{s_3})\, ds_3 ds_2 ds_1$$

and, by induction,

$$f(y_1) = f(y_0) + \sum_{k=1}^{[p]} \frac{1}{k!}\big\{V(\boldsymbol{\Lambda}_{ts}^{\star k})f\big\}(x) + \int \big\{V(\boldsymbol{\Lambda}_{ts}^{\star([p]+1)})f\big\}(y_{[p]+1})\, \mathbf{1}_{0 \leqslant s_1 \leqslant \cdots \leqslant s_{[p]+1}} ds$$

$$= \big\{V(\mathbf{X}_{ts})f\big\}(y_0) + \Theta_f\big(|t-s|^a\big) + \int \big\{V(\boldsymbol{\Lambda}_{ts}^{\star([p]+1)})f\big\}(y_{[p]+1})\, \mathbf{1}_{0 \leqslant s_1 \leqslant \cdots \leqslant s_{[p]+1}} ds$$

The $\Theta_f\big(|t-s|^a\big)$ term comes from the fact that $\mathbf{\Lambda} \in \mathcal{H}_N$ and $\mathbf{X} \in \mathcal{H}_N$ are build from one another using the restriction of the maps $\log_\star$ and $\exp_\star$, so $\mathbf{X}_{ts} \in \mathcal{H}_N$ coincides with $\sum_{i=0}^N \frac{1}{n!} \mathbf{\Lambda}_{ts}^{\star n}$ up to some terms in $\mathcal{H}\backslash\mathcal{H}_N$ of size $O\big(|t-s|^a\big)$. The conclusion follows then from that explicit representation of the Taylor remainder. $\qquad\qquad\triangleright$

It follows from Lemma 5 that a path z is a solution in Davie' sense iff

$$z_t = \exp\big(V(\mathbf{\Lambda}_{ts})\big)(z_s) + O\big(|t-s|^{a'}\big),$$

for some exponent $1 < a' \leqslant \frac{\gamma}{p}$. The end of the proof of Theorem 2 is then identical to the proof of Theorem 1.

Remarks − 1. Write ε_{ts}^f for the remainder $f \circ \mu_{ts} - V(\mathbf{X}_{ts})f$. One can see directly that the μ_{ts} form a C^1-approximate flow from the identity

$$
\begin{aligned}
\mu_{tu} \circ \mu_{us} &= \big\{V(\mathbf{X}_{tu})\mathrm{Id}\big\} \circ \mu_{us} + \varepsilon_{tu}^{\mathrm{Id}} \circ \mu_{us} \\
&= V(\mathbf{X}_{us})V(\mathbf{X}_{tu})\mathrm{Id} + \varepsilon_{us}^{V(\mathbf{X}_{tu})\mathrm{Id}} + \varepsilon^{\mathrm{Id}_{tu}} \circ \mu_{us} \\
&= V(\mathbf{X}_{ts})\mathrm{Id} + \varepsilon_{us}^{V(\mathbf{X}_{tu})\mathrm{Id}} + \varepsilon_{tu}^{\mathrm{Id}} \circ \mu_{us} \\
&= \mu_{ts} + \varepsilon_{ts}^{\mathrm{Id}} + \varepsilon_{us}^{V(\mathbf{X}_{tu})\mathrm{Id}} + \varepsilon_{tu}^{\mathrm{Id}} \circ \mu_{us}.
\end{aligned}
$$

One then knows from Theorem 1 in [2] that the rough differential equation (2.11) has a unique solution flow, associated with the C^1-approximate flow μ. This provides a direct proof that the rough differential equation (2.11) has a unique solution flow.

2. Here as in Section 1, working with Hölder rough paths or rough paths controlled by some more general control $w(s,t)$ does not make any difference. Working with branched rough paths over an infinite dimensional space E does not make any difference either, as long as one takes care of working with a symmetric system of cross seminorms generating the topology of E, and complete accordingly the different algebraic tensor spaces $E^{\otimes k}$, such as emphasized in [10], Section 3.

References

[1] Bailleul, I., Flows driven by Banach space valued rough paths. *Séminaire de Probabilités*, Vol. XLVI:195–205, 2015.

