Branched Rough Path Notes

December 13, 2024

My notes on the theory of branched rough paths mainly based on the paper of Hairer and Kelly and the paper of Gubinelli.

1 Rooted Labelled Trees

For higher order rough paths, it is useful to introduce tree notations to keep track of the algebraic and analytic conditions for the extra terms arising from the Taylor expansion.

Definition 1.1 Denote $\mathcal{T}$ for the set of all rooted and labeled trees and $\mathcal{F}$ for the set of the associated forests, i.e. $\mathcal{F}$ is the free abelian monoid generated buy $\mathcal{T}$ and $f \in \mathcal{F}$ is a finite collection of rooted labeled trees (where we allow multiple copies of the same tree).

Notation wise, we use greek letters $\tau, \sigma, \varrho, \ldots$ for trees in $\mathcal{T}$ and $f, g, h \ldots$ for forests in $\mathcal{F}$.

Some examples of rooted and labeled trees are

$$\bullet^a \,,\; {}^b\!\!\underset{a}{\overset{}{\bullet}}\,,\; {}^b\!\!\underset{a}{\overset{c}{\vee}}\,,\; {}_b\!\!\underset{a}{\overset{c}{\bullet}}\,,\; {}_b\!\!\underset{a}{\overset{c\;d}{\vee}}\,,\; {}^c\!\!\underset{b}{\overset{d}{\vee}}{}_a\,,\; {}^d\!\!\underset{b}{\overset{c}{\bullet}}{}_a\,,\; \cdots$$

We write $f = \prod_{i \in I} \tau_i \in \mathcal{F}$ as a commutative product where $\tau_i \in \mathcal{T}$ and I is a finite index set. We allow $\mathcal{T}$ to contain the empty tree and we denote it by $\mathbf{1}$ since it is the multiplicative identity in $\mathcal{F}$. We introduce the rooting operation $[-]_a : \mathcal{F} \to \mathcal{T}$ where for $f \in \mathcal{F}$, we define $[f]_a$ for the tree obtained by joining all the trees in f to the root a, e.g.

$$[\mathbf{1}]_a = \bullet^a \;\text{ and }\; [\bullet^a\; {}^c\!\!\underset{b}{\overset{}{\bullet}}]_d = {}^a\!\!\underset{d}{\overset{d}{\vee}}{}_b\,.$$

For $\tau \in \mathcal{T}$, denote f_τ for the number of copies of τ in f. Thus, we can write $f = \prod_{\tau \in \mathcal{T}} \tau^{f_\tau}$ where f_τ is zero for all but finitely many terms. Namely, $\tau \in f$ if and only if $f_\tau \neq 0$. For $\tau \in \mathcal{T}$, the order of τ: $|\tau|$ is simply the number of vertices of τ and for $f \in \mathcal{F}$, $|f| = \sum_{\tau \in \mathcal{T}} f_\tau |\tau|$. We also write $\#f = \sum_{\tau \in \mathcal{T}} f_\tau$.

We denote $\langle \mathcal{F} \rangle$ for the free vector space generated by $\mathcal{F}$ so $\langle \mathcal{F} \rangle$ is a unital commutative algebra with identity $\mathbf{1}$. We define the inner product of $\langle \mathcal{F} \rangle$ by setting

$$\langle f, f' \rangle = \mathbf{1}_{f=f'} \tag{1.1}$$

`eq:inner-product`

for all $f, f' \in \mathcal{F}$ and extending linearly.

We define the coproduct $\Delta : \langle \mathcal{F} \rangle \to \langle \mathcal{F} \rangle \otimes \langle \mathcal{F} \rangle$ by first defining Δ on $\mathcal{F}$ inductively and then extending linearly. In particular, we set

- $\Delta \mathbf{1} = \mathbf{1} \otimes \mathbf{1}$.
- For $f \in \mathcal{F}$, $\Delta f = \prod_{h \in \mathcal{T}} (\Delta h)^{f_h}$.
- For $f \in \mathcal{F}$, $\Delta [f]_a = [f]_a \otimes \mathbf{1} + (\text{id} \otimes [-]_a) \Delta f$.

Here are some examples:

- $\Delta \, {\cdot}^a = {\cdot}^a \otimes \mathbf{1} + \mathbf{1} \otimes {\cdot}^a$.
- $\Delta \, {}^b\mathord{\vardiamond}_a = {}^b\mathord{\vardiamond}_a \otimes \mathbf{1} + {\cdot}^b \otimes {\cdot}^a + \mathbf{1} \otimes {}^b\mathord{\vardiamond}_a$.
- $\Delta \, {}^b\mathbf{V}_a^c = {}^b\mathbf{V}_a^c \otimes \mathbf{1} + {\cdot}^b {\cdot}^c \otimes {\cdot}^a + {\cdot}^c \otimes {}^b\mathord{\vardiamond}_a + {\cdot}^b \otimes {}^c\mathord{\vardiamond}_a + \mathbf{1} \otimes {}^b\mathbf{V}_a^c$.