[2] Bailleul, I., Flows driven by rough paths. *Revista Mat. Iberoamericana*, 31(3):901–934, 2015.

[3] Bailleul, I., A flows-based approach to rough differential equations. https://perso.univ-rennes1.fr/ismael.bailleul/files/M2Course.pdf.

[4] Bahouri, H. and Chemin, J.-Y. and Danchin, R., Fourier analysis and nonlinear partial differential equations. *Grundlehren der mathematischen Wissenschaften*, **343**, Springer.

[5] Baudoin, F., Diffusion Processes and Stochastic Calculus. *Ems Textbooks in Mathematics*, 2014.

[6] Boutaib, Y. and Gyurkó, L. and Lyons, T. and Yang, D., Dimension-free Euler estimates of rough differential equations *Rev. Roumaine Math. Pures Appl.*, 2014.

[7] Brault, A. and Lejay, A., The nonlinear sewing lemma II: Lipschitz continuous formulation. arXiv:1810.11988, 2018.

[8] Bruned, Y. and Chevyrev, I. and Friz, P. and Preiss, R., A Rough Path Perspective on Renormalization. arXiv:1701.01152, 2017.

[9] Cass, T. and Driver, B.and Lim, N. and Litterer, C., On the integration of weakly geometric rough paths. *J. Math. Soc. Japan*, **68**(4):1505–1524, 2016.

[10] Cass, T. and Weidner, M., Tree algebras over topological vector spaces in rough path theory. arXiv:1604.07352v2, 2017.

[11] Cayley, A., On the analytical forms called trees. *Am. J. Math.*, (**4**)(1/4):266-268, 1881.

[12] Chartier, Ph. and Hairer, E. and Vilmart, G., Algebraic structures of B-series. *Found. Comput. Math.*, **10**(4):407–427, 2010.

[13] Chen, K.T., Integration of paths, geometric invariants and a generalized Baker-Hausdorff formula. *Ann. Math.*, **65**(1):163-178, 1957.

[14] Connes, A. and Kreimer, D., Hopf algebras, renormalization and noncommutative geometry. *Comm. Math. Phys.*, **191**(1):203–242, 1998.

[15] Curry, Ch. and Ebrahimi-Fard, K. and Manchon, D. and Munthe-Kaas, H., Planarly branched rough paths and rough differential equations on homogeneous spaces. arXiv:1804.0851v3, 2018.

[16] Davie, S., Differential equations driven by rough paths: and approach via discrete approximations. *Appl. Math. Res. Express*, **2**, 2007.

[17] Friz, P. and Hairer, M., A Course on Rough Paths, *Spinger, Universitext*, 2014.

[18] Gubinelli, M., Controlling rough paths. *J. Funct. Anal.*, **16**:86–140, 2004.

[19] Gubinelli, M., Ramification of rough paths. *J. Diff. Eq.*, **248**(4):693–721, 2010.

[20] Hairer, M. and Kelly, D., Geometric versus non-geometric rough paths. *Ann. Inst. H. Poincaré, Probab. Statist.*, **51**(1):207–251, 2015.

[21] Lyons, T., Differential equations driven by rough signals. *Revista Mat. Iberoamaricana*, **14**(2):215–310, 1998.

[22] Caruana, M. and Levy, Th. and Lyons, T. Differential equations driven by rough paths. *Lect. Notes Math.*, **1908**, Springer, 2007.

[23] Milnor, J. Morse theory. *Annals of Mathematics studies*, Princeton University Press, **151**, 1963.

[24] Reutenauer, C., Free Lie algebras. *London Math. Soc. Monographs New Series*, 1993.

[25] Sweedler, M., Hopf algebras. *Math. Lect. Notes*, Benjamin INC., 1969.

[26] Young, L.C., An inequality of Hölder type, connected with Stieljes integration. *Acta Math.*, **67**:251–282, 1936.

• I. Bailleul - Institut de Recherche Mathematiques de Rennes, 263 Avenue du General Leclerc, 35042 Rennes, France. ismael.bailleul@univ-rennes1.fr