Geometrically, the coproduct of a tree corresponds to all the sum of all the ways of splitting the tree into two parts with one cut with the part containing the root placed on the right hand side of the tensor product.

2   Hopf Algebra

The algebra $\langle \mathcal{F} \rangle$ equipped with the coproduct Δ forms what is known as a Hopf algebra (in particular, it is the Connes-Kreimer Hopf algebra). We in this section provide a informal introduction to Hopf algebras.

Before we proceed with defining Hopf algebra, we first need to introduce *bialgebras*. Suppose we have an algebra $\mathcal{H}^*$ acting on another algebra $\mathcal{H}^*$ via the action

$$\langle \cdot, \cdot \rangle : \mathcal{H}^* \otimes \mathcal{H} \to \mathbb{R}.$$

Informally speaking, a bialgebra is then $\mathcal{H}$ equipped with a *coproduct* $\Delta : \mathcal{H} \to \mathcal{H} \otimes \mathcal{H}$ which encodes this pairing and in particular, preserves the product structure of its *coalgebra* $\mathcal{H}^*$. More precisely, Δ is dual to the product operator ∇^* of $\mathcal{H}^*$ in the sense that

$$\langle \nabla^*(f \otimes g), h \rangle = \langle f \otimes g, \Delta h \rangle$$

where $\nabla^* : \mathcal{H}^* \otimes \mathcal{H}^* \to \mathcal{H}^*$ is such that $\nabla^*(x, y) = xy$. Thus, in some sense, once we have the coproduct Δ, we can forget about $\mathcal{H}^*$ and work solely with $\mathcal{H}$.

A bialgebra $(\mathcal{H}, \Delta)$ is *graded* if it has the decomposition

$$\mathcal{H} = \bigoplus_{n \in \mathbb{N}} \mathcal{H}_{(n)}$$

where $H_{(n)}$ are vector spaces such that for all $n, m \in \mathbb{N}$,

$$\mathcal{H}_{(n)} \cdot \mathcal{H}_{(m)} \subseteq \mathcal{H}_{(n+m)}, \text{ and } \Delta \mathcal{H}_{(n)} \subseteq \bigoplus_{p+q=n} \mathcal{H}_{(p)} \otimes \mathcal{H}_{(q)}.$$

Denoting $\mathcal{F}_{(n)} = \{f \in \mathcal{F} : |f| = n\}$, taking $\mathcal{H}_{(n)} = \langle \mathcal{F}_{(n)} \rangle$, we observe that $\mathcal{F}$ is a graded bialgebra.

While the coproduct preserves the product structure of $\mathcal{H}^*$, we would also like a map which preserves the units (invertible elements) of $\mathcal{H}^*$. This is achieved with the *antipole* map $S : \mathcal{H} \to \mathcal{H}$ which is required to satisfy

$$\nabla(S \otimes \mathrm{id})\Delta h = \nabla(\mathrm{id} \otimes S)\Delta h = \langle \mathbf{1}, h \rangle \mathbf{1} \qquad (2.1) \quad \boxed{\texttt{eq:antipole}}$$

for all $h \in \mathcal{H}$, where ∇ is the product operator of $\mathcal{H}$ and

$$\langle \mathbf{1}, h \rangle = \begin{cases} 1 & \text{if } h = \mathbf{1}, \\ 0 & \text{otherwise.} \end{cases}$$

This condition is motivated by the fact that, if S is such that for any unit $f \in \mathcal{H}^*$, $S^*f = f^{-1}$ where $S^* : \mathcal{H}^* \to \mathcal{H}^*$ is the dual of S, then

$$\langle f \otimes f, (S \otimes \mathrm{id})\Delta h \rangle = \langle \nabla^*(S^*f \otimes f), h \rangle = \langle S^*f \cdot f, h \rangle = \langle \mathbf{1}, h \rangle.$$

On the other hand,

$$\langle f \otimes f, \langle \mathbf{1}, h \rangle \mathbf{1} \otimes \mathbf{1} \rangle = \langle \mathbf{1}, h \rangle.$$

Hence, in order for S to preserve the unit structure of $\mathcal{H}^*$, it is necessary to require $\nabla(S \otimes \text{id})\Delta h = \nabla\langle \mathbf{1}, h\rangle \mathbf{1} \otimes \mathbf{1} = \langle \mathbf{1}, h\rangle$.

With the antipole motivated, a Hopf algebra is then simply a bialgebra equipped with an antipole.

Proposition 2.1 *A graded bialgebra satisfying $\mathcal{H}_{(0)} = \mathbb{R}$ is automatically a Hopf algebra.*

Thus, as $\langle \mathcal{F}_{(0)}\rangle = \mathbb{R}$, it follows $\langle \mathcal{F}\rangle$ is a Hopf algebra. The antipole S of $\langle \mathcal{F}\rangle$ can be computed explicitly by using the identity (2.1). We give some examples:

- $S\mathbf{1} = \mathbf{1}$.
- $0 = \nabla(S \otimes \text{id})\Delta \bullet^a = S\bullet^a + \bullet^a$. Thus, $S\bullet^a = -\bullet^a$.
- $0 = \nabla(S \otimes \text{id})\Delta\, ^b\mathfrak{l}_a = S\, ^b\mathfrak{l}_a + (S\bullet^b)\bullet^a + \, ^b\mathfrak{l}_a$. Thus, $S\, ^b\mathfrak{l}_a = -\, ^b\mathfrak{l}_a + \bullet^a\bullet^b$.
- $0 = \nabla(S \otimes \text{id})\Delta\, ^b\mathsf{V}_a^c = S\, ^b\mathsf{V}_a^c + (S\bullet^b\bullet^c)\bullet^a + (S\bullet^c)\, ^b\mathfrak{l}_a + (S\bullet^b)\, ^c\mathfrak{l}_a + \, ^b\mathsf{V}_a^c$. Thus, $S\, ^b\mathsf{V}_a^c = -\, ^b\mathsf{V}_a^c + \bullet^c\, ^b\mathfrak{l}_a + \bullet^b\, ^c\mathfrak{l}_a - \bullet^a\bullet^b\bullet^c$.

Equipping $\langle \mathcal{F}\rangle$ with the inner product as defined by Equation (1.1), we can view $\langle \mathcal{F}\rangle$ acting on itself via the action $\langle f, \cdot\rangle$ for all $f \in \langle \mathcal{F}\rangle$. Thus, the underlying bialgebra of $(\langle \mathcal{F}\rangle, \Delta, S)$ can be viewed as $\mathcal{H} = \langle \mathcal{F}\rangle$ with itself as its own coalgebra $\mathcal{H}^* = \langle \mathcal{F}\rangle$ (with a different multiplication). We now give a description of the algebraic structure of $\mathcal{H}^* = \langle \mathcal{F}\rangle$.

By the property of the coproduct, if $\mathcal{H}^* = \langle \mathcal{F}\rangle$ has product $\nabla^*(f, g) = f * g$, then for any $f, g \in \mathcal{F}$ and $h \in \mathcal{F}$, we have the defining property for ∇^*:

$$\langle f * g, h\rangle = \langle \nabla^*(f, g), h\rangle = \langle f \otimes g, \Delta h\rangle.$$

The product $*$ is known as the *convolution product* and we observe that for $f, g \in \mathcal{F}$, $\langle f * g, h\rangle = 1$ if and only if $f \otimes g$ is a term of Δh. For trees, the convolution can be explicitly described by

$$\tau_1 * \tau_2 = \tau_1\tau_2 + \tau_1 *_t \tau_2$$

for $\tau_1, \tau_2 \in \mathcal{T}$ where $\tau_1 *_t \tau_2$ denotes the sum of all trees obtained by attaching τ_1 to a vertex of τ_2.

$\langle \mathcal{F}\rangle$ equipped with $*$ forms an associative (*non-commutative*) algebra and we denote its coproduct by $\delta : \langle \mathcal{F}\rangle \to \langle \mathcal{F}\rangle \otimes \langle \mathcal{F}\rangle$ such that

$$\langle \delta f, h_1 \otimes h_2\rangle = \langle f, h_1 h_2\rangle.$$

Thus, if $\tau \in \mathcal{T}$, then, using Sweedler's notation,

$$1 = \langle \tau, \tau \cdot \mathbf{1}\rangle = \langle \delta\tau, \tau \otimes \mathbf{1}\rangle = \sum \left\langle \tau^{(1)}, \tau\right\rangle\left\langle \tau^{(2)}, \mathbf{1}\right\rangle.$$

Namely, there exists a unique component i of $\delta\tau$ for which $\tau^{(1)i} = \tau$ and $\tau^{(2)i} = \mathbf{1}$.

The tensor algebra $T(\mathbb{R}^d)$ can be identified in $\langle \mathcal{F}\rangle$ by identifying e_a by $\bullet^a$, $e_b \otimes e_a$ by $\, ^b\mathfrak{l}_a$, $e_c \otimes e_b \otimes e_a$ by $\, ^b\mathfrak{l}_a^{\,c}$ and so on for $a, b, c = 1, \ldots, d$. Namely, denoting $\iota : T(\mathbb{R}^d) \hookrightarrow \langle \mathcal{F}\rangle$ for this inclusion, $\iota(T(\mathbb{R}^d))$ corresponds to the set of all trees with a single branch.

Definition 2.2 Let $\mathcal{H}$ be a Hopf algebra with coalgebra $\mathcal{H}^*$. Denote $\hom(\mathcal{H}, \mathbb{R}) \subseteq \mathcal{H}^*$ for the the space of characters (i.e. $\mathbb{R}$-valued algebra homomorphisms) of $\mathcal{H}$.

For all (by an abuse of notation) $f = \langle f, \cdot\rangle \in \hom(\mathcal{H}, \mathbb{R})$, as f is multiplicative, it follows that

$$\langle f, h_1 h_2\rangle = \langle f, h_1\rangle\langle f, h_2\rangle = \langle f \otimes f, h_1 \otimes h_2\rangle.$$

On the other hand, denoting δ for the coproduct on the coalgebra $\mathcal{H}^*$ (the cocoproduct?), we have

$$\langle \delta f, h_1 \otimes h_2\rangle = \langle f, h_1 h_2\rangle.$$

Thus, we can identify elements of $\hom(\mathcal{H}, \mathbb{R})$ with $f \in \mathcal{H}^*$ which satisfies $\delta f = f \otimes f$. More precisely,

$$\hom(\mathcal{H}, \mathbb{R}) \simeq \{f \in \mathcal{H}^* : \delta f = f \otimes f\} =: G(\mathcal{H}).$$

Definition 2.3 $G(\mathcal{H})$ forms a group and is known as the *Butcher group*.

Proposition 2.4 *For all $f \in G(\mathcal{H})$, $f^{-1} = S^* f$ with S^* being the adjoint of the antipole.*

3   Branched Rough Paths

Finally, we can now define branched rough paths. Defining $\mathcal{H}^*_{(n)} = \langle \mathcal{F}_{(n)} \rangle$ where the notation
* indicates we are working in the dual space, we define

$$G_N(\mathcal{H}) = \frac{G(\mathcal{H})}{G(\mathcal{H}) \cap \bigoplus_{n \geq N+1} \mathcal{H}^*_{(n)}}$$

where the fraction indicates the quotient group (it is not difficult the see that the denominator is
indeed a normal subgroup). Namely, we identify the forests in $G(\mathcal{H})$ which has order greater
than N with the identity element.

Definition 3.1 For a map $\mathbf{X} : [0, T] \to G_N(\mathcal{H})$, we denote $\mathbf{X}_{st} = \mathbf{X}_s^{-1} * \mathbf{X}_t$. $\mathbf{X}$ is a γ-Hölder
branched rough path if

$$\sup_{s < t \in [0,T]} \frac{|\langle \mathbf{X}_{st}, f \rangle|}{|t - s|^{\gamma |f|}} < \infty$$

for all $f \in \mathcal{F}_N$.

For a branched rough path $\mathbf{X}$, we define its path component X^a by

$$\delta X^a_{st} = \langle \mathbf{X}_{st}, \bullet^a \rangle.$$

We observe that the path components of a γ-Hölder branched rough path is γ-Hölder. More-
over, Chen's condition (that is $\mathbf{X}_{st} = \mathbf{X}_{su} * \mathbf{X}_{ut}$) is automatically satisfied by branching rough
paths.

For general trees $\tau = [f]_a$, we interpret $\langle \mathbf{X}_{st}, \tau \rangle = \int_s^t \langle \mathbf{X}_{sr}, f \rangle \mathrm{d} X^a_r$ (note that here, the right
hand side is strictly formal and is simply a notation for the left hand side). Thus,

$$\langle \mathbf{X}_{st}, \bullet^a \rangle = \int_s^t \mathrm{d} X^a_r = \delta X^a_{st}, \qquad \left\langle \mathbf{X}_{st}, {}^b\!\!\mathbf{I}_a \right\rangle = \int_s^t \int_s^r \mathrm{d} X^b_u \mathrm{d} X^a_r,$$

$$\left\langle \mathbf{X}_{st}, {}^b\mathbf{V}^c_a \right\rangle = \int_s^t \left(\int_s^r \mathrm{d} X^b_u \right) \left(\int_s^r \mathrm{d} X^c_u \right) \mathrm{d} X^a_r,$$

and so on. These form a basis should we repeatedly apply Taylor expansion of a solution to an
RDE.

Definition 3.2 A path $\mathbf{Y} : [0, T] \to \mathcal{H}_{N-1} = \bigoplus_{n=0}^{N-1} \mathcal{H}_{(n)}$ is a $\mathbf{X}$-controlled rough path for
some γ-Hölder branched rough path $\mathbf{X}$ if

$$R^f_{st} = \langle f, \mathbf{Y}_t \rangle - \langle \mathbf{X}_{st} * f, \mathbf{Y}_s \rangle \tag{3.1}$$

`eq:remainder`

satisfies

$$\sup_{s < t \in [0,T]} \frac{|R^f_{st}|}{|t - s|^{(N-|f|)\gamma}} < \infty$$

for all $f \in \mathcal{F}_{N-1}$. We denote $Y_t = \langle \mathbf{1}, \mathbf{Y}_t \rangle$ for the path component of $\mathbf{Y}$. Moreover, for $f \in \mathcal{F}_{N-1}$, denote $\partial^f_X Y_t = \langle f, \mathbf{Y}_t \rangle$. We call $\partial^f_X Y_t$ the f-th Gubinelli derivative of $\mathbf{Y}$ against $\mathbf{X}$.

We now discuss the case where there is only one label and we omit it from the notations.

This definition corresponds to the usual definition of level 2 controlled rough paths in the
following sense: assuming we are in 1-dimension and $N = 2$, testing against $h = \mathbf{1}$, we obtain

$$R^{\mathbf{1}}_{st} = Y_t - \langle \mathbf{X}_{st} * \mathbf{1}, \mathbf{Y}_s \rangle = Y_t - \langle \mathbf{X}_{st}, \mathbf{Y}_s \rangle.$$

Now, as

$$\langle \mathbf{X}_{st}, \mathbf{Y}_s \rangle = \sum_{f \in \mathcal{F}_1} \langle f, \mathbf{Y}_s \rangle \langle \mathbf{X}_{st}, f \rangle = Y_s \langle \mathbf{X}_{st}, \mathbf{1} \rangle + \langle \cdot, \mathbf{Y}_s \rangle \langle \mathbf{X}_{st}, \cdot \rangle$$

$$= Y_s + \langle \cdot, \mathbf{Y}_s \rangle \delta X_{st}$$

where the last equality follows as $\mathbf{X}_{st} \in \hom(\mathcal{H}, \mathbb{R})$ so $\langle \mathbf{X}_{st}, \mathbf{1} \rangle = 1$. Thus, the Gubinelli derivative of Y against X in this case is $Y_s' = \partial_X^{\cdot} Y_s = \langle \cdot, \mathbf{Y}_s \rangle$. More generally, for any N, we find that

$$R_{st}^{\mathbf{1}} = \delta Y_{st} - \sum_{f \in \mathcal{F}_{N-1} \backslash \{\mathbf{1}\}} \partial_X^f Y_s \langle \mathbf{X}_{st}, f \rangle$$

$$= \delta Y_{st} - Y_s' \delta X_{st} - \sum_{f \in \mathcal{F}_{N-1} \backslash \{\mathbf{1}, \cdot\}} \langle f, \mathbf{Y}_s \rangle \langle \mathbf{X}_{st}, f \rangle$$

For the sake of computation, we introduce the following short hands. For all $f \in \mathcal{F}_N$,

- $\Delta_r f = \Delta f - \mathbf{1} \otimes f$,
- for $g, h \in \mathcal{F}_N$, set $c_r(f, g, h) = \langle g \otimes h, \Delta_r f \rangle$. Namely,

$$\Delta_r f = \sum_{g, h \in \mathcal{F}_N} c_r(f, g, h) g \otimes h.$$

<div style="border:1px solid;padding:2px;display:inline-block">eq:c'-path</div> **Lemma 3.3** *For $f \in \mathcal{F}_N$, $c_r(f, g, h) = \sum_l \langle f, l \rangle c_r(l, g, h)$.*

Proof.

$$\sum_{g, h} \left(\sum_l \langle f, l \rangle c_r(l, g, h) \right) g \otimes h = \sum_l \langle f, l \rangle \sum_{g, h} c_r(l, g, h) g \otimes h$$

$$= \sum_l \langle f, l \rangle \Delta_r l = \Delta_r \left(\sum_l \langle f, l \rangle l \right) = \Delta_r f.$$

$\square$

Using these notations, we have the equivalent formulation of the remainder term.

<div style="border:1px solid;padding:2px;display:inline-block">prop:remainder</div> **Proposition 3.4** *For a γ-Hölder rough path $\mathbf{X}$ and a path $\mathbf{Y} : [0, T] \to \mathcal{H}_{N-1}^*$, for all $f \in \mathcal{F}_{N-1}$,*

$$R_{st}^f = \delta(\partial_X^f Y)_{st} - \sum_{g, h \in \mathcal{F}_{N-1}} c_r(g, h, f) \partial_X^g Y_s \langle \mathbf{X}_{st}, h \rangle$$

where R^f is defined as in Equation (3.1).

Proof. Indeed,

$$R_{st}^f = \partial_X^f Y_t - \langle \mathbf{X}_{st} * f, \mathbf{Y}_s \rangle = \partial_X^f Y_t - \langle \mathbf{X}_{st} \otimes f, \Delta \mathbf{Y}_s \rangle$$

$$= \partial_X^f Y_t - \langle \mathbf{X}_{st} \otimes f, \mathbf{1} \otimes \mathbf{Y}_s \rangle - \langle \mathbf{X}_{st} \otimes f, \Delta_r \mathbf{Y}_s \rangle$$

$$= \delta(\partial_X^h Y)_{st} - \sum_{h, l} c_r(\mathbf{Y}_s, h, l) \langle \mathbf{X}_{st} \otimes f, h \otimes l \rangle$$

$$= \delta(\partial_X^h Y)_{st} - \sum_{h, l} \sum_g c_r(g, h, l) \partial_X^g Y_s \langle \mathbf{X}_{st}, h \rangle \langle f, l \rangle$$

$$= \delta(\partial_X^h Y)_{st} - \sum_{g, h} c_r(g, h, h) \partial_X^g Y_s \langle \mathbf{X}_{st}, h \rangle$$

where the penultimate equality follows from Lemma 3.3. $\square$

Proposition 3.5 *Let* $\mathbf{Y} : [0, T] \to \mathcal{H}_{N-1}^*$ *be a controlled rough path against the γ-Hölder rough path* $\mathbf{X}$*. Then, denoting the two parameter process*

$$I_{st}^a = \sum_{f \in \mathcal{F}_{N-1}} \langle f, \mathbf{Y}_s \rangle \langle \mathbf{X}_{st}, [f]_a \rangle$$

we have that

$$\|\delta I\|_{(N+1)\gamma} \lesssim \sum_{f \in \mathcal{F}_{N-1}} \|R^f\|_{(N-|f|)\gamma} \|\mathbf{X}^{[f]}\|_{(|f|+1)\gamma}$$

where we denote $\mathbf{X}^f = \langle \mathbf{X}, f \rangle$*.*

Proof. Firstly, we observe that, for all $f \in \mathcal{F}_{N-1}$,

$$\begin{aligned}
\langle \mathbf{X}_{st}, [f] \rangle - \langle \mathbf{X}_{su}, [f] \rangle &= \langle \mathbf{X}_{su} \otimes \mathbf{X}_{ut}, (\mathrm{id} \otimes [-]) \Delta f \rangle \\
&= \langle \nabla(\mathbf{X}_{su} \otimes [-]^* \mathbf{X}_{ut}), f \rangle \\
&= \langle \mathbf{X}_{su}, f \rangle \langle \mathbf{X}_{ut}, [f] \rangle.
\end{aligned}$$

Thus, adopting Einstein's summation convention, we have

$$\begin{aligned}
\delta I_{sut} &= \langle f, \mathbf{Y}_s \rangle (\langle \mathbf{X}_{st}, [f] \rangle - \langle \mathbf{X}_{su}, [f] \rangle) - \langle f, \mathbf{Y}_u \rangle \langle \mathbf{X}_{ut}, [f] \rangle \\
&= (\langle f, \mathbf{Y}_s \rangle \langle \mathbf{X}_{su}, f \rangle - \langle f, \mathbf{Y}_u \rangle) \langle \mathbf{X}_{ut}, [f] \rangle \\
&= -R_{su}^f \langle \mathbf{X}_{ut}, [f] \rangle + \underbrace{(\langle f, \mathbf{Y}_s \rangle \langle \mathbf{X}_{su}, f \rangle - \langle \mathbf{X}_{su} * f, \mathbf{Y}_s \rangle)}_{?} \langle \mathbf{X}_{ut}, [f] \rangle \\
&= \cdots
\end{aligned}$$

$\square$

As a consequence of the above proposition, in the case where $(N + 1)\gamma > 1$, we may apply the sewing lemma to I_{st} to obtain

$$\int_s^t Y_r \mathrm{d}\mathbf{X}_r := \lim_{|\mathcal{P}| \to 0} \sum_{(u,v) \in \mathcal{P}} I_{uv}$$

with $\mathcal{P}$ being a partition of $[s, t]$ and we have the remainder inequality

$$\left| \int_s^t Y_r \mathrm{d}\mathbf{X}_r - \sum_{f \in \mathcal{F}_{N-1}} \langle f, \mathbf{Y}_s \rangle \langle \mathbf{X}_{st}, [f] \rangle \right|$$

$$\lesssim \left(\sum_{f \in \mathcal{F}_{N-1}} \|R^f\|_{(N-|f|)\gamma} \|\mathbf{X}^{[f]}\|_{(|f|+1)\gamma} \right) |t - s|^{(N+1)\gamma}$$

We note that at this point $\int_s^{\cdot} Y_r \mathrm{d}\mathbf{X}_r$ takes value in $\mathbb{R}$ and is notably not a tree/forest path. As we would like to talk about rough differential equations, namely we would want to phrase a fixed point argument involving rough integrals, we need to extend the $\mathbb{R}$-valued path $\int_s^{\cdot} Y_r \mathrm{d}\mathbf{X}_r$ to a forest path $\int_s^{\cdot} \mathbf{Y}_r \mathrm{d}\mathbf{X}_r$. This is achieved by specifying $\partial_X^f \left(\int_s^{\cdot} Y_r \mathrm{d}\mathbf{X}_r \right) = \langle f, \int_s^{\cdot} \mathbf{Y}_r \mathrm{d}\mathbf{X}_r \rangle$ for each $f \in \mathcal{F}_{N-1}^*$ with

- $\partial_X^{\mathbf{1}} \int_s^{\cdot} \mathbf{Y}_r \mathrm{d}\mathbf{X}_r = \int_s^{\cdot} Y_r \mathrm{d}\mathbf{X}_r$.
- For $f \in \mathcal{F}_{\leq N-2}$, we set $\partial_X^{[f]} \int_0^{\cdot} \mathbf{Y}_r \mathrm{d}\mathbf{X}_r = \partial_X^f \mathbf{Y}_{\cdot\cdot}$.
- For $f_1 \cdot f_2 \in \mathcal{F}_{\leq N-1}$, $\partial_X^{f_1 \cdot f_2} \int_0^{\cdot} \mathbf{Y}_r \mathrm{d}\mathbf{X}_r = 0$.

4   Composition with smooth functions

Recall that, in the case where $N = 2$, for $(X, \mathbb{X}) \in \mathcal{C}^{\alpha}([0, T], V)$, $(Y, Y') \in \mathcal{D}_X^{2\alpha}([0, T], W)$ and $\varphi \in C_b^2(W, \bar{W})$, we have

$$(\varphi(Y), \varphi(Y)') = (\varphi(Y), D\varphi(Y)Y') \in \mathcal{D}_X^{2\alpha}([0, T], \bar{W}).$$

This fact can be generalized for branched rough paths resulting in the following.

`prop:comp` **Proposition 4.1** *Let* $\mathbf{X}$ *be a level-N γ-Hölder rough path and* $\mathbf{Y}$ *be a d-dimensional controlled rough path against* $\mathbf{X}$. *Then, for any* $\varphi \in C_b^{N-1}(\mathbb{R}^d, \mathbb{R}^d)$, *defining* $\varphi(\mathbf{Y})$ *by*

$$\varphi(\mathbf{Y}) = \sum_{f \in \mathcal{F}_{N-1}} \partial_X^f \varphi(\mathbf{Y}) f$$

where

$$\partial_X^f \varphi(\mathbf{Y}) = \sum_{\tau_1 \cdots \tau_{\#f} = f} \frac{1}{(\#f)!} D^{\#f} \varphi(Y)[(\partial_X^{\tau_i} Y)_{i=1}^{\#f}]$$

for any $f \in \mathcal{F}_{N-1}$, *we have that* $\varphi(\mathbf{Y})$ *is also a controlled rough path against* $\mathbf{X}$.

In the level $N = 2$ case, we have that

$$\|R_{st}^{\varphi(\mathbf{Y})}\| \leq \sup_{\lambda \in [0,1]} \frac{1}{2} \|D^2 \varphi((1-\lambda)Y_s + \lambda Y_t)\| \|Y_{st}\|^2 + \|D\sigma(x_s) R_{st}^{\mathbf{Y}}\|,$$

I suspect the following generalization holds true (maybe require $\mathbf{Y}$ to satisfy the RDE $d\mathbf{Y}_t = \varphi(\mathbf{Y}_t)d\mathbf{X}_t$):

Proposition 4.2 *For any* $\tau \in \mathcal{F}_{N-2}$, *we have*

$$\|R_{st}^{\varphi(\mathbf{Y}),\tau}\| \leq \frac{\|\delta Y_{st}\|^{N-|\tau|}}{(N-|\tau|)!} \sup_{\lambda \in [0,1]} \|D^{N-|\tau|} \partial_X^\tau \varphi(\lambda \mathbf{Y}_s + (1-\lambda)\mathbf{Y}_t)\|$$
$$+ \text{ terms of regularity } (N-|\tau|)\gamma$$

where the extra terms involve $R_{st}^{\mathbf{Y},\mathbf{1}}$, $\langle \mathbf{X}_{st}, \sigma \rangle$ *for some* $\sigma \in \mathcal{F}_{N-1}$.

This is certainly the case for $\tau = \mathbf{1}$:

`lem:crone` **Lemma 4.3** *For any* $f, g \in \mathcal{F}_{N-1} \setminus \{\mathbf{1}\} =: \mathcal{F}_{N-1}^+$, $c_r(f, g, \mathbf{1}) = \langle f, g \rangle$.

`lem:remain-Y-one` **Lemma 4.4** *Assuming that* $\mathbf{Y}$ *satisfies the RDE* $d\mathbf{Y}_t = \varphi(\mathbf{Y}_t)d\mathbf{X}_t$, *then, denoting* $\mathcal{T}_{N-1}^{+,\cdot} := \mathcal{T}_{N-1} \setminus \{\mathbf{1}, \cdot\}$, *we have that*

$$\varphi(Y_s)\langle \mathbf{X}_{st}, \cdot \rangle = \delta(Y)_{st} - R_{st}^{\mathbf{Y},\mathbf{1}} - \sum_{\tau \in \mathcal{T}_{N-1}^{+,\cdot}} \partial_X^\tau Y_s \langle \mathbf{X}_{st}, \tau \rangle$$
$$= \delta(Y)_{st} - R_{st}^{\mathbf{Y},\mathbf{1}} - \sum_{f \in \mathcal{F}_{N-2}^+} \partial_X^f \varphi(Y_s) \langle \mathbf{X}_{st}, [f] \rangle$$

Proposition 4.5 *Taking* $\mathbf{X}, \mathbf{Y}$ *as above, we have*

$$R_{st}^{\varphi(\mathbf{Y}),\mathbf{1}} = \delta(\varphi(Y))_{st} - \sum_{k=1}^{N-1} \frac{1}{k!} D^k \varphi(Y_s)[(\delta Y_{st})^{\otimes k}] - R_{st}^{\sharp}$$

where $R^{\sharp}$ *has regularity* $\geq N\gamma$.

Proof. By Proposition 3.4 and Lemma 4.3, we have that

$$R_{st}^{\varphi(\mathbf{Y}),\mathbf{1}} = \delta(\varphi(Y))_{st} - \sum_{f \in \mathcal{F}_{N-1}^+} \partial_X^f \varphi(Y_s)\langle \mathbf{X}_{st}, f\rangle$$

$$= \delta(\varphi(Y))_{st} - \sum_{k=1}^{N-1} \partial_X^{(\cdot)^k}\varphi(Y_s)\langle \mathbf{X}_{st}, \cdot\rangle^k - \sum_{f \notin \{(\cdot)^k\}_{k=1}^{N-1}} \partial_X^f \varphi(Y_s)\langle \mathbf{X}_{st}, f\rangle.$$

On the other hand, by Proposition 4.1,

$$\partial_X^{(\cdot)^k}\varphi(Y_s) = \frac{1}{k!}D^k\varphi(Y_s)[(\partial_X^\cdot Y_s)^{\otimes k}] = \frac{1}{k!}D^k\varphi(Y_s)[(\varphi(Y_s))^{\otimes k}].$$

So, combining with Lemma 4.4, we have that

$$\partial_X^{(\cdot)^k}\varphi(Y_s)\langle \mathbf{X}_{st}, \cdot\rangle^k = \frac{1}{k!}D^k\varphi(Y_s)[(\varphi(Y_s)\langle \mathbf{X}_{st}, \cdot\rangle)^{\otimes k}]$$

$$= \frac{1}{k!}D^k\varphi(Y_s)\left[\left(\delta Y_{st} - R_{st}^{\mathbf{Y},\mathbf{1}} - \sum_{\tau \in \mathcal{T}_{N-1}^{+,\cdot}}\partial_X^\tau Y_s\langle \mathbf{X}_{st}, \tau\rangle\right)^{\otimes k}\right]$$

$$= \frac{1}{k!}D^k\varphi(Y_s)[(\delta Y_{st})^{\otimes k}] + S_1 + S_2$$

where

$$S_1 := \frac{1}{k!}\sum_{\tau_1,\dots,\tau_k \in \mathcal{T}_{N-1}^{+,\cdot}} D^k\varphi(Y_s)[\partial_X^{\tau_1} Y_s\langle \mathbf{X}_{st}, \tau_1\rangle \otimes \cdots \otimes \partial_X^{\tau_k} Y_s\langle \mathbf{X}_{st}, \tau_k\rangle],$$

and

$$S_2 := \frac{1}{k!}\sum_{m=0}^{k-1}\sum_{\tau_1,\dots,\tau_m \in \mathcal{T}_{N-1}^{+,\cdot}} D^k\varphi(Y_s)[(R_{st}^{\mathbf{Y},\mathbf{1}})^{\otimes(k-m)}\otimes\partial_X^{\tau_1} Y_s\langle \mathbf{X}_{st}, \tau_1\rangle\otimes\cdots\otimes\partial_X^{\tau_m} Y_s\langle \mathbf{X}_{st}, \tau_m\rangle].$$

Now, by Proposition 4.1, we observe that S_1 can be written as

$$S_1 = \sum_{\cdot \notin f \in \mathcal{F}_{k(N-1)}} \mathbf{1}_{\{\#f=k\}} \sum_{\tau_1\cdots\tau_k=f} \frac{1}{(\#f)!}D^k\varphi(Y_s)[\partial_X^{\tau_1} Y_s \otimes \cdots \otimes \partial_X^{\tau_k} Y_s]\prod_{i=1}^k \langle \mathbf{X}_{st}, \tau_i\rangle$$

$$= \sum_{\cdot \notin f \in \mathcal{F}_{k(N-1)}} \mathbf{1}_{\{\#f=k\}}\partial_X^f\varphi(Y_s)\langle \mathbf{X}_{st}, f\rangle.$$

Thus, we have

$$R_{st}^{\varphi(\mathbf{Y}),\mathbf{1}} = \delta(\varphi(Y))_{st} - \sum_{k=1}^{N-1}\frac{1}{k!}D^k\varphi(Y_s)[(\delta Y_{st})^{\otimes k}]$$

$$- \sum_{\cdot \notin f \in \mathcal{F}_{k(N-1)}\setminus\mathcal{F}_{N-1}} \mathbf{1}_{\{\#f=k\}}\partial_X^f\varphi(Y_s)\langle \mathbf{X}_{st}, f\rangle - S_2$$

as desired. □

5 Relation with regularity structures

Branched rough paths can be viewed as a special case of regularity structures:

- $\mathcal{H}_{(n)}$ corresponds to the graded algebra T appearing in regularity structures.
- Branched rough path $\mathbf{X}$ corresponds to models.
- Controlled rough path $\mathbf{Y}$ corresponds to modelled distributions.
- The map $\mathbf{Y}_s \mapsto (R_{st}^{\mathbf{Y},\tau})_\tau - \mathbf{Y}_t$ is the change of base point map Γ_{st}.
- The sewing lemma (and consequently rough integration) corresponds to the reconstruction theorem